Unsafe Daughters: Sexual Violence in Public Spheres and Intrahousehold Stated Preference for Sons *

Tista Mukherjee † Sukanta Bhattacharya‡ Ishita Mukhopadhyay §

Abstract

Increasing non-partner sexual violence against young women is a serious concern among parents with girl children. Against this backdrop, we provide theory and evidence on how intrahousehold reported preference for sons varies in response to the increasing risk of sexual harassment against women in public spheres. We argue that the lack of female safety in a locality increases both implicit and explicit costs of raising girls shifting parental preference in favor of boys. Using a cost-benefit framework, we first show that intrahousehold son preference is formed when girls face a higher risk of non-partner sexual violence. To estimate the effect, we matched self-reported fertility preference data from women who are in their child-bearing age to the administrative sexual crime records at the district and state level from India. Employing state-level proportions of men above the minimum legal drinking age (MLDA) as an instrumental variable (IV) for sexual crimes, we find that the lack of

*We are indebted to Joseph Flavian Gomes, Krisztina Kis-Katos, Sebastian Vollmer, Marcela Ibañez Diaz, Andreas Fuchs, Ritwik Banerjee, Somdeep Chatterjee, Manisha Chakrabarty, Anirban Mukherjee for their insightful comments and suggestions on the earlier drafts. We also thank participants at the Shadow Economics Conference,Tallinn; Development Economics Seminar, University of Göttingen; Departmental Seminar, University of Calcutta; Research Scholars' Workshop, Jadavpur University; Winter School, Delhi School of Economics; Econometric Society Asia Meeting, Indian Institute of Technology, Bombay. All remaining errors are our own. Declarations of interest: none.

†Corresponding author; Faculty of Business and Economics, University of Göttingen; Postal Address: Platz der Göttinger Sieben 3, 37073, Göttingen, Germany; Email: tista.mukherjee@wiwi.uni-goettingen.de

‡Department of Economics, University of Calcutta; Email: sbheco@caluniv.ac.in

§Department of Economics, University of Calcutta; Email: imeco@caluniv.ac.in

female safety outside home, significantly increases women's stated preference for sons. We define difference-in-differences (DID) and regression discontinuity (RD) specifications as complementary estimation strategies, and document that a plausibly exogenous gang rape case in Kolkata (2012) had a positive causal impact on the stated son preference among young women from West Bengal. Ensuring women's safety in public spheres thus appears as a crucial policy choice to reduce gender bias in preferences within households.

Keywords: Sexual Violence, Safety, Stated Son Preference, Instrumental Variable, Difference-in-Differences, Regression Discontinuity, India

JEL Codes: D13, D91, J13, J16, J18, O12

1 Introduction

For more than a hundred years, son preference has been an endemic phenomenon in many societies. Despite rapid modernization and economic growth, the desire to have sons is persistent in countries like India, China, and South Korea (Das Gupta et al., 2003). Such favoritism towards males is predominantly explained by the deep-rooted and sticky cultural forces that shape strong disincentive to raise daughters and motivate parents to favor sons (Das Gupta et al., 2003; Jayachandran, 2015) [1]. For example, many cultures practice patrilocality in which a married woman ceases to be a member of the natal family and joins her husband's family. Under such a system, where parents coreside with sons, returns to investments in sons are higher (Jayachandran, 2015). Ebenstein (2014) shows that the male-to-female sex ratios are positively correlated with the rates of coresidence between adult sons and their parents both across and within countries. Similarly, dowry, a transfer of parental property at the marriage of a daughter is a widely practiced custom. Dowry seems to be a financial burden on a girl's parents and is often cited as a motivation for son preference (Arnold et al., 1998; Bhalotra et al., 2020b; Das Gupta et al., 2003; Jayachandran, 2015). Until recently, evidence on the impact of dowry on son preference was primarily anecdotal. A recent study by Bhalotra et al. (2020b) established the causal impact of dowry on son-preferring behavior in India. Using variation in gold price that acts as an exogenous shock to the burden of dowry, the study finds an increase in girl relative to boy mortality in the neonatal and infant periods. In this study, we explore the lack of female safety outside home, another burden on households with girls but largely overlooked in the existing literature, as an explanation for son preference. Sexual violence (SV) against women is a manifestation of patriarchal norms that assign lower status to women compared to men (Jayachandran, 2021). Such incidents have now become a daily reality for many young women around the world. A systematic meta-analysis reveals that in 2010,

[1] Plausible explanations include patrilocality, patrilineality, dowry, funeral procedures, desire to protect women's chastity, old age security etc.

7·2 percent of women worldwide had ever experienced non-partner sexual violence (Abrahams et al., 2014). India in particular, has experienced an alarming increase in the reported rapes over the last few years (Iyer et al., 2012). The infamous *Kathua* rape case of 2018 that brutally shattered the life of an eight-year-old minor girl from Jammu was just one of many examples that sparked widespread anger and made headlines in domestic and international press (Livemint, 2023). In a survey of young women and girls in Delhi, nearly 75 percent reported having experienced some form of sexual violence in their neighborhoods (Women and ICRW, 2013). Increasing sexual harassment against young women is a cause of concern among parents and pervasive fear of rape is cited as a cause of parents' daughter aversion (Deeksha, 2020). In this study, we provide theory and evidence that non-partner sexual crimes against women in public spaces motivate son preference within families. Conceptually, the increasing risk of sexual violence should increase perceived as well as the actual cost of raising a girl, particularly in a setting where the purity of a woman's body is highly prized. For instance, with the increasing cases of sexual harassment, parents typically endure significant psychological costs such as anxiety, fear etc. Notably, in India, deep bias against women's sexual rights results in immense shame for a rape victim's family. A large number of cases go unreported due to the lack of executive will, dismal conviction rates, fears of police harassment, and social ostracism (Ghosh, 2013). In the presence of ineffective enforcement mechanisms as well as high reputation costs associated with SV, it is plausible that parents need to incur several economic costs to protect girl children. Hence in an unsafe location for girls, the safety cost for a boy is far lower and it may motivate gender bias in preferences for children within households.

The theoretical background of our conjecture is based on a simple cost-benefit framework where parents optimally choose the preferred numbers of boys and girls. They consider both pecuniary and non-pecuniary benefits of children. Children's identity-neutral costs include the efforts that parents exert to raise them. On the other hand, there exist certain women-

specific safety costs since girl children are vulnerable to sexual harassment. These costs typically include supervision costs to protect the female children and psychological costs such as fear and trauma stemming from the stigma that a conservative society attaches to such incidents. Certainly, these costs are higher in societies where women's chastity is highly prized. Using this simple structure, we show that son preference is formed when girl children are vulnerable to sexual violence.

Using data from India, we then empirically estimate the role of local SV in shaping son preference within households. To establish the quantitative link between SV and stated son preference (SSP), we use data from various sources. First, we collected individual-level data on SSP from the nationally representative Indian Human Development Survey (IHDS) (2011-12). Further, we collected district and state-level records on the previous year's reported sexual violence against women from official crime statistics published by the National Crime Records Bureau (NCRB) (2010). For the estimates of the female population and all other confounding factors, we use data from the nearest Census (2011). In line with our prediction, baseline findings suggest that there exists a significant positive association between locality-specific SV and intrafamily SSP. More specifically, we find that an additional case of SV per thousand women in a district is associated with a 3 percentage point increase in a woman's SSP. The corresponding estimate for state-level sexual crimes is 6 percentage points. To investigate the underlying mechanism, we estimate the association between local SV and the absolute quantity of desired daughters. Aligned with our conjecture, the results show that an increase in the number of sexual crimes reduces the absolute number of preferred girls. We also collect data on gender-neutral crimes from NCRB (2010) and conduct various placebo checks that lend support for our conjecture that the positive association between SV and SSP is primarily driven by the women-specific cost of SV. To address the issues of measurement error and unobserved heterogeneity at the collective level, we exploit minimum legal drinking age (MLDA) across states as a plausible source of quasi-experimental variation in SV against women and conduct an Instrumental Variable (IV) analysis. The IV estimators

further strengthen the positive association between the two. Next, we exploit variations in enumeration dates and the states of residence for all participants of IHDS (2011-12) and employ a difference-in differences (DID) framework to estimate the causal impact of an infamous gang rape case in Kolkata (2012). Estimates from the alternative strategy is similar and suggest that the event significantly increased reported son preference among women from West Bengal. Moreover, the implications are similar when we employed regression discontinuity (RD) approach to rule out the possibility of estimation-bias due to the presence of spatial spillovers.

Our study adds to the limited literature that establishes the theoretical as well as the quantitative association between the risk of non-partner sexual violence and the behavioral responses within households (Sarkar, 2021; Song et al., 2022; Theerthaana and Sheik Manzoor, 2019). There exists some anecdotal evidence on the role of intention to protect daughters' chastity in determining son preference (Arnold et al., 1998; Robitaille, 2013). The key argument is that the parents of sons do not have to bear the cost of chaperonage (Arnold et al., 1998). A recent study by Sarkar (2021) shows that living in a locality with a high perceived risk of sexual violence against girls is associated with a higher probability of early marriage among young women. The study also investigates the mechanisms and finds that a strong incentive to protect daughters' chastity acts as the key driver of such sub-optimal outcome within the family. Only a few studies to the best of our knowledge have attempted to investigate the link between sexual harassment against women and son or daughter preference within families (Song et al., 2022; Theerthaana and Sheik Manzoor, 2019). For instance, Theerthaana and Sheik Manzoor (2019) found that safety concerns are associated to gender disappointment with girl children in India. Interestingly a recent study by Song et al. (2022) found a negative association between perceived risk of female safety and SSP in the Hubei province of China. Further, the study finds that the values of sons are reinforced by such risk perception in a gendered imbalanced society which largely outweighs the negative association between female safety

concerns and SSP.

