Educational Hypogamy and Female Employment in Rural India [*]

Punarjit Roychowdhury[†]
Shiv Nadar University, Delhi NCR and GLO
Gaurav Dhamija
Indian Institute of Technology Hyderabad

November 29, 2022

Abstract

Educational hypogamy - the practice of men marrying women who are more educated than themselves - has been increasing in rural India over the last two decades. Can this explain rural India's declining female labor force participation rate (FLFPR)? We examine this question by testing whether women in hypogamous marriages are less likely to participate in the labor force than women in non-hypogamous marriages in rural India. This could be the case since women in hypogamous marriages are viewed as 'gender norm deviant' which is likely to cause their marriage quality to be worse than that of women in non-hypogamous marriages. This might make participation in labor force costlier for the former than the latter. To estimate the causal relationship between hypogamy and women's labor force participation, we employ a non-parametric bounds approach. We find that, indeed, compared to women in non-hypogamous marriages, women in hypogamous marriages are significantly less likely to participate in the labor force. Further, we provide suggestive evidence that this is likely because marriage quality of women in hypogamous marriages are relatively worse. Overall, therefore, our results suggest the rise in hypogamy is likely an important reason for the decline in FLFPR in rural India.

JEL: J12, J16, J22, O12
Keywords: Female Labor Force Participation Rate, Hypogamy, India, Partial Identification, Women.

[*]The authors are very grateful to Bharti Nandwani and Shubhro Sarkar, participants at the inaugral ICG Annual Conference held on 26-27 August 2022 at Goa, and seminar particpants at IGIDR and IIFT Kolkata for helpful comments and suggestions. Roychowdhury had carried out a part of the research while he was visiting IGIDR Mumbai in fall 2022 He gratefully acknowledges the excellent hospitality extended to him by IGIDR during his stay. The data used are publicly available. They, along with all code for data cleaning and analysis, can also be obtained from the corresponding author upon request. All remaining errors are our own.

[†]Corresponding Author. Department of Economics, School of Humanities and Social Sciences, Shiv Nadar University, NH91, Tehsil Dadri, Greater Noida, Uttar Pradesh 201314, India. Email: punarjitroychowdhury@gmail.com.

1 Introduction

Over the past two decades, rural India has witnessed a secular decline in female labor force participation rate (FLFPR). As evident from Figure 1 plotted using data from different rounds of the National Sample Survey (NSS) and Periodic Labor Force Survey (PLFS) 2017-18, from 33% in 1993-94, FLFPR in rural India fell to 25% in 2004-05 and further to 18% in 2017-18. Analysis of NSS data reveals mainly currently married women, who are 85% of rural women in the 25–65 age group, are the ones driving this decline in labor force participation (Afridi et al., 2018). For urban India, over the same time period, FLFPR has been more or less stagnant.

The decline in FLFPR in rural India is in contrast with the global trends,[1] and is somewhat puzzling since over the past two decades, along with substantial economic growth in the country, there has been a marked rise in female education and a fall in fertility. Figure 2, plotted using the same surveys used to plot Figure 1, clearly shows the proportion of illiterate women fell from close to 65% at the end of the last century to around 42% in 2017-18, whereas over the same period, the proportion of women with middle school education and proportion of women with secondary school (or higher) education increased from 9.6% to 17.8% and from 7.7% to 22.7% respectively. Figure 3 shows fertility rates for rural India plotted based on data from different rounds of the National Family Health Survey (NFHS). From 3.7 in 1992-93, the fertility rate has fallen to 2.4 in 2015-16.

A large literature has come into being in the last one decade or so which has examined several potential causes for the decline in rural FLFPR ranging from a decline of suitable job opportunities to increasing quantity of schooling to increasing returns to home production (see Section 1.1). However, considerable interest in this topic continues to displayed by researchers and policymakers. In light of this, in this paper, we propose and empirically examine a novel and somewhat 'unorthodox' cause of the decline in rural FLFPR: the increase

[1] As shown by Blundell and MaCurdy (1999), there have been large gains by women in labor force participation over time in most developed countries.

in educational hypogamy in rural India over time.

