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

Gaurav Dhamija†

Shiv Nadar University

Punarjit Roychowdhury‡

Indian Institute of Management, Indore

October 1, 2018

Abstract

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. In this paper, we provide the first causal analysis of the impact of women's age at marriage on prevalence of domestic violence using newly available nationally representative household data from India. We use 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. Our findings confirm the relevance of policies that seek to prevent child marriages or delay marriages of women in reducing the prevalence of domestic violence in a developing country.

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

*The authors are very grateful to Daniel Millimet and Tarun Jain for comments and suggestions. All remaining errors are our own.

†Department of Economics, Shiv Nadar University. email: gd343@snu.edu.in. NH91, Tehsil Dadri, Greater Noida, Uttar Pradesh 201314, India.

‡Corresponding Author. Department of Economics, Indian Institute of Management Indore. email: punarjitroychowdhury@gmail.com. Prabandh Shikhar, Rau-Pithampur Road, Indore, Madhya Pradesh 453556, India.

"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[1]

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 (WHO), 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.[2] 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 an article published in *The Washington Post* (February 22, 2018) in the US alone this cost is about $460 billion annually.[3] In this paper, we provide the first *causal* analysis of the impact of women's age at marriage on their exposure to domestic violence, and more specifically spousal or intimate

[1] http://www.bbc.com/news/world-asia-india-29708612

[2] http://apps.who.int/iris/bitstream/handle/10665/77432/WHO_RHR_12.36_eng.pdf;jsessionid=6CADDFE06611AE8566C228C40575BE46?sequence=1

[3] https://www.washingtonpost.com/opinions/the-cost-of-domestic-violence-is-astonishing/2018/02/22/f8c9a88a-0cf5-11e8-8b0d-891602206fb7_story.html?noredirect=on&utm_term=.9c1983316e34

partner violence (IPV).[4] We use newly available nationally representative data from India, where according to a *BBC* report (October 24, 2014), one incident of domestic violence is reported in every five minutes (which, of course, is only a fraction of how much actually occurs).[5] 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, and be less resistive to domestic violence because they are young. This makes them '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).[6] This limits their options outside marriage and the economic and social resources at the women's disposal (Chowdhury et al. 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 and 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 women who marry late may be more vulnerable to mistreatment. Overall, thus,

[4] Although technically spousal violence or IPV is a subset of domestic violence, we shall use the terms domestic violence and IPV interchangably throughout the paper since three-quarters of violence against women is intimate (Aizer 2010).

[5] http://www.bbc.com/news/world-asia-india-29708612

[6] For example, Field and Ambrus (2008) find that an additional year of marriage delay leads to an increase in schooling by 0.22 years for women in Bangladesh.

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, 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. 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 victims of 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. As noted by Sekhri and Debnath (2014), variation in the age at menarche generates a quasi-random difference in the age at which a girl enters the marriage market. This instrument is motivated by the observation that has been made by sociologists and anthropologists that parents become extremely anxious to get their daughters married once they have reached menarche, partly to avert any unwanted pregnancies (Caldwell et al. 1983; Srinivas 1984).

We note up-front that ideally one would perhaps be better off using administrative data instead of survey data for studying domestic violence, since administrative data such as police reports or hospital records are objective measures of violence and are not subject to self-reporting bias. However, as noted by Erten and Keskin (2018), this kind of information is very likely to be flawed, especially in a developing country context. This is because in such countries only a selected group of women has access to hospitals or police stations after they experience a violent episode. Moreover, it is even more difficult, if not impossible, to capture the extent of emotional violence using administrative reports. By using a carefully designed survey that includes self-reported information on physical, sexual and emotional violence against women, we are able to examine the effect of marriage timing on different forms of domestic violence that are otherwise impossible to observe.

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 rules 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 significant at 5% level of significance. However, the effect of women's age at marriage on sexual violence and emotional violence are not statistically significant. We show that our results are robust to alternative non-linear methods of estimation. Further, 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 a subsample of women for whom we should not find any such effect. The results of this test suggest that our instrument is likely to satisfy the exclusion restriction, and thus increase our confidence in the empirical strategy employed.

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 et al., 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 e.g. Oshiro et al. 2011; Santhay 2011; Speizer and Pearson 2011; Nasrullah et al. 2014; Rahman et al. 2014; Yount et al., 2017). These studies, however, fail to establish a causal relationship by accounting for the potential omitted variable bias or measurement error. As such, our work is the first causal analysis of the relationship between women's age at marriage and domestic violence.

Additionally, our study contributes to the literature that looks at the impact of women's marital age on their wellbeing measured along various socioeconomic dimensions including

schooling, health, labor market outcomes, and human capital of women's children in developing countries like Bangladesh and India (see for e.g. Field and Ambrus 2008; Sekhri and Debnath 2014; Chari et al. 2017; Dhamija and Roychowdhury 2018). Given that child marriage and early marriage are issues of deep concern in developing countries,[7] this study by focusing on domestic violence is likely to extend our understanding of the effects of early marriage of women. Our findings 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 dataset used. Section 3 presents the econometric model and empirical strategy. Results are presented in the section 4. The last section concludes.

2 Data

The data come from the fourth round of National Family Health Survey (NFHS-4) of India, 2015-16. 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.[8] 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

[7] The mean marital age of women in India was 19.3 years according to the 2011 Census data. Moreover, an article in the *The Wire* (June 1, 2016) states that in India as many as 102 million girls (30% of the female population) were married before 18 in 2011 even though the Prohibition of Child Marriage Act states that a girl in India cannot marry before she turns 18. In Bangladesh 46.1% of women between the ages of 15 and 19 were married between 2003 and 2005. Corresponding figures for some other poor countries for the same period were: 42% in Chad, 32.9% in Malawi, 50.4% in Mali, 38.2% in Mozambique and 31.7% in Nigeria (data from Demographic and Health surveys). For the developed countries, the average age of marriage for women is much higher. For instance, the mean age of marriage of women in the US was 26.9 years in 2011 (Pew Research Foundation, 2011), for Germany it was 30.9 years, and for Sweden 33.3 years.