Our study broadly relates to different strands of literature. First, it adds to the existing knowledge on the factors determining SSP (Asadullah et al., 2021; Chung and Gupta, 2007; Gaudin, 2011; Koolwal, 2007; Mavisakalyan and Minasyan, 2023; Mughal et al., 2023; Pande and Astone, 2007; Robitaille, 2013, 2020; Song et al., 2022). Earlier studies have shown that SSP is determined by the women's characteristics such as education, access to media, access to credit, sex-composition of the existing children, age at marriage (Chung and Gupta, 2007; Mughal et al., 2023; Pande and Astone, 2007; Robitaille, 2013); household's characteristics including family structure, caste, religion, and wealth (Gaudin, 2011; Pande and Astone, 2007). At a more aggregate level, cultural norms such as dowry, patrilocality, old age security, women's status in the locality, and the extent of urbanization are some well-documented forces that shape son preference (Chung and Gupta, 2007; Pande and Astone, 2007; Robitaille, 2020). Son bias can also be a behavioral consequence of conflict. Mavisakalyan and Minasyan (2023) for example, provide evidence that the fear of conflict is associated with a greater preference for sons as males are traditionally viewed as defenders. On the contrary, Koolwal (2007) presents evidence that the increased earnings opportunities for girls can lower household preference for sons, as measured by the household's average reported ideal number of sons relative to ideal number of children. Reported son preference has a sizable impact on the intrafamily resource allocation among children [2] and infant girl mortality within households [3] (Clark, 2000; Ebert and Vollmer, 2022; Lin and Adserà, 2013; Robitaille and Chatterjee, 2018). We also add to the growing literature highlighting the adverse consequences of non-partner sexual crimes on several dimensions of women's socio-economic status (Borker, 2017; Chakraborty et al., 2018; Chakraborty and Lohawala, 2021; Sarkar, 2021). For instance, Borker (2017) examines the

[2]Lin and Adserà (2013) show that a mother's reported son preference correlates with a larger gender gap in the hours of housework among children in India. Ebert and Vollmer (2022) show that the reported child-specific son preference acts as a penalty in early mental functions for unwanted girls.

[3]Robitaille and Chatterjee (2018) show that parental son preference is associated with higher sex-selective abortions and infant mortality of girl children in India.

impact of the perceived risk of street harassment on women's choice of college, and finds that the concern for safety translates into 20 percent lower expected post-college earnings in India. Similarly, both actual and perceived sexual crimes against women have a deterrence impact on women's labor force participation in India (Chakraborty et al., 2018; Chakraborty and Lohawala, 2021). In terms of magnitude, this safety penalty is comparable to the motherhood penalty, the most cited factor explaining the stagnancy of women's work (Chakraborty and Lohawala, 2021).

The major contribution of our study is threefold. Firstly, it underlines the role of SV, an external cultural constraint against girls in explaining the disproportionate preference for sons within families. While the existing literature primarily accounts kinship system as the key cultural force disincentivizing parents from raising daughters, we show that the lack of women's safety outside home, another expression of patriarchal norms, imposes a burden on families with girl children and reinforces the traditional institution of son bias. However, the presence of unobserved heterogeneity at the collective level acts as the key identification challenge to establish the one-way causal link between the two. We averted this issue in two ways. First, we exploit a quasi-experimental variation in sexual crimes against women and employ IV estimators. Second, we define a meaningful DID framework to estimate the causal impact of a violent sexual crime that took place in Kolkata (2012). Hence the second major contribution of the study lies in the novelty of our econometric approach. Finally, the study highlights a more straightforward avenue through which policy can reduce gender inequality within households by ensuring women's safety and reducing the cost of raising daughters.

The rest of this paper is organized as follows. The next section presents the analytical framework and our key testable prediction. The empirical evidence on the association between reported cases of SV and SSP is presented in Section 3. Section 4 provides evidence of the causal impact of a violent gang rape case in Kolkata (2012) on the reported son preference. Section 5 concludes.

2 Analytical Framework

To characterize the effect of SV on son preference, we construct a simple one-period household model where parents optimally choose the desired numbers of boys and girls. They receive both pecuniary and non-pecuniary benefits from offspring. On the other hand, parents also incur several costs to raise them. However, there exist some additional costs of raising girls since they are vulnerable to sexual violence in public spaces. Based on these assumptions, the net utility from children is defined as follows:

$$U(n_m, n_f, \theta) = n^\beta + (\alpha_m n_m + \alpha_f n_f)y - [\frac{c}{2}(n_m^2 + \theta n_f^2) + \phi n_m n_f] \tag{1}$$

where n_m and n_f are the numbers of desired sons and daughters respectively. n^β represents the expected intrinsic benefit from children. Such payoff is child's identity-neutral and typically originates from the emotional satisfaction parents receive from children. However, we assume $\beta < 1$ implying that the non-pecuniary benefit is decreasing in the number of children. The argument is straightforward. The first child delivers the highest emotional satisfaction (Bulatao, 1981). The term $(\alpha_m n_m + \alpha_f n_f)y$ captures the material payoff from offspring. Let α_m (α_f) be the share of a boy's (girl's) potential income (y) that parents expect to receive. Here the child may provide labor in the family business or can participate in the formal labor market. We assume $\alpha_m > \alpha_f$, since in a patriarchal society, parents expect to live with and receive financial support from sons. $C > 0$ denotes the gender-neutral expected costs of raising children such as financial expenditures, opportunity costs from cutting down labor market opportunities and leisure hours to provide child-care etc. Additionally, there are certain safety costs denoted by $\theta > 1$, that parents incur for daughters since young girls are vulnerable to the risk of sexual harassment. Such costs include potential expenditures on supervision and the fear associated with such events. A victim's family is highly likely to be socially ostracized, particularly in a society that attaches more weight to women's chastity. Similar to Becker and Lewis (1973), we assume interdependent marginal cost which means

that the additional cost of a boy is increasing in the number of girls and vice versa. The term $\phi < C$ captures this cost.

Now the F.O.Cs of parents' optimization problem are given by the following expressions:

$$\frac{\delta U}{\delta n_m} = \left(\beta n^{\beta-1} + \alpha_m y\right) - \left(C n_m + \phi n_f\right) \leq 0 \tag{2}$$

$$\frac{\delta U}{\delta n_f} = \left(\beta n^{\beta-1} + \alpha_f y\right) - \left(C\theta n_f + \phi n_m\right) \leq 0 \tag{3}$$

with the equality holding for $n_m^* > 0$ and $n_f^* > 0$ respectively. The first and the second bracketed terms in each expression represent the marginal benefit and marginal cost of having an extra child of a particular kind. Notice that the marginal benefit from a son is always higher than that of a daughter. Also notice that $n_f^* = n_m^* = 0$ can not be the solution to the parents' problem since the marginal benefit of having any child is then infinitely large given $\beta < 1$. At the optimum, if $n_f^* > 0$ and $n_m^* = 0$, $\frac{\delta U}{\delta n_m} > \frac{\delta U}{\delta n_f} = 0$ since $\alpha_m > \alpha_f$ and $C\theta > \phi$. Hence, at the optimum, $n_m^* > 0$. Now if $n_f^* = 0$, it must be the case that

$$\frac{\delta U}{\delta n_m} = \beta {n_m^*}^{\beta-1} + \alpha_m y - C n_m^* = 0$$

and

$$\frac{\delta U}{\delta n_f} = \beta {n_m^*}^{\beta-1} + \alpha_f y - \phi n_m^* \leq 0$$

which together imply

$$n_m^* \leq \frac{(\alpha_m - \alpha_f) y}{C - \phi}$$

Hence, at the optimum $n_f^* = 0$ (along with $n_m^* > 0$) if and only if

$$\beta \left[\frac{(\alpha_m - \alpha_f) y}{C - \phi}\right]^{\beta-1} \leq \frac{(\phi\alpha_m - C\alpha_f)\, y}{C - \phi}$$

Thus for $n_f^* > 0$, the parameters need to satisfy the following condition which we call

Condition A:

$$\beta\left[\frac{(\alpha_m-\alpha_f)y}{C-\phi}\right]^{\beta-1} > \frac{(\phi\alpha_m - C\alpha_f)\,y}{C-\phi} \tag{4}$$

We are interested in the case $n_f^* > 0$. For the rest of this section, we assume that the parameters are such that (4) is satisfied.

For $n_f^* > 0, n_m^* > 0$, we must have

$$n_m^* = \frac{C\theta-\phi}{C-\phi}n_f^* + \frac{(\alpha_m-\alpha_f)y}{C-\phi} \tag{5}$$

$$n_f^* = \frac{C-\phi}{C\theta-\phi}n_m^* - \frac{(\alpha_m-\alpha_f)y}{C\theta-\phi} \tag{6}$$

Now substituting the value of ${n^*}_m$ from Equation 5 in $n = n_m + n_f$, we can derive $n^* = f(n_f)$. This gives

$$n^* = \frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m-\alpha_f)y}{C-\phi} \tag{7}$$

Further, by substituting the values of n^* from Equation 7, and ${n^*}_m$ from Equation 5 in Equation 6, we can derive ${n^*}_f = g(\theta)$, where θ is our key parameter of interest from the following expression:

$$\beta[\frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m-\alpha_f)y}{C-\phi}]^{\beta-1} + \alpha_f y = \frac{C^2\theta-\phi^2}{C-\phi}n_f + \phi\frac{(\alpha_m-\alpha_f)y}{C-\phi}$$

$$\Leftrightarrow \beta[\frac{C(1+\theta)-2\phi}{C-\phi}n_f^* + \frac{(\alpha_m-\alpha_f)y}{C-\phi}]^{\beta-1} = \frac{C^2\theta-\phi^2}{C-\phi}n_f^* + \frac{(\phi\alpha_m - C\alpha_f)y}{C-\phi}$$

Differentiating this expression w.r.t. θ will give our first prediction relating to the effect of SV on the preferred quantity of daughters (${n^*}_f$).

$$\frac{\delta {n^*}_f}{\delta \theta} = -Cn_f[\frac{C + \beta(1-\beta)[\frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m - \alpha_f)y}{C-\phi}]^{\beta-2}}{(C^2\theta - \phi^2) + [C(1+\theta) - 2\phi]\beta(1-\beta)[\frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m - \alpha_f)y}{C-\phi}]^{\beta-2}}] \tag{8}$$

Recalling the key parameter restrictions ($\beta < 1, C > \phi, \theta > 1, \alpha_m > \alpha_f$) we can claim that the $R.H.S < 0$. Hence our first comparative static result is along the expected line. An increase in the risk of SV against girls increases the marginal cost of raising daughters and subsequently reduces the optimal quantity of preferred girls. Lemma 1 describes this scenario.