Educational hypogamy occurs when in a marital relationship the wife's educational attainment systematically exceeds that of her husband. As evident from Figure 4 plotted using data from the NFHS, the proportion of such marital relationships in rural India has increased from 7% in 1992-93 to 20.6% in 2015-16 to a further 24.1% in 2019-21. When we look at marriage patterns of different birth cohorts of women surveyed in NFHS and India Human Development Survey (IHDS) 2011-12, a similar increasing trend of hypogamy is visible (see Figure 5).

As noted by Lin et al. (2020), the rise in educational hypogamy in rural India is primarily an outcome of the large gains made by women in education for whom the partner pools are often restricted by certain characteristics like caste and kinship network or by the requirement that marriage has to be consanguineous marriage (in which blood relatives marry),[2] and *not* of reversals in gender norms. As noted by Chakrabarti and Biswas (2012), Kishor and Gupta (2004), Weitzman (2014) and many others, gender inequality is pervasive, and patriarchal norms have still a stronghold in the Indian society. In fact, arranged marriages – which are profoundly rooted in patriarchal culture – are also still near universal especially in the rural parts of the country (Allendorf and Pandian, 2016).[3,4]

How can the rise in educational hypogamy explain the decline of FLFPR in rural India? We hypothesize that women in hypogamous marriages have less likelihood to participate in the labor force than women in non-hypogamous marriages.[5] Since, educational hypogamy (or the proportion of marriages which are hypogamous) is rising, this would mean fewer married women participating in the labor force over time. Consequently, the decline in

[2] As shown by Lin et al. (2020), in consanguineous marriages, more prevalent in South India, it is 20% more likely for a woman to marry beneath her educational level.

[3] In light of this, Lin et al. (2020) remarks, traditional marriage power dynamics have accomodated the rise in hypogamy in India.

[4] In the Appendix, we plot the proportion of men with regressive gender attitudes in rural India based on the NFHS data (see Figure A1). We observe that the proportion of men with regressive gender attitudes have more or less remained stagnant between 2005 and 2021 suggesting that the rise in hypogamy is unlikely a result of changing gender norms.

[5] Throughout the paper hypogamous (non-hypogamous) marriages refer to marriages in which hypogamy (does not) occur with respect to couples' education.

married women's labor force participation rate, and hence overall FLFPR, in rural India should not be surprising (recall, the fall in labor force participation has occurred mainly among currently married women in rural India).

The basis of our hypothesis is as follows. In deeply patriarchal societies like those prevalent in rural India, women are expected to marry those with higher education (or earning power) than themselves (Therborn, 2004; Lin et al., 2020). Hypogamy, clearly, undermines such beliefs and norms about gender roles and women in hypogamous marriages are seen as gender norm-deviant (Macmillan and Gartner, 1999; Bertrand et al., 2015; Baland and Ziparo, 2017; Bernard et al., 2020).[6] This may cause them to face repercussions in form of marital conflict, stress and tension, and severely reduce the quality of marriage.[7] Such conflicts, stress and tension could increase manifold and ultimately lead to marital destabilization if they work outside home as opposed to staying home since that further violates patriarchal norms and prescriptions such as "men should be earning and women should be taking care of the home" or "a woman's place is in the home". Anticipation of this might cause many women in hypogamous women to simply not join the labor force or withdraw from the labor force since divorce is still a taboo in rural India and marital destabilization is extremely costly for rural Indian women.

For women in non-hypogamous marriage, working outside, of course, also is violation of gender norms. However, since their marital quality is likely to be better than women

[6]Hypogamy is in complete contrast with the widespread Indian concept *Anuloma* which is a Sanskrit term used in *Manusmriti* (the famous ancient legal text and constitution of Hinduism) to describe an union between a high born man and a woman of a lower standing (by birth) relative to the respective man. As per Hindu scriptures, *Anuloma* marriages or unions are accepted historically in the Indian society. On the other hand, the reverse union called *Pratiloma* marriages, where a high born woman unites with a man of low birth (relative to the woman) was condemned. Manu bitterly criticizes and condemns these unions which were considered as 'going against the hair or grain'and holds them responsible for the degeneration of the parties involved, subsequent to the union. Hypogamy also goes against the doctrine of *Kanyadan*, which is described by Parry (1979, p. 200) as a doctrine in which "the virgin (kanya) is. . .a meritorious gift made to somebody of superior status".