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

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 all the 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 information on menstrual hygiene and related topics, including age at menarche, was collected from the women in the age group of 15-25 years. Note, to answer questions specifically on domestic violence, only one eligible woman per household was randomly selected.

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."[9] For instance, as noted previously, only one woman per household was selected for the interviews.[10] 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

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

[10] 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.

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.[11]

The domestic violence measures include binary variables on whether a woman had ever experienced any kind of less severe physical violence, severe physical violence, sexual violence and emotional violence from her intimate partner (spouse). Less severe physical violence is measured by acts of pushing, shaking, throwing something, twisting arm, pulling hair, slapping, punching with partner's fist or something else. Severe physical violence is measured by acts of kicking, beating, choking, burning, threatening or attacking with any kind of weapon. Sexual violence is measured by 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 each of the four categories of domestic violence, for a particular women, the binary variable takes a value one if the woman reported to have faced at least any one kind of the underlying violences.

Of course, one could argue that when studying domestic violence it would have been ideal to use administrative data such as police reports or hospital records since these are objective measures of violence and are not subject to self-reporting bias. However, as noted by Erten and Keskin (2018, p. 67), "this kind of information is likely to be flawed, especially in a developing country setting in which only a selected group of women has access to hospitals or police stations after they experience a violent episode." Using surveys from 24 countries in the DHS program, Palermo et al. (2014) in fact show that only seven percent of women who experienced such violence made a formal report that would be captured in administrative data (e.g., police, medical, or social services). For India in particular, the situation seems

[11]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.

to be more grim. According to a recent report of the *The Times of India* (December 6, 2017), comparison of the data from the NFHS-3 and the National Crime Records Bureau of India reveals that "less than 0.17% of women who face marital violence actually filed a case under Section 498A IPC [the section of the Indian Penal Code that deals with domestic violence]." This self selection in reporting domestic violence makes the use of administrative data unattractive. Moreover, it is even more difficult, if not impossible, to capture the exposure to emotional violence using administrative records.

For the analysis of the effect of women's age at marriage on the prevalence of domestic violence, we restrict ourself to the women who have non-missing information on the different categories of domestic violence, whose marital age is not less than 5 years[12] 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.[13]

Table 1 provides descriptive statistics of our analytical sample. In our sample, 25% of the women have faced less severe physical violence, 6% have faced severe physical violence, 6% have faced sexual violence, and 11% have faced emotional violence.[14] 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 proportion of women exposed to different types of domestic violence by their age at marriage.

[12]In our sample, year of marriage was known for 99.73 percent (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 of marriage (reported only for those cases where the year of marriage is unknown) available in the data set.

[13]The normal menarcheal age is between 10 and 15 years. However, menarcheal age as low as 9 years is not unusual (see for e.g. https://timesofindia.indiatimes.com/city/goa/Girl-talk-Menarche-now-at-8-9-years/articleshow/34169175.cms). 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 pathologic 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.

[14]The proportion of women who have faced at least one of the four kinds of domestic violence in our sample is 28%.

3 Empirical strategy

3.1 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 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. 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.

While estimating equation (1), we exclude various determinants of domestic violence such as educational attainment from the estimation, as these variables are potentially endogenous variables that could be influenced by a woman's decision about her timing of marriage. That is, these variables themselves could be the reasons why age at marriage affects women's exposure to domestic violence. We thus condition on only exogenous variables. 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.[15][16]

The second issue relates to the accuracy of the reported age of marriage. In the NFHS 2014-15, 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, partly to avert any unwanted pregnancies (Caldwell et al. 1983; Srinivas 1984; Chari et al. 2017). 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.[17] In what follows, we discuss our IV strategy in detail.

[15] 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.

[16] 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.

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

3.2 Instrumental variable strategy

The IV approach involves estimating a two stage 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 first stage is given by the equation (2), and equation (3) is the structural equation. 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 (TSLS)) and cluster standard errors at the district level.[18]

3.3 Validity of the instrumental variable

In this section, we perform several checks to test the validity of the instrumental variable. First, we examine whether age at menarche predicts women's age at marriage which is the endogenous regressor. In line with the findings of Field and Ambrus (2008) in context of Bangladesh, and that of Sekhri and Debnath (2014) and Chari et al. (2017) in context of India, we find that age at menarche is significantly correlated with women's age at marriage. The results from the regression of women's age at marriage on age at menarche are pre-

[18] Later we use alternative non-linear methods 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 (2009, 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."

sented in Table 2. Column (1) reports the coefficient of age at menarche without additional controls. The value of the coefficient is 0.216, and it is statistically significant at 1% level of significance. The F-Statistic for the regression model is 77.63. These results eliminate concerns about 'weak instruments' Additionally, Figure 4 also presents the kernel density estimate of women's age at marriage by menarcheal age groups (early and late menarche)[19] revealing that the distributions of women's age at marriage is positively related to age at menarche.

Next, we examine the potential threats to the validity of this instrument. Medical literature suggests that 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 e.g. Stathopolu et al. (2003) note that acute malnutrition could result in stunting) and their labor market prospects, in turn reducing their options 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, we include adult height in the regression in column (2). As noted by Chari et al. (2017), if height is a sufficient statistic for health investments and if undernutrition that affects menarche is also severe enough to result in stunting, then conditioning on height is likely to eliminate any confounding factor related to health investments that affect both menarche and marriage conditions. We find that inclusion of height as an additional control changes the point estimates slightly (the standard errors remain unchanged). Even if height is not a sufficient statistic for health, since it is closely related to health (Strauss and Thomas 1989), the fact that controlling for height has very small effects on our results suggests that they are not driven by unobserved health inputs that also affect age at menarche.