Lemma 1 *Suppose Condition A holds. Then, n_f^* is strictly decreasing in θ.*

Now we can prove our main conjecture related to the marginal effect of SV on son preference with the following simple steps. First, recall the expression of ${n^*}_m$ from Equation 5. Differentiating both sides w.r.t. θ, we get:

$$\frac{\delta {n^*}_m}{\delta \theta} = \frac{c\theta - \phi}{C - \phi}[\frac{\delta {n^*}_f}{\delta \theta}] + \frac{Cn_f}{C - \phi} \tag{9}$$

Rearranging this expression, we can write

$$\frac{\delta {n^*}_m}{\delta \theta} > 0 \leftrightarrow \frac{\delta {n^*}_f}{\delta \theta} > -\frac{Cn_f}{C\theta - \phi} \tag{10}$$

Now substituting the value of $\frac{\delta {n^*}_f}{\delta \theta}$ from Equation 8 in the above expression, and simplifying, we can derive our key prediction related to the risk of SV and son preference. Let us first write the precise expression.

$$-Cn_f[\frac{C + \beta(1-\beta)[\frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m - \alpha_f)y}{C-\phi}]^{\beta-2}}{(C^2\theta - \phi^2) + [C(1+\theta) - 2\phi]\beta(1-\beta)[\frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m - \alpha_f)y}{C-\phi}]^{\beta-2}}] > -\frac{Cn_f}{C\theta - \phi}$$

$$\implies$$

$$\frac{C+\beta(1-\beta)[Z]^{\beta-2}}{(C^2\theta-\phi^2)+[C(1+\theta)-2\phi]\beta(1-\beta)[Z]^{\beta-2}} < \frac{1}{C\theta-\phi}$$

where $Z = \frac{C(1+\theta)-2\phi}{C-\phi}n_f + \frac{(\alpha_m-\alpha_f)y}{C-\phi} > 0$. Further simplifying and rearranging, we get

$$[C^2\theta - C\phi] + [C\theta-\phi]\beta(1-\beta)Z^{(\beta-2)} < [C^2\theta-\phi^2] + [C(1+\theta)-2\phi]\beta(1-\beta)Z^{(\beta-2)} \quad (11)$$

Now given $C > \phi$, we can claim that $C^2\theta - C\phi < C^2\theta - \phi^2$ and $C\theta - \phi < C(1+\theta) - 2\phi$. Hence the above inequality holds with certainty. This gives our second comparative static result. The marginal effect of SV on the optimal quantity of preferred boys is positive and we describe this in lemma 2.

Lemma 2 *Suppose Condition A holds. Then, n_m^* is strictly increasing in θ.*

We can now write the precise expression for stated son preference which is the ratio of the number of desired sons to the total preferred children: $\frac{n^*{}_m}{n^*} = \frac{n^*{}_m}{n^*{}_m + n^*{}_f} = \frac{1}{\frac{n^*{}_f}{n^*{}_m}+1}$. From the above two lemmas, we can easily claim that this expression is strictly rising in θ. This gives our key proposition:

Proposition 1 *The increasing risk of sexual violence against girls shapes parental preference for sons, i.e. $\frac{\delta}{\delta\theta}\left[\frac{n_m^*}{n^*}\right] > 0$.*

The results are along the expected line. If everything else remains the same, a marginal increase in the risk of sexual violence against girls increases the cost of raising daughters and shapes preference in favor of boys. Against the backdrop of rising sexual crimes against women around the world, our model generates some testable implications on how intrafamily preference for children responds to such events.

3 Empirical Evidence: Reported Sexual Crimes and SSP

3.1 Data Sources

To test the key prediction, we compile data from various sources. The data on individual-level SSP are collected from the Indian Human Development Survey (2011-12). It is a nationally representative, multi-topic survey of $42,152$ households from 384 districts, 1420 villages, and 1042 urban neighborhoods across India. It covers all states and union territories of the country except the two: Andaman/Nicobar and Lakshadweep (Desai and Vanneman, 2015). In particular, we exploit the women's schedule that interviews all married women, in each household and captures detailed information on their socio-economic characteristics, gender relations, marriage history, fertility preferences etc. Data on SV against women are collected from the official statistics on crimes published by the National Crime Records Bureau (NCRB) (2010). We matched individual-level data on reported son preference with the previous year's sexual crime records at the district and state levels. Further, we use district and state-level female population estimates from the nearest Census (2011). Using these data, we calculate female-per-capita SV, which is our key independent variable of interest. We also collected data from the Census (2011) on various district and state-level characteristics such as shares of urban and *Hindu* populations, shares of *Scheduled Caste (SC)* women, female literacy rates, child sex ratios (F/M) as they are used as confounding factors in the regressions. Finally, for the instrumental variable (IV) analysis, we collect state-level data on the minimum legal drinking age (MLDA) from the laws published by state excise departments.

3.2 Sample and Measures

3.2.1 Stated Son Preference (SSP)

Our key outcome of interest, SSP is measured (based on Anukriti et al. (2022); Bhalotra et al. (2020a); Clark (2000); Koolwal (2007); Lin and Adserà (2013); Rastogi and Sharma (2022); Robitaille (2020)) by the ratio of preferred number of sons to the total number of preferred kids if the respondent could start over the fertility history. Values range from 0 to 1. However, it must be noted that the measure is based on the women's answers at the time of the interview. Hence the survey captures the reported fertility preference with the underlying assumption that a woman's response does not vary with her age. To relax such a strong assumption, we restrict our sample only to the women who are in their reproductive years at the time of the survey. Provided that a woman's early reproductive years are around the age of 20, confining our sample to the age cohort around 20 allows us to estimate the effect of local SV on SSP under a much weaker assumption that a woman's reported fertility preference in her reproductive years reflects her preferences around the time she had her first child. To enhance the credibility of our estimates further, we control for the sex-composition of existing children along with other characteristics of the woman in all models. Our final analytical sample consists of 12,102 married women in the age range of 18 to 32 years who are in their childbearing years during the survey. We exclude observations from a few states such as Sikkim, Arunachal Pradesh, Nagaland, Mizoram, Manipur, and Tripura due to the following reasons: first, IHDS does not identify districts for these states; second, small sample sizes from these states make the data non-representative at the state level. However, it is worth noting that the total observations from these states constitute less than one percent of the sample of married women in the same age range. The average ideal proportion of sons in our sample is .55 reflecting the presence of son bias within a household [4]. The data from the Demographic and Health Survey (DHS) shows a roughly similar distribution of SSP among women from India (Bhalotra et al., 2020a). Figure 1, presents the distribution of the

[4]Table A1 in the appendix reports the summary of women's characteristics.

absolute preference for children by households' perceived fear of sexual harassment in the neighborhood. A quick look suggests a positive correlation between fear of SV and the desire for boys. More precisely, households that report frequent sexual crimes against unmarried girls within the locality also express a greater desire for sons. We further exploit a much cleaner measure of reported son preference as an alternative outcome variable. IHDS asks each woman about the gender preference for an additional child. We exclude those with no gender bias in preference for the extra child and construct a binary indicator that measures whether a woman's preferred additional child would be a boy. Among 4,309 respondents within this sub-sample, 80 percent state that the preferred extra child would be a boy.

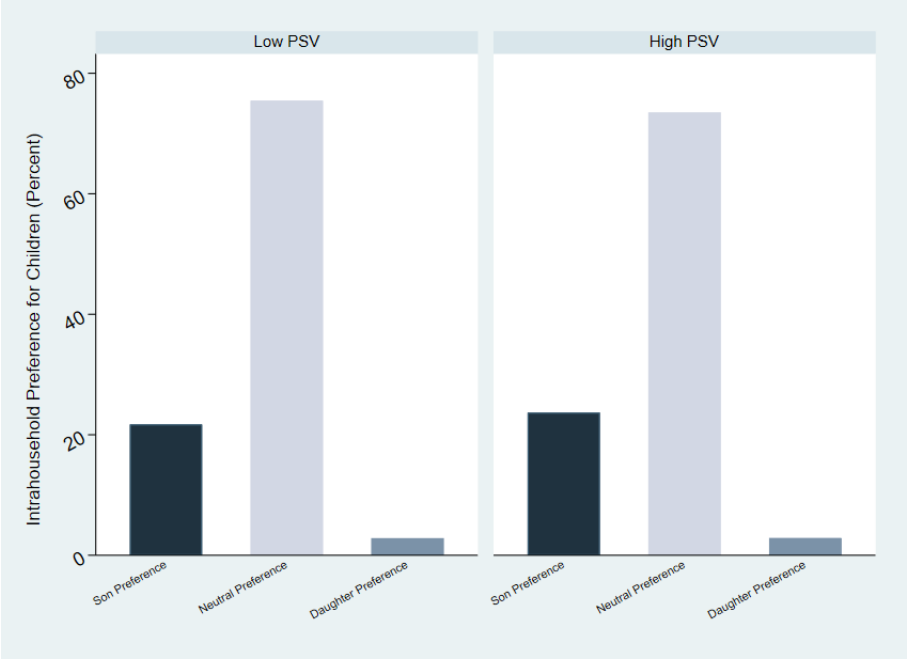


Figure 1: Perceived Risk of SV and Reported Son Preference

Source: Own calculations from IHDS (2011-12). The figure presents the distribution of the absolute reported preference for sons and daughters among Indian married women by the family's perceived sexual violence (PSV) within locality. Son(daughter) preference measures whether the preferred number of boys(girls) exceeds that of girls(boys).

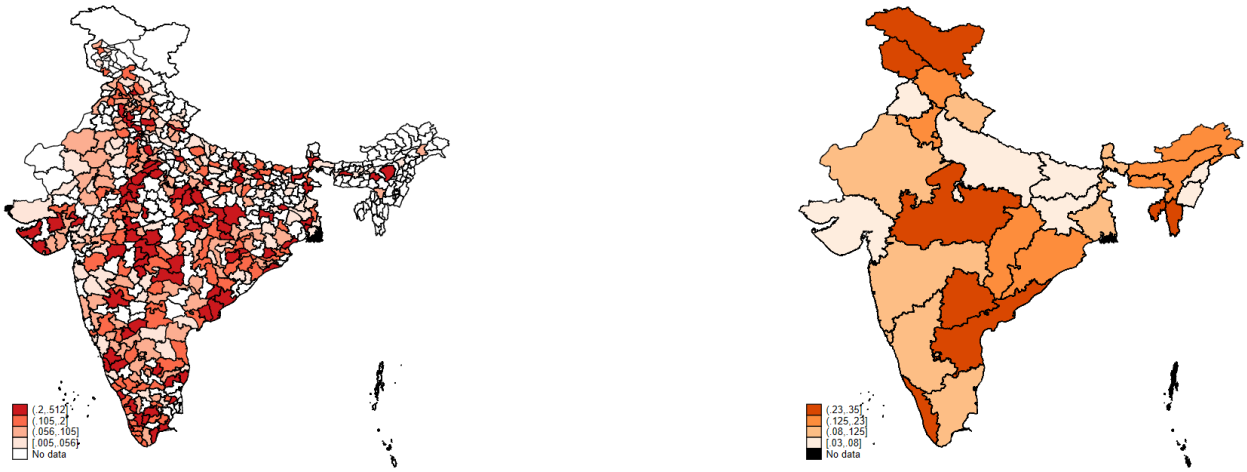
3.2.2 Sexual Violence (SV)

Our key independent variable SV, is the sum of reported rape (R), molestation (M) and sexual harassment (SH) cases per thousand women in the district or state recorded in the

previous year (2010). The precise expression of SV is:

$$SV = \frac{R + M + SH}{FP} * 1000$$

where FP represents the district or state-level female population estimates from the closest decennial census (2011). A district on average had around 14 sexual crimes per hundred women. This estimate is consistent with Chakraborty and Lohawala (2021). Figure 2 shows the distribution of SV at the district and state levels [5].



