The causal impact of women's age at marriage on domestic violence in India[*]

Punarjit Roychowdhury[†]

University of Nottingham, UK

Gaurav Dhamija[‡]

Indian Statistical Institute, Delhi Centre

August 4, 2020

Abstract

We examine the causal effect of women's age at marriage on prevalence of domestic violence using newly available household data from India. We employ an empirical strategy that utilizes variation in age at menarche to obtain exogenous variation in women's age at marriage. We find robust evidence that a one-year delay in women's marriage causes a significant decline in physical violence, although it has no impact on sexual or emotional violence. Further, we provide suggestive evidence that the effect of women's marital age on physical violence arises because older brides, as compared to younger brides, are more educated and are married to more educated men. Overall, our findings underscore the importance of better enforcement of existing social policies that seek to delay marriages of women, as well as formulation of newer interventions, to reduce the prevalence of domestic violence in developing countries.

JEL: J12, J16, O12, O15
Keywords: Age at Marriage, Domestic Violence, India, Instrumental Variables, Women.

[*]The authors are very grateful to the associate editor, three anonymous reviewers, Daniel Millimet, and session participants at the NEUDC Conference 2018 held at the Cornell University, New York, 14th Annual Conference on Economic Growth and Development held at the Indian Statistical Institute Delhi, and 2nd International Conference on South Asian Economic Development at the South Asian University, New Delhi for comments and suggestions. All remaining errors are our own.

[†]Corresponding Author. Industrial Economics Division, Nottingham University Business School, University of Nottingham, Jubilee Campus, Wollaton Road, Nottingham NG8 1BB, United Kingdom. email: punarjitroychowdhury@gmail.com.

[‡]Centre for research on Economics of Climate, Food, Energy and Environment (CECFEE), Economics and Planning Unit, Indian Statistical Institute 7, S. J. S. Sansanwal Marg, New Delhi, Delhi 110016, India. email: gaurav.dhamija5@gmail.com.

The causal impact of women's age at marriage on domestic violence in India

Abstract

We examine the causal effect of women's age at marriage on prevalence of domestic violence using newly available household data from India. We employ an empirical strategy that utilizes variation in age at menarche to obtain exogenous variation in women's age at marriage. We find robust evidence that a one-year delay in women's marriage causes a significant decline in physical violence, although it has no impact on sexual or emotional violence. Further, we provide suggestive evidence that the effect of women's marital age on physical violence arises because older brides, as compared to younger brides, are more educated and are married to more educated men. Overall, our findings underscore the importance of better enforcement of existing social policies that seek to delay marriages of women, as well as formulation of newer interventions, to reduce the prevalence of domestic violence in developing countries.

JEL: J12, J16, O12, O15
Keywords: Age at Marriage, Domestic Violence, India, Instrumental Variables, Women.

"My husband came into the room, locked the door. He turned up the music so that no one could hear us outside. Then he took out his belt and started to hit me. He kept whipping me for the next 30 minutes...As he was doing this, he warned me that I shouldn't make a sound, I shouldn't cry, I shouldn't scream, because if I did, he was going to hit me even harder. He was hitting me with his belt, his hands... soon he began to choke me. He was just so angry."

– Experience of a 19 year old woman, Aditi (name changed), one of the millions of victims of domestic violence in India (BBC, 2014)

1 Introduction

Domestic violence is a global pandemic that affects one in three women in their lifetime. According to a study by the World Health Organization (2012), partner violence is the most common form of violence in women's lives and is far greater than assaults or rape by strangers, acquaintances or any other perpetrators in both developing and developed countries. Women who suffer domestic violence experience serious health consequences including injury, emotional distress, suicidal thoughts, physical symptoms of severe illness, absence from work, alcohol and substance abuse, sexually transmitted diseases and unintended pregnancies (Campbell, 2002; Coker et al., 2002; Ackerson and Subramanian, 2008; Ellsberg et al., 2008). The cost of domestic violence to an economy in terms of victim's suffering, medical bills, lost productivity, judicial expenditures and the lost productivity from the incarcerated offender is massive. For example, according to Lomborg and Williams (2018) in the US alone this cost is about $460 billion annually.

In this paper, we examine the role of a potential determinant of domestic violence: *women's age at marriage*. We specifically ask: Does women's age at marriage have a *causal* effect on their exposure to domestic violence, and more specifically, spousal or intimate partner violence (IPV)?[1] If so, in which direction and to what extent? We use newly available

[1] Although technically spousal violence or IPV is a subset of domestic violence, we shall use the terms

household data from India, where according to a *BBC* report (2014), one incident of domestic violence is reported in every five minutes (which, of course, is only a fraction of how much actually occurs). We find that a year of delay in women's marriage causes a significant reduction in (non-sexual) physical violence, but does not impact sexual or emotional violence.

In theory, the causal impact of women's age at marriage on domestic violence could be either negative or positive. On one hand, women who marry early are likely to be unassertive, naive, socially isolated, experience severe psychological depression and have lesser bargaining power within marriage (Field and Ambrus, 2008; Nour, 2009; Le Strat et al., 2011; Chari et al., 2017). This makes them less resistive to domestic violence and hence 'safer' to be victimized. They are also likely to be less educated since early marriage often interrupts the accumulation of formal education for women due to family responsibilities (Field and Ambrus, 2008). This limits their options outside marriage and the economic and social resources at the women's disposal (Chowdhury, 2004), negatively influencing their empowerment within marriage (Farmer and Tiefenthaler, 1996; Stevenson and Wolfers, 2006; Aizer, 2010; Hidrobo and Fernald, 2013; Erten and Keskin, 2018; Yount et al., 2018). Both these factors would suggest a negative relationship between women's age at marriage and domestic violence.

On the other hand, although women who marry later might be more able to advocate for their preferences in the spousal household, have greater access to social capital, have greater bargaining power and consequently be more resistive to domestic violence, they might face a stronger backlash from their partners (Field et al., 2016). Moreover, since education is positively correlated to age at marriage and more education leads to greater availability of economic resources, women who marry late may experience violence or threats of violence from their spouses who might want to control these resources (Bloch and Rao, 2002; Eswaran and Malhotra, 2011; Bobonis et al., 2013). These two factors, taken together, suggest that

domestic violence and IPV interchangably throughout the paper since three-quarters of violence against women is intimate (Aizer, 2010).

women who marry late may be more vulnerable to mistreatment. Overall, thus, the causal effect of women's age at marriage on prevalence of domestic violence is *a priori* ambiguous.

To examine the relationship between women's age at marriage and domestic violence, we use data from the National Family Health Survey of India (NFHS), 2015-16. This survey includes detailed information on the prevalence of domestic violence, gender role, health, and marriage market indicators. As noted by Golder et al. (2016), the National Family Health Survey collects information on domestic violence with utmost caution following both Indian and international guidelines (more specifically the WHO ethical guidance for research on domestic violence against women, 2001, for the ethical collection of data on violence). We focus on four types of domestic violence against women: less severe physical violence, severe physical violence, sexual violence, and emotional violence (we discuss each category in detail later in the data section).

The main empirical challenge in identifying the causal effect of age at marriage on prevalence of domestic violence is that marriage age might be endogenous due to omitted variables. For instance, according to classic patriarchy, women are expected to marry young to exchange obedience for protection from men (e.g., Kabeer, 1988; Alam, 2007; Yount and Li, 2010), and to respect men's authority to punish disobedience (Feldman, 2010; Yount and Li, 2010). Thus, those women who come from families that strictly follow such patriarchal norms are likely to get married early as well as be more tolerant of, and hence exposed to greater domestic violence. Such unobserved characteristics of women's natal family could in theory drive the relationship between women's age at marriage and domestic violence. Unobserved ability of women might also be correlated with marital age and domestic violence. Specifically, more able women might get married late as well as be less vulnerable to domestic violence. This might be perhaps due to the positive correlation between ability and labor market prospects. This could be also because women of higher ability might choose to marry into households relatively late only after their earnings potential is fully revealed and these households might be systematically different (perhaps better in terms of prevalence of

domestic violence) from the average household. In addition to omitted variables, of course, endogeneity could also arise due to potential measurement error in age at marriage.

To address the issue of endogeneity and estimate the causal effect of women's age at marriage on domestic violence, we employ the empirical strategy proposed by Field and Ambrus (2008), who instrument women's age at marriage by their age at menarche. The rationale for the instrument is that in patriarchal socieites like India, parents become extremely anxious to get their daughters married once they have reached menarche.[2] This is primarily to avoid the brides having any pre-marital sexual experiences (or unwanted pregnancies) that would cast doubt on their 'purity'– a quality which is perceived to protect the honor of the bride's family as well as that of the family receiving the bride (Caldwell et al., 1983; Srinivas, 1984; Wahhaj, 2015). Consequently, the variation in the age at menarche generates a quasi-random difference in the age at which a girl enters the marriage market. This instrument has recently been used by Sekhri and Debnath (2014), Chari et al. (2017) and Asadullah and Wahhaj (2019) among a few others.

Our results are compelling. The ordinary least squares (OLS) results for the full sample indicate that a year of delayed marriage of women is associated with a reduction in all types of domestic violence considered. However, as noted above these effects are not necessarily causal but instead could arise due to unobserved factors. To distinguish causation from correlation, we use the instrumental variable (IV) two stage least squares approach. The first stage results for the IV are strong and rule out any concerns of weak instruments. The main IV results indicate a strong negative effect of women's age at marriage on less severe and severe forms of physical violence. Specifically, based on our preferred specification, we find that a delay in women's marriage by a year causes less severe physical violence to decrease by 7 percentage points and severe physical violence to decrease by 4 percentage points. Both these effects are statistically significant at 5% level of significance. However, the effect of women's age at marriage on sexual violence and emotional violence are not

[2]The NFHS 2015-16 data show that almost 30% women get married within 3 years of menarche and 60% within 5 years of menarche.

statistically significant. We carry out a battery of robustness checks and faslification tests to assess the robustness of our results. Our results, reassuringly, survive all these tests. Further, we examine some potential mechanisms driving our results, and provide suggestive evidence that the effect of women's marital age on physical violence arises because older brides (women who get married relatively late), as compared to younger brides (women who get married relatively early), are more educated and are married to more educated men. Overall, our findings bolster the relevance of policies that seek to delay marriages of women in reducing the prevalence of domestic violence in developing countries.[3]

Our study contributes to the growing body of literature that examine various possible determinants of domestic violence from a causal perspective. While the existing studies in this literature have looked at factors such as education (Erten and Keskin, 2018), income (Rivera et al., 2006; Angelucci, 2008; Bobonis et al., 2013) and intrahousehold bargaining power (Stevenson and Wolfers, 2006; Aizer, 2010) that could potentially explain the prevalence of domestic violence, none of them focus on the relationship between women's age at marriage and domestic violence. The studies that do look at how early marriage (or child marriage) impacts domestic violence, mostly report a negative correlation between them (see for example Oshiro et al., 2011; Santhya, 2011; Speizer and Pearson, 2011; Nasrullah et al., 2014; Rahman et al., 2014; Yount et al., 2016). These studies, however, fail to establish a causal relationship by accounting for the potential omitted variable bias or measurement error. Our work contributes to this literature by providing the first piece of causal evidence on the impact of women's age at marriage on domestic violence using a large-scale microdata.

Additionally, our study relates to the literature that looks at the impact of women's marital age on their economic well-being measured along various dimensions including schooling, health, labor market outcomes, gender norms and human capital of women's children in developing countries like Bangladesh and India (see for example Field and Ambrus, 2008;

[3] Our results are subject to a caveat. As we discuss in Section 3, the data on age at menarche are available for the women aged 15-25 in the survey. Thus, we had to restrict our analysis to women of this age group only.

Sekhri and Debnath, 2014; Chari et al., 2017; Asadullah and Wahhaj, 2019; Sunder, 2019; Dhamija and Roychowdhury, 2020). Given that child marriage and early marriage continue to be issues of deep concern in developing countries, this study by focusing on the relationship between women's age at marriage and domestic violence is likely to extend our understanding of the socioeconomic consequences of early marriage of women. Our findings, thus we believe, are likely to be useful for governments and policymakers in assessing the relevance and effects of policies that seek to delay marriages of women in developing countries.

The rest of the paper unfolds as follows. In section 2 we discuss the background and context of our study. In section 3 we discuss the dataset used. Section 4 presents the econometric model and empirical strategy. Results are presented in the section 5. The last section concludes.

2 Women's Age at Marriage and Domestic Violence in India

2.1 Women's Age at Marriage in India

The average age at marriage of women in India is 19.3 years according to the 2011 Census data. This is significantly lower compared to women's age at marriage in most developed countries, and also several developing countries. For example, as per the World Bank estimates, among the developed countries, the average age at marriage of women in the US was 26.9 years in 2011, in Germany 30.9 years, and in Sweden 33.3 years, and among the developing countries, the average age at marriage of women in Pakistan was 23.1 years, in Nepal 20.7 years, in Phillipines 23.4 years and in Indonesia 21.8 years.[4] More worrisome is that early marriage or child marriage is rampant in India. According to the UNICEF (2014), close to 30% of Indian girls are married before their 15th birthday, and almost 1 in

[4] https://datacatalog.worldbank.org/search?query=%20mean-age-first-marriage-female

3 child brides worldwide are in India. This is despite the Prohibition of Child Marriage Act (PCMA) 2006 clearly stating that the minimum legal age of marriage in India is 18 years for girls and 21 years for boys with no exceptions, and that child marriage is a cognizable and non-bailable offence.

As per the National Family Health Survey 2015-16 data, child marriage can be seen across India but it is far higher in rural than in urban areas. Further, girls from poorer families and excluded communities–scheduled castes and tribes–are more likely to marry at a younger age. According to UNICEF (2016), the states with the highest prevalence of child marriage (50% and above) are Bihar, Rajasthan, Jharkhand, Uttar Pradesh, West Bengal, Madhya Pradesh, Andhra Pradesh and Karnataka. However, even in states with overall lower prevalence of child marriage, there are often pockets of high prevalence.

While child marriage can happen to both boys and girls, the practice mostly affects girls. Girls often get married early because of pressure from parents and relatives, poverty, gender norms and lack of alternatives. As noted by Jensen and Thornton (2003) and UNICEF (2016), families may be unwilling to postpone their daughter's marriage due to the high premium placed on female virginity and fears of loss of sexual purity. Limited access to quality education and families' prioritization of boys' rather than girls' education–in part because of limited job opportunities–contribute to perpetuate the practice. Law enforcement to prohibit child marriage is also relatively weak. Limited detailed knowledge on how to apply laws and little understanding of the consequences of the laws, as well as limited trust in institutions enforcing them, undermines the implementation of the PCMA 2006.

2.2 Domestic Violence in India

Domestic violence has been recognized as a criminal offence under Indian Penal Code 498-A since 1983. However, it was only after the enactment of the Protection of Women from Domestic Violence Act 2005 (PWDVA), which came into effect in 2006, that civil protections could be provided to victims of domestic violence. The definition of domestic violence

under PWDVA is comprehensive: it includes all forms of physical, emotional, verbal, sexual, and economic violence, and covers both actual acts of such violence and threats of violence. Further, the PWDVA recognizes marital rape and covers dowry related harassment as forms of abuse. The Act requires the appointment of protection officers to assist victims, and also recognizes the importance of collaboration between the government and external organizations for protecting women. Although the PWDVA was primarily meant to provide protection from domestic violence for wives and female live-in partners at the hands of husbands and male live-in partners or their relatives, it has been extended to also protect women living in a household, such as sisters, widows, or mothers. However, despite the PWDVA, domestic violence continues to be a major challenge and a threat to women's lives in India.