As argued by Field and Ambrus (2008), sudden changes in diet might also impact maturation. Sekhri and Debnath (2014) in this context note that, agriculture and agriculture-

[19]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.

related activities, that employ majority of the Indians, are extremely dependent on weather. Extreme weather conditions such as droughts and floods 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. We control for this possibility in our first stage regression. In column (3), in addition to height, we add age of the women to account for extreme weather events at the time of birth. Moreover, we also include controls for spousal age in column (3). The point estimates and standard errors are similar across columns (2) and (3). We condition all subsequent results on women's height, women's age, and spousal age.

It is thought that hard physical labor during childhood can have a negative effect on children's health and lead to a delay in menarche (Pellerin-Massicotte et al. 1997). 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. However, as argued by Sekhri and Debnath (2014), the children who work in India are not involved in hard physical work such as construction. They note that detailed data on child labor collected from northern India show that more than 99 % of working girls of age 6 to 14 are engaged in domestic work while 0.001 % of them work for wage (Basu et al. 2010). As such, strenuous physical labor during early childhood is unlikely to render our instrument endogenous. Nevertheless, 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), a variable capturing domestic violence in the women's natal household, 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* (November 25, 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."[20] 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. It is also likely that domestic violence is more prevalent in households belonging to the lower end of the income distribution. So the indicator for domestic violence in women's natal family is also likely to serve as an additional proxy for the economic status of the women's natal household. Finally, caste should also serve as an additional proxy of women's natal family economic status. As evident from the results reported in Column (4), 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.[21]

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

[20] https://www.economist.com/news/special-report/21731491-parents-lose-control-over-marriage-indian-society-shaking-marriage-india

[21] 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.

still highly statistically significant and the first stage F statistic is sufficiently high.

The final 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. 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.[22] 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 an upward trend in the relationship between schooling and age at menarche. 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. Third, we explore the relationship between years of schooling of women, age at menarche, and marriage age using a regression framework. Results are reported in Table 3. We find that age at menarche has a positive and significant impact on years of schooling when we do not control for age at marriage. However, when we control for age at marriage, menarcheal age no longer significantly affects

[22] 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.

educational attainment of women (in fact, the coefficient of age at menarche is now close to zero). This suggests that conditional on age at marriage, menarcheal age does not have an effect on educational attainment. Thus, all the evidences suggest that not controlling for educational attainment of women is unlikely to confound our analysis.[23]

4 Results

4.1 OLS results

The OLS estimates of the effect of women's age at marriage on domestic violence are presented in Table 4. 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 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 3 percentage points, severe physical violence by 0.9 percentage points, sexual violence by 0.8 percentage points, and emotional violence by 1.4 percentage points respectively. These effects are statistically significant at 1% level of significance. When we include controls for only demographic char-

[23]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 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 regression, which must be assumed to be uncorrelated to menarcheal age, for their second stage parameter estimates to be consistent.

acteristics, 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 has family norms and/or ability, we use the IV approach.

4.2 IV results

We next turn to the IV results in Table 5. 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 8 percentage point decline in the probability of less severe physical violence, 4 percentage point decline in the probability of severe physical violence, 2 percentage point decline in the probability of sexual violence, and 4 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% level of significance. The effect of women's age at marriage on sexual violence is significant at 5% level.

When we include controls for demographic characteristics, these estimates change slightly: a one year delay in women's marriage now leads to a 7 percentage point decline in probability of less severe physical violence, a slightly over 4 percentage point decline in probability of severe physical violence, a 1 percentage point decline in probability of sexual violence, and a 3 percentage point decline in probability of emotional violence. However, now, although the effects of women's age at marriage on less severe physical violence and severe physical violence are statistically significant at 5% level of significance, the effects on sexual violence and emotional violence are no longer statistically significant.

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 and severe forms of physical violence remain almost unchanged compared to the magnitude of that obtained from the specifications that include only demographic controls. Specifically, a one year delay in marriage of women causes the probability of less severe and severe physical violence to decrease by 7 percentage points and 4 percentage points respectively. These effects continue to remain statistically significant at 5% level of significance as well. The effects of women's age at marriage on sexual violence and emotional violence, however, are now much smaller in magnitude (around 0.4 percentage points) compared to the specifications that included demographic characteristics but not district fixed effect. Moreover, these effects continue to remain statistically insignificant.[24]

Thus, our results indicate that a one year increase in women's age at marriage nationwide would reduce the prevalence of less severe physical violence from 25% of women to 18%, and that of severe physical violence from 6% to 2%. 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,[25] 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 73 million to 53 million, and the number of women exposed to severe physical violence to fall from 18 million to 6 million.

In sum, thus, our IV results suggest that a year of delay in marriage causes a significant

[24]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.

[25]http://www.censusindia.gov.in/vital_statistics/srs_report/9chap%202%20-%202011.pdf

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.[26]

4.3 Robustness Checks

4.3.1 Alternative methods 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 maximum likelihood method (IV-Probit) as well as 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 Panels A and B of Table 6.

Our results remain qualitatively unchanged. Specifically, our preferred specifications (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

[26] The IV approach used in this paper addresses the potential bias due to omitted variables like ability of women, social norms, discounting factor, etc. It also accounts for endogeneity due to measurement error in age at marriage. However, if the variables capturing domestic violence are measured with error (in the sense that women do not always correctly report their exposure to domestic violence), and if this measurement error is non-random, then the bias induced by this non-classical measurement error may not be addressed by our IV approach. Although the NFHS, following international guidelines, takes several kinds of precautions to avoid misclassification error in the domestic violence variables, in principle, the chances of false positives or false negatives cannot be fully ruled out. To assess whether there are significant self-reporting error in domestic violence data, Aguero and Frisancho (2017) develop a new method to deal with self reporting bias. They implement an indirect questioning technique which provides further anonymity to the respondents and compare the prevalence rates of physical and sexual IPV estimated by this method to that obtained from direct questions from the DHS. In particular, they apply the methodology of list experiments (e.g., Blair and Imai, 2012; Karlan and Zinman, 2012; Glynn, 2013) as well as DHS direct questions to a sample of women of Lima, Peru. On average, they find no significant differences in reporting of physical and sexual violence across direct and indirect methods. However, for the subsample of women who completed college education, they find some evidence of misreporting. If one is willing to assume that these results are externally valid or is willing to extrapolate these results to other developing countries, it might be useful to re-estimate our baseline IV regression model by excluding those women who have completed at least 15 years of education. We carry out this exercise, and find that the effect of age at marriage on less severe and severe forms of physical violence continue to remain negative and statistically significant. The results of these regressions are not reported in the paper, but are available from the authors upon request.