[7]This line of argument is line with Akerlof and Kranton's (2000) seminal work on how social identity influences economic outcomes. Akerlof and Kranton (2000) propose that every individual belongs to some social category and these social categories are associated with different prescribed behaviors. If individuals deviate from the prescribed behavior of the social category to which they belong, that could be inherently costly since violating prescriptions may devalue others' social identity. In our case, the social groups are man and woman, and these groups are associated with specific behavioral prescriptions.

in hypogamous marriage, such violations are unlikely to significantly increase the chances of marital destabilization. In light of this, women in non-hypogamous marriage are likely to have higher odds of joining the labor force and lower odds of withdrawing than their counterparts. In sum, thus, the likelihood of participation in labor force is likely to lower for women in hypogamous marriages than women in non-hypogamous marriages.[8]

To examine whether women in hypogamous marriages are really less likely to participate in the labor force than women in non-hypergamous marriages, we use data from 2005-06, 2015-16 and 2019-21 rounds of the NFHS and IHDS 2011-12. Both NFHS and IHDS are widely used pan-India household-level surveys, and are extremely rich sources of information on individual education, labor market outcomes, health, household assets etc. Importantly, NFHS and IHDS, both, provide information based on which several vital indicators of labor force participation of women can be constructed (e.g., the information provided in the NFHS allows us to construct variables indicating whether or not a woman was employed when the survey was being conducted, whether or not that employment was paid, etc.). Further, NFHS and IHDS allows us to construct the indicator of educational hypogamy which indicates whether or not in a marital relationship the wife is more educated than the husband.

Identifying the causal effect of educational hypogamy on women's likelihood of participation in the labor force is not straightforward because of the usual problem of unobserved confounders (which include not only unobserved ability and skill but also unobserved personality traits like motivation, patience, self-control and confidence, etc.; see Section 3) which cannot be addressed using traditional causal approaches like instrumental variables (IV) or difference-in-differences (DID) without invoking very strong assumptions. To circumvent the identification problem, we use a nonparametric bounding approach (Manski, 1995; Manski

[8]There could be another indirect reason why women in hypogamous marriages have lower labor force participation rate than women in non-hypogamous marriage. It is widely known that marital conflict and tension engenders poor mental health (Grzywacz and Bass, 2003; Bilodeau, Marchand and Demers, 2020). To the extent that poor mental health hampers individuals' employment (Layard, 2017), women in hypogamous marriage could be relatively less likely to participate in the labor force due to marital stress induced poor mental health.

and Pepper, 2000; Pepper, 2000).[9]

The nonparametric bounding approach for conducting inference on parameters in econometric models recognizes that identification is not an all-or-nothing concept, and that models that do not point identify parameters of interest can (and typically do) contain important information about these parameters (Tamer, 2010). Employing this approach we provide sharp bounds on the average treatment effect (ATE) of hypogamy on women's labor force participation, when marriage type is non-random. These bounds typically require weaker (nonparametric) assumptions than those typically employed in traditional instrumental variable (IV) based methods. However, as a consequence of having weaker identification assumptions, we obtain bounds rather than point estimates. Nonetheless, the bounds reveal exactly what can be learned under different assumptions concerning the nature of the selection process. Tamer (2010, p. 168) summarizes the advantages of this approach: "This partial identification approach favors the principle that inference—and conclusions and actions—based on empirical models with fewer suspect assumptions is more robust, hence more sensible and believable. Stronger assumptions will lead to more information about a parameter, but less credible inferences can be conducted."