(a) Sexual Violence District Wise

(b) Sexual Violence State Wise

Figure 2: Sexual Violence in India

Source: Own calculations from NCRB (2010) and Census (2011). Panel (a) and (b) present district and state-wise sexual crimes (rape, molestation, and sexual harassment) per thousand women.

3.3 Estimation

Our baseline specification to assess the association between SV in the district and SSP is as follows:

$$Y_{ihd} = \beta_0 + SV_d\beta_1 + \beta_2 X'_{ihd} + \gamma_r + \epsilon_{ihd} \tag{12}$$

[5]Table A2 in appendix reports the summaries of district and state characteristics.

where Y_{ihd} is the SSP of a woman i, belonging to the household h in the district d. SV_d denotes sexual crimes against women in the district d. The parameter of interest β_1 gives us the association between SV in the district and a woman's SSP. The vector X'_{ihd} includes individual, household, and district-level confounding factors. More specifically, we control for district-level female literacy rates, shares of urban and *Hindu* populations, child sex ratios (F/M), and the shares of *Scheduled caste (SC)* women. Individual-level controls include the respondent's age, years of education, second-order polynomials of age and education, total number of children, sex-composition of existing children (whether the respondent has more sons than daughters), employment status, age at marriage, son dependence (whether the respondent expects to receive financial support from son), relationship with spouse (whether household decisions are taken jointly), and perceived spousal violence in the community (whether wife-beating is common within community if a woman neglects the home or the children). Household-level controls [6] include the family's religion (whether *Hindu*), caste (whether Upper-Caste), wealth measured by the asset index (based on principal component score of household assets), access to media (whether female members watch television regularly), location of residence (rural/urban), highest education (in years), and the number of adult members. γ_r reflects the region-fixed effects (North, North-West, West, East, Central-North, Central-South, South) capturing all time-invariant unobserved characteristics across regions that jointly determine the desire for sons and sexual crimes against women. The Estimation strategy is OLS. Robust standard errors in all models are clustered across districts.

The following equation shows the augmented specification that estimates the association between SV in the state and SSP:

$$Y_{ihs} = \alpha_0 + SV_s\alpha_1 + \alpha_2 X'_{ihs} + \gamma_r + \epsilon_{ihs} \tag{13}$$

where Y_{ihs} is the SSP of a woman i, belonging to the household h in the state s. SV_s denotes

[6] A summary of household characteristics is reported in Table A1 in the appendix.

sexual crimes against women in the state s. The parameter of interest α_1 gives us the quantitative link between SV in the state and the woman's SSP after controlling all state, individual, household factors, and the region-fixed effects. The estimation strategy is OLS and all robust standard errors are clustered across states.

3.4 Results: Baseline

Estimates from the baseline specifications are reported in Table 1. Panel A presents the association between SV at the district level and a woman's SSP. Column 1 shows the relationship between the two variables of interest without controlling other factors except the geographic fixed effects. The result is along the expected line and the coefficient of SV is positive and statistically significant. Columns 2 and 3 report the coefficients after incorporating district and individual-level controls respectively. With the inclusion of these factors, the coefficients of SV are marginally higher in magnitude. Finally, column 4 gives us the estimates from the full model after incorporating household characteristics. The coefficient of SV is positive significant and suggests that an additional case of SV per thousand women in the district is associated with a 3 percentage point increase in a woman's SSP. The relationship between SV at the state level and SSP is reported in panel B. The coefficient of SV in column 4 corresponds to the full model and suggests that an additional case of SV per thousand women in the state is associated with a 6 percentage point increase in a woman's SSP. We further conduct some sub-sample analyses to estimate the association across different groups. First, we report the estimated coefficient for less wealthy households [7] (Column 5). The implication is qualitatively similar and the coefficient suggests that an increase in SV in a district or state significantly increases reported son bias among poorer households. Similarly, we find a significant positive association between SV and SSP among the lower-caste women (Column 6).

[7]Economic status is measured by whether the household lies below the poverty line.

	Dependent Variable: SSP (Desired Son Ratio)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	*+Dist/State*	*+Ind*	*+HH*	*Poor*	*Lower Caste*
	Panel A: District-Level					
SV	.015	.024*	.031**	.030**	.049*	.024*
	(.014)	(.013)	(.014)	(.014)	(.027)	(.013)
Constant	.541***	.550***	.478***	.491***	.511***	.528***
	(.005)	(.017)	(.042)	(.043)	(.048)	(.050)
N	12,102	12,102	11,320	11,124	2,496	8,358
R Squared	.021	.030	.056	.063	.051	.054
	Panel B: State-Level					
SV	.070**	.062***	.062**	.060**	.050*	.102***
	(.032)	(.019)	(.023)	(.025)	(.026)	(.030)
Constant	.534***	.583***	.529***	.539***	.606***	.680**
	(.006)	(.138)	(.116)	(.156)	(.119)	(.156)
N	12,102	12,102	11,320	11,124	2,496	8,358
R Squared	.022	.026	.056	.057	.050	.055
Geographic Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Dist/State Controls	No	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
HH Controls	No	No	No	Yes	Yes	Yes

Table 1: Baseline: SV and SSP

Source: Own calculations from IHDS (2011-12) and NCRB (2010).

Notes: OLS estimation. Sample consists of married women in the age range of 18 to 32. Figures in parentheses are robust standard errors clustered at the district/state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

3.5 Robustness

3.5.1 Alternative Measure of SSP

The estimates from the baseline specifications corresponding to our alternative measure are reported in Table 2. We employ linear probability models to estimate the association between SV and SSP. Using step-wise regressions, panel A presents the association between SV in the district and the likelihood that a woman's gender preference for the extra child is a boy. Column 4 reports the results using the full set of controls. The estimated coefficient suggests that an additional incident of SV per thousand women in a district is associated with a 39 percentage point increase in the probability that a woman's preferred additional child would be a boy. The estimates for state-level sexual crimes are reported in panel B. The first row of column 4, suggests that one additional incident of SV per thousand women in a state is associated with a 70 percentage point increase in the probability that a woman's preferred additional child would be a boy. In each panel, columns 5 and 6 report the results from the sub-sample analyses. All estimates are aligned with the results from subsection 3.4, and suggest that increasing SV in a district or state is associated with higher reported son preference among the poorer and lower-caste women (Columns 7,8).

3.5.2 SV and Daughter Preference

Our analytical framework suggests that increasing SV acts as a burden on families with female children, and hence reduces the optimal number of preferred daughters [8]. To exploit this empirically, we regress the number of preferred daughters on district and state-specific sexual offenses using all controls specified in subsection 3.3. Table 3 reports the OLS estimates. Panel A and B report the results w.r.t. the district and state-level crimes respectively. In each panel, column 1 includes no controls except the geographic fixed effects. Signs of the coefficients of SV are along the expected line. Columns 2 and 3 sequentially add district or state characteristics and individual factors respectively. With

[8] See Lemma 1. The proof is given in Equation 8.

	Dependent Variable: Gender Preference for Additional Child (Boy)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	*+Dist/State*	*+Ind*	*+HH*	*Poor*	*Lower Caste*
	Panel A: District-Level					
SV	.254**	.364***	.355**	.387***	.205	.239*
	(.106)	(.105)	(.145)	(.103)	(.163)	(.119)
Constant	.541***	2.39**	2.40***	2.70**	3.73**	2.46*
	(.082)	(1.03)	(1.15)	(1.13)	(1.31)	(1.34)
N	4,309	4,309	4,055	3,980	1,000	2,960
R Squared	.102	.112	.145	.173	.182	.215
	Panel B: State-Level					
SV	.332**	.658***	.680***	.695***	.670***	.602***
	(.132)	(.116)	(.108)	(.106)	(.116)	(.132)
Constant	.378***	3.12**	4.58***	4.24***	4.84***	5.21**
	(.116)	(1.56)	(1.68)	(1.72)	(2.00)	(2.09)
N	4,309	4,309	4,055	3,980	1,000	2,960
R Squared	.076	.112	.156	.190	.167	.209
Geographic Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Dist/State Controls	No	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
HH Controls	No	No	No	Yes	Yes	Yes

Table 2: Robustness: Alternative Measure of SSP

Source: Own calculations from IHDS (2011-12) and NCRB (2010).

Notes: Linear probability models. Sample consists of married women in the age range of 18 to 32 excluding those who reported a gender-neutral preference for an extra child. Figures in parentheses are robust standard errors clustered at the district/ state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

these inclusions, the coefficients of SV become statistically significant. Finally, column 4 in each panel shows the estimates from the full model after including household controls. Results suggest that increasing SV in a district or state has a significant deterrence effect on a woman's absolute daughter preference. These findings corroborate our key testable prediction suggesting that SV in a locality increases the cost of raising girls which translates into a reduction in preference for daughters.

3.5.3 Placebo Checks

From our theoretical standpoint, the positive association between SV and SSP is led by the increasing safety costs for girls. Hence, gender-neutral violent crimes such as murder, kidnapping etc. are unlikely to drive such response since those offenses are not supposed to impose women-specific costs on the family. To examine this, we conduct various placebo tests and re-estimate the baseline specifications using the above-mentioned crimes as the regressors. Table 4 reports the OLS estimates. Panel A and B report the results for murder and kidnapping respectively. Column 1 includes no controls except the geographic fixed effects. In columns 2 and 3, we sequentially add district or state-level characteristics and individual factors respectively. Column 4 shows the results from the full specification after adding household characteristics. We found no significant association between gender-neutral violent crimes such as murder or kidnapping and reported son preference. Further, we regress SSP on property crimes such as stealing to examine whether the positive association between SV and SSP is explained by the argument of protection. The key idea is that if SV is supposed to increase the value of sons since males are traditionally considered as the defenders, we should expect a similar association between son bias and other property crimes in the locality that raise a demand for protecting resources within households. Panel C of Table 4 reveals no such association which rules out the argument related to protection and strengthens the validity of our conjecture that the risk of SV is associated with safety costs for daughters motivating son-preferring behavior within household.