According to the National Family Health Survey of India 2015-16, 33% of ever-married Indian women age 15-49 have experienced physical, sexual, or emotional spousal violence. The most common type of spousal violence is physical violence (30%), followed by emotional violence (14%). Seven percent of ever-married women report to have experienced spousal sexual violence. Of ever-married women who have experienced spousal physical or sexual violence, one-fourth report experiencing physical injuries, including 8% who have had dislocations, sprains, eye injuries, or burns and 5% percent who have had deep wounds, broken teeth, broken bones, or any other serious injury. Women's experience of any spousal physical, sexual, or emotional violence in India varies greatly by state, from 4% of women in Sikkim and 7% in Himachal Pradesh to close to 50% of women in Andhra Pradesh, Bihar, Telangana, and Tamil Nadu and 55% in Manipur. In most states, however, women in rural areas are more likely (36%) than women in urban areas (28%) to experience one or more forms of spousal violence.

3 Data

The data come from the NFHS 2015-16 (NFHS-4). The NFHS, a nationwide cross-section demographic health survey for India, provides information on various topics such as population demographics, health and nutrition for India. It is conducted by the International Institute for Population Sciences (IIPS) in Mumbai administered under the Ministry of Health and Family Welfare (MoHFW), Government of India, and is a part of the global Demographic Health Survey (DHS) program.[5] The NFHS-4 survey was conducted between January 2015 and December 2016, and covered 601,509 households located throughout India. The sample is drawn using stratified random sampling (see IIPS and ICF, 2017 for more details on the survey methodology).

The NFHS-4 administered a separate woman's questionnaire to collect information from all eligible women aged 15-49 years in the surveyed households. This questionnaire included questions on a variety of topics such as background characteristics, reproduction, prevalence of hysterectomy, menstrual hygiene, family planning, contacts with community health workers, maternal, child health, breast-feeding, nutrition, marriage, sexual activity, fertility preferences, husband's background, women's work, women's empowerment, HIV/AIDS, other health issues and domestic violence. The total number of women surveyed were 699,686.

Two things are worth noting in this context. First, the NFHS collected information on menstrual hygiene and other topics related to menstruation, including age at menarche, from the surveyed women in the age group of 15-25 years only. Second, the domestic violence information was collected for 79,729 women. These were the women who were from households belonging to the State module of the NFHS (which was implemented for a subsample of 15% of the households) since the domestic violence questionnaire was administered to a randomly selected woman from each household which was a part of the State module.[6]

[5] The DHS surveys for all countries are available at https://dhsprogram.com/

[6] Some households in the State module did not have eligible women who could answer questions on domestic violence. Also, in some households in the State module domestic violence questionnaire could not be administered since privacy could not be obtained or due to other issues.

The questions on domestic violence provide detailed information on physical, sexual, and emotional violence. Collecting valid and reliable data on domestic violence, however, poses serious challenges due to the sensitivity of the issue and the consequent difficulties in collecting correct and complete information, maintaining ethical concerns, ensuring safety of the respondent and interviewer, as well as protecting the women who disclose violence. However, as noted by Golder et al. (2016, p. 2), "all these issues are well addressed in the NFHS surveys. It follows both Indian and international guidelines, viz. WHO ethical guidance for research on domestic violence against women, 2001, for the ethical collection of data on violence."[7] For instance, as noted previously, only one woman per household was selected for the interviews.[8] Selecting only one woman for the domestic violence module even when there are more women eligible for interview, allows the interviewed respondent to keep the information confidential. Next, there was no one else in the room when the interviews were conducted. Further, the respondents were informed that their answers would be kept confidential and would not be told to anyone else and that no one else in the household would be asked these questions. Note, the domestic violence module was specially designed to allow the interviewer to continue the interview only if privacy was obtained. If privacy could not be obtained, the interviewer was instructed to skip the module, thank the respondent, and end the interview.[9]

The NFHS classifies domestic violence into four broad categories: less severe physical violence, severe physical violence, sexual violence, and emotional violence. Less severe physical violence includes acts of pushing, shaking, throwing something, twisting arm, pulling hair, slapping, punching with partner's fist or something else. Severe physical violence includes acts of kicking, beating, choking, burning, threatening or attacking with any kind of weapon.

[7] See http://www.who.int/gender/violence/womenfirtseng.pdf

[8] In households with more than one eligible woman, the woman administered the module was randomly selected through a specially designed sample selection procedure based on the "Kish Grid" which was built into the household questionnaire.

[9] For more on specificities about collection of data on domestic violence in NFHS, see NFHS data documentation (p. 496) available at http://rchiips.org/nfhs/NFHS-3%20Data/VOL-1/Chapter%2015%20-%20Domestic%20Violence%20(468K).pdf. Also see NFHS surveyor training manual (p. 8) at http://rchiips.org/nfhs/Manuals/DV_Training_Manual.pdf.

Sexual violence includes forced sexual acts, forced sexual relations resulting from the fear of what the partner would do otherwise, and humiliating sexual acts. Finally, emotional violence includes activities which caused women to face humiliation, insult, various kinds of threats from their partners to hurt the women or her closed ones.

For all the four categories of domestic violence, there is a binary variable for each underlying violence which takes a value one for the woman if she reports to have faced the underlying violence in the *last twelve months*, and zero otherwise (e.g., for less severe domestic violence, there is binary variable which takes a value one if the husband has ever pushed the woman in the last twelve months, zero otherwise; likewise there are binary variables corresponding to acts of shaking, throwing something, twisting arm, etc.). Additionally, for all the four categories of domestic violence, there is also binary variable for each underlying violence which takes a value one for the woman if she reports to have *ever* faced the underlying violence, and zero otherwise.

For our analysis, we create four binary variables – one for each of the four categories of domestic violence. For a given category of domestic violence, our binary variable takes a value one for the woman if she reports to have faced at least any one kind of the underlying violences in the *last twelve months*, and zero otherwise. We use information pertaining to domestic violence exposure in the last twelve months and refrain from using the information on postmarriage lifetime exposure to domestic violence to create our main outcome variables since using outcome variables based on postmarriage lifetime exposure to domestic violence might create a mechanical relationship between women's marital age and their exposure to domestic violence.[10]

The total number of ever-married women with valid information on domestic violence for whom we also have information on age at menarche is 10,468 (these are the women in the age group 15-25 years in the sample of women for whom we have valid information on domestic violence). For our analysis, we restrict ourselves to the women who have non-

[10] We thank a referee and the associate editor of this journal for highlighting this issue.

missing information on the different categories of domestic violence, whose marital age is not less than 5 years and menarcheal age is between 9 and 21 years, have valid information on age, spousal age, height, family attributes like caste, wealth, indicator for violence between parents in her natal home and place of residence (rural/urban), leaving us with a sample of 9,343 women.[11]

Table 1 provides descriptive statistics of our analytical sample. In our sample, 21% of the women have faced less severe physical violence, 5% have faced severe physical violence, 6% have faced sexual violence, and 10% have faced emotional violence.[12] The average age at marriage of women is 18.23 years and average age at menarche is 13.57 years. Figures 1 and 2 graph the distribution of the age at marriage and age at menarche respectively for our sample. Figure 3 graphs the scattered plot of percentage of women exposed to each category of domestic violence by age at marriage.

4 Empirical Strategy

4.1 Econometric Model

To examine the impact of women's age at marriage on their exposure to domestic violence, we begin by estimating the following econometric model:

$$y_i = \beta_0 + \beta_1 MarriageAge_i + \beta_2 X_i + \varepsilon_i \tag{1}$$

where y_i denotes a particular category of domestic violence against woman i, $MarriageAge_i$ denotes the woman's age at (first) marriage, X_i denotes the vector of individual and household level controls, and ε_i is the idiosyncratic error term that includes unobserved attributes like ability, social norms, discount rate, etc. Our parameter of interest is the coefficient β_1 which captures the effect of women's age at marriage on their exposure to domestic violence.

[11] See Appendix A for additional details about the analytical sample.

[12] The proportion of women who have faced at least one of the four kinds of domestic violence in our sample is 25%.

If we obtain $\beta_1 < (>)\ 0$, this indicates that women's age at marriage has a negative (positive) impact on the probability of their exposure to domestic violence. Note, while estimating equation (1), we exclude women's educational attainment and indicators of bargaining power (which are likely to be endogenous) from the estimation since educational attainment and bargaining power could be channels through which women's age at marriage affects domestic violence. Consequently, the estimated coefficient β_1 should be interpreted as the *total* effect of women's age at marriage on domestic violence.

We could have consistently estimated β_1 via OLS and interpreted it as causal effect of women's age of marriage on the level of domestic violence if, conditioning on exogenous characteristics, age at marriage was uncorrelated with unobservable determinants of physical, sexual and emotional violence against women (or more formally, $\mathbb{E}[MarriageAge\cdot \varepsilon|\ X] = 0$). However, such an assumption may be violated for several reasons. First, omitted variables may affect both the age at marriage of the women and probability of physical, sexual and emotional violence. For instance, classic patriarchy norms require women to marry young to exchange obedience for protection from men (e.g. Kabeer 1988; Alam 2007; Yount and Li 2010), and to respect men's authority to punish disobedience. Thus, those women who come from families that strictly follow such patriarchal norms are likely to get married early as well as believe that husbands can be justified in beating their wives – a belief that places them at higher risk for domestic violence. Unobserved ability of women might also be correlated with marital age and domestic violence. Specifically, more able women might get married late as well as be less victims of domestic violence since they are likely to have more bargaining power and more outside options of divorcing and economically supporting themselves or re-entering the marriage market after the divorce. Both these instances suggests that $\mathbb{E}[MarriageAge\cdot \varepsilon \mid X] \neq 0$. As a result, OLS estimates would be biased and inconsistent.[13][14]

[13] Note that both examples suggest that OLS estimates are likely to be biased downwards. In the first example, $\mathbb{E}[MarriageAge\cdot \varepsilon \mid X] < 0$ and the coefficient of (unobserved) patriarchy would be positive. In the second example, $\mathbb{E}[MarriageAge\cdot \varepsilon \mid X] > 0$, and the coefficient of (unobserved) ability would be negative.

[14] In principle, there might be other potential omitted variables which are not orthogonal to age of marriage of the women and might be correlated with their exposure to domestic violence.

The second issue relates to the accuracy of the reported age of marriage. In the NFHS 2015-16, age at marriage was self reported. Inaccurate reports would generate measurement error in the explanatory variable. This could attenuate the estimates of the coefficient of interest. To address these concerns, we follow an instrument variable (IV) approach. We use age of menarche as an instrument for women's age at marriage. This instrument is motivated by the observation that has been made by sociologists and anthropologists that parents become extremely anxious to get their daughter married once she has reached menarche. This is primarily to avoid the brides having any pre-marital sexual experiences (or unwanted pregnanicies) that would cast doubt on their 'purity'– a quality which is perceived to protect the honor of the bride's family as well as that of the family receiving the bride (Caldwell et al., 1983; Srinivas, 1984; Wahhaj, 2015). As noted by Field and Ambrus (2008), a significant portion of the variation in timing of menarche is random, rendering it a good instrument for the age at marriage.[15] In what follows, we discuss our IV strategy in detail.

4.2 IV Strategy

We estimate a two stage IV model which is specified as follows:

$$MarriageAge_i = \alpha_0 + \alpha_1 MenarcheAge_i + \alpha_2 X_i + \eta_i \quad (2)$$

$$y_i = \beta_0 + \beta_1 MarriageAge_i + \beta_2 X_i + \varepsilon_i \quad (3)$$

The women's age at marriage, $MarriageAge_i$, is instrumented by $MenarcheAge_i$, their age at menarche, and y_i are the different categories of domestic violence against woman i. As above, X_i denotes a vector of individual and household level controls such as the woman's age, height, wealth, place of residence (urban/rural), spousal age, caste and district fixed effects. We use a standard two stage estimation procedure (i.e., two stage least squares

[15] Studies of twins have found that random genetic variation is the single largest source of variations in menarche (see for example Kaprio et al., 1995)

(TSLS)) and cluster standard errors at the district level.[16]

4.3 Examining IV Validity and Threats to Identification

In this section, we examine whether age at menarche can be used as a valid instrumental variable for women's age at marriage. The results from the regression of women's age at marriage on age at menarche are presented in Table 2. We begin by examining whether age at menarche predicts women's age at marriage which is the endogenous regressor in absence of any control variables. The results are reported in Column (1). As evident, age at menarche is significantly correlated with women's age at marriage. The value of the estimated coefficient is 0.216 and it is statistically significant at 1% level of significance. The F-Statistic for the regression model is 77.63. Additionally, Figure 4 also presents the kernel density estimate of women's age at marriage by menarcheal age groups (early and late menarche)[17] revealing that the distributions of women's age at marriage is positively related to age at menarche.[18]

Next, we examine the potential threats to the validity of this instrument by controlling for various factors in the first stage regression that might potentially be determining the age at which women attain menarche as well as be directly affecting their marriage timing. First, severe malnutrition in early childhood might result in delayed onset of menarche (Sekhri and Debnath, 2014). Exposure to severe malnutrition could potentially also affect long term health of the women (for example Stathopolu et al. (2003) note that acute malnutrition could result in stunting) and their labor market prospects, in turn reducing their options

[16] Later we use an alternative non-linear method of estimation to assess the robustness of our baseline results. However, for our baseline analysis we use a linear approach since, as noted by Wooldridge (2010), "this procedure [IV-TSLS] is relatively straightforward and might provide a good estimate of the average effect." Angrist and Pischke (2008, p. 107) also argue "...while a nonlinear model may fit the CEF (conditional expectation function) for LDVs (limited dependent variable models) more closely than a linear model, when it comes to marginal effects, this probably matters little. This optimistic conclusion is not a theorem [but]...it seems to be fairly robustly true."

[17] The early menarche group consists of those women who attained menarche at the age of 14 or earlier. The late menarche group consists of those women who attained menarche after the age of 14.

[18] In Figure A1 in the Appendix, we also graph the scattered plot of average age at marriage by age at menarche with a linear fitted line. The graph shows clear evidence of significant positive relationship between age at menarche and age at marriage.

outside marriage. This suggest that malnutrition, by affecting long term health, could make women more vulnerable to physical, sexual and emotional violence**.** Consequently, as a proxy for severe malnutrition in childhood, following Chari et al. (2017), we include adult height in the regression in Column (2). In addition, column (2) also include controls for women's age and spousal age. We find that inclusion of these controls change the point estimate of the coefficient of age at menarche slightly (the standard errors remain almost unchanged).