Our results are compelling. Relying on fairly weak assumptions regarding the selection, for all the four indicators of labor force participation, we find that the bounds on the ATE of educational hypogamy on different indicators of labor force participation are strictly negative and statistically significant indicating that hypogamy reduces women's participation in the labor force. For example, for the NFHS sample, we find educational hypogamy reduces the women's likelihood of being employed at the time of the survey by at least 7-8% (4-5% if we consider only paid employment), and of being employed in the twelve months preceding the survey by at least 8-10% (6-7% if we consider only paid employment). We further show that women in hypogamous marriages are more likely to face emotional stress within marriage

[9] For notable extensions and applications of this approach, see Kreider and Pepper (2007, 2008), Gundersen and Kreider (2008, 2009), Molinari (2008, 2010), Kreider and Hill (2009), de Haan (2011), Gundersen et al. (2012), Kreider et al. (2012), Millimet and Roy (2015), Cygan-Rehm et al. (2017) and Roychowdhury and Dhamija (2022).

compared to women in non-hypogamous marriages indicating that marriage quality of the former is worse than that of the latter. This lends supports to our proposed theory that women in hypogamous marriage are relatively less likely to participate in the labor force because their marriage quality is worse than women in non-hypogamous marriages which makes working outside the home costlier for the former than the latter.

1.1 Literature

As noted previously, our work contributes to the growing body of empirical literature that attempts to explain the declining FLFPR in rural India. Five potential explanations for the puzzling decline in rural FLFPR can be found in this literature. Firstly, women in rural areas are now pursuing higher education and are therefore not available for the labor force (Chowdhury 2011; Rangarajan et al. 2011; Neff et al., 2012). Secondly, increases in education in rural India has led to an increase in women's relative returns to home production compared with market production, thus, adversely affecting their participation in the labor force in rural India (Afridi et al., 2018; Afridi et al., 2022). Thirdly, household incomes have risen in rural areas due to higher wage levels, which has taken the pressure off of women to seek employment (Himanshu 2011; Rangarajan et al. 2011; Neff et al., 2012). Fourthly, the decline in FLFPR is due to an overall decline in suitable short- and long-term employment opportunities in rural areas (World Bank 2010; Chowdhury 2011; Mazumdar and Neetha 2011; Chatterjee et al., 2014; Mehrotra and Parida, 2017). Finally, the decline in the rural FLFPR is due to constraints imposed by social norms regarding the gendered division of labor (the 'requirement' of women contributing a certain number of hours to home production) (Afridi et al., 2022). Evidently, to date, no study in this literature has focused on the rising hypogamy has a potential cause of the decline in rural FLFPR.

Additionally, our work is related to at least two strands of literature. Firstly, it is related to the literature that looks as how norms and cultural factors that perpetuate gender inequality (including gender gaps in employment) in developing countries. This literature has

examined several contributors to gender inequality that derive from context-specific features including dowry system (Bloch and Rao, 2002; Alfano, 2017; Bhalotra et al., 2020; Sekhri and Storeygard, 2014), old-age support norm (Ebenstein and Leung, 2010), patrileanility (Deininger et al., 2013; Anderson and Genicot, 2015), importance of sons in religious rituals (Chakraborty and Kim, 2010; Jayachandran, 2017), norms about gender roles in labor market (Boserup, 1970; Alesina, 2013) etc. The findings of this literature indicate that gender inequality does not stem from just the economic environment one faces, but also on cultural background (see Jayachandran (2015) for a review of literature on cultural factors of gender inequality, and Jayachandran (2021) for a review of literature on how norms perpetuate female unemployment).

Secondly, it is related to a growing literature in economics of family and marriage that show whenever women outcompete their husbands in terms of earnings power or economic status, they are subject to 'male backlash' which in turn affects their lives in several ways. Bertrand et al. (2015), for instance, find that the marriages where the wife is the primary breadwinner are less happy and less stable. In experimental data, Ratliff and Oishi (2013) finds that men's self-esteem is lower when their partner succeeds. Using Danish data, Pierce et al. (2013) find that men who are out-earned by their wives experience higher sexual and mental illness. Bertrand (2019) shows that boys' gender norms, more than girls', appear to be positively influenced by the role model they find in their working mother, especially if she is also the primary breadwinner in the household; however, role model effect for boys associated with mother's work and relative economic power in the household is lessened in more gender conservative environments. In a recent study, Roychowdhury and Dhamija (2022), using Indian data, find that women whose economic status equals or exceeds that of their husbands are more likely to face domestic violence than their counterparts.