	Dependent Variable: No of Preferred Daughters					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	*+Dist/State*	*+Ind*	*+HH*	*Poor*	*Lower Caste*
	Panel A: District-Level					
SV	-.124	-.180*	-.174**	-.189**	-.168**	-.230**
	(.085)	(.095)	(.086)	(.085)	(.064)	(.099)
Constant	1.07***	1.13***	1.40***	1.39***	1.11***	1.23***
	(.028)	(.110)	(.104)	(.106)	(.199)	(.137)
N	12,138	12,138	11,349	11,153	2,502	8,381
R Squared	.023	.027	.131	.150	.101	.154
	Panel B: State-Level					
SV	-.153	-.334***	-.338***	-.336***	-.419***	-.230**
	(.143)	(.085)	(.086)	(.099)	(.100)	(.118)
Constant	1.07***	1.14***	1.25***	1.28***	.993***	1.13***
	(.038)	(.276)	(.267)	(.228)	(.265)	(.198)
N	12,138	12,138	11,349	11,153	2,502	8,381
R Squared	.023	.031	.139	.152	.105	.159
Geographic Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Dist/State Controls	No	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
HH Controls	No	No	No	Yes	Yes	Yes

Table 3: SV and Daughter Preference

Source: Own calculations from IHDS (2011-12) and NCRB (2010).

Notes: OLS estimation. Sample consists of married women in the age range of 18 to 32. Figures in parentheses are robust standard errors clustered at the district/ state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

	Dependent Variable: SSP (Desired Son Ratio)							
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
	Baseline	*+Dist*	*+Ind*	*+HH*	*Baseline*	*+State*	*+Ind*	*+HH*
Panel A: Murder	*District*				*State*			
Murder Rate	.198	.201	.234	.199	.308	.142	.194	.230
	(.226)	(.231)	(.267)	(.291)	(.380)	(.288)	(.280)	(.284)
Constant	.441***	.417***	.457***	.438***	.541***	.614***	.552***	.565***
	(.007)	(.010)	(.126)	(.129)	(.005)	(.115)	(.120)	(.123)
N	12,102	12,102	11,320	11,124	12,102	12,102	11,320	11,124
R Squared	.026	.030	.043	.050	.021	.034	.056	.056
Panel B: Kidnapping	*District*				*State*			
Kidnapping Rate	020	.067	.108	.094	-.061**	.041	.021	.013
	(.035)	(.112)	(.124)	(.147)	(.026)	(.040)	(.039)	(.041)
Constant	.446***	.471***	.462***	.451***	.546***	.576***	.557***	.511***
	(.090)	(.124)	(.125)	(.138)	(.023)	(.117)	(.126)	(.132)
N	12,102	12,102	11,320	11,124	12,102	12,102	11,320	11,124
R Squared	.020	.032	.049	.055	.022	.048	.056	.057
Panel C: Thieving	*District*				*State*			
Thieving Rate	-.008	.012	-.113	.143	-.020**	-.004	.021	.005
	(.010)	(.129)	(.136)	(.139)	(.008)	(.016)	(.019)	(.020)
Constant	.411***	.476***	.433***	.502***	.546***	.609***	.562***	.574***
	(.015)	(.124)	(.132)	(.157)	(.023)	(.132)	(.147)	(.153)
N	12,102	12,102	11,320	11,124	12,102	12,102	11,320	11,124
R Squared	.021	.027	.050	.054	.022	.025	.057	.058
Geographic Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dist/State Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes
HH Controls	No	No	No	Yes	Yes	Yes	Yes	Yes

Table 4: Placebo Checks

Source: Own calculations from IHDS (2011-12) and NCRB (2010).

Notes: OLS estimation. Sample consists of married women in the age range of 18 to 32. Figures in parentheses are robust standard errors clustered at the district/ state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

3.5.4 IV Estimation

The key empirical challenge associated with our baseline specifications is the presence of omitted variables at the collective level. For example, pre-existing patriarchal social norms simultaneously determine sexual violence in the locality and desire for boys within households. For instance, northern India follows a more conservative kinship system leading to greater gender inequality in the north (Das Gupta et al., 2003; Dyson and Moore, 1983). On the other hand, level of economic development simultaneously influences the attitude towards gender-based violence and preference for boys. For example, cross-country evidence suggests that the fraction of women believing wife-beating is justified is negatively correlated with per-capita GDP (Jayachandran, 2015). Similarly, Chung and Gupta (2007) showed that the economic development that helped normalize national child sex ratios in South Korea worked largely by driving normative and behavioral changes within the society as a whole rather than just through improving people's socioeconomic circumstances, and increasing women's economic value is a manifestation of changing social norms. To address omission bias, we adopt the instrumental variable (IV) approach. More specifically, we exploit state-level policies governing alcohol access that create a quasi-experimental variation in sexual offenses across states. In India, state-level laws regulating the minimum legal drinking age (MLDA) generate substantial variation in the availability of commercially produced alcohol across people of different ages. Since our measure of sexual violence is for the period 2010, we consider state-wise alcohol policies up to 2010. From 2000 to 2010, alcohol policies across states remained stable. Gujarat for example prohibited the sale of alcohol. Among other states where drinking was allowed, MLDA varied from 18 to 25 years. Tamilnadu was the only state that raised MLDA from 18 to 21 years in 2004. Such stability rules out the possibility that other time-varying state characteristics might have influenced the legal rules related to alcohol consumption. Despite there exists significant law evasion, men who are legally allowed to drink are substantially more likely to consume alcohol (Luca et al.,

2019). Moreover, recent studies have found that a stricter alcohol law in India significantly lowers the crimes against women but not the other forms of crime (Luca et al., 2015, 2019). The causal impact of the drinking-age law on sexual violence is also documented in other countries (Gatley et al., 2017) [9].

MLDA policies in India also create a variation in the fraction of men legally allowed for alcohol consumption across states. Similar to Chakraborty and Lohawala (2021), we exploit this variation and construct our IV which is the fraction of men legally eligible to drink across states. It is worth mentioning that this measure could also vary across districts. However, since the data from IHDS is not representative at the district level, we construct a valid measure of the male population above the legally allowed drinking age at the state level. Let MP_s and $MLDA_s$ denote the total male population and minimum legal drinking age in a state s respectively. Then the instrument (Z_s) is calculated as follows:

$$Z_s = \frac{\sum_{i=1}^{MP_s} I(Age_i > MLDA_s)}{MP_s} * 1000$$

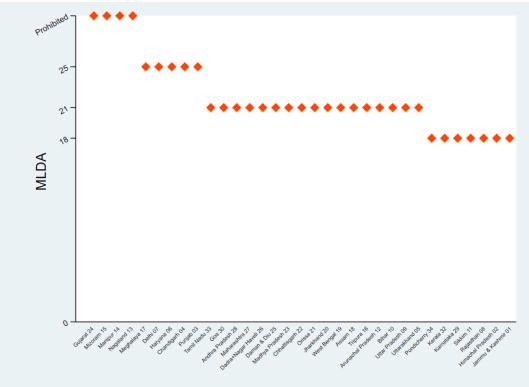
Panel (a) and (b) of Figure 3 show the state-wise distributions of MLDA and the share of men eligible to drink respectively. A quick look at these figures suggests that stricter alcohol policies (for instance Gujarat where drinking is completely prohibited) are associated with relatively lower sexual offenses.

Next, we estimate the following equations using two-stage least squares.

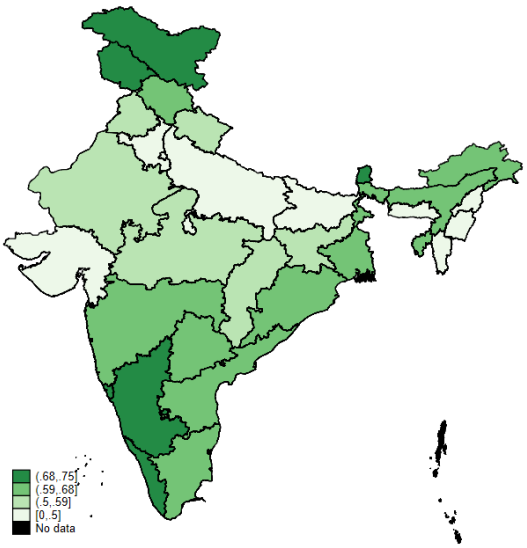
$$Y_{ihs} = \delta_0 + SV_s\delta_1 + \delta_2 X'_{ihs} + \theta_r + \epsilon_{ihs} \tag{14}$$

$$SV_s = \gamma_0 + Z_s\gamma_1 + \gamma_2 X'_s + \phi_r + \omega_s \tag{15}$$

[9] Ertan Yörük and Yörük (2015) document that young adults tend to drink up to 2.1 days more once they are granted legal access to alcohol at age 21. Although the discrete jump in alcohol consumption at the MLDA is associated with an increase in the sexual activity by up to 7.8 percentage points, it does not have a significant impact on the probability of engaging in risky sexual behaviors among young adults



(a) MLDA across States



(b) State-Wise Shares of Men Eligible to Drink

Figure 3: Minimun Legal Drinking Age and Fractions of Drinking-age Men across States in India

Source: State excise departments and IHDS (2011-12). The figures present the minimum legal drinking age and the fractions of men legally allowed to drink across states in India.

The key identifying assumption of the IV estimation is that the law regulating alcohol access does not affect SSP directly whereas it has a significant correlation with sexual violence. Figure 4 presents the first-stage correlation between the drinking-age male population and SV across states. Aligned with our expectation, the association between the two is positive implying that a higher share of the drinking-age male population is associated with higher incidents of sexual violence in a state.

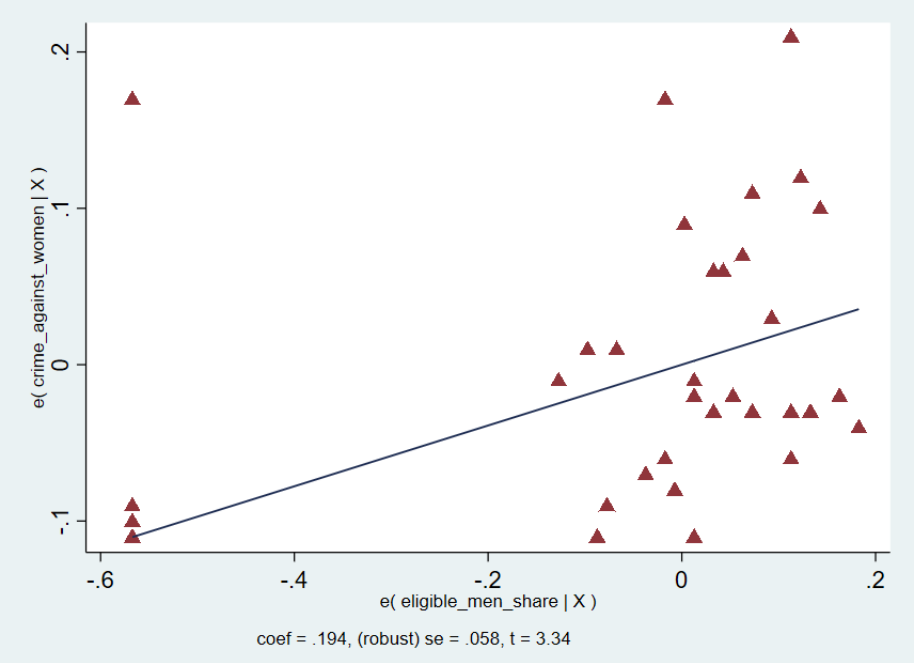


Figure 4: First-Stage Correlation

Source: Own calculations from IHDS (2011-12), NCRB (2010). The figure reports the correlation between fraction of drinking-age men and sexual violence in a state.