Second, it is thought that physical labor during childhood can have a negative effect on children's health and lead to a delay in menarche (Pellerin-Massicotte et al., 1987). Thus women who end up marrying late may also be less healthy, and this could have a direct effect on her emotional ability to resist domestic violence or her divorce-based outside options. To address this concern it would be ideal to include controls for economic status of women's natal family such as parental education and income. However, unfortunately, we do not have information on these variables in our dataset. To circumvent this issue, we include controls for wealth level of women's spousal household (more specifically, indicators for which quintile of the wealth distribution the women's spousal household belongs) and a set of indicator variables for caste. The inclusion of the first variable can be justified on the grounds that a woman is likely to get married into a family which belongs to a more or less similar economic status as her natal family. As noted in a recent article in *The Economist* (2017), "the idea that the best marriage partner is someone with the same family background and belonging to precisely the same social group seems to be rooted in the [Indian] subcontinent." As such, it is likely that the women's natal family belongs to the same quintile of the wealth distribution to which the women's spousal household belongs, and hence the wealth variables are likely to serve as good proxies for the economic status of women's natal family. Finally, as noted by Nayar (2007), in the Indian context, caste might be considered as a proxy for socioeconomic status and poverty. Consequently, we include caste affiliation of women as an additional control as it is likely to serve as a good control for their natal family economic status. As evident from the results reported in Column (3), the inclusion of the proxies for women's

natal family characteristics as additional controls does not change the point estimates of the coefficient of age at menarche significantly.[19]

Third, it is noted by some studies that age of menarche is influenced by inter-parental violence that the women face in their childhood (see for example Henrichs et al., 2014). The argument is that inter-parental violence – by acting as a stressor in childhood – may have biological influences on endocrine development resulting in early menarche. Since, experience of inter-parental violence during childhood might also affect women's attitude towards domestic violence when she gets married, not controlling for this might render the age of menarche variable endogenous. As such, we report regression results that include a dummy variable indicating whether a woman reports to have had experienced any kind of inter-parental violence before marriage (in addition to the control variables used in Column (3)) in Column (4). Reassuringly, the point estimate of the coefficient of age of menarche does not change significantly.

Fourth, age at menarche might also be potentially endogenous due to geographical factors such as temperature, rainfall, altitude, etc. (Field and Ambrus, 2008; Chari et al., 2017). To address this issue, we control for place of residence (whether the household resides in an urban or a rural locality) and use district fixed effects to account for spatial variation in exposure to environmental factors that affect menarche. Note, we are able to control for district of residence of the married woman, and not her natal district since we do not have any information about the location of her natal family. This again, however, is not likely to be a problem because in India most marriages occur within the same district, so the district of residence of the married woman is also likely also her natal district (Fulford, 2015). The results of the specification that include geographic controls, in addition to the controls included in Column (4), is presented in Column (5). The coefficient of age at menarche is still highly statistically significant and the first stage F-statistic is sufficiently high.

[19] Note, although not caste, but spouse's wealth level may be endogenous to marriage. For instance, parents who are in a hurry to marry their daughters might have a lower reservation quality of spouse, as reflected in their wealth. However, this is unlikely to cause the IV estimate of the effect of women's age at marriage on domestic violence inconsistent since age at menarche is unlikely to be correlated with spouse's wealth.

The next concern that we need to address is whether our instrument is exogenous given that we are not controlling for education which is a potential determinant of women's exposure to domestic violence. One might argue that a woman's educational attainment as measured by her years of schooling, is correlated with her age at menarche. More specifically, menarche itself might be a barrier to schooling (as often cited in the popular media). If this is the case, then leaving out education from the set of control variables will violate the condition that $\mathbb{E}[MenarcheAge \cdot \varepsilon | X] = 0$, and the IV regressions will not yield consistent estimates of the parameters of interest.

While this is possible, Field and Ambrus (2008) in their seminal paper provide robust evidence that menarcheal age has no direct impact on women's schooling using data from Bangladesh. Oster and Thornton (2011) although document a statistically significant effect of menstruation on school attendance for girls in Nepal, this effect is extraordinarily small. Specifically, they estimate that girls miss a total of only 0.4 days in a 180 day school year (although 47 percent of the girls in their study reported missing some school due to menstruation in the past year). Further, Oster and Thornton (2011) show that improved sanitary technology has no effect on reducing this small gap: girls who randomly received sanitary products were no less likely to miss school during their period. Grant et al. (2013) conduct a study in Malawi to examine the individual and the school level factors associated with menstruation-related school absenteeism. In line with the findings of Field and Ambrus (2008) and Oster and Thornton (2011), they find no evidence that menstrual periods account for female absenteeism. Thus, even though it is often believed that menstruation causes girls to be absent from school, these findings indicate that in reality it is unlikely to be the case.

Nevertheless, to address the concern that our instrument might potentially be endogenous due to omission of schooling from our model, we do the following. First, we plot the average years of schooling of women by different menarcheal age in Figure 5. We find no evidence of upward trend in the relationship between schooling and age at menarche of women (barring the few who attained menarche at 21 years of age). Second, we present the kernel density

estimate of women's years of schooling by terciles of menarcheal age in Figure 6. The figure reveals that the population distributions, and not just averages, are remarkably similar across all subsamples. This is not what we would have expected to find if menarcheal age was correlated with years of schooling. This suggests that not controlling for educational attainment of women is unlikely to confound our analysis. We provide further assessment of the possible correlation between women's age at menarche and educational attainment as well as several other robustness tests in the section on robustness checks.[20] [21]

5 Results

5.1 OLS Results

The OLS estimates of the effect of women's age at marriage on domestic violence are presented in Table 3. Columns (1), (4), (7) and (10) report the coefficient of age at marriage from the regression equations where we do not include controls for demographic characteristics or district fixed effects. Columns (2), (5), (8) and (11) report the coefficient of age at marriage from the regressions where we include controls for demographic characteristics but not district fixed effects. Finally, Columns (3), (6), (9) and (12) report the coefficient of age at marriage from the regressions where we include controls for demographic characteristics as well as district fixed effects. While these estimates are not causal, nevertheless they are likely to serve as useful benchmarks with which we would be able to compare our IV estimates.

Examining the results of regression models without any demographic controls or district

[20] In Figure A2 in the Appendix, we also graph the scattered plot of average years of schooling by age at menarche with a linear fitted line excluding the outliers. The graph does not show any evidence of significant positive relationship between age at menarche and years of schooling.

[21] Note, Sekhri and Debnath (2014) and Chari et al. (2017) also implicitly assume that age of menarche is not correlated with women's education. Both the papers investigate the impact of marital age of the mother on child health and education outcomes. Marital age is instrumented by menarcheal age, but mother's education is not controlled for. Given that mother's education is conjectured to a determinant of child outcomes, mother's education becomes of the part of the error term in the second stage regressions, which must be assumed to be uncorrelated to menarcheal age, for their second stage parameter estimates to be consistent.

fixed effects, we find that a year of delay in marriage is associated with a decrease in the probability of women's exposure to less severe physical violence by 2.2 percentage points, severe physical violence by 0.8 percentage points, sexual violence by 0.7 percentage points, and emotional violence by 1.3 percentage points respectively. These effects are statistically significant at 1% level of significance. When we include controls for only demographic characteristics, and controls for demographic characteristics as well as district fixed effects, the estimates of the coefficients of age at marriage on different categories of domestic violence remain roughly unchanged. Overall, thus, the OLS results appear to be suggesting that the net effect of women's age at marriage on domestic violence is negative. To examine whether this effect is causal or purely arises due to omitted characteristics such as family norms and/or ability, we use the IV approach.

5.2 IV Results

We next turn to the IV results in Table 4. Based on the specifications in which we do not include controls for demographic characteristics and district fixed effects, we find that a delay in marriage of women by a year leads to a 6.7 percentage point decline in the probability of less severe physical violence, 3.7 percentage point decline in the probability of severe physical violence, 1.6 percentage point decline in the probability of sexual violence, and 2.8 percentage point decline in the probability of emotional violence. The effects of women's age at marriage on less severe physical violence, severe physical violence and emotional violence are statistically significant at 1% , 1% and 5% level of significance respectively. However, the effect of women's age at marriage on sexual violence is not statistically significant.

When we include controls for demographic characteristics, these estimates change slightly: a one year delay in women's marriage now leads to a 6.1 percentage point decline in probability of less severe physical violence, a slightly over 4 percentage point decline in probability of severe physical violence, a 0.5 percentage point decline in probability of sexual violence, and a 1.7 percentage point decline in probability of emotional violence. However, now, al-

though the effects of women's age at marriage on less severe physical violence and severe physical violence are statistically significant at 5% and 1% level of significance respectively, the effect on emotional violence is no longer statistically significant. The effect of women's age at marriage on sexual violence continues to remain statistically insignificant.

Our preferred IV specifications are the ones that are reported in Columns (3), (6), (9) and (12). Based on our preferred specifications, we find that the magnitude of the effect of women's age at marriage on less severe physical violence increases slightly and the magnitude of the effect of women's age at marriage on severe physical violence remains almost unchanged compared to the magnitudes of the corresponding effects obtained from the specifications that include only demographic controls. Specifically, controlling for demographic characteristics and district fixed effects, a one year delay in marriage of women causes the probability of less severe and severe physical violence to decrease by approximately 7 percentage points and 4 percentage points respectively. These effects are statistically significant at 5% level of significance. The effects of women's age at marriage on sexual violence and emotional violence continue to remain small and statistically insignificant.[22]

These results indicate that a one year increase in women's age at marriage nationwide would reduce the prevalence of less severe physical violence from 21% of women to 14%, and that of severe physical violence from 5% to 1%. If one is willing to extrapolate these result from our sample to the entire India, the implications of our finding are extremely striking. Given that female population in India as per the 2011 Census is 586 million of whom 50% are married (Government of India, 2013), our findings imply that a nationwide delay in women's age at marriage by a year would cause the number of women exposed to less severe physical violence to fall from 62 million to 41 million, and the number of women exposed to severe

[22] The magnitude of the coefficients of age at marriage in the regressions with less severe and severe physical violence as outcome variables are quite large relative to the population average. One can argue that this might be due to how we have defined our outcome variables. To check whether the size of the estimated coefficients are sensitive to our baseline definition of outcome variables, in the Appendix, we follow Kling et al.'s approach (2007) and create z-score based indices of domestic violence and repeat our baseline analysis. We continue to find large and economically significant effect of age at marriage on physical violence. This suggests that size of the estimated coefficients are not driven by our definition of the outcome variables.

physical violence to fall from 15 million to 3 million.

In sum, thus, our IV results suggest that a year of delay in marriage causes a significant reduction in women's exposure to less severe as well as severe forms of physical violence, but has no impact on sexual violence or emotional violence.[23]

Why does age at marriage not affect sexual and emotional violence? A possible reason for us not finding any effect of women's age at marriage on sexual and emotional violence but finding statistically significant effect of age at marriage on physical violence is the following. Sexual violence and emotional violence, unlike physical violence, are hard to recognize to start with. The problem of under-recognition of sexual and emotional violence is, in fact, likely to be more pronounced for younger brides. Younger brides, given their greater disconnect with the society and lack of experience in relationships (due to early marriage) might take "being forced to have sex with husband" (sexual violence) or "being mentally tortured" (emotional violence) as normal. As such, even though younger brides might actually be subject to higher sexual and emotional violence than older brides, many of them might not recognize that, and therefore end up under-reporting their exposure to sexual and emotional violence. Had the younger brides really been able to recognize sexual and emotional violence, we might have obtained a significant negative relationship between

[23]It is worth noting that the IV estimates of age at marriage, in general, are larger than the corresponding OLS estimates. This might be because of omitted factors like classical patriarchy or ability of women. As discussed previously, if the omitted factor is classical patriarchy, the covariance between the omitted factor and marriage age would be negative and the coefficient of unobserved patriarchy should be positive implying the sign of the bias to be negative. For the case of omitted ability, the covariance is likely to be positive and the coefficient of unobserved ability should be negative again rendering the sign of the bias as negative. IV estimates could be larger than OLS estimates might be due to measurement error in age at marriage as well. Measurement error in marriage will tend to attenuate the OLS coefficients but not the IV ones. Further, as pointed out by Chari et al. (2017), it is also important to note that the local average treatment effect interpretation of an instrumental variable estimate implies that we are estimating the causal effect of marriage and for the subpopulation whose marriage timing is affected by the instrument, i.e., menarche. It is possible that causal effects for this subpopulation are larger than those for the population as a whole which might be the reason why we find the coefficient estimates from the IV regressions to be larger than those from the OLS regressions.

age at marriage and sexual and emotional violence.[24][25]

5.3 Robustness Checks

To assess the robustness of our results, we carry out a battery of robustness tests. We show that our results are robust to an alternative non-linear method of estimation, inclusion of survey year fixed effects, inclusion of birth year fixed effects, use of alternative measures of domestic violence, inclusion of square of age and spousal age at marriage in the set of control variables, removing observations with top and bottom 1 percentile of age at menarche and age at marriage, and using a proxy variable to control for women's exposure to family violence during childhood.

Additionally, to assess the validity of our instrument, we also perform a falsification test by trying to find a systematic reduced form effect of age at menarche on domestic violence among the subsample of women who got married before attaining menarche. The results of this exercise indicate that age at menarche has no effect on domestic violence (which is what we expect if age at menarche affects domestic violence through age at marriage), and suggest that our instrument is likely to satisfy the exclusion restriction.

Further, we employ two carefully designed tests to rule out potential concerns regarding our IV being endogenous due to measurement error in age at menarche and omitted educational attainment. Reassuringly, our main results survive both these tests, and thus increase our confidence in the empirical strategy employed.

Finally, we examine whether the effect of women's age at marriage on domestic violence

[24] That under-recognition is a potential issue at least with sexual violence has been also noted by Raj et al. (2010) and Rahman et al. (2014). For physical violence, however, this problem does not arise. Physical violence is well-defined, and even younger brides can easily say whether they were physically abused by their husbands or not.

[25] Using a simple two variable IV regression model in which emotional/sexual violence is regressed on age at marriage, and age at marriage is instrumented by age at menarche, it can be shown that if emotional/sexual violence is underreported (due to under-recognition) and this under-reporting varies inversely with age at marriage, the IV estimate of the coefficient of age at marriage obtained using the underreported emotional/sexual violence data is higher than the IV estimate of the coefficient of age at marriage that would be obtained if one had access to emotional/sexual violence data without misreporting ('true' IV estimate). This suggests that if sexual/emotional is underreported, it is possible for one to find the IV estimate of age at marriage to be zero (or even positive), when in fact the true IV estimate is negative.

varies by caste and spousal age. We find no evidence of significant heterogeneity. We relegate the discussion and results of the robustness tests to the Appendix (Figures A1-A3 and Tables A1-A11).

5.4 Discussion of Underlying Mechanisms

Our results indicate that age at marriage negatively affects women's exposure to less severe and severe physical violence. In this section we examine four potential channels through which this effect might be operating, namely, women's education, participation in workforce, bargaining/decision power within households, and husbands' education and labor market outcomes. It is possible that older brides, compared to younger brides, could be facing less physical violence either because they are more educated, they are more likely to participate in the workforce, they have greater bargaining/decision power or because their husbands have better educational and labor market outcomes (i.e., they are married well).

To examine the channels, we regress women's educational outcomes, an indicator for women's for women's workforce participation, indicators for women's bargaining/decision power within households, and husbands' educational and labor market outcomes on women's age at marriage.[26] The results are presented in Table 5. In all regressions, women's age at marriage is instrumented by their age at menarche, and the full set of controls, including district fixed effects, are used.

We find that age at marriage has a positive effect on the probability of women being literate, women completing primary education and years of education (or years of schooling). The estimated coefficients are not only statistically significant, but also large in terms of magnitudes. For example, the results for years of education completed indicates that a delay in marriage by one year causes women, on average, to stay in school for slightly more than a year.[27] Additionally, we find strong evidence that women's age at marriage has a positive

[26] Table A12 in the Appendix present the definitions and summary statistics of all these additional outcome variables.