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.[27]

4.3.2 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 7. Columns (1), (3), (5) and (7) report the OLS coefficients of age at menarche from the regressions based on the subsample of women who got married before attaining menarche. For comparison purpose, we also present the coefficients of age at menarche from the regressions based on the subsample of women who did not get married before attaining menarche in Columns (2), (4), (6) and (8).

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. This implies that there exists no systematic relationship between

[27] 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.

age at menarche and domestic violence for 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 (see Columns (2) and (4)). 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.

4.3.3 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 marriage 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).[28] We graph 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 explore this issue, ideally we would like to follow Field and Ambrus (2008) and compare 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 plot the distribution of women's age at marriage and menarche age by their report of whether they have witnessed domestic violence in natal home in Figure 7. It is likely that the households in which domestic violence is (not) prevalent are households that belong to the (higher) lower socioeconomic status, and these households might have a preexisting preference for earlier (later) marriage of their daughters. So for women who come from households in which they have witnessed domestic violence between

[28] 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".

their parents, their age at marriage might be lower compared to the others. However, the age at menarche for these women should not be different from the others. Reassuringly, we find that the distribution of women's age at marriage differs across the two subsamples, but not the age at menarche. This provides suggestive evidence that the measurement error in age at marriage and age at menarche are unlikely to be correlated.

As an additional check, we also use data from the Indian Human Development Survey (IHDS) 2012, which is nationally representative household level survey conducted between 2011 and 2012 by the National Council of Applied Economic Research (NCAER) and the University of Maryland and plot the distribution of women's marriage age and menarche age for those with mothers who never attended school, and those with mothers who had some level of schooling in Figure 8. As evident, we find that the distribution of women's age at marriage differs across the two subsamples, but not the age at menarche. This indicates that the sampled women are unlikely to link the two events – marriage and menarche. If we are willing to extrapolate these results for women in India in general, measurement error in age at marriage and menarche age are unlikely to be correlated for the women in our sample (note that, our sample was administered only a couple of years after the IHDS, and it is not very different from IHDS in terms of breadth and scope).

5 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 nationally representative 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 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.

Our findings confirm the relevance of conditional cash transfer programs and other social policies that seek to delay marriages of women in India (e.g. "Kanyashree Prakalpa" program in West Bengal, "Apni Beti Apni Dhan" program in Haryana, etc.) in reducing the prevalence of domestic violence. 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. In addition, it would also be interesting to examine whether age at marriage impacts physical violence by itself or whether the causal effect is mediated through some specific channel such as education. This would be useful for fine-tuning the existing marital delay policies as well as for designing newer and more innovative ones, in order to reduce domestic violence.

References

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

[2] Aguero, J.M., and Frisancho, V. (2017). "Misreporting in Sensitive Health Behaviors and its Impact on Treatment Effects: An Application to Intimate Partner Violence," IDB Working Paper No. 853.

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

[4] 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.

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

[6] Basu, K., Das, S., and Dutta, B. (2010). "Child labor and household wealth: Theory and empirical evidence of an inverted-U," *Journal of Development Economics*, 91, 8–14.

[7] Blair, G., and Imai, K. (2012). "Statistical Analysis of List Experiments," *Political Analysis*, 20, 47–77

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

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

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

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

[12] Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement Error in Survey data, In: Arrow, K., Intriligator, M. (Eds.), *Handbook of Econometrics*, vol. 5, 3705–3843, North Holland.

[13] Caldwell, J., Reddy, P., and Caldwell, P. (1983). "Causes of marriage change in South India," *Population Studies*, 37, 343–361.

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

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

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

[17] Chowdhury, F.D. (2009). "Theorizing patriarchy: The Bangladesh context," *Asian Journal of Social Science*, 37, 599–622.

[18] Coker, A. L., Davis, K. E., Arias, I., Desai, S., Sanderson, M., Brandt, H. M., and 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.

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

[20] Ellsberg, M., Jansen, H. A., Heise, L., Watts, C. H., and 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.

[21] Erten, B., and 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.

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

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

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

[25] 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.

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

[27] Frisch, R. E., Von Gotz-Welbergen, A., McArthur, J. W., Albright, T., Witschi, J., Bullen, B., Burnholz, J., Reed, R. B., and 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.

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

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

[30] Glynn, A. N. (2013). "What Can We Learn with Statistical Truth Serum?: Design and Analysis of the List Experiment," *Public Opinion Quarterly*, 77, 159–172.

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

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

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

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

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

[36] Karlan, D., and Zinman, J. (2012). "List Randomization for Sensitive Behavior: An Application for Measuring Use of Loan Proceeds," *Journal of Development Economics*, 98, 71–75.

[37] Must, A., Phillips, S. M., Naumova, E. N., Blum, M., Harris, S., Dawson-Hughes, B., and 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.

[38] Nasrullah, M., Zakar, R., and 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.

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

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

[41] Palermo, T., Bleck, J., and Peterman, A. (2014). "Tip of the Iceberg: Reporting and Genderbased Violence in Developing Countries," *American Journal of Epidemiology*, 179, 602– 612.

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

[43] Rahman, M., Hoque, A., Mostofa, G., and 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.