Results from the two-stage least square regressions are reported in Table 5 [10]. In Panel A, the outcome variable is the desired son ratio. Column 4 reports the estimates from the full specification. The IV estimate replicates our findings from baseline models suggesting that an increase in sexual violence in a state significantly increases the reported son preference of a woman in her childbearing age. In columns 5 and 6, we report IV estimates for the two sub-groups. In line with our baseline results, a similar effect of SV is found for less wealthy and lower-caste women. Panel B reports the estimates for our alternative outcome measure i.e gender preference for the extra child. The coefficient of interest from Column 4 corroborates the findings from Panel A. Hence for both outcomes, the IV estimates are in line with the baseline findings.

[10]Two-stage regressions and baseline regressions include the same controls (see subsection 3.3).

	Dep Var: SSP (Desired Son Ratio)					
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	*+Dist/State*	*+Ind*	*+HH*	*Poor*	*Lower Caste*
SV	.345**	.194**	.205**	.215**	.446**	.192*
	(.165)	(.088)	(.094)	(.100)	(.160)	(.104)
Constant	.485***	.534***	.528***	.472***	.437***	.525***
	(.019)	(.123)	(.134)	(.135)	(.174)	(.209)
N	12,102	12,102	11,320	11,124	2,496	8,358
R Squared	.013	.023	.030	.055	.054	.052
	Panel B: Alternative Dep Var (Gender Preference for Extra Child)					
SV	.564	.657	.812*	.994**	.945**	.922*
	(.414)	(.428)	(.431)	(.438)	(.458)	(.471)
Constant	.485***	.534***	.528***	.472***	.437***	.525***
	(.019)	(.123)	(.134)	(.135)	(.174)	(.209)
N	4,309	4,309	3,726	3,660	998	2,747
R Squared	.121	.134	.176	.187	.201	.188
Geographic Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State Controls	No	Yes	Yes	Yes	Yes	Yes
Individual Controls	No	No	Yes	Yes	Yes	Yes
HH Controls	No	No	No	Yes	Yes	Yes

Table 5: Two-Stage Least Square Estimation

Source: Own calculations from IHDS (2011-12) and NCRB (2010).

Notes: Two-stage least square estimation. Sample consists of married women in the age range of 18 to 32. Figures in parentheses are robust standard errors clustered at the state level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4 SV and SSP: Quasi-experimental Evidence from an Infamous Gang Rape Case in Kolkata (2012)

Cases of sexual violence are largely under-reported due to the lack of judicial support and fear of social exclusion (Ghosh, 2013; Iyer et al., 2012). To address such measurement error associated with the administrative records of sexual crimes, we complement the previous estimation strategy with a quasi-experimental difference-in-differences (DID) framework similar to (Brodeur et al., 2021; McCoy et al., 2020; Metcalfe et al., 2011; Montalvo, 2011) and show how the intrahousehold reported preference for son responded to a plausibly exogenous gang rape event. In the last decade (2010-20), India witnessed some of the worst rape cases in the country's history. Park Street gang rape (Kolkata, 2012) is one of the most horrific Indian rape cases that shook the entire state of West Bengal (WB) (Times, 2015). On the night of February 5, 2012, a woman was gang-raped at gunpoint in a moving car and was thrown out. A few days later, she mustered the courage to complain and faced several forms of harassment by the police. However, the event received widespread media coverage from February 15, the day when the medical examination report at a state-run hospital confirmed that the woman was raped, and the case was transferred to the Detective Department of Kolkata Police for formal investigation (TOI, 2012; Anandabazar, 2012). We consider this date for the analysis, not the event date, as the psychological effects of the case may have started becoming apparent as soon as the news hit the public.

4.1 Data and Estimation

To establish the causal impact of that brutal gang rape event, we primarily exploit the two unique features of IHDS (2012). First, it gives us information on a woman's state of residence at the time of the survey. Second, it records the enumeration date for each participant allowing us to access the additional dimension i.e. time. The survey was primarily administered between October 2011 and March 2013. Moreover, the interview dates were

randomly chosen by the survey administrators. These features allow us to identify the women who were relatively more exposed to the media coverage of the event and were interviewed after the news reached to public. Our analytical sample for this study consists of 2,471 married women in their fertile age (18 to 32) from WB and its neighboring states including Bihar, Jharkhand, Odisha, and Assam which share a common border with WB [11], who completed the interview before September 1, 2012 [12]. The treated group consists of 741 married women in the age range of 18 to 32 from WB and the control group includes the women belonging to the same age cohort from the states that share a common border with WB. SSP is measured by the ideal proportion of sons and gender preference for the extra child (see subsection 3.2). To estimate the causal impact of the event, we employ the following DID specification:

$$Y_{iht} = \phi_0 + \phi_1 Treated_s + \phi_2 Treated_s * Post_t + \phi_3 X'_{ih} + \theta_D + \omega_{1d} + \omega_{2w} + \epsilon_{iht} \qquad (16)$$

where Y_{iht} is the SSP of woman i, a member of household h, and belonging to the interview date t. $Post_t$ is an indicator variable taking a value equal to one for a woman in our sample who was surveyed after the day when the event received widespread media exposure. $Treated_s$ is a variable taking value equal to one if a woman's state of residence is WB at the time of survey. The parameter of interest ϕ_2 gives us the causal impact of the incident on a woman's SSP under the assumption that, in the absence of that brutal gang rape event, preference for sons would have evolved similarly in WB and its neighbor states. The vector, X'_{ih} includes the standard set of individual and household-level controls specified in subsection 3.3. θ_D, ω_{1d}, and ω_{2w} are the district, day-of-month, and week-of-year-fixed

[11] Summary of all variables across treated and control groups are given in Table A3 in appendix.

[12] To the best of our knowledge, there were no other major cases of sexual violence during that time frame that could confound the effects of the Park Street rape case. Most notably, the brutal *Nirbhaya* rape case took place in Delhi on December 16th, 2012 (PTI, 2020). However, the way we set the sampling window can rule out the possible confounding effects of the above-mentioned event. More importantly, the next major event of sexual violence in WB took place at the Kamduni village, on June 7, 2013 (Gupta, 2015) which does not belong to the time frame that we consider.

effects respectively. Robust standard errors in all models are clustered at the district-week level.

4.2 Results

4.2.1 Baseline

Results from the DID estimation are reported in Table 6. Panel A and B report the causal impact of the Park Street gang rape case on average changes in SSP w.r.t. the two sets of sampling windows respectively. In each panel, row 1 reports the coefficients of interest. In each panel, column 1 includes no controls except the district and the day-of-month-fixed effects. Column 2 incorporates individual characteristics. Finally, column 3 gives the estimates from the full model after including the household characteristics and week-fixed effects. From panel A, we conclude that the event increased SSP by 9 percentage points among women from WB. Columns 4 and 5 report the coefficients from the sub-sample analyses. The event significantly increased reported son preference among young women from poorer households. To check the sensitivity of our baseline results, we re-estimate Equation 16 for a smaller sampling window: from October 1, 2011 to June 30, 2012. Panel B reports the DID estimates w.r.t. this specification. Results are qualitatively similar and suggest that the violent rape case significantly increased SSP among respondents from the treated cohort.

Results from the DID estimation for the alternative measure of SSP are shown in Table 7. Panel A and B report the causal impact of the Park Street gang rape case on average changes in the reported gender preference for extra child w.r.t. the two sets of sampling windows. In each panel, column 3 gives the estimates from the full model after including all individual and household characteristics, district, day-of-month, and week-of year fixed effects. From panel A, we conclude that the event increased the likelihood of reporting a son as the preferred extra child by 9 percentage points among respondents from the treated group. To check the

sensitivity of the results, we re-estimate Equation 16 for a smaller sampling window: from October 1, 2011 to June 30, 2012. Panel B reports the DID estimates w.r.t. this specification. Results are qualitatively similar and suggest that the violent rape case significantly increased the likelihood that the gender preference for an extra child would be a boy among young women from the WB.

	Dependent Variable: SSP (Desired Son Ratio)				
	(1)	(2)	(3)	(4)	(5)
	Baseline	*+Ind*	*+HH*	*Poor*	*Lower Caste*
	Panel A: Broader Window (October 1, 2011 - August 31, 2012)				
Treated*Post	.114***	.071**	.089**	.157***	.027
	(.022)	(.027)	(.038)	(.033)	(.031)
Constant	.566***	.439***	.442***	.625***	.462***
	(.005)	(.048)	(.053)	(.163)	(.062)
N	2,470	2,445	2,388	743	1,742
R Squared	.036	.093	.108	.138	.137
	Panel B: Smaller Window (October 1, 2011 - June 30, 2012)				
Treated*Post	.105***	.116***	.084**	.149***	.065*
	(.030)	(.032)	(.037)	(.037)	(.033)
Constant	.481***	.521***	.428***	.650***	.506***
	(.010)	(.032)	(.060)	(.183)	(.029)
N	2,023	1,999	1,958	614	1,435
R Squared	.060	.093	.105	.121	.110
District Fixed Effects	Yes	Yes	Yes	Yes	Yes
Day-of-Month Fixed Effects	Yes	Yes	Yes	Yes	Yes
Week-of-Year Fixed Effects	No	No	Yes	Yes	Yes
Individual Controls	No	Yes	Yes	Yes	Yes
HH Controls	No	No	Yes	Yes	Yes

Table 6: SV and SSP: DID Estimates

Source: Own calculations from IHDS (2011-12)

Notes: DID estimation. Sample consists of married women in the age range of 18 to 32 who completed the interview before September 1, 2012. Treated group: married women from WB in the age range 18 to 32 who completed the interview before September 1, 2012. Post is a dummy variable taking a value equal to one, if the respondent was interviewed after February 15, 2012, the day when the rape case received widespread media coverage. Figures in parentheses are robust standard errors clustered at the district-week level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Dep Var: Gender Preference for Extra Child (Boy)			
	(1)	(2)	(3)
	Baseline	*+Ind*	*+HH*
	Panel A: Broader Window		
Treated*Post	.274**	.367**	.259*
	(.151)	(.153)	(.145)
Constant	.566***	.439***	.442***
	(.005)	(.048)	(.053)
N	989	977	944
R Squared	.142	.178	.254
	Panel B: Smaller Window		
Treated*Post	.387***	.354***	.332**
	(.112)	(.132)	(.158)
Constant	.481***	.521***	.428***
	(.010)	(.032)	(.060)
N	765	754	730
R Squared	.117	.168	.199
District Fixed Effects	Yes	Yes	Yes
Day-of-Month Fixed Effects	Yes	Yes	Yes
Week-of-Year Fixed Effects	No	No	Yes
Individual Controls	No	Yes	Yes
HH Controls	No	No	Yes

Table 7: SV and SSP: DID Estimates

Source: Own calculations from IHDS (2011-12)

Notes: DID estimation. Sample consists of married women in the age range of 18 to 32 who completed the interview before September 1, 2012. Treated group: married women from WB in the age range 18 to 32 who completed the interview before September 1, 2012. Post is a dummy variable taking a value equal to one, if the respondent was interviewed after February 15, 2012, the day when the rape case received widespread media coverage. Figures in parentheses are robust standard errors clustered at the district-week level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

4.2.2 Trend Analysis

Figure 5 provides a graphical representation of the trend analysis from the event study estimates. Each point represents the estimated effect of the event for several time points relative to the media exposure date. The reference group for each estimate includes respondents in the age range from 18 to 32 residing in all neighbor states except WB who completed the interview 40 to 60 days before the cut-off date. The plotted line demonstrates an abrupt increase in the estimates following the incident. Moreover, the effect persisted thereafter.