[27] This, in other words, mean that if marriage is delayed by a year, an average woman continues going to

effect on their husbands' years of education. Specifically, our results indicate that a delay in marriage by one year causes women, on average, to get married to men who have 0.77 more years (or around 9 more months) of schooling. We also find that a delay in age at marriage positively impacts women's probability to have access to a bank account which is one of metrics we use to measure women's bargaining/decision power. However, this effect is statistically significant at only 10% level of significance. We do not find any evidence of causal effect of women's age at marriage on the other bargaining power measures (i.e., say in healthcare and say in large household purchases). Also we fail to find any causal effect of women's age at marriage on the probability of workforce participation or on the labor outcomes of their husbands (i.e., the probability of workforce participation and probability of working in a white-collar job). Altogether, these findings suggest that a delay in age at marriage reduces less severe and severe physical violence mainly because older brides, compared to younger brides, are likely to be more educated and are likely to be married to more educated men.

Before concluding this section, it is worth emphasizing that the fact that, compared to the "less educated" younger brides, the "more educated" older brides are likely to face lesser domestic violence despite the older brides not having a higher likelihood of participation in workforce than the younger brides, suggests that education in itself acts as a deterrent to domestic violence. This is perhaps because the options outside marriage that are available to a more educated woman (irrespective of whether she is employed or not) are likely to significantly higher than that available to a less educated woman. Also, it is likely that more educated women might have easier access to divorce or be less concerned about social taboos surrounding divorce than less educated women.

school for that full year and drops out of school right after marriage.

6 Conclusion

Domestic violence affects one in three women in their lifetime. It remains a crucial problem with adverse health and economic consequences in both developed and developing countries. The cost of domestic violence to an economy in terms of victim's suffering, medical expenses, lost productivity and judiciary expenses is massive. In this paper, we examine the causal impact of age at marriage on domestic violence against women using newly available household data from India. We focus on four types of domestic violence against women: less severe physical violence, severe physical violence, sexual violence, and emotional violence. The main empirical challenge in identifying the causal effect of age at marriage on prevalence of domestic violence is that marriage age might be endogenous due to omitted variables and/or measurement error. To address this issue, we use an empirical strategy that utilizes variation in age at menarche to obtain exogenous variation in women's age at marriage. We find that a one-year delay in marriage of women causes a significant decline in less-severe and severe forms of physical violence but has no impact on sexual or emotional violence. Further, we show that the effect of women's marital age on physical violence operates mainly through the channels of own education and husbands' education, i.e., a delay in age at marriage reduces the probability of physical violence primarily because older brides, as compared to younger brides, are more educated and are married to more educated men.

Our findings underscore the importance of better enforcement of existing social policies that seek to delay marriages of women (e.g. "Kanyashree Prakalpa" program in West Bengal, "Apni Beti Apna Dhan" program in Haryana, etc.), as well as formulation of newer and more innovative interventions, to reduce the prevalence of domestic violence in India. Future work should focus on testing that external validity of our findings by replicating our study for not only other developing nations, but also for developed nations since domestic violence is a major public health issue worldwide.

References

Ackerson, L.K. & Subramanian, S. (2008). Domestic violence and chronic malnutrition among women and children in India, *American Journal of Epidemiology*, 167, 1188–1196.

Aizer, A. (2010). The Gender Wage Gap and Domestic Violence, *American Economic Review*, 100, 1847–1859.

Alam, S. (2007). Islam, culture, and women in a Bangladesh village, In V.J. Cornell (Ed.), *Voices of Islam*, 35–53. Westport, CT: Praeger Publishers.

Angelucci, M. (2008). "Love on the Rocks: Domestic Violence and Alcohol Abuse in Rural Mexico." *B.E. Journal of Economic Analysis and Policy*, 8, Article 29.

Angrist, J.D., & Pischke, J.S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Asadullah, M.N., & Wahhaj, Z. (2019). Early marriage, social networks and the transmission of norms. *Economica*, 86, 801-831.

BBC (2014). "100 Women 2014: Violence at home is India's 'failing'," BBC (29 October, 2014). Retrieved from https://www.bbc.co.uk/news/world-asia-india-29708612

Bloch, F. & Rao, V. (2002). Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India, *American Economic Review*, 92, 1029–1043.

Bobonis, Gustavo J., Gonzalez-Brenes, M., & Castro, R. (2013). PublicTransfers and Domestic Violence: The Roles of Private Information and Spousal Control, *American Economic Journal: Economic Policy*, 5, 179–205.

Caldwell, J., Reddy, P., & Caldwell, P. (1983). Causes of marriage change in South India, *Population Studies*, 37, 343–361.

Campbell, J.C. (2002). Health Consequences of Intimate Partner Violence, *Lancet*, 359, 1331–1336.

Chari, A.V., Heath, R., Maertens, A., & Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India, *Journal of Development Economics*, 127, 42–55.

Chowdhury, F.D. (2004). The socio-cultural context of child marriage in a Bangladeshi village, *International Journal of Social Welfare*, 13, 244–253.

Coker, A. L., Davis, K. E., Arias, I., Desai, S., Sanderson, M., Brandt, H. M., & Smith, P. H. (2002). Physical and mental health effects of intimate partner violence for men and women. *American Journal of Preventive Medicine*, 23, 260–268.

Dhamija, G., & Roychowdhury, P. (2020). Age at Marriage and Women's Labour Market Outcomes in India. *Journal of International Development*, 32, 342-374.

Ellsberg, M., Jansen, H. A., Heise, L., Watts, C. H., & Garcia-Moreno, C. (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, 1165–1172.

Erten, B., & Keskin, P. (2018). For Better or for Worse?: Education and the Prevalence of Domestic Violence in Turkey, *American Economic Journal: Applied Economics*, 10, 64–105.

Eswaran, M., & Malhotra, N. (2011). Domestic Violence and Women's Autonomy in Developing Countries: Theory and Evidence, *Canadian Journal of Economics*, 44, 1222–1263.

Farmer, A., & Tiefenthaler, J. (1996). Domestic Violence: The Value of Services as Signals, *American Economic Review*, 86, 274–279.

Feldman, S. (2010). "Shame and honour: The violence of gendered norms under conditions of global crisis," *Women's Studies International Forum*, 33, 305–315.

Field, E., & Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh, *Journal of Political Economy*, 116, 881–930.

Field, E., Pande, R., Rigol, N., Schaner, S., and Moore, C.T. (2016). On Her Account: Can Strengthening Women's Financial Control Boost Female Labor Supply?, Working Paper.

Fulford, S. L. (2015). Marriage migration in India: Vast, varied, and misunderstood, Working Paper 820, Boston College.

Golder, S., Agrawal, N., Bhogal, R.K., Tada, P.R., Menon, R.R., Thekkudan, J., Thomas, M., & Laddha. R. (2016). *Measurement of Domestic Violence in NFHS Surveys and Some Evidence*, Oxfam India.

Government of India (2013). *Sample Registration System Statistical Report 2011.* Report No. 1 of 2013, Office of the Registrar General, Ministry of Home Affairs, New Delhi. Available at https://censusindia.gov.in/vital_statistics/SRS_Reports.html

Grant, M., Lloyd, C., & Mensch, B. (2013). Menstruation and School Absenteeism: Evidence from Rural Malawi, *Comparative Education Review*, 57, 260–284.

Henrichs, K. L., McCauley, H. L., Miller, E., Styne, D. M., Saito, N., & Breslau, J. (2014). Early menarche and childhood adversities in a nationally representative sample, *International Journal of Pediatric Endocrinology*, 2014, 14.

Hidrobo, M., & Fernald, L. (2013). Cash Transfers and Domestic Violence, *Journal of Health Economics*, 32, 304–319.

International Institute for Population Sciences (IIPS) & ICF (2017). *National Family Health Survey (NFHS-4), 2015-16: India.* IIPS, Mumbai.

Jensen, R., and Thornton, R. (2003). "Early female marriage in the developing world," *Gender & Development*, 11, 9–19.

Kabeer, N. (1988). Subordination and struggle: Women in Bangladesh, New Left Review I, 168, 95–121.

Kaprio, J., Rimpelä, A., Winter, T., Viken, R. J., Rimpelä, M., & Rose, R. J. (1995). Common genetic influences on BMI and age at menarche, *Human Biology*, 739–753.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects, *Econometrica*, 75, 83–119.

Le Strat, Y., Dubertret, C., & Le Foll, B. (2011). Child Marriage in the United States and Its Association With Mental Health in Women, *Pediatrics*, 128, 524–530.

Lomborg, B., & Williams, M.A. (2018). "The cost of domestic violence is astonishing," The Washington Post (22 February 2018). Retrived from https://www.washingtonpost.com/opinions/the-cost-of-domestic-violence-is-astonishing/2018/02/22/f8c9a88a-0cf5-11e8-8b0d-891602206fb7_story.html?noredirect=on

Nasrullah, M., Zakar, R., & Zakar, M. Z. (2014). Child marriage and its associations with controlling behaviors and spousal violence against adolescent and young women in Pakistan, *Journal of Adolescent Health*, 55, 804–809.

Nayar, K.R. (2007). Social exclusion, caste & health: a review based on the social determinants framework, *Indian Journal of Medical Research*, 126, 355–363.

Nour, N. W. (2009). Child Marriage: A Silent Health and Human Rights Issue, *Review of Obstetric Gynecology*, 2, 51–56.

Oshiro, A., Poudyal, A. K., Poudel, K. C., Jimba, M., & Hokama, T. (2011). Intimate partner violence among general and urban poor populations in Kathmandu, Nepal, *Journal of Interpersonal Violence*, 26, 2073–2092.

Oster, E., & Thornton, R. (2011). Menstruation, Sanitary Products, and School Attendance: Evidence from a Randomized Evaluation, *American Economic Journal: Applied Economics*, 3, 91–100.

Pellerin-Massicotte, J., Brisson, G.R., St.-Pierre, C., Rioux, P., & Rajotte, D. (1987). Effect of Exercise on the Onset of Puberty, Gonadotropins, and Ovarian Inhibin, *Journal of Appllied Physiology*, Vol. 63, pp. 1165–1173.

Rahman, M., Hoque, A., Mostofa, G., & Makinoda, S. (2014). Association between adolescent marriage and intimate partner violence: A study of young adult women in Bangladesh, *Asia-Pacific Journal of Public Health*, 26, 160–168.

Raj, A., Saggurti, N., Lawrence, D., Balaiah, D., & Silverman, J. G. (2010). Association between adolescent marriage and marital violence among young adult women in India, *International Journal of Gynecology & Obstetrics*, 110, 35–39.

Rivera, L., Hernández, B., & Castro, R. (2006). "Asociación entre la violencia de pareja contra las mujeres de las zonas urbanas en pobreza extrema y la incorporación al Programa Oportunidades." In *El Programa Oportunidades examinado desde el género*, edited by María de la Paz López and Vania Salles, 69–94. Mexico City: Colegio de Mexico.

Santhya, K. G. (2011). Early marriage and sexual and reproductive health vulnerabilities of young women: A synthesis of recent evidence from developing countries, *Current Opinion in Obstetrics and Gynecology*, 23, 334–339.

Sekhri, S., & Debnath, S. (2014). Intergenerational consequences of early age marriages of girls: effect on children's human capital, *Journal of Development Studies*, 50, 1670–1686.

Speizer, I. S., & Pearson, E. (2011). Association between early marriage and intimate partner violence in India: A focus on youth from Bihar and Rajasthan, *Journal of Interpersonal Violence*, 26, 1963–1981.

Srinivas, M. (1984). *Some Reflections on Dowry.* Oxford University Press, New Delhi, India.

Stathopolu, E., J. Antony H., & David C. (2003). Difficulties with Age Estimation of Internet Images of Southeast Asians, *Child Abuse Review*, 12, 46–57.

Stevenson, B., & Wolfers, J. (2006). Bargaining in the Shadow of the Law: Divorce Laws and Family Distress, *Quarterly Journal of Economics*, 121, 267–288.

Sunder, N. (2019). Marriage Age, Social Status and Intergenerational Effects in Uganda, *Demography*, 56, 2123-2146.

The Economist (2017). "Marriage in India is becoming less traditional," *The Economist* (23 November 2017). Retrieved from https://www.economist.com/news/special-report/21731491-parents-lose-control-over-marriage-indian-society-shaking-marriage-india

UNICEF (2014). *Ending Child Marriage: Progress and prospects*, UNICEF, New York. Retrieved from https://www.unicef.org/media/files/Child_Marriage_Report_7_17_LR..pdf

UNICEF (2016). Reducing child marriage in India: a model to scale up results. Retrieved from https://www.girlsnotbrides.org/wp-content/uploads/2016/12/Centre-for-Budget-and-Policy-Studies-India-report.pdf

Yount, K. M., & Li, L. (2010). Domestic violence against married women in Egypt, *Sex Roles*, 63, 332–347.

Yount, K. M., Crandall, A., Cheong, Y. F., Osypuk, T. L., Bates, L. M., Naved, R. T., & Schuler, S. R. (2016). Child marriage and intimate partner violence in rural Bangladesh: a longitudinal multilevel analysis, *Demography*, 53, 1821-1852.

Yount, K. M., Crandall, A., & Cheong, Y. F. (2018). Women's Age at First Marriage and Long-Term Economic Empowerment in Egypt, *World Development*, 102, 124–134.

Wahhaj, Z. (2015). An economic model of early marriage, *Journal of Economic Behavior & Organization*, 152, 147-176.

World Health Organization (2012). World Health Organization. (2012). Understanding and addressing violence against women: Intimate partner violence (No. WHO/RHR/12.36). World Health Organization. Retrieved from http://apps.who.int/iris/bitstream/handle/10665/77432/WHO_RHR_12.36_eng.pdf;jsessionid=6CADDFE06611AE8566C228C40575BE46?sequenc