[44] Raj, A., Saggurti, N., Lawrence, D., Balaiah, D., and 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.

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

[46] 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.

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

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

[49] Speizer, I. S., and 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.

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

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

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

[53] Strauss, J., and Thomas, D. (1989). "Health, nutrition, and economic development," *Journal of Economic Literature*, 36, 766–817.

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

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

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

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

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

Figure 1. Distribution of women's age at marriage

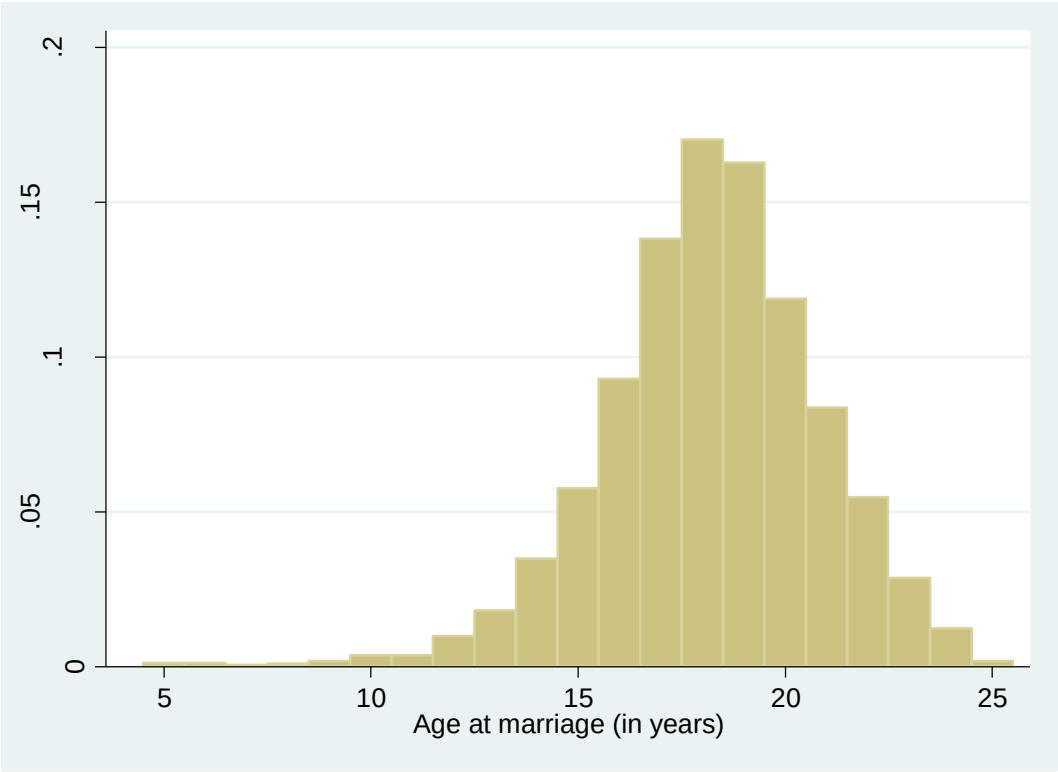


Figure 2. Distribution of age at menarche

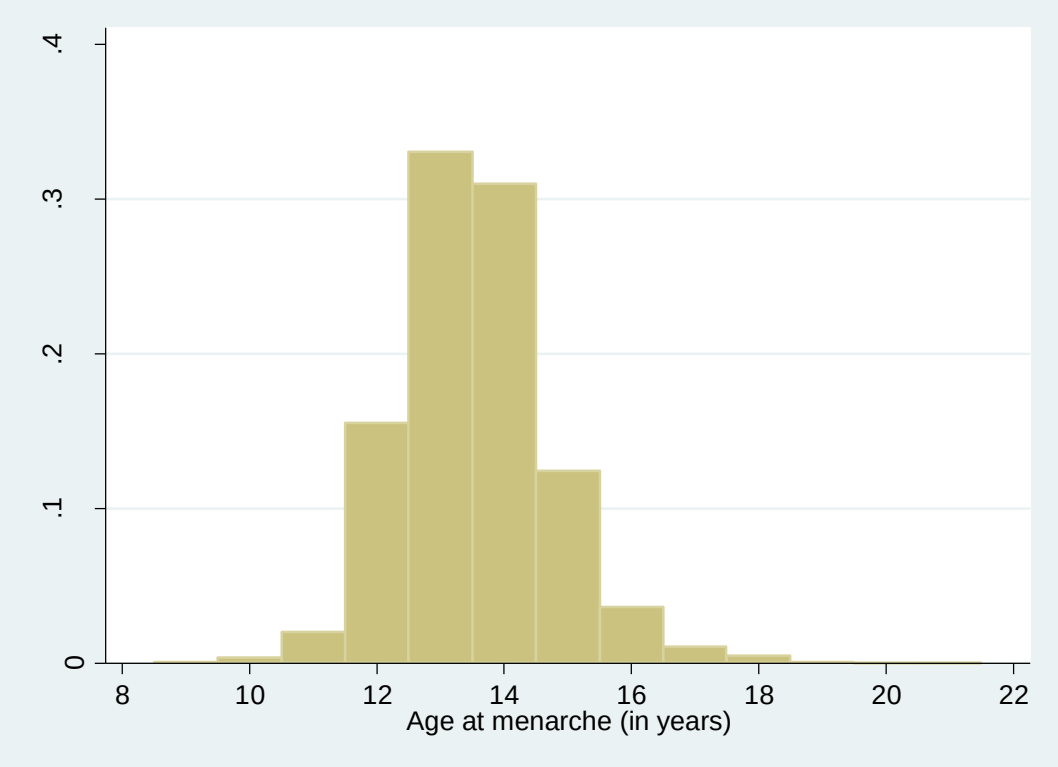
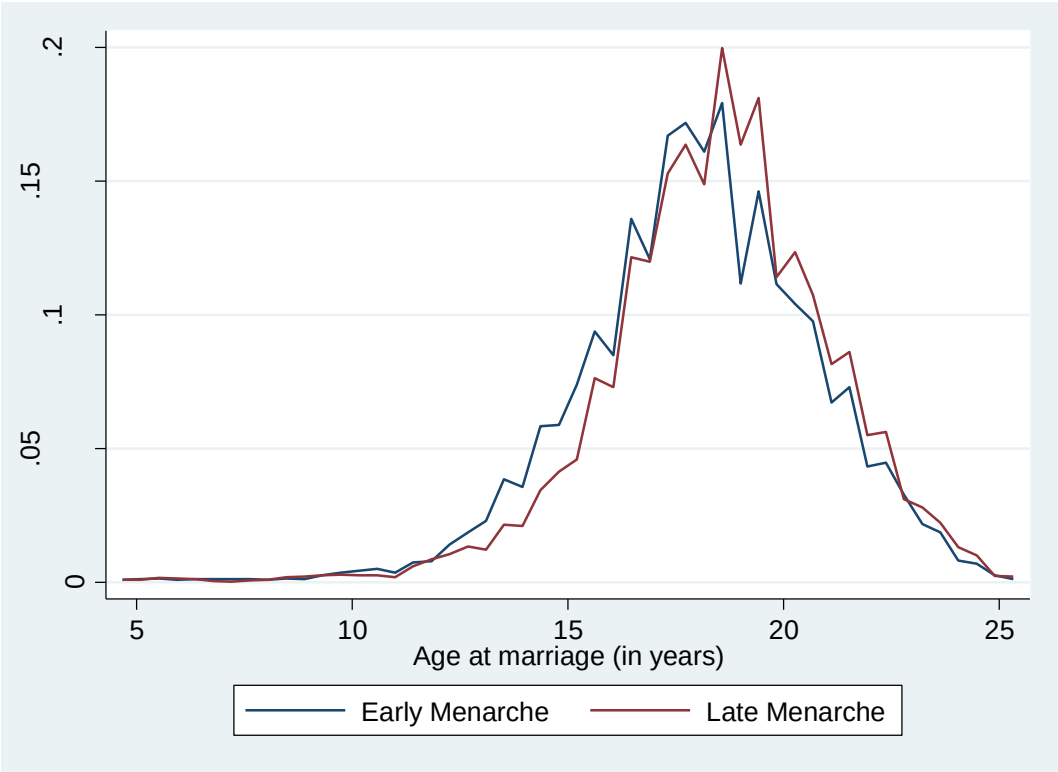