The rest of the paper unfolds as follows. Section 2 discusses the data. The empirical strategy is presented in Section 3. In section 4 we discuss the results. The last section concludes.

2 Data

2.1 NFHS and IHDS

For our analysis, the data come from the third, fourth and fifth rounds of NFHS of India (NFHS 2004-05, NFHS 2015-16, NFHS 2019-20), and IHDS 2011-12. The NFHS is a nationally representative household demographic and health survey for India. It 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. The NFHS 2004-05 was conducted between November 2005 and August 2006, and covered 109,041 households located throughout India. The NFHS 2015-16 survey was conducted between January 2015 and December 2016, and covered 601,509 households located throughout India. The NFHS 2019-21 was conducted between June 2019 and April 2021, and covered 636,669 households located throughout India. In every round, the sample was drawn using stratified random sampling. All rounds of the NFHS survey are publicly available at the DHS website.[10]

The IHDS 2011-12, like the NFHS, is also a nationally representative multitopic household survey. It was conducted by the National Council for Applied Economic Research (NCAER) in New Delhi and University of Maryland (Desai et al. 2015) between November 2011 and October 2012, and covered 42,152 households located throughout India. The survey covered all the states and union territories of India (except Andaman and Nicobar, and Lakshadweep which account for less than 0.05 percent of India's population). The sample was drawn using stratified random sampling. The data are publicly available from the Data Sharing for Demographic Research program of the Inter-university Consortium for Political and Social Research (ICPSR).[11]

All the three NFHS rounds administered a separate woman's questionnaire to collect in-

[10] https://dhsprogram.com/Countries/Country-Main.cfm?ctry_id=57&c=India&Country=India&cn=&r=4
[11] http://www.icpsr.umich.edu/icpsrweb/DSDR/studies/36151

formation on all women aged 15-49 in the sampled households. The questionnaire included questions on background characteristics, family planning, nutrition, marriage, sexual activity, husband's background, women's work, and women's empowerment, domestic violence, etc.[12] In line with the NFHS, the IHDS also administered a separate health and education questionnaire for women. The questionnaire included questions on marriage, work force participation, health investments, etc. But this was administered to only one randomly chosen ever-married women above the age of 15 from each surveyed household.

This study examines the impact of educational hypogamy on women's labor force participation. Therefore the covariate of interest (or treatment variable) is educational hypogamy. Based on years of schooling information from the three rounds of the NFHS and IHDS, this is measured using a binary variable which takes a value one if the married woman's years of schooling completed is higher than that her husband's; zero otherwise. The set of the outcomes consist of four binary employment indicators: currently employed in any work (i.e., whether or not a woman is employed in any work when the survey was being conducted), employed in any work in last one year (i.e., whether or not a woman was employed in any work during the twelve months preceding the survey), currently employed in paid work (i.e., whether or not a woman is employed in paid work when the survey was being conducted)), and employed in paid work in last one year (i.e., whether or not a woman was employed in paid work during the twelve months preceding the survey).[13] If a woman is currently employed in any (paid) work, the first (third) variable takes a value one; if not, zero. If a woman was employed in any (paid) work anytime during the twelve months preceding the survey, the second (fourth) variable takes a value one; if not, zero. Note, only the NFHS data allowed us to create all the four employment indicators; with IHDS data, we could create only the second and third employment indicator.

A couple of things are worth noting here. First, given that the focus of our study is

[12] However, questions on certain topics like domestic violence and menstrual hygeine were restricted to a subset of the eligible women.

[13] Paid work indicates work for which women get payment in form of cash or kind.

women's labor force participation, the natural choice of data should have been NSS Employment and Unemployment Surveys since they are specifically designed to measure employment and work force participation of individuals. We could not use these surveys because they provide information on educational attainment of individuals in broad categories only; the exact number of years of schooling completed is not provided.[14] This naturally renders these surveys unsuitable for our purpose since whether a woman is more educated than her husband or not (i.e., a marital relationship is hypogamous or not) cannot be always ascertained based on such data.