Figure 1. Distribution of women's age at marriage

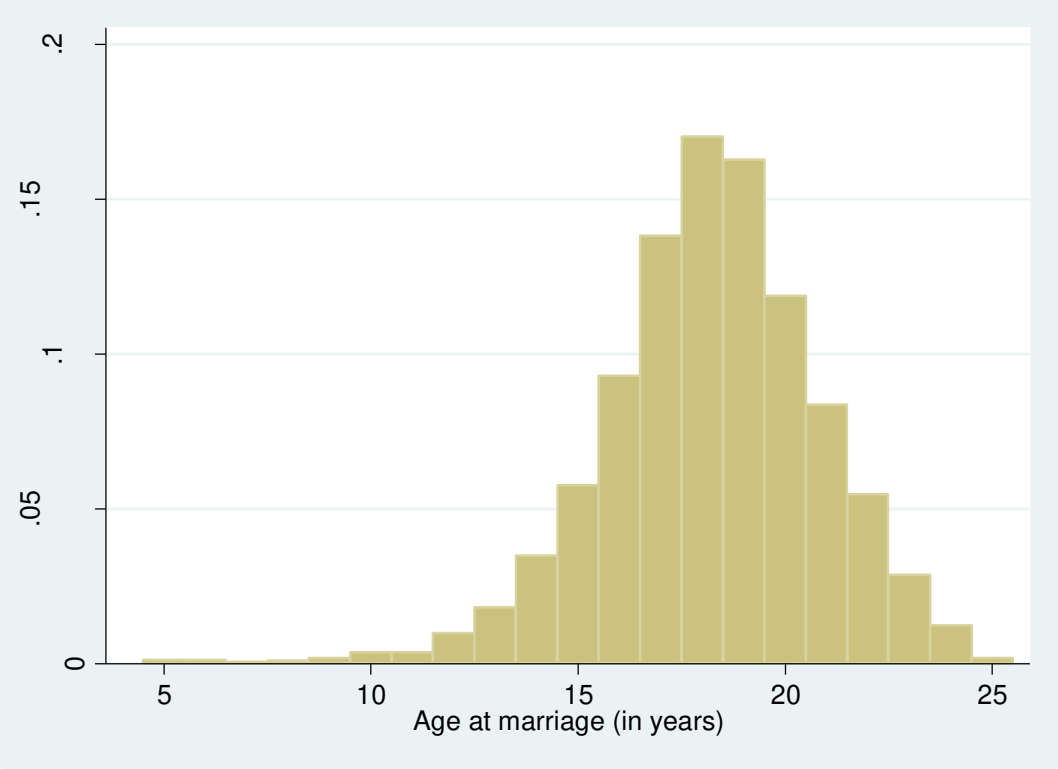


Figure 2. Distribution of age at menarche

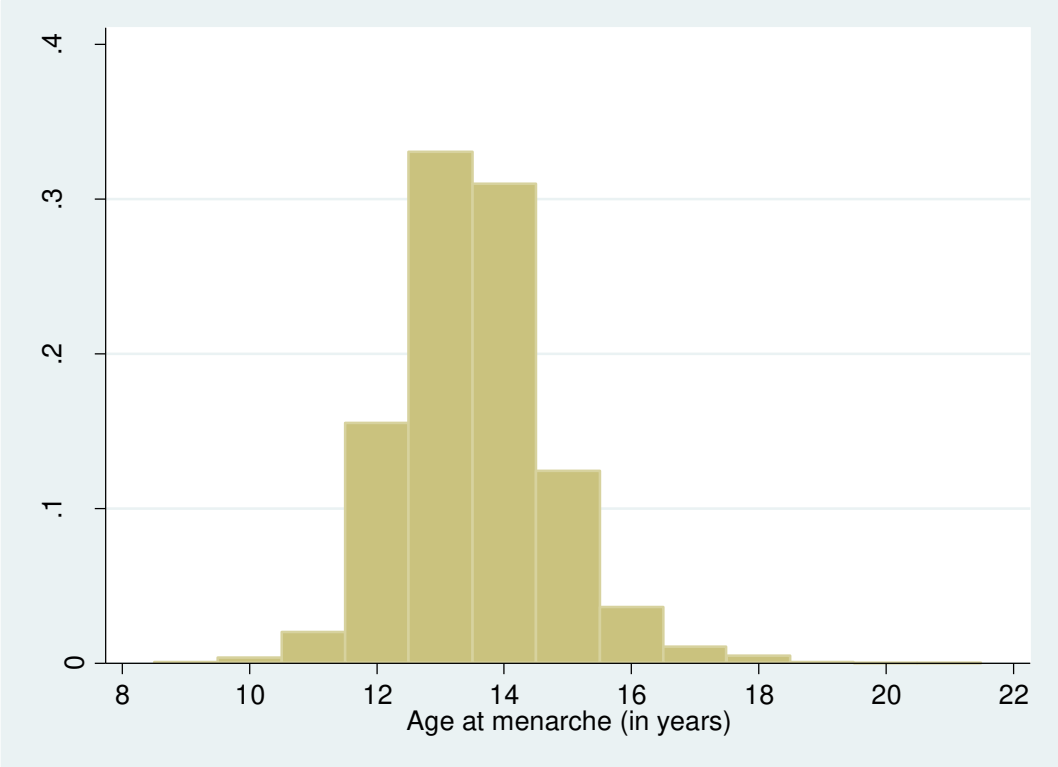


Figure 3. Prevalence of domestic violence by age at marriage

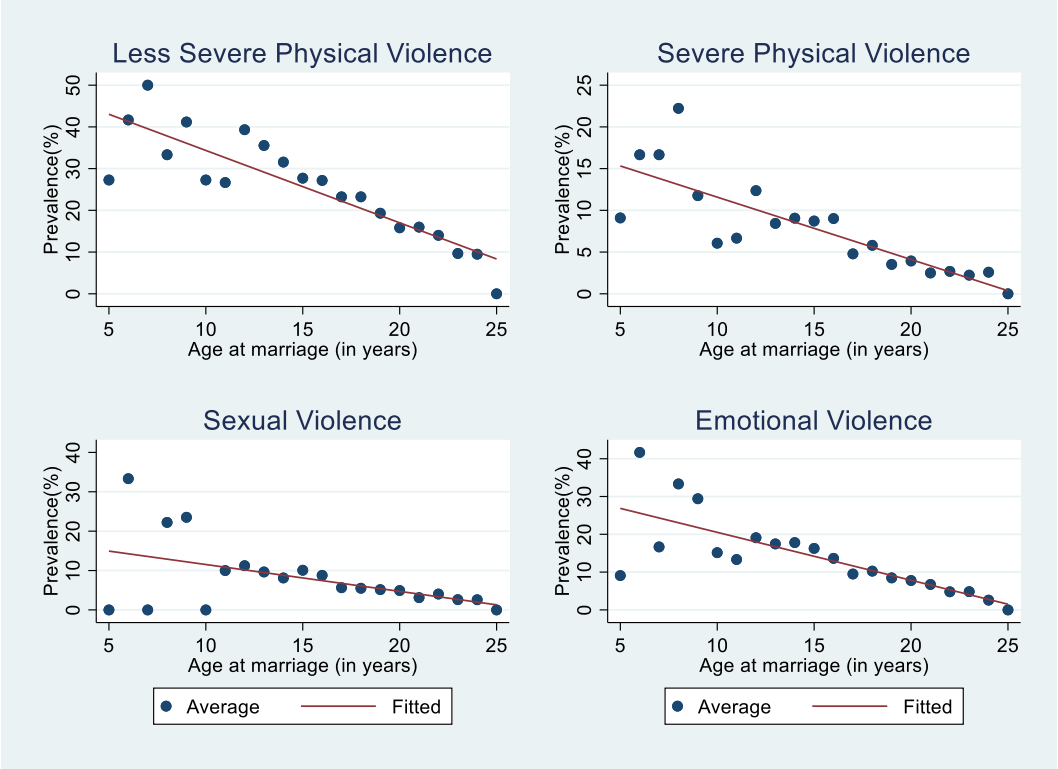
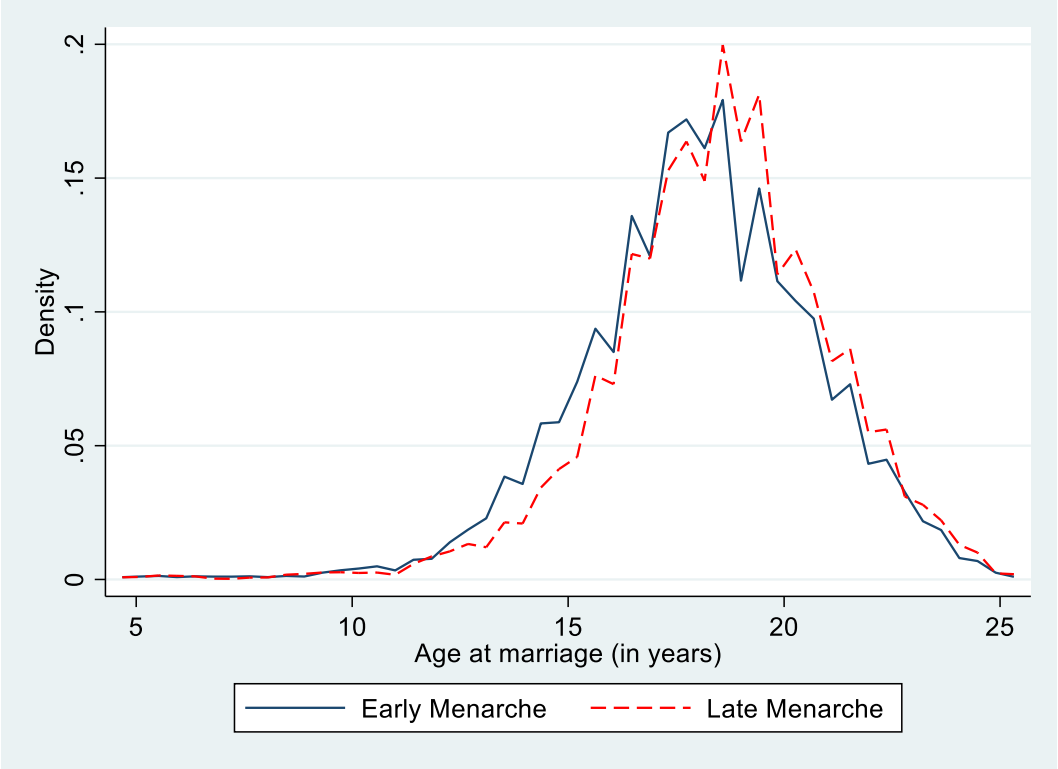


Figure 4. Distribution of women's age at marriage by age at menarche



Notes: Early menarche group includes those women who attained menarche before 14 years of age. Late menarche group includes the rest.

Figure 5. Relationship between women's average years of schooling and age at menarche

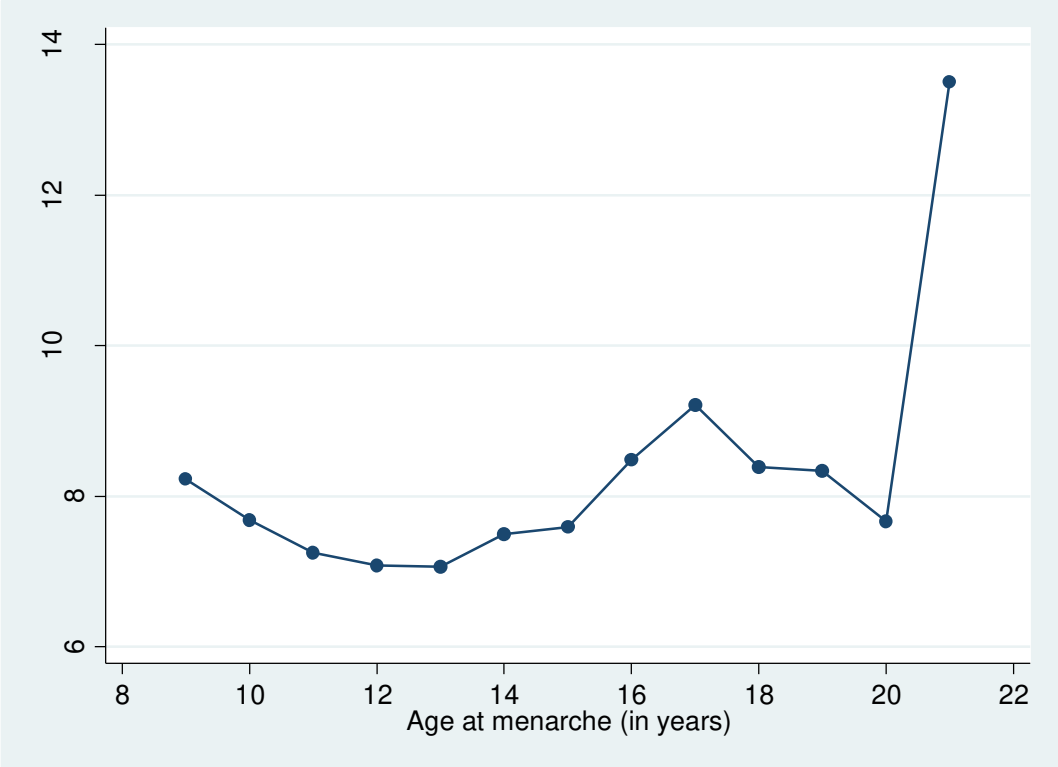


Figure 6. Kernel density estimates of women's years of schooling by terciles of age at menarche

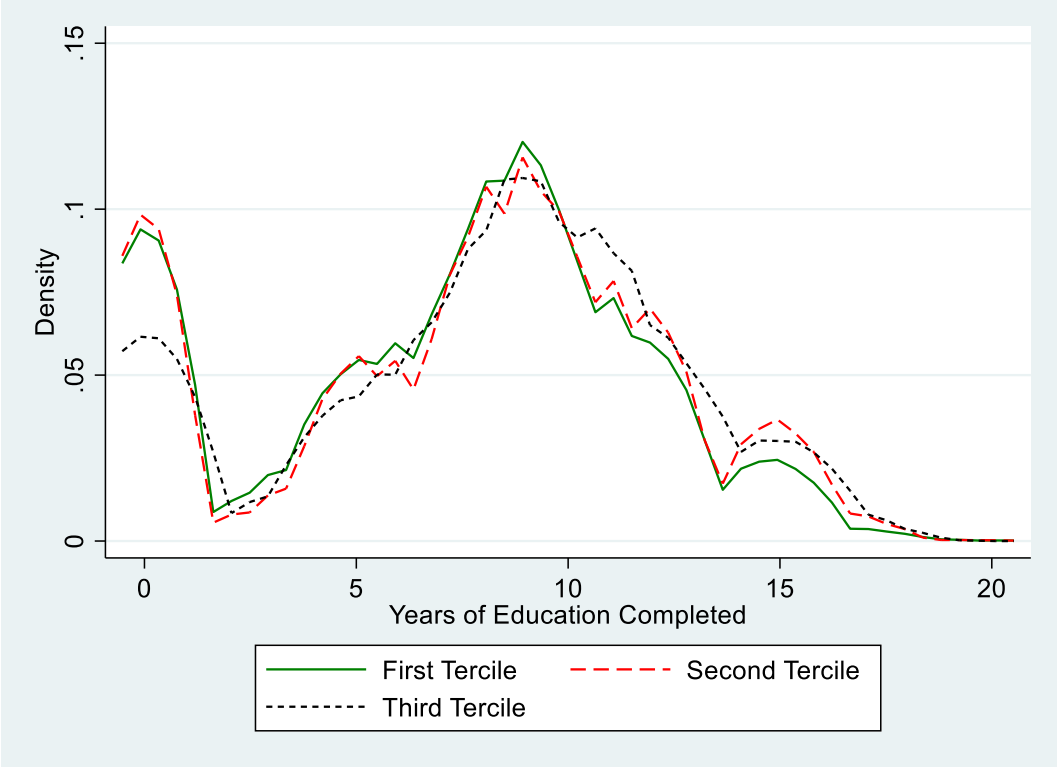


Table 1. Summary statistics

	Mean	SD
Domestic Violence Outcomes		
Less Severe Physical Violence	0.21	0.41
Severe Physical Violence	0.05	0.22
Sexual Violence	0.06	0.23
Emotional Violence	0.10	0.30
Demographic Characteristics		
Age at Marriage	18.23	2.63
Age at Menarche	13.57	1.21
Age	21.65	1.99
Spousal age	26.47	4.39
Height (in cm)	151.75	5.96
Years of Education Attained	7.39	4.63
Wealth Indicators		
Poorest	0.22	0.42
Poorer	0.25	0.43
Middle	0.23	0.42
Richer	0.18	0.38
Richest	0.12	0.33
Caste Indicators		
Scheduled Caste (SC)	0.21	0.41
Scheduled Tribe (ST)	0.18	0.39
Other Backward Caste (OBC)	0.43	0.50
Other Castes	0.17	0.38
Seen domestic violence among parents	0.22	0.41
Place of Residence (=1 if Urban)	0.77	0.42
N	9,343	

Table 2. OLS estimates of the effect of age at menarche on age at marriage

	[1]	[2]	[3]	[4]	[5]
Age at Menarche	0.216***	0.161***	0.128***	0.128***	0.127***
	(0.024)	(0.023)	(0.022)	(0.022)	(0.024)
F-statistic	77.63	193.78	131.31	131.31	27.44
Observations	9,343	9,343	9,343	9,343	9,343