Figure 3. Prevalence of domestic violence by age at marriage

Notes: Early marriage group includes those women who got married before they were 19 years old. Late marriage group includes the rest.

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

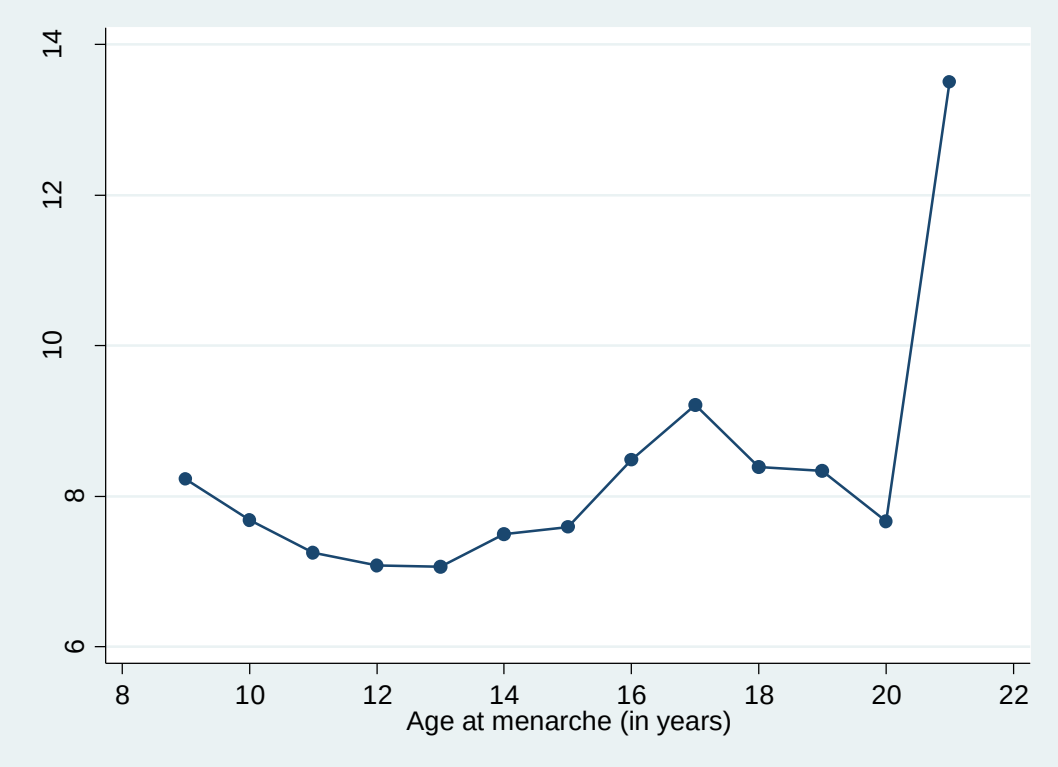


Figure 6. Kernel density estimates of women's years of schooling by terciles of age at menarche