Second, using couples' *observed* educational attainment to classify marriages as hypogamous or non-hypogamous, in general, has a risk since educational level at the time of the survey might not be a true reflection of the partners' educational levels at the time of union formation as they may have obtained further qualifications since then (and hence the treatment variable might not reflect whether or not a marital relationship is actually hypogamous or not). For India in particular this, however, is unlikely to be a cause of concern. This is because Indian women generally do not pursue further formal education after getting married given that patrilocal extended families are still prevalent in India (Allendorf, 2013; Lin et al., 2020). Further, most Indian men do not go for tertiary education (the average years of schooling of men (married men) in NFHS 2019-21 is 8.81 (7.89) years) and hence are likely to drop out of educational system before reaching the marital age. This suggests that education of both men and women in India are likely to become 'fixed' before marriage. As such, using couples' *observed* educational attainment to classify marriages as hypogamous or non-hypogamous should not be problematic in the Indian context

[14]For instance, in the NSS Employment Unemployment Survey 2011-12, individuals' educational attainment is reported as: not literate (01), literate without formal schooling: EGS/ NFEC/ AEC (02), TLC (03), others (04); literate: below primary (05), primary (06), middle (07), secondary (08), higher secondary (10), diploma/certificate course (11), graduate (12, postgraduate and above (13).

2.2 Analytical Sample

Our analytical sample consists of 108,894 women from the three rounds of NFHS and 24,120 women from IHDS. These are the women who have non-missing and valid information for all the outcome variables and covariate of interest or treatment variable.

Table 1 presents the descriptive statistics of the analytical sample. As evident from the table, in the NFHS sample, around 32% married women report to be employed in any work when the survey was being conducted, 40% report to be employed in any work in the twelve months preceding the survey, 25% employed in paid work when the survey was being conducted, and 31% employed in paid work in the twelve months preceding the survey.[15] In the IHDS sample, 52% married women report to be employed in any work in the twelve months preceding the surveys, while 27% report to be employed in paid work when the survey was being conducted.[16] Of all the married women in NFHS (IHDS), 21% (17%) were in marriages which could be described as hypogamous (in terms of couples' education). In the NFHS (IHDS) sample, average age and years of schooling of women is 33 (35) and 5 (4) years respectively, while husbands' average age and schooling is 37 (40) years and 7 (6) years respectively.

On average, women in the NFHS (IHDS) sample are 152 (151) cm tall. 77% of women in the NFHS are from Hindu households, around 78% are from non-upper caste households

[15] These are the employment statistics computed based on the samples of all the three NFHS rounds taken together.

[16] Clearly, the employment rate of women in the twelve months preceding the survey is much higher in the IHDS sample than the NFHS sample. In fact, women's employment rate in IHDS is also higher than that computed from the NSS Employment Unemployment Surveys. This is a widely discussed issue in the literature focusing on women's labor force participation in India (see Dhamija and Roychowdhury, 2020). As suggested by Desai (2017), this difference most likely has to do with how participation in labor market of women has been captured in IHDS as compared to other surveys like the NSS and NFHS. Specifically, Desai (2017) note:

> "Unlike the NSSO, the IHDS collects data on both income and employment in a single module. Thus, it first asks whether the household owns or cultivates land, then asks about season-wise production, and finally asks who engaged in farm work. Similarly, for wage and salary work, it lists every single paid activity that individuals undertake, regardless of the number of days they work. This allows for a greater capture of fragmented and multiple activities. As a result, IHDS work participation rates for women are higher than the NSS participation rates, but those for men are comparable."

and 39% reside in southern or eastern India; for IHDS, the respective figures are 84%, 74% and 35%. Finally, the average family size of women in both the samples is 5.7.