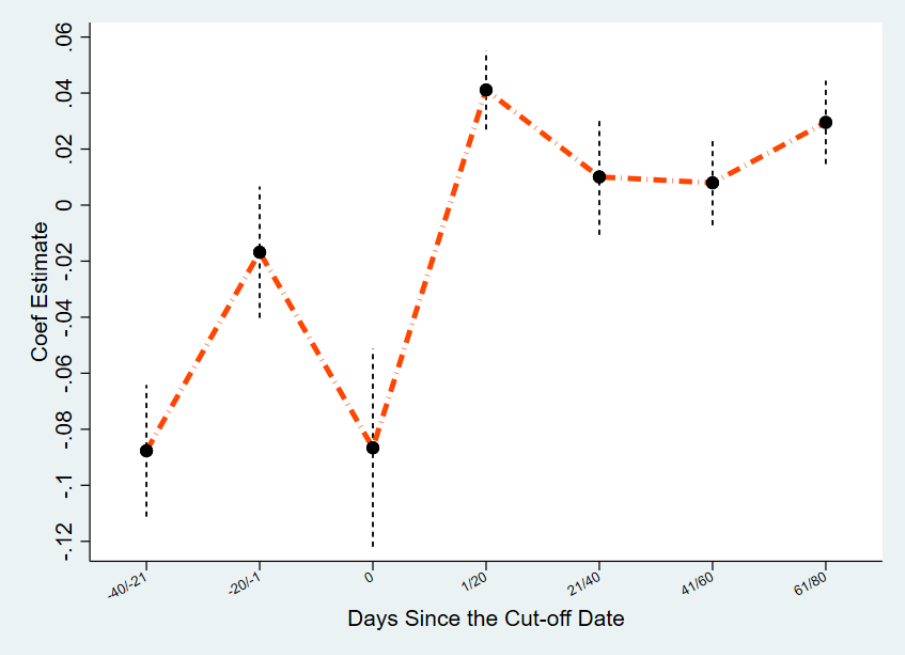


Figure 5: Trend Analysis: Results from Event Study

Source: Own calculations from IHDS (2011-12). The figure reports the estimated coefficients from the event study. The sample consists of married women in the age range of 18 to 32 who completed the interview before September 1, 2012.

4.2.3 Estimates from Regression Discontinuity

However, the key threat associated with our above-mentioned estimation strategy is the existence of spatial spillovers. To address this concern, we restrict the sample to women participants from WB and explore the impact of the Park Street rape case on SSP in a

regression discontinuity framework similar to (McCoy et al., 2020) that exploit time as the running variable. Our estimating equation is of the form:

$$Y_{iht} = \rho_0 + \rho_1 DaysSinceEvent_t + \rho_2 DaysSinceEvent^2{}_t + \rho_3 DaysSinceEvent_t * Post_t + \rho_4 DaysSinceEvent^2{}_t * Post_t + \rho_5 Post_t + \rho_6 X'_{ih} + \theta_n + \omega_{1d} + \omega_{2w} + \epsilon_{iht} \quad (17)$$

where Y_{iht} is the SSP of woman i, belonging to household h, and interview-date t. $DaysSinceEvent_t$ denotes the number of days between the enumeration date of each respondent i and the date of media exposure of the attack. $Post_t$ is an indicator variable taking a value equal to one for any individual in our sample who was surveyed after the day when the event received widespread media exposure, and zero otherwise. The parameter of interest is ρ_5, which gives us the discrete jump in SSP among women from WB following the cut-off date under the assumption that all other time-varying factors vary continuously around the specified date. We also add the quadratic polynomial of the term $DaysSinceEvent_t$ and its interaction with $Post_t$ in our RD specification. The vector, X'_{ih} includes all individual and household-level controls specified in subsection 3.3. θ_n, ω_{1d}, ω_{2w} denote the neighborhood, day-of-month, and the week-of-year-fixed effects respectively. In all models, robust standard errors are clustered at the neighborhood-week level.

RD estimates are reported in Table 8. Column 1 presents the baseline result without including any controls except the neighborhood and the day-of-month-fixed effects. The coefficient of interest is positive and statistically significant. In columns 2 and 3, we sequentially add the individual characteristics, week-fixed effects, and household-level factors respectively. The magnitude of the coefficients increases and the estimated effect becomes statistically significant at the 1 percent level. Overall, the results suggest that the effect of the brutal Park Street gang rape was confined to WB and there was a 30 percentage point jump in the average SSP among young women from WB following the

event. It is worth noting that there is a severe loss of data for this analysis [13]. However, the stability of the coefficients across all models ensures that the degree of precision is not compromised.

	Dep Var: SSP (Desired Son Ratio)		
	(1)	(2)	(3)
	Baseline	*+Ind*	*+HH*
Post	.258**	.304***	.305***
	(.117)	(.088)	(.089)
Constant	.342***	.454***	.361***
	(.089)	(.155)	(.176)
N	741	710	715
R Squared	.168	.231	.221
Neighborhood Fixed Effects	Yes	Yes	Yes
Day-of-Month Fixed Effects	Yes	Yes	Yes
Week-of-Year Fixed Effects	No	Yes	Yes
Individual Controls	No	Yes	Yes
HH Controls	No	No	Yes

Table 8: Estimates from Regression Discontinuity Approach

Source: Own calculations from IHDS (2011-12).

Notes: RD estimation. Sample consists of married women from WB in the age range of 18 to 32 who completed the interview before September 1, 2012. Post is a dummy variable taking a value equal to one, if the respondent was interviewed after February 15, 2012, the day when the rape case received widespread media coverage. Figures in parentheses are robust standard errors clustered at the neighborhood-week level. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5 Conclusion

A sharp rise in sexual violence against women has threatened parents with girl children in many countries including India. Under this scenario, our study is motivated to explore how reported preference for sons responds to the increasing risk of sexual violence in the locality. We rationalize parents preferring boys in response to the rising sexual violence against women in the neighborhood and such outcome is primarily driven by the increasing

[13] To avoid further data loss, we replaced missing values of some control variables such as son dependence, and spousal relationship with the mean values of the variables.

expected cost of raising daughters. Exploiting several quasi-experimental estimation strategies and validating our findings with various robustness checks, we provide systematic evidence that an increase in sexual violence in a district or state significantly increases the intrafamily reported preference for boys.

A vast literature has primarily focused on the kinship system as the key cultural force driving gender inequality within households in countries such as India, China, and South Korea. Norms including dowry, patrilocality, funeral procedures, and old age security are thus the most cited determinants of gender bias within families. However, there is surprisingly limited evidence on the role of safety concerns outside home, an added external cultural constraint against women in determining pro-male bias. Our study fills this gap in the literature and provides systematic causal evidence linking these two crucial factors associated with gender inequality. Our study also highlights the role of formal institutions in reducing gender inequality through curtailing the risk of non-partner sexual harassment in public spaces. In general, the role of institutional intervention is critical when it comes to addressing the issue of intrafamily gender inequality. For example, it is well-documented that one of the most gender-progressive policies that allow women's entitlement to property rights reinforced parental son bias in India instead of eliminating it [14]. Similarly, There exists stylized evidence that the legal restrictions on sex-selective abortion led to postnatal discrimination against girls as reflected in their educational attainment [15]. The key reason why some policies fail to bring changes in the desired direction is that they face drastic social norm backlash. More specifically, policies often have very little to offer when son bias is an outcome of parental cost-benefit exercise fundamentally guided by the rigid patriarchal social and kinship norms. However, our study shows a more straightforward avenue through which policy can reduce gender bias in fertility preferences within households by reducing the expected cost of raising daughters through ensuring safety outside home. Future research in this field can help identify whether safety concern is associated with revealed son-preferring behaviors within

[14] See (Bhalotra et al., 2020a).

[15] see (Rastogi and Sharma, 2022)

households.

Appendix

Table A1: Summary: Individual and Household Characteristics

	Mean	SD	N
	Panel A: Women's Characteristics		
Stated Son Preference (Desired Ratio)	.551	.108	12,102
Gender Preference of Extra Child (Boy)	.789	.402	4,309
Age	26.40	3.87	12,102
Education (in years)	6.30	4.73	12,101
Age at Marriage	18.18	3.15	12,097
Number of Children	1.46	1.15	12,095
More Son among Existing Children	.354	.478	11,998
Son Dependence	.725	.410	11,420
Relationship with Spouse	.701	.421	12,102
Perceived Spousal Violence	.424	.494	12,091
Employment Status	.209	.387	12,102
	Panel B: HH Characteristics		
Number of Adults	3.45	1.68	12,102
Highest Education (in years)	8.84	4.46	12,102
Urban	.304	.460	12,102
Religion (whether Hindu)	.826	.379	12,102
Caste (whether Upper-Caste)	.247	.431	12,083
Economic Status (Poor)	.221	.414	12,097
Access to Media (TV)	.591	.492	12,020
Wealth (Rich)	.389	.486	11,993

Source: Own calculations from IHDS (2011-12).
Note: Sample consists of married women in the age range of 18 to 32.