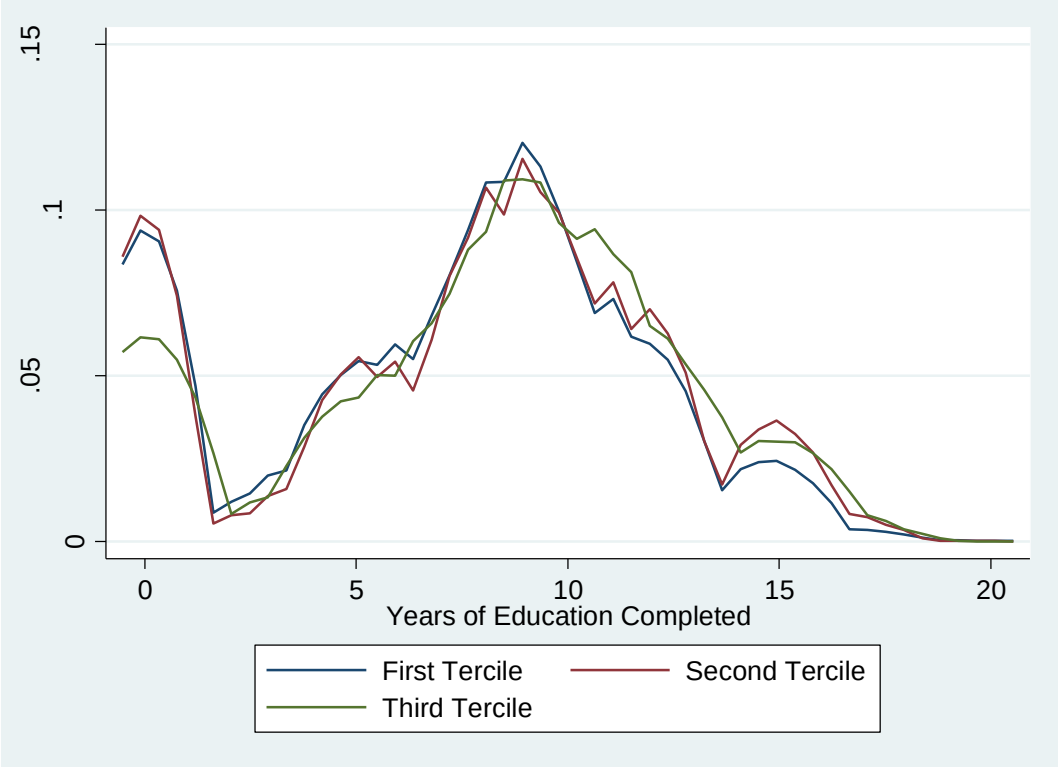


Figure 7. Kernel density estimates of women's age at marriage and age at menarche by domestic violence status in natal household

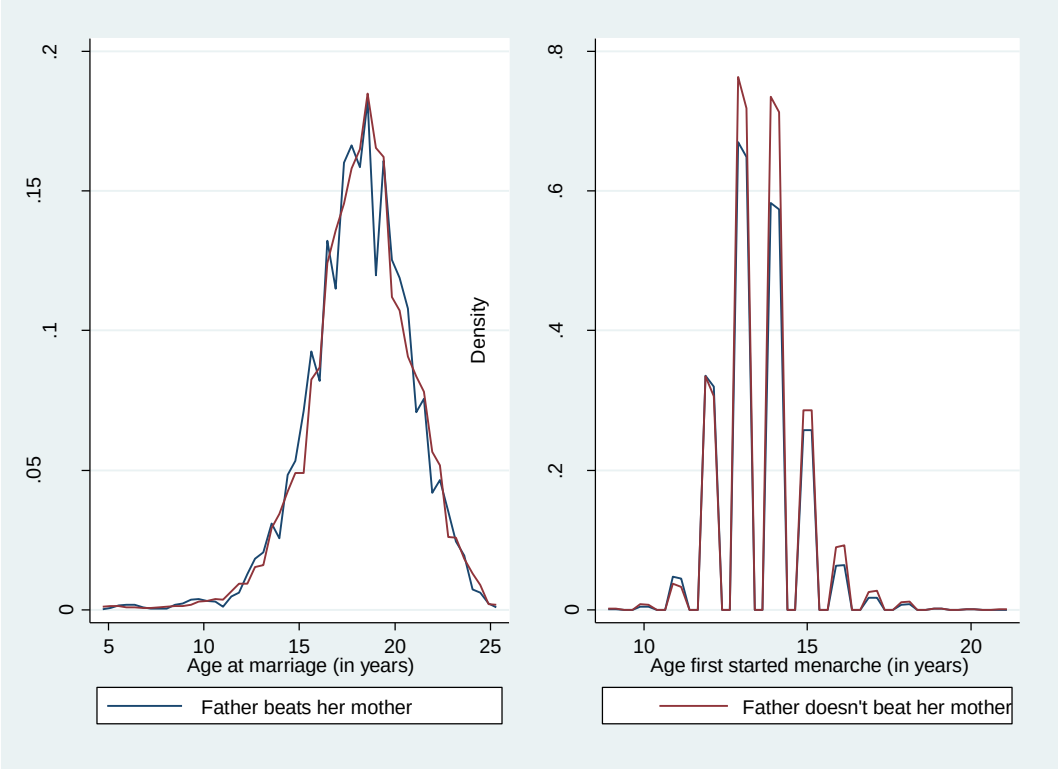


Figure 8. Kernel density estimates of women's age at marriage and age at menarche by educational attainment of mothers, IHDS 2012 data