Notes: Estimation via OLS. The outcome variable is women's age at marriage. Regression reported in column (1) does not include any controls. Regression reported in column (2) control for women's height, age, and spousal age. In column (3), controls include women's height, age, spousal age, wealth dummies, and women's caste affiliation. In column (4), controls include women's height, age, spousal age, wealth dummies, women's caste affiliation, and indicator for whether women have seen domestic violence among their parents. In column (5), we include place of residence (i.e., rural or urban locality) and district fixed effects in addition to all controls used in column (4). Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 3. OLS estimates of the effect of age at marriage on domestic violence

	Less Severe Physical Violence			Severe Physical Violence			Sexual Violence			Emotional Violence		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Age at Marriage	-0.022***	-0.019***	-0.018***	-0.008***	-0.007***	-0.006***	-0.007***	-0.006***	-0.005***	-0.013***	-0.012***	-0.011***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Demographic Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
District Fixed Effects	N	N	Y	N	N	Y	N	N	Y	N	N	Y
Observations	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343

Notes: Estimation via OLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 4. IV estimates of the effect of age at marriage on domestic violence

	Less Severe Physical Violence			Severe Physical Violence			Sexual Violence			Emotional Violence		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Age at Marriage	-0.067***	-0.061**	-0.070**	-0.037***	-0.044***	-0.043**	-0.016	-0.005	0.008	-0.028**	-0.017	0.002
	(0.018)	(0.028)	(0.030)	(0.010)	(0.017)	(0.019)	(0.010)	(0.017)	(0.019)	(0.012)	(0.020)	(0.023)
Demographic Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
District Fixed Effects	N	N	Y	N	N	Y	N	N	Y	N	N	Y
First stage F statistic	77.63	32.83	27.44	77.63	32.83	27.44	77.63	32.83	27.44	77.63	32.83	27.44
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	89.13	37.56	35.22	89.13	37.56	35.22	89.13	37.56	35.22	89.13	37.56	35.22
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Observations	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 5. Causal mechanisms: IV estimates of the effect of women's age at marriage on education, labor market participation, decision making power and husbands' characteristics

	Women's Characteristics							Husband's Characteristics		
	Literate	Completed Primary Education	Years of education completed	Workforce participation	Say in health care	Say in large household purchase	Bank account access	Years of education completed	Workforce participation	White Collar Occupation
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Age at Marriage	0.046*	0.079***	1.054***	0.018	0.004	-0.010	0.065*	0.774***	0.027	0.012
	(0.026)	(0.028)	(0.294)	(0.031)	(0.035)	(0.036)	(0.039)	(0.298)	(0.020)	(0.035)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
First stage F statistic	27.44	27.44	27.44	27.44	27.44	27.44	27.44	27.56	26.60	25.44
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen Paap rK LM statistic	35.22	35.22	35.22	35.22	35.22	35.22	35.22	35.30	34.19	32.26
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	9,316	9,343	9,343	9,343	9,343	9,343	9,343	9,316	9,251	8,710

Notes: Estimation via TSLS. The estimated coefficients of age at marriage in the columns (1) - (7) show the effect on women's characteristics and columns (8) - (9) on women's husband characteristics. Columns (1) - (3) show the effect of women's age at marriage on different educational outcomes. The estimated coefficient of age at marriage in column (4) shows the effect of age at marriage on the probability of workforce participation. The estimated coefficients of age at marriage in columns (5) - (7) show the effect of age at marriage of different indicators of bargaining/decision making power of women within households. The estimated coefficients of age at marriage in columns (8) - (10) show the effect of women's age at marriage on her husbands' educational attainment, work force participation, and occupation (white collar occupation vs others) respectively. The women's outcome variables are defined as follows: Literate takes a value one (literate) if the woman is able to read a whole sentence, and zero otherwise; Completed Primary Education takes a value one if the woman reports to have completed 5th or higher grade, and zero otherwise; Years of education completed is a continuous variable which records the years of education completed by the woman; Workforce participation takes a value one if the woman reports that she has worked in the last one year, and zero otherwise; Say in health care takes a value one if the woman reports that she herself or jointly with her husband makes the decision about her healthcare, and zero otherwise otherwise; Say in large household purchase takes a value one if the woman reports that she herself or jointly with her husband makes decisions about large household purchases, and zero otherwise; Bank account access takes a value one if the woman reports that she has a bank or savings account that she uses, and zero otherwise. The husbands' outcome variables are defined as follows: Years of education completed is a continuous variable which records the years of education completed by the woman's husband; Workforce participation takes a value one if the husband reports that he is currently working, and zero otherwise; White Collar Occupation takes a value one if the husband reports that he works in professional, technical, managerial, clerical, sales, and services sector and zero otherwise (if employed in agriculture, skilled and unskilled manual work). Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation, and place of residence. Sample size in column (1) reduces by 27 observations as literacy could not be completed due to unavailability of literacy card with required language and blindness. Sample size in column (8) and (9) reduces by 27 and 92 observations respectively as due to unavailability of information on completion of education and work force participation. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Appendix

The causal impact of women's age at marriage on domestic violence in India

[FOR ONLINE PUBLICATION ONLY]

1 Data Assembly Details

The following details about the analytical sample are worth noting. First, in our sample, year of marriage was known for 99.73% (9318 out of 9343) observations. We combine this information with the year of birth information to get the age of marriage. For the remaining cases where the year of marriage was unavailable, we use the age at marriage (reported only for those cases where the year of marriage is unknown) available in the data set.

Second, the fact that our sample includes women with menarche age between 9 and 21 years demands some explanation. The normal menarcheal age is between 10 and 15 years. However, menarcheal age as low as 9 years is not unusual (The Times of India, 2014). Similarly, menarcheal age above 15 years, and in fact, as high as 20-21 years is also not biologically impossible. Delayed puberty may be constitutional or due to pathological causes (Blondell et al., 1999). Undernourishment during childhood is, in fact, one major reason for delayed menarche. Also, intense physical activity during childhood may delay menarcheal age. In this context, based on a survey of dancers and athletes, Frisch et al. (1980) and Frisch et al. (1981) note that dancers and athletes who began their training at ages 9 or 10 years still had not menarche at ages 18–20 years.

2 Robustness Checks

2.1 Alternative Method of Estimation

While our baseline results are obtained using the IV-TSLS approach, it is worthwhile to check the sensitivity of our findings to using an alternative non-linear method of estimation since our outcome variables are binary in nature. Towards that end, we repeat our analysis using a Probit approach. Specifically, we estimate Probit models using the control function (CF) approach proposed by Rivers and Vuong (1988), and later developed by Blundell and Powell (2004) and Wooldridge (2010, 2015). Results are reported in Table A1.

In line with the baseline results, the results of this exercise (those which include both demographic controls as well as district fixed effects) indicate that a delay in women's marriage by a year causes the probability of their exposure to less severe and severe physical violence to fall significantly. The effect of women's age at marriage on sexual as well as emotional violence continue to remain statistically insignificant as before. This is reassuring, and indicates that our results our robust to the choice of estimation method.[1]

2.2 Falsification Test

Our IV strategy rests on the assumption that the women's age at marriage is the only channel through which age at menarche affects prevalence of domestic violence (in other words, the exclusion restriction is valid). If this assumption is correct, then a significant relationship between age at menarche and domestic violence should not exist when we restrict our sample to the women who got married before attaining menarche because menarche could not have impacted their marriage timing.

To assess the validity of the IV estimates, we undertake this falsification test: we test the reduced form effect of age at menarche on the different forms of domestic violence for the subsample of women who got married before attaining menarche. Results of this test are reported in Table A2.

As evident, the coefficients of age at menarche from the regressions based on the subsample that includes only those women who got married before attaining menarche turn out to be statistically insignificant (without as well as with district fixed effects). This implies that there exists no systematic relationship between age at menarche and domestic violence for

[1] When we include district fixed effects in Columns (3), (6), (9) and (12), our sample size reduces since Stata drops observations from several districts for which the districts perfectly predicts the failure or success. (i.e., for those districts no women reports to have faced domestic violence or all women have reported to face domestic violence). While IV-TSLS can produce consistent estimates of the parameters even with several districts in which there is no variation in the outcome variable, MLE cannot do so and hence these districts need to be dropped. However, the fact that thse observations are dropped just means they are not contributing any information to help identify the other parameters in the model. Implicitly, these observations are also not helping us estimate the coefficients beyond the fixed effects in our baseline IV-TSLS model either. So, the results across the two are still comparable.

this subsample of women. The coefficients of age at menarche from the regressions with less severe violence and severe violence as outcome variables, on the other hand, are statistically significant for the subsample of women who got married after attaining menarche. This is consistent with our IV results that women's age at marriage has a significant causal impact on less severe and severe forms of physical violence.

In sum, thus, the results of this falsification exercise suggest that our instrument is likely to satisfy the exclusion restriction, and therefore increases our confidence in the empirical strategy that we have used.

2.3 Addressing Concerns About Measurement Error in Age at Menarche

We have noted that women's age at marriage can be subject to reporting bias. In a similar vein, one could raise concerns about measurement error in the age at menarche. If age at menarche contains measurement error, this might cause the IV estimates of the coefficient of age at marriage to be inconsistent.

While recall error in age at menarche is possible since we use self-reported survey data, Must et al. (2002) provide compelling evidence to show that this is unlikely to be a reason for severe concern. They use the US Newton Girls Study (1965–1975), a prospective study of development in a cohort of girls followed through menarche, to assess the accuracy and precision of recall of several early menstrual characteristics. In 1998–1999, around 60% of the original 793 Newton Girls Study participants completed a mailed questionnaire to assess the accuracy of recall for age and body size at menarche, usual cycle length during the first 2 years, and age at regularity. They found recalled and original age at menarche to be highly correlated. Original mean menarcheal age did not differ significantly from recalled mean menarcheal age. On average, women recalled their menarche as being 0.95 months (i.e., less than a month) earlier than their original menarche. In fact, in context of India, recall error in age at menarche is likely to be even less of a concern since Garg et al. (2001) and Sharma

et al. (2006) note that menarche is a major event for girls in India, and girls of both low and high caste report knowing little or nothing about menstruation before it began, but afterwards learning of taboos about eating and mobility during menstrual periods. These changes in lifestyle imply that respondents are likely to recall its timing with fair degree of accuracy (Chari et al., 2017).[2] We have already graphed the distribution of reported age at menarche in Figure 2. It does not show any heaping at key ages (e.g. school leaving ages) that might be suggestive of significant recall error.

Note, even if age at menarche contains measurement error, this will cause the IV estimate of the coefficient of age at marriage to be inconsistent only if reporting bias in age of marriage is correlated with that in age at menarche. This might be the case if respondents use the former as a point of reference to recollect the latter. To examine this issue, Field and Ambrus (2008) suggest comparing the distribution of reported age of marriage and age at menarche for two subsamples of women: (i) women with mothers who never attended school, and (ii) women with mothers who had at least some schooling before the onset of puberty. The idea here is to isolate a group of families who have a preexisting preference for later marriage unrelated to their daughter's maturation. Since menarche is exogenous to this preference, a significant difference in reported age of onset across these types would suggest either recall bias or strategic misreporting.

However, the data that we use do not have information on the educational attainment of women's mothers. As an alternative, we compare the distribution of age at marriage and menarche age of women from two southern states of India – Kerala and Tamil Nadu – to that of women from nine major non-southern states including Haryana, Punjab, Delhi, Madhya Pradesh, Gujarat, Bihar, Jharkhand, West Bengal and Odisha. We do this because Kerala and Tamil Nadu are ranked first and second respectively among all the Indian states in terms of sex ratios (the number of females per 1000 of males) as per the 2011 Indian census. On

[2] Ellis (2004, 921) based on a survey also note, "both adolescent girls and adult women are generally willing and able to report accurately on their ages at menarche...and retrospective reports may be more reliable than those obtained during puberty".

the other hand, the selected non-southern Indian states have significantly lower sex ratios. Since the sex ratio is generally thought to reflect the level of patriarchy (and women's social disadvantage) prevalent in the society (with higher sex ratios indicating lower prevalence of patriarchy) (Bhalotra et al., 2017), and that level of patriarchy is negatively correlated to age at marriage, the age at marriage of women from Kerala and Tamil Nadu is likely to be significantly higher than the age at marriage of women from the selected non-southern states. In terms of the distribution of age at marriage, therefore, the distribution of age at marriage of women from Kerala and Tamil Nadu should lie to the right of that from the non-southern states. However, in absence of recall error in age at marriage being correlated with recall error in age at menarche, the distributions of age at menarche across these two subsamples of women should not be different.

We plot the distributions of age at marriage and age at menarche for these two subsamples of women in Figure A3. Reassuringly, we find that the distribution of women's age at marriage differs across the two subsamples, but not the distribution of age at menarche. This provides suggestive evidence that even if there is recall or reporting error in age at menarche, that is unlikely to be correlated with age at marriage.

2.4 Further Examination of the Correlation between Menarche Age and Educational Attainment

As discussed previously, our IV estimates of the effect of women's age at marriage on domestic violence could be inconsistent if age at menarche is correlated with educational attainment. While previous literature as well as our graphical analysis (presented previously) suggests that this is unlikely to be the case, in this section we reexamine the issue. Note that the main reason why educational attainment could be impacted by women's age at menarche is as follows: once a girl starts menstruating, she might stop attending school, eventually dropping out of it. This might be because of lack of sanitation facilities in school or because social norms dictate that a girl who attains puberty should no longer be going to school. As such,

the women whose educational attainment might have been impacted with age at menarche, they must have had dropped out of school soon after they had reached that age (i.e., age at menarche). It is presumable that if a woman was enrolled in school for a significant period of time post-menarche (say, at least more than a year), her schooling attainment is unlikely to have been determined by her age at menarche. This is because if she could have made it to school right after attaining menarche despite lack of adequate sanitation facilities in her school or social taboos, it is unlikely that these factors later on would suddenly serve as impediments for her to attend school. As such, if we could find a subsample of women who had been enrolled in school for at least a year or two after attaining menarche, for this subsample of women at least, age at menarche is unlikely to have affected their educational attainment, and hence the IV estimate of the effect of age at marriage on domestic violence based on this subsample would likely be consistent.

However, finding this subsample of women is difficult given our data. This is because although we have information on age at menarche and years of education of women, we do not have information on the age at which they dropped out of school. To circumvent this problem, we try to back out the school leaving age of women from their reported years of education. Specifically, we assume that a women starts her formal schooling at 6 years of age (which according to the World Bank is the minimum age for starting formal schooling for most countries in the world including India[3]) and add the total number of years of education to this. This gives us the lower bound of the age at which the women have dropped out of school. We call this the lower bound because it is perfectly possible that a woman had started schooling at an age higher than 6.[4] Further, there might have been gaps in her schooling after she starts going to school. In fact, she might be even repeating grades. For example, suppose a woman reports that she has 5 years of education. As per our estimate, her school leaving age is around 11 years. However, if this woman had started schooling at 8 years of age and further stops going to school for one year after 2 years of starting to go to

[3] https://data.worldbank.org/indicator/SE.PRM.AGES

[4] Although it is very unlikely that a woman begins her formal schooling before she turns 6.

school, her true age when she drops out of school was actually around 14 years (she starts going to school when she is 8 years old, goes to school till she is 10 years, takes a gap of a year, returns back to school and goes to school for another three years).

Using this method, we calculate the lower bound of the age of women at which they had dropped out of school in our sample. We then create a subsample consisting of those women for whom the difference between the school dropout age that we calculate and the age at menarche is at least one year. In other words, these are the women who were in school for at least one year post-menarche. Note that, these women form the subset of all the women who were enrolled in school for at least one-year post menarche. This is because, given that we consider the lowest possible age at which the women had dropped out of school, we exclude those women who might have been in school for at least one year post-menarche, but as per our calculation of her school leaving age, they had dropped out of school before completion of a year after menarche.