3 Empirical Framework

To examine the causal relationship between educational hypogamy and likelihood of married women's labor force participation, we focus on the partial identification of the Average Treatment Effect (ATE). It represents the effect of educational hypogamy on likelihood of participation in the labor market for a randomly chosen married woman from the entire population. To proceed, we define the conditional ATE as

$$\Psi(1,0 \mid X \in \Omega) = P[Y(1) = 1 | X \in \Omega] - P[Y(0) = 1 | X \in \Omega] \tag{1}$$

where Y is the *realized* labor force participation outcome (which is binary in nature), $Y(1)$ denotes the *potential* labor force participation of a woman if she were to have been in a hypogamous marriage, $Y(0)$ denotes the analogous outcome if the woman were to have been in a non-hypogamous marriage, and $X \in \Omega$ denotes conditioning on observed covariates whose values lie in the set Ω. Thus, the ATE reveals how the mean outcome would differ if all women were in hypogamous marriages versus the mean outcome if all women were in non-hypogamous marriages. In our analysis, $Y = 1$ denotes that the woman has participated/participates in the labor, and $Y = 0$ otherwise.

In our analysis, we simplify the notation by suppressing the conditioning on subpopulations of interest captured in X. In the usual regression framework, researchers attempt to "correctly" choose a set of control variables for which the exogenous selection assumption applies. Inevitably, however, there is much debate about whether the researcher omitted "important" explanatory variables. In contrast, conditioning on covariates in our approach serves *only* to define subpopulations of interest as there are no regression orthogonality conditions to be satisfied (recall that we are *not* estimating a regression model). The problem

is well-defined regardless of how the subpopulations are specified (Pepper, 2000).

The main identification problem that arises when assessing the impact of hypogamy on women's labor force participation is the following: the potential outcome $Y(1)$ is (unobserved) counterfactual for all women who are in non-hypogamous marriage, while $Y(0)$ is (unobserved) counterfactual for all women who are in hypogamous marriage. In other words, for any given woman, only one of two potential outcomes is observed. This is referred to as the *selection problem*. Using the Law of Total Probability, this identification problem can be elaborated as follows:

$$P[Y(1) = 1] = P[Y(1) = 1|H = 1]P(H = 1) + P[Y(1) = 1|H = 0]P(H = 0) \qquad (2)$$

where $H = 1$ denotes that a woman is in a hypogamous marriage, and $H = 0$ otherwise. If we observe the actual marriage type of the women, the sampling process identifies $P(H = 1)$ and $P(H = 0)$ and the expected potential outcome conditional on the outcome being observed, $P[Y(1) = 1|H = 1]$. However, the sampling process cannot reveal the mean outcome for those women who are in non-hypogamous marriage, $P[Y(1) = 1|H = 0]$. Thus, $P[Y(1) = 1]$ is not point-identified by the sampling process alone. Absent other information, this value could lie anywhere between 0 and 1. A similar result follows for $P[Y(0) = 1]$.

In light of the identification problem outlined above, we derive bounds on the ATE under minimal and transparent assumptions. In order to derive the bounds in the absence of nonparametric identification of the ATE, we use various assumptions related to the nature of selection process discussed below.

Assumption 1. *No Selection Assumption*

A natural starting point is to ask what can be learned in the absence of any assumptions invoked to address the selection problem (see Manski 1995; Pepper 2000). Following Manski's (1995) terminology, we refer to this case as the case of worst-case bounds.

In the absence of any assumption on the selection into the treatment, we can assume that

the missing counterfactuals $P[Y(1) = 1|H = 0]$ and $P[Y(0) = 1|H = 1]$ must lie within $[0, 1]$ as they represent latent probabilities. Using this information on the missing counterfactuals, we can bound the individuals components of the ATE, $P[Y(1) = 1]$ and $P[Y(0) = 1]$, as follows:

$$P(Y = 1, H = 1) \leq P[Y(1) = 1] \leq P(Y = 1, H = 1) + P(H = 0) \tag{3}$$

$$P(Y = 1, H = 0) \leq P[Y(0) = 1] \leq P(H = 1) + P(Y = 1, H = 0) \tag{4}$$

Each of the terms in these bounds is identified by the observed data. Taking the difference between the upper bound on $P[Y(1) = 1]$ and the lower bound on $P[Y(0) = 1]$ obtains a sharp upper bound on ATE, and analogously a sharp lower bound (Manski, 1995):

$$UB_{ATE} = P(Y = 1, H = 1) - P(Y = 1, H = 0) + P(H = 0). \tag{5}$$

$$LB_{ATE} = P(Y = 1, H = 1) - P(Y = 1, H = 0) - P(H = 1) \tag{6}$$

However, as evident these bounds have a width equals unity and includes zero. Hence, it is not possible to sign the ATE in this scenario. To be able to make any meaningful inference regarding the ATE, therefore, the bounds need be narrowed by making some assumption(s) about the relationship between marriage type and FLFPR. Towards that end, we consider the identifying power of two types of monotonicity assumptions: monotone treatment selection (MTS) and a monotone instrumental variable (MIV) restriction.

Assumption 2. *Monotone Treatment Selection (MTS)*

The MTS (Manski and Pepper, 2000) assumption assumes that the expected potential outcomes move in a particular direction *conditional on treatment assignment* (i.e., when individuals are compared *across* the treatment as well as the control group). In our context, we

assume that women in hypogamous marriages are potentially more likely to participate in the labor force than women in non-hypogamous marriages conditional on treatment assignment (i.e., holding treatment status fixed). More formally:

$$P[Y(1) = 1|H = 1] \geq P[Y(1) = 1|H = 0] \tag{7}$$

$$P[Y(0) = 1|H = 1] \geq P[Y(0) = 1|H = 0]. \tag{8}$$

Before going on to discuss the basis of the MTS assumption, a clarificatory comment is in order. At a first glance, the MTS assumption might appear to be somewhat inconsistent with the theory linking hypogamy and women's likelihood of labor force participation proposed in Section 1. However, that is actually not the case: the MTS assumption is absolutely consistent with the theory proposed. The proposed theory suggests that women in hypogamous marriages are likely to have a actual lower labor force participation rate than women in non-hypogamous marriages due to the difference in treatment status (i.e., hypogamous versus non-hypogamous marriage). Thus, it essentially makes a prediction about how women's actual likelihood of labor force participation would change when her *treatment status changes*. The MTS assumption, on the other hand, suggests that, *holding treatment status fixed*, if we consider the potential outcomes of women in the two groups, women in the treatment group are potentially more likely to participate in the labor force than women in the control group.

We believe the MTS assumption is plausible in our setting because women in hypogamous marriages are likely to be advantaged compared with women in non-hypogamous marriages across several observed socioeconomic dimensions (see Table A1 in the Appendix). For example, women in hypogamous marriages have a significantly higher educational attainment and smaller family size than women in non-hypogamous marriages. Further, women in hypogamous marriages are more likely to belong to upper castes and less likely belong to

scheduled castes, scheduled tribes and OBCs. All these could make women in hypogamous marriages, conditional on treatment status, more likely to participate in the labor market than women in non-hypogamous marriages. The positive correlation between education and labor force participation, as well as the negative correlation between family size and labor force participation is widely documented in labor literature, and hence requires no further explanation. Belonging to upper castes as opposed to lower castes could impact the likelihood of rural women's participation in the labor force positively because of caste-based discrimination which continues to be an endemic feature of labor markets in India (Thorat et al., 2021).

In addition to being advantanged in terms of observed socioeconomic characteristics like education, family size and caste, women in hypogamous marriages are also likely to be advantaged compared with women in non-hypogamous marriages across several unobserved attributes as well. It is well-known that unobserved ability and skills as well as unobserved personality traits (the so-called 'non-cognitive skills') like motivation, confidence, self-control, patience, etc. are positively correlated with educational attainment (Hakimi et al., 2011). These unobservables are also likely be positively correlated with labor market outcomes of women (Fletcher, 2013) thereby making women in hypogamous marriages more likely to participate in the labor force than women in non-hypogamous marriages.

Under the MTS assumption, the bounds on ATE as derived in Kreider et al. (2012) are:

$$UB_{ATE} = \frac{P(Y=1,H=1)}{P(H=1)} - \frac{P(Y=1,H=0)}{P(H=0)}. \tag{9}$$

$$LB_{ATE} = P(Y=1,H=1) - P(Y=1,H=0) + P(H=1) \tag{10}$$

Assumption 3. *Monotone