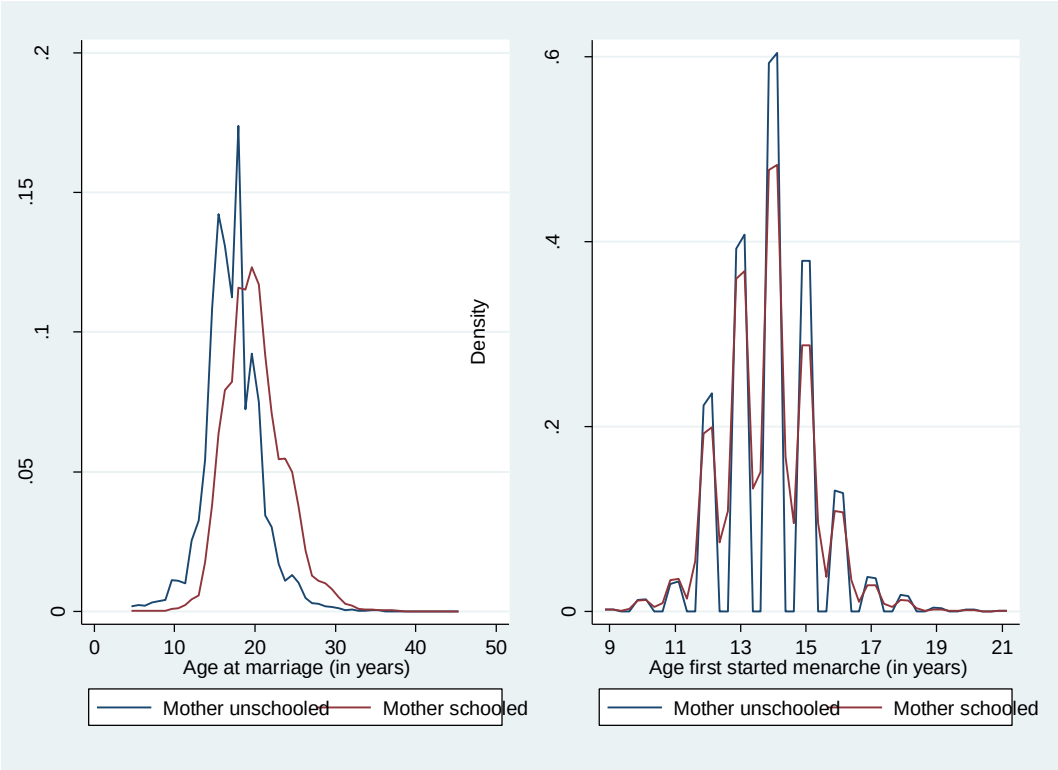


Table 1. Summary Statistics

	Mean	SD
Domestic Violence Outcomes		
Less Severe Physical Violence	0.25	0.43
Severe Physical Violence	0.06	0.24
Sexual Violence	0.06	0.25
Emotional Violence	0.11	0.31
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***	0.194***	0.161***	0.128***	0.127***
	(0.024)	(0.024)	(0.023)	(0.022)	(0.024)
F-statistic	77.63	67.51	193.78	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. In column (2) regression we include women's height as a control. In column (3) regression the control variables are women's height, age, and spousal age. In column (4), controls include women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, and women's caste affiliation. In column (5), we include 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 menarche and age at marriage on years of schooling

	[1]	[2]
Age at Menarche	0.059*	0.004
	(0.036)	(0.034)
Age at Marriage		0.433***
		(0.019)
Observations	9,343	9,343

Notes: Estimation via OLS. The outcome variable is women's years of schooling completed. In both regressions, we include controls for women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, and women's caste affiliation. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 4. 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	-0.026***	-0.024***	-0.022***	-0.009***	-0.008***	-0.008***	-0.008***	-0.007***	-0.006***	-0.014***	-0.013***	-0.012***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
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, and women's caste affiliation. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 5. IV estimates of the effect of age at marriage on domestic violence

	Less Severe Domestic Violence			Severe Domestic Violence			Sexual Violence			Emotional Violence		
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
Age at Marriage	-0.076***	-0.068**	-0.068**	-0.037***	-0.043**	-0.042**	-0.022**	-0.013	0.004	-0.037***	-0.028	-0.004
	(0.019)	(0.030)	(0.031)	(0.010)	(0.017)	(0.019)	(0.011)	(0.018)	(0.020)	(0.013)	(0.021)	(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	65.41	30.40	27.42	65.41	30.40	27.42	65.41	30.40	27.42	65.41	30.40	27.42
	[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, and women's caste affiliation. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 6. Estimates of the effect of age at marriage on domestic violence, Alternative estimation methods

Panel A. IV Probit Estimates: 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	-0.066***	-0.062***	-0.070***	-0.049***	-0.056**	-0.074***	-0.024*	-0.014	0.013	-0.037***	-0.025	-0.007
	(0.012)	(0.021)	(0.023)	(0.014)	(0.022)	(0.027)	(0.015)	(0.021)	(0.031)	(0.014)	(0.023)	(0.026)
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	8,504	9,343	9,343	5,367	9,343	9,343	5,471	9,343	9,343	7,081

Panel B. Probit Estimates using the 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	-0.035***	-0.031***	-0.082**	-0.014***	-0.008***	-0.074**	-0.015***	-0.012***	0.017	-0.020***	-0.018***	-0.006
	(0.004)	(0.005)	(0.035)	(0.002)	(0.003)	(0.032)	(0.002)	(0.003)	(0.034)	(0.003)	(0.004)	(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	8,504	9,343	9,343	5,367	9,343	9,343	5,471	9,343	9,343	7,081

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, and women's caste affiliation. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.

Table 7. 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]
Age at Menarche	0.0035	-0.012***	-0.038	-0.007***	0.010	-0.000	-0.027	-0.003
	(0.059)	(0.004)	(0.045)	(0.002)	(0.041)	(0.002)	(0.047)	(0.003)
Observations	317	9,026	317	9,026	317	9,026	317	9,026

Notes: Estimation via OLS. The outcome variables are different categories of domestic violence. Regressions reported in columns (1), (3), (5) and (7) are based on the subsample of women who got married before attaining menarche. Regressions reported in columns (2), (4), (6) and (8) are based on the rest of the women. All specifications include demographic controls such women's height, age, spousal age, wealth dummies, indicator for whether women have seen domestic violence among their parents, and women's caste affiliation, as well as district fixed effects. Standard errors reported in the parentheses are clustered at the district level. ***p < 0.01, **p <0.05, *p < 0.1.