For this subsample of women, we estimate our baseline regression model instrumenting age at marriage by age at menarche. The results are reported in columns (1) – (4) of Table A3. We also estimate the regression model for the subsample of women who have been in school for at least 2 years post-menarche (i.e., the difference between the lower bound of their school leaving age and age at menarche is at least 2 years). For this subsample, age at menarche is more unlikely to be have determined the educational attainment of the women. These results are reported in columns (5) – (8). Reassuringly, we find that for both these subsamples, the effect of age at marriage on less severe and severe forms of physical violence are negative as well as statistically significant. This is in consonance with our baseline results. This suggests that not including educational attainment in our baseline specification is unlikely to have rendered our estimates of the effect of age at marriage on less severe and severe forms of physical violence to be inconsistent.

2.5 Inclusion of Survey Year Fixed Effects

The data collection period for NFHS 2015-16 is quite long (January 2015 to December 2016). As such there might be possible changes over time which, if not accounted for, might contaminate our results. To address this issue we include survey year fixed effects as additional controls in Table A4. As evident, our results are robust to inclusion of survey year fixed effects.

2.6 Alternative Measures of Domestic Violence

In our baseline analysis, we measure each of the four categories of domestic violence using binary variables that take a value of one if the woman reported having experienced at least one kind of violence in the relevant class of violent acts. To check the sensitivity of our results to the baseline measure of domestic violence, we use two alternative measures of domestic violence. First, we construct a new measure of the four types of domestic violence by taking a simple average of the binary indicators for all the underlying acts of violence. Second, we construct z-scores for each act of violence underlying each of the four types of domestic violence using the mean and standard deviation of the variable, and then construct four domestic violence indices using simple average of these z-scores (Kling et al., 2007; Erten and Keskin, 2018). Reassuringly, the results presented in Table A5 indicate that the baseline results are not sensitive to our definition of domestic violence.

2.7 Non-Linearities in Age

Given the definition of domestic violence measures in our baseline analysis, one might suspect a non-linear relationship to exist between age and each category of domestic violence. In order to control for this possibility, we include square of age as an additional control in the final specification of our model. Our results (reported in Table A6) are robust to inclusion of this additional control.

2.8 Removing Outliers

Our analytic sample includes women who attained menarche at 21 years of age and women who got married when they were 5 years old. However, attaining menarche at 21 years of age and getting married at 5 years of age might appear to be a bit implausible, and perhaps these might represent enumeration error. In order to address this concern, we carry out our baseline analysis excluding the top and bottom one percentile of age at marriage and age at menarche of our analytic sample. The results are reported in Table A7. As can be seen, the results based on the trimmed sample are qualitatively similar to the baseline results.

2.9 Birth Year Fixed Effects

As argued by Field and Ambrus (2008), sudden changes in diet might impact maturation. Sekhri and Debnath (2014) in this context note that, agricultural activities, that employ majority of the Indians, are to a large extent dependent on weather. Extreme weather conditions in the women's year of birth might adversely affect household income resulting in transitory but severe malnutrition. Therefore, females born during these extreme weather events may experience delayed age at menarche as they are more likely to be malnourished. These women may also be more susceptible to domestic violence due to having less options outside marriage (in the labor market) due to being malnourished. To control for this possibility, we use birth year fixed effects as additional controls. The results are presented in Table A8. Reassuringly, the results remain unchanged.

2.10 Control for Spousal Age at Marriage

The baseline analysis uses current spousal age as one of the controls. In Table A9, we re-estimate the baseline regression model using spousal age at marriage instead of current spousal age as a control. Since spousal age at marriage is not available in the survey, we calculate spousal age at marriage using women's age at marriage, women's current age and

spousal current age. The results are reported in Table A9. They are qualitatively similar to the baseline results.

2.11 Control for Proxy for Family Violence

It is possible that violence against women prior to menarche (or during childhood) is correlated with age at menarche as well as with age at marriage. Ideally, therefore, one should control for this possibility to obtain consistent estimates of the effect of women's age at marriage on domestic violence using age at menarche as an IV. Unfortunately, in the data that we have used, there is no information about women's exposure to violence during childhood. However, we think, this is unlikely to be a major issue in our case because of our use of a control variable that captures whether a woman has witnessed inter-parental violence in her childhood. We believe this variable is likely to proxy for whether a woman herself has faced violence during childhood and reduce concerns of omitted variable bias. Our assertion is based on existing studies which show that women's exposure to violence during childhood is highly correlated to whether they witness inter-parental violence or conflict at home. As shown in a study by Edleson (1999), between 30 to 66 per cent of children suffer direct abuse when living with domestic violence. More recently, a UNICEF report based on multiple studies on the impact of domestic violence on lives of children also notes, "Children who grow up in a violent home are more likely to be victims of child abuse" (UNICEF 2009, p.3).

However, we also carry out another exercise to assess this issue. We check the sensitivity of our main results to inclusion of an additional control which we believe is also a good proxy of whether a woman exposed violence during childhood at her natal home. This variable indicates whether a woman has faced violence from mother, father, sister, brother or other relative from the time she was 15 years old; it takes a value 1 if she has faced violence from either of these family members, and zero otherwise. This variable is likely to be a good proxy for whether a woman herself has faced violence during childhood because it is likely that a woman who reports that she was physically abused by her parents or other family members

when she was 15 years old has also experienced violence as a child. It is highly unlikely that a woman, who was never physically abused as a child, would suddenly experience violence when she is 15 years or older. Reassuringly, the results of the regression that includes this variable as an additional control (Table A10), are in line with the baseline results.[5]

2.12 Heterogeneity Analysis

Does the effect of age at marriage on domestic violence vary by spousal age and caste? To examine whether the effect of age at marriage on domestic violence varies by spousal age, we create a binary variable indicating whether the spousal age is less than the median spousal age in our sample (= 1) or otherwise (= 0). We interact the age at marriage with this binary variable and use this interaction term as an additional regressor in our baseline regression model. Instrumenting this by the interaction between age at menarche and the binary variable that we created for spousal age, we re-estimate the baseline regressions (of course, as before, we continue to instrument age at marriage by age at menarche). The results of this analysis are reported in the first four columns of Table A11. The results indicate that the impact of age at marriage on different categories of domestic violence does not differ across the two groups of women, i.e., those who are married to spouses whose age is less than the median spousal age and those who are married to spouses whose age is greater than or equal to the median spousal age. In other words, the effect of marriage age on domestic violence does not vary by spousal age.

In order to examine whether the effect of age at marriage varies by caste, we create a binary variable indicating whether a woman belongs to a household with lower caste (Schedule Caste, Schedule Tribe, Other Backward Caste) or other caste. As before, we interact the age at marriage with this binary variable and use this interaction term (instrumented

[5] The results reported are based only on the subsample of women who were married after they were 15 years old. Including women who were married before 15 years of age would potentially mean that we are including women who might have report be exposed to (natal) family violence after their marriage which is appears to be slightly implausible. Nevertheless, we also carried out the same regressions for the full sample. The results for the full sample are very similar to the results reported in Table A10.

by the interaction between age at menarche and the binary indicator for caste) as an additional regressor. The results presented in the last four columns of Table A11 do not provide any substantial evidence of the effect of age at marriage on domestic violence to be varying by caste (at best, there is only some weak evidence that the impact of age at marriage on emotional violence is different for women belonging to lower caste relative to upper caste).

2.13 Further Evidence on Causal Mechanism

To provide further evidence that it is the women's educational is a mechanism driving our results, we carry out our main analysis for the women who are uneducated and educated separately. If indeed it is the educational channel that mainly drives our results, we should find evidence of a causal relationship between women's age at marriage and physical violence for the second subsample of women only.

Table A13 reports the results of this exercise. As evident, we find no evidence of causal relationship between women's age at marriage and physical violence for the uneducated sample; the coefficient of women's age at marriage in all the regressions are statistically and economically insignificant. In contrast, for the educated sample of women, the coefficient of women's age at marriage in the regressions with less severe physical violence and physical severe violence as outcome variables are large and economically significant; in fact, in terms magnitude, these coefficients are similar to the corresponding coefficients obtained based on the full sample (see Table 4). However, the standard errors of these coefficients are now very large suggesting that these coefficients are measured imprecisely. This, however, is not surprising at all since by restricting our analysis to only women who are educated, we lose close to 30% of the sample which severely reduces the variation in the data resulting in mechanical inflation of the standard error of the estimated coefficients.

These results provide additional evidence that it is the educational channel through which age at marriage affects women's exposure to domestic violence.

References

Bhalotra, S., Clots-Figueras, I., & Iyer, L. (2018). Pathbreakers? Women's electoral success and future political participation, *Economic Journal*, 128, 1844-1878.

Blondell, R.D., Foster, M.B., & Dave, K.C. (1999). Disorders of Puberty, *American Family Physician*, 60, 209–218.

Blundell, R., & Powell, J. (2004). Endogeneity in Semiparametric Binary Response Models, *Review of Economic Studies*, 71, 655–79.

Chari, A.V., Heath, R., Maertens, A., & Fatima, F. (2017). The causal effect of maternal age at marriage on child wellbeing: Evidence from India, *Journal of Development Economics*, 127, 42–55.

Edleson, J. (1999) 'Children's witnessing of adult domestic violence', *Journal of Interpersonal Violence*, 14, 839-870.

Ellis, Bruce J. (2004). Timing of Pubertal Maturation in Girls: An Integrated Life History Approach, *Psychological Bulletin*, 130, 920–58.

Erten, B., & Keskin, P. (2018). For better or for worse?: Education and the prevalence of domestic violence in turkey. *American Economic Journal: Applied Economics*, 10, 64-105.

Field, E., & Ambrus, A. (2008). Early marriage, age of menarche, and female schooling attainment in Bangladesh, *Journal of Political Economy*, 116, 881–930.

Frisch, R. E., Wyshak, G., and Vincent, L. (1980). Delayed menarche and amenorrhea in ballet dancers, *New England Journal of Medicine*, 303, 17–19.

Frisch, R. E., Von Gotz-Welbergen, A., McArthur, J. W., Albright, T., Witschi, J., Bullen, B., Burnholz, J., Reed, R. B., & Hermann, H. (1981). Delayed menarche and amenorrhoea of college athletes in relation to age of onset of training, *Journal of the American Medical Association*, 246, 1559-1563.

Garg, S., Sharma, N., & Sahay, R. (2001). Socio-cultural aspects of menstruation in an urban slum in Delhi, India, *Reproductive Health Matters*, 9, 16–25.

Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects, *Econometrica*, 75, 83–119.

Must, A., Phillips, S. M., Naumova, E. N., Blum, M., Harris, S., Dawson-Hughes, B., & Rand, W. M. (2002). Recall of early menstrual history and menarcheal body size: after 30 years, how well do women remember?, *American Journal of Epidemiology*, 155, 672-679.

Rivers, D., & Vuong, Q. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models, *Journal of Econometrics*, 39, 347–366.

Sekhri, S., & Debnath, S. (2014). Intergenerational consequences of early age marriages of girls: effect on children's human capital, *Journal of Development Studies*, 50, 1670–1686.

Sharma, N., Vaid, S., & Manhas, A. (2006). Age at Menarche in two caste groups (Brahmins and Rajputs) from rural areas of Jammu, *Anthropologist*, 8 (1), 55–57.

The Times of India (2014). "Girl talk: Menarche now at 8-9 years," *The Times of India*, 24 April 2014. Available at https://timesofindia.indiatimes.com/city/goa/Girl-talk-Menarche-now-at-8-9-years/articleshow/34169175.cms

UNICEF. (2009). "Behind Closed Doors: The Impact of Domestic Violence on Children," United States of America, United Nations Secretary-General's Study on Violence Against Children.

Wooldridge, J.M. (2010). *Econometric analysis of cross section and panel data*. MIT Press.

Wooldridge, J. (2015). Control Function Methods in Applied Econometrics, *Journal of Human Resources*, 50, 420–445.

Figure A1. Average Age at Marriage of Women by Age at Menarche

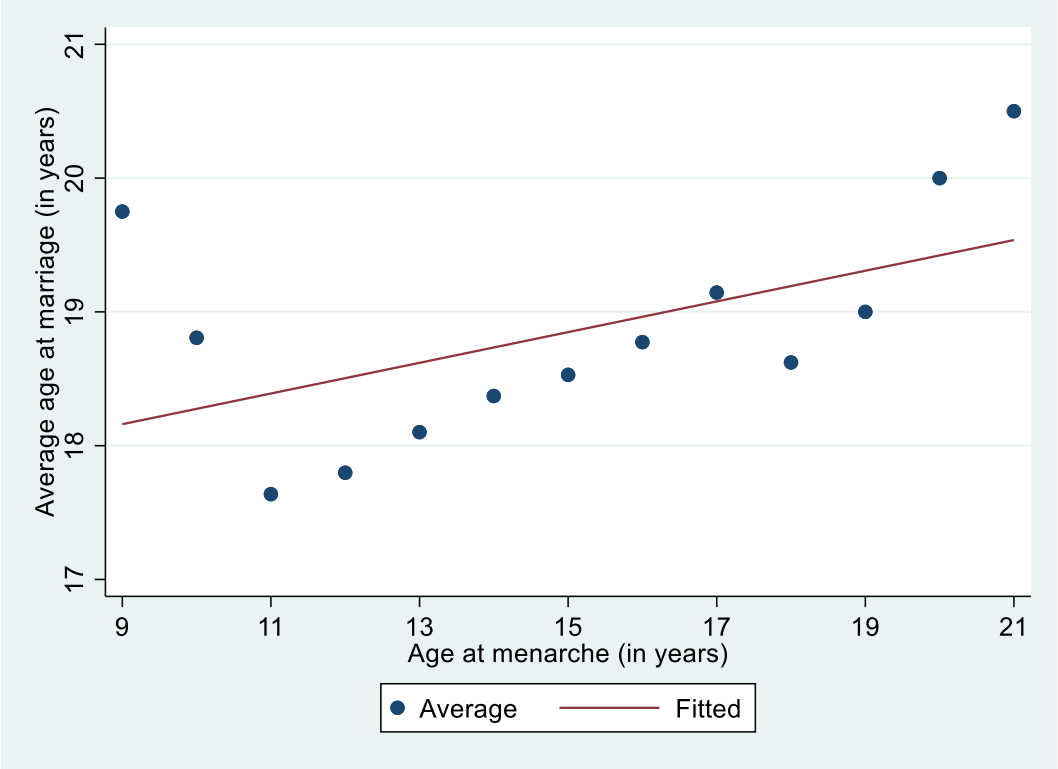


Figure A2. Average Years of Education Completed by Women by Age at Menarche (excluding outliers)