Table A2: Summary: District and State Characteristics

	Panel A: District Characteristics		
	Mean	SD	N
Sexual Violence (per 1000 women)	.140	.106	384
Share of Hindu	.801	.198	384
Share of Scheduled Caste (SC) Women	.169	.077	384
Urban Population	.342	.229	384
Female Literacy Rate	.652	.121	384
Child Sex Ratio (F/M)	.917	.042	384
Murder (per 1000 persons)	.031	.028	384
Kidnapping (per 1000 persons)	.082	.024	384
Stealing (per 1000 persons)	.401	.514	384
	Panel B: State Characteristics		
Sexual Violence (per 1000 women)	.134	.084	27
Share of Hindu	.672	.273	27
Share of Scheduled Caste (SC) Women	.126	.083	27
Urban Population	.374	.213	27
Female Literacy Rate	.699	.107	27
Child Sex Ratio (F/M)	.925	.038	27
Murder (per 1000 persons)	.028	.010	27
Kidnapping (per 1000 persons)	.071	.074	27
Stealing (per 1000 persons)	.422	.547	27
Share of Men Eligible to Drink	.531	.215	27

Source: Own calculations from Census (2011), NCRB (2010).

Table A3: Summary: Treated vs. Control Cohort

Women's Characteristics	*Treated Cohort*			*Control Cohort*		
	Mean	SD	N	Mean	SD	N
Stated Son Preference (Desired Ratio)	.545	.159	741	.553	.126	1,730
Gender Preference of Extra Child (Boy)	.778	.415	307	.840	.366	683
Age	25.60	4.01	741	26.29	4.00	1,730
Education (in years)	5.96	4.10	741	5.43	4.76	1,730
Age at Marriage	17.65	2.96	741	17.99	3.01	1,730
Number of Children	1.42	1.00	741	1.64	1.05	1,730
More Son among Existing Children	.365	.481	725	.360	.480	1,726
Son Dependence	.697	.459	741	.734	.442	1,730
Relationship with Spouse	.724	.427	741	.712	.411	1,730
Perceived Spousal Violence	.517	.500	740	.373	.483	1,727
Employment Status	.212	.377	740	.197	.389	1,730
HH Characteristice	*Treated Cohort*			*Control Cohort*		
Number of Adults	3.09	1.44	741	3.36	1.68	1,730
Highest Education (in years)	7.23	4.49	741	7.62	4.87	1,730
Urban	.373	.484	741	.247	.431	1,730
Religion (whether Hindu)	.692	.461	741	.839	.366	1,730
Caste (whether Upper-Caste)	.289	.500	741	.217	.376	1,727
Economic Status (Poor)	.267	.442	741	.330	.470	1,730
Access to Media (TV)	.598	.463	739	.571	.490	1,699
Wealth (Rich)	.349	.421	732	.367	.486	1,713

Source: Own calculations from IHDS (2011-12).

Note: Sample consists of married women in the age range of 18 to 32 in WB and its neighboring states including Bihar, Jharkhand, Odisha, and Assam which share a common border with WB, who completed the interview before September 1, 2012.

References

Abrahams, N., Devries, K., Watts, C., Pallitto, C., Petzold, M., Shamu, S., and GarcÍa-Moreno, C. (2014). Worldwide prevalence of non-partner sexual violence: a systematic review. *The Lancet*, 383(9929):1648–1654.

Anandabazar (2012). Dhorshitar shonge oshohojogita, obhijukto poolish. *Anandabazar Patrika*.

Anukriti, S., Bhalotra, S., and Tam, E. H. (2022). On the quantity and quality of girls: Fertility, parental investments and mortality. *The Economic Journal*, 132(641):1–36.

Arnold, F., Choe, M. K., and Roy, T. K. (1998). Son preference, the family-building process and child mortality in india. *Population studies*, 52(3):301–315.

Asadullah, M. N., Mansoor, N., Randazzo, T., and Wahhaj, Z. (2021). Is son preference disappearing from bangladesh? *World Development*, 140:105353.

Becker, G. S. and Lewis, H. G. (1973). On the interaction between the quantity and quality of children. *Journal of political Economy*, 81(2, Part 2):S279–S288.

Bhalotra, S., Brulé, R., and Roy, S. (2020a). Women's inheritance rights reform and the preference for sons in india. *Journal of Development Economics*, 146:102275.

Bhalotra, S., Chakravarty, A., and Gulesci, S. (2020b). The price of gold: Dowry and death in india. *Journal of Development Economics*, 143:102413.

Borker, G. (2017). Safety first: Perceived risk of street harassment and educational choices of women. *Job Market Paper, Department of Economics, Brown University*, pages 12–45.

Brodeur, A., Clark, A. E., Fleche, S., and Powdthavee, N. (2021). Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of public economics*, 193:104346.

Bulatao, R. A. (1981). Values and disvalues of children in successive childbearing decisions. *Demography*, 18(1):1–25.

Chakraborty, T. and Lohawala, N. (2021). Women, violence and work: Threat of sexual violence and women's decision to work. *IZA Discussion Paper, No:14372*.

Chakraborty, T., Mukherjee, A., Rachapalli, S. R., and Saha, S. (2018). Stigma of sexual violence and women's decision to work. *World Development*, 103:226–238.

Chung, W. and Gupta, M. D. (2007). The decline of son preference in south korea: The roles of development and public policy. *Population and development review*, 33(4):757–783.

Clark, S. (2000). Son preference and sex composition of children: Evidence from india. *Demography*, 37(1):95–108.

Das Gupta, M., Zhenghua, J., Bohua, L., Zhenming, X., Chung, W., and Hwa-Ok, B. (2003). Why is son preference so persistent in east and south asia? a cross-country study of china, india and the republic of korea. *The Journal of Development Studies*, 40(2):153–187.

Deeksha, T. (2020). Why indian parents don't want a girl child? *The Times of India*.

Desai, S. and Vanneman, R. (2015). India human development survey-ii (ihds-ii). *University of Maryland and National Council of Applied Economic Research, New Delhi*, pages 1–45.

Dyson, T. and Moore, M. (1983). On kinship structure, female autonomy, and demographic behavior in india. *Population and development review*, pages 35–60.

Ebenstein, A. (2014). Patrilocality and missing women. *Available at SSRN 2422090*.

Ebert, C. and Vollmer, S. (2022). Girls unwanted–the role of parents' child-specific sex preference for children's early mental development. *Journal of Health Economics*, 82:102590.

Ertan Yörük, C. and Yörük, B. K. (2015). Alcohol consumption and risky sexual behavior among young adults: evidence from minimum legal drinking age laws. *Journal of Population Economics*, 28:133–157.

Gatley, J. M., Sanches, M., Benny, C., Wells, S., and Callaghan, R. C. (2017). The impact of drinking age laws on perpetration of sexual assault crimes in canada, 2009–2013. *Journal of Adolescent Health*, 61(1):24–31.

Gaudin, S. (2011). Son preference in indian families: absolute versus relative wealth effects. *Demography*, 48(1):343–370.

Ghosh, S. (2013). The treatment of sexual violence in india. *Warwick Student L. Rev.*, 3:41.

Gupta, J. (2015). Kamduni rape trial comes to an end after two-and-a-half year. *The Times of India*.

Iyer, L., Mani, A., Mishra, P., and Topalova, P. (2012). The power of political voice: women's political representation and crime in india. *American Economic Journal: Applied Economics*, 4(4):165–193.

Jayachandran, S. (2015). The roots of gender inequality in developing countries. *economics*, 7(1):63–88.

Jayachandran, S. (2021). Social norms as a barrier to women's employment in developing countries. *IMF Economic Review*, pages 1–20.

Koolwal, G. B. (2007). Son preference and child labor in nepal: The household impact of sending girls to work. *World Development*, 35(5):881–903.

Lin, T.-c. and Adserà, A. (2013). Son preference and children's housework: The case of india. *Population Research and Policy Review*, 32:553–584.

Livemint (2023). Hathras gang-rape case: A look at the timeline as up court sets 3 accused free. *Livemint*.

Luca, D. L., Owens, E., and Sharma, G. (2015). Can alcohol prohibition reduce violence against women? *American Economic Review*, 105(5):625–629.

Luca, D. L., Owens, E., and Sharma, G. (2019). The effectiveness and effects of alcohol regulation: evidence from india. *IZA Journal of Development and Migration*, 9(1):1–26.

Mavisakalyan, A. and Minasyan, A. (2023). The role of conflict in sex discrimination: The case of missing girls. *Economic development and cultural change*, 71(2):443–484.

McCoy, S. J., McDonough, I. K., and Roychowdhury, P. (2020). The impact of terrorism on social capital: evidence from the 2015 charlie hebdo paris shooting. *Oxford bulletin of economics and statistics*, 82(3):526–548.

Metcalfe, R., Powdthavee, N., and Dolan, P. (2011). Destruction and distress: using a quasi-experiment to show the effects of the september 11 attacks on mental well-being in the united kingdom. *The Economic Journal*, 121(550):F81–F103.

Montalvo, J. G. (2011). Voting after the bombings: A natural experiment on the effect of terrorist attacks on democratic elections. *Review of Economics and Statistics*, 93(4):1146–1154.

Mughal, M., Javed, R., and Lorey, T. (2023). Female early marriage and son preference in pakistan. *The Journal of Development Studies*, pages 1–21.

Pande, R. P. and Astone, N. M. (2007). Explaining son preference in rural india: the independent role of structural versus individual factors. *Population Research and Policy Review*, 26(1):1–29.

PTI (2020). Chronology of events in nirbhaya case. *The Hindu*.

Rastogi, G. and Sharma, A. (2022). Unwanted daughters: the unintended consequences of a ban on sex-selective abortions on the educational attainment of women. *Journal of Population Economics*, 35(4):1473–1516.

Robitaille, M.-C. (2013). Determinants of stated son preference in india: are men and women different? *The Journal of Development Studies*, 49(5):657–669.

Robitaille, M.-C. (2020). Conspicuous daughters: Exogamy, marriage expenditures, and son preference in india. *The Journal of Development Studies*, 56(3):630–647.

Robitaille, M.-C. and Chatterjee, I. (2018). Sex-selective abortions and infant mortality in india: The role of parents' stated son preference. *The Journal of Development Studies*, 54(1):47–56.

Sarkar, S. (2021). Local crime and early marriage: Evidence from india. *WIDER Working Paper, No: 2021/37*.

Song, R., Li, S., and Eklund, L. (2022). Can risk perception alter son preference? evidence from gender-imbalanced rural china. *The Journal of Development Studies*, 58(12):2566–2582.

Theerthaana, P. and Sheik Manzoor, A. (2019). Gender disappointment in india: Sem modeling approach. *Archives of women's mental health*, 22:593–603.

Times, H. (2015). Park street gang rape case: Three accused found guilty. *Hindutstan Times*.

TOI (2012). Woman raped in car at gunpoint in kolkata, cops jeer. *The Times of India*.

Women and ICRW (2013). Unsafe: An epidemic of sexual violence in delhi's public spaces: Baseline findings from the safe cities delhi programme.