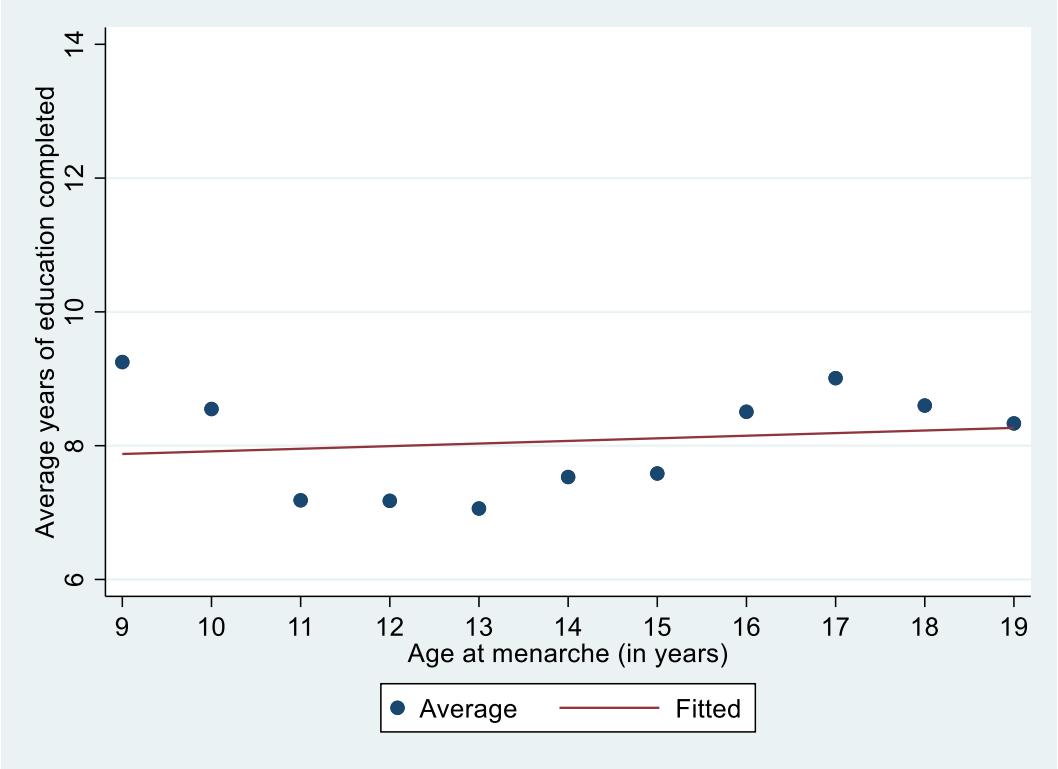
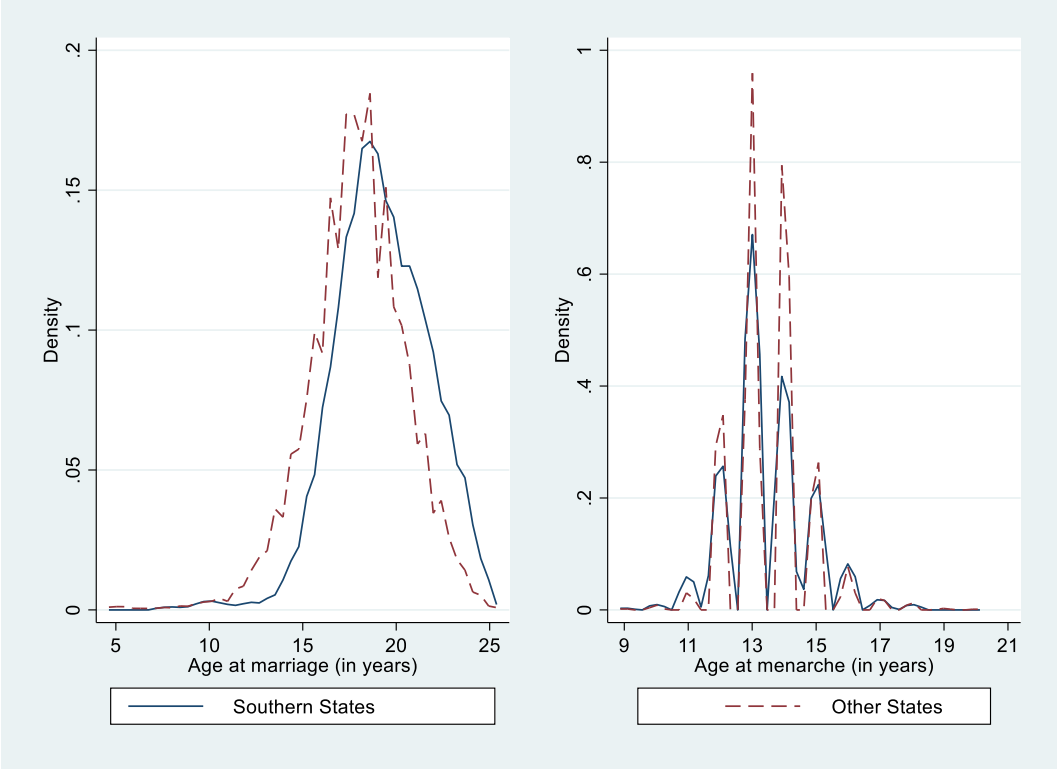


Figure A3. Kernel density estimates of women's age at marriage and age at menarche by states



Notes: Southern states include Kerala and Tamil Nadu. Other States include Haryana, Gujarat, Punjab, Delhi, Madhya Pradesh, Bihar, Jharkhand, Odisha, and West Bengal.

Table A1. Estimates of the effect of age at marriage on domestic violence: Probit estimates using control function approach (Marginal Effects)

	Less Severe Physical Violence			Severe Physical Violence			Sexual Violence			Emotional Violence		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Age at Marriage	-0.031***	-0.024***	-0.084**	-0.013***	-0.009***	-0.083**	-0.014***	-0.009***	0.020	-0.019***	-0.017***	0.001
	(0.004)	(0.005)	(0.033)	(0.002)	(0.002)	(0.034)	(0.002)	(0.003)	(0.034)	(0.003)	(0.003)	(0.031)
Demographic Controls	N	Y	Y	N	Y	Y	N	Y	Y	N	Y	Y
District Fixed Effects	N	N	Y	N	N	Y	N	N	Y	N	N	Y
Observations	9,343	9,343	8,431	9,343	9,343	5,035	9,343	9,343	5,174	9,343	9,343	6,776

Notes: The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. Sample size in columns (3), (6), (9), and (12) is smaller due to perfect prediction within a group. ***p < 0.01, **p <0.05, *p < 0.1.

Table A2. Falsification test

	Less Severe Physical Violence				Severe Physical Violence				Sexual Violence				Emotional Violence			
	Marriage before Menarche		Marriage not before Menarche		Marriage before Menarche		Marriage not before Menarche		Marriage before Menarche		Marriage not before Menarche		Marriage before Menarche		Marriage not before Menarche	
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]	[13]	[14]	[15]	[16]
Age at Menarche	-0.003	-0.023	-0.010***	-0.011***	0.013	-0.022	-0.007***	-0.007***	-0.003	0.001	-0.001	0.000	0.015	-0.027	-0.005*	-0.002
	(0.016)	(0.058)	(0.004)	(0.004)	(0.012)	(0.045)	(0.002)	(0.002)	(0.011)	(0.031)	(0.002)	(0.003)	(0.016)	(0.047)	(0.003)	(0.003)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N	Y
Observations	317	317	9,026	9,026	317	317	9,026	9,026	317	317	9,026	9,026	317	317	9,026	9,026

Notes: Estimation via OLS. The outcome variables are different categories of domestic violence. Regressions reported in columns (1), (2), (5), (6), (9), (10), (13), and (14) are based on the subsample of women who got married before attaining menarche. Regressions reported in columns (3), (4), (7), (8), (11), (12), (15), and (16) are based on the rest of the women. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A3. IV estimates of the effect of age at marriage on domestic violence: For a subsample of women who have been in school post-menarche

	In school for at least one-year post-menarche				In school for at least two years post-menarche			
	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]	[1]	[2]	[3]	[4]
Age at Marriage	-0.137***	-0.037*	0.007	-0.009	-0.124**	-0.037	-0.003	0.009
	(0.050)	(0.022)	(0.023)	(0.032)	(0.052)	(0.024)	(0.021)	(0.034)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
First stage F statistic	12.52	12.52	12.52	12.52	10.85	10.85	10.85	10.85
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.001]	[p=0.001]	[p=0.001]	[p=0.001]
Kleibergen Paap rK LM statistic	17.97	17.97	17.97	17.97	16.67	16.67	16.67	16.67
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Observations	3,804	3,804	3,804	3,804	2,867	2,867	2,867	2,867

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A4. IV estimates of the effect of age at marriage on domestic violence with survey year fixed effects

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.070**	-0.043**	0.009	0.002
	(0.030)	(0.019)	(0.019)	(0.022)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
Survey Year Fixed Effects	Y	Y	Y	Y
First stage F statistic	27.64	27.64	27.64	27.64
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	35.44	35.44	35.44	35.44
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Observations	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A5. IV estimates of the effect of age at marriage on using alternative measures of domestic violence

	Average of binary indicators for all the underlying acts of violence				Average of z-scores computed based on binary indicators for all the underlying acts of violence			
	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Age at Marriage	-0.045**	-0.014*	0.002	-0.004	-0.147**	-0.063	0.005	-0.018
	(0.018)	(0.008)	(0.012)	(0.014)	(0.060)	(0.065)	(0.067)	(0.061)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
First stage F statistic	27.44	27.44	27.44	27.44	27.44	27.44	27.44	27.44
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen Paap rK LM statistic	35.22	35.22	35.22	35.22	35.22	35.22	35.22	35.22
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Outcome variables in first four columns are constructed by taking a simple average of each act of violence. In the next four columns, we use simple average of z-scores constructed using the mean and standard deviation of each act of violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A6. IV estimates of the effect of age at marriage on domestic violence controlling for non-linear effects of age

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.072**	-0.046**	0.008	0.002
	(0.031)	(0.020)	(0.020)	(0.023)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
First stage F statistic	25.91	25.91	25.91	25.91
	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen Paap rK LM statistic	33.20	33.20	33.20	33.20
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, square of age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A7. IV estimates of the effect of age at marriage on domestic violence after excluding the outliers

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.057*	-0.044*	0.025	-0.023
	(0.032)	(0.026)	(0.026)	(0.032)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
First stage F statistic	17.78	17.78	17.78	17.78
	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen Paap rK LM statistic	22.00	22.00	22.00	22.00
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	8,735	8,735	8,735	8,735

Notes: Estimation via TSLS. Final sample excludes the top and bottom one percentile of age at marriage and age at menarche. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, squared of age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A8. IV estimates of the effect of age at marriage on domestic violence with birth year fixed effects

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.076**	-0.048**	0.007	0.002
	(0.032)	(0.021)	(0.020)	(0.024)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
Birth Year Fixed Effects	Y	Y	Y	Y
First stage F statistic	24.17	24.17	24.17	24.17
	[0.000]	[0.000]	[0.000]	[0.000]
Kleibergen Paap rK LM statistic	31.21	31.21	31.21	31.21
	[0.000]	[0.000]	[0.000]	[0.000]
Observations	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A9. IV estimates of the effect of age at marriage on domestic violence controlling for spousal age at marriage

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.083**	-0.053**	0.013	0.005
	(0.039)	(0.024)	(0.024)	(0.029)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
First stage F statistic	20.46	20.46	20.46	20.46
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	25.47	25.47	25.47	25.47
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Observations	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age at marriage, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A10. IV estimates of the effect of age at marriage on domestic violence including indicator for family violence: Subsample of women married after the age of 15

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]
Age at Marriage	-0.152*	-0.128**	0.024	0.006
	(0.090)	(0.065)	(0.051)	(0.060)
Demographic Controls	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y
First stage F statistic	6.99	6.99	6.99	6.99
	[p=0.003]	[p=0.003]	[p=0.003]	[p=0.003]
Kleibergen Paap rK LM statistic	8.66	8.66	8.66	8.66
	[p=0.003]	[p=0.003]	[p=0.003]	[p=0.003]
Observations	8,092	8,092	8,092	8,092

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, indicator for whether women were ever physically hurt by natal family members, women's caste affiliation, and place of residence i.e. rural or urban. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A11. Heterogeneity analysis: IV estimates of the effect of age at marriage and its interaction with spousal age and caste on domestic violence

	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Age at Marriage	-0.070**	-0.043**	0.009	0.002	-0.065	-0.056**	0.002	-0.047
	(0.030)	(0.019)	(0.019)	(0.023)	(0.040)	(0.027)	(0.021)	(0.031)
Age at Marriage*Spousal Age at Marriage Less than the Median Age	0.001	0.000	-0.000	-0.001				
	(0.001)	(0.000)	(0.000)	(0.001)				
Age at Marriage*Low Caste					-0.006	0.015	0.008	0.058*
					(0.040)	(0.026)	(0.021)	(0.031)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
First stage F statistic (age at marriage)	20.10	20.10	20.10	20.10	13.78	13.78	13.78	13.78
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
First stage F statistic (age at marriage*spousal age)	36394.99	36394.99	36394.99	36394.99				
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]				
First stage F statistic (age at marriage*low caste)					16.80	16.80	16.80	16.80
					[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Kleibergen Paap rK LM statistic	36.40	36.40	36.40	36.40	36.43	36.43	36.43	36.43
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]
Observations	9,343	9,343	9,343	9,343	9,343	9,343	9,343	9,343

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation and place of residence. First four columns include interaction between age at marriage and a binary variable indicating spousal age less than the median spousal age i.e. 26 years. We instrument this by interaction between age at menarche and a binary variable indicating whether spousal age is less than the median spousal age. Next four columns include interaction between age at marriage and a binary variable indicating a woman belongs to low caste i.e. Schedule Caste, Schedule Tribe, Other Backward Caste. We instrument this interaction between age at marriage and a binary variable indicating a woman belongs to low caste. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table A12. Summary statistics: Additional outcome variables

	N	Mean	SD
Women's characteristics			
Literate	9316	0.71	0.46
(=1 if the woman is able to read)			
Completed Primary Education	9343	0.76	0.43
(=1 if the woman has completed 5th or higher grade)			
Years of education completed	9343	7.39	4.63
Workforce participation	9343	0.19	0.39
(=1 if the woman reports to have worked in the last one year)			
Say in health care	9343	0.70	0.46
(=1 if the woman reports that decisions about her healthcare are made by herself or jointly with her husband)			
Say in large household purchase	9343	0.67	0.47
(=1 if the woman reports that decisions about large household purchases are made by herself or jointly with her husband)			
Bank account to access	9343	0.42	0.49
(=1 if the woman reports that she has a bank or savings account that she uses)			
Husband's characteristics			
Years of education completed	9316	8.23	4.49
Workforce participation	9251	0.94	0.23
(=1 if the husband reports that he is currently working)			
White Collar Occupation	8710	0.30	0.46
(=1 if the husband reports that he works in professional, technical, managerial, clerical, sales, and services sector)			

Table A13. IV estimates of the effect of age at marriage on domestic violence: Uneducated vs educated women

	Sample of Uneducated Women				Sample of Educated Women			
	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence	Less Severe Physical Violence	Severe Physical Violence	Sexual Violence	Emotional Violence
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Age at Marriage	0.003	-0.019	0.013	0.007	-0.112*	-0.049	0.013	0.015
	(0.028)	(0.016)	(0.016)	(0.021)	(0.059)	(0.036)	(0.035)	(0.044)
Demographic Controls	Y	Y	Y	Y	Y	Y	Y	Y
District Fixed Effects	Y	Y	Y	Y	Y	Y	Y	Y
First stage F statistic	35.27	35.27	35.27	35.27	8.00	8.00	8.00	8.00
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.005]	[p=0.005]	[p=0.005]	[p=0.005]
Kleibergen Paap rK LM statistic	44.92	44.92	44.92	44.92	10.64	10.64	10.64	10.64
	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.000]	[p=0.001]	[p=0.001]	[p=0.001]	[p=0.001]
Observations	2,741	2,741	2,741	2,741	6,575	6,575	6,575	6,575

Notes: Estimation via TSLS. The outcome variables are different categories of domestic violence. Demographic controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, women's caste affiliation, and place of residence. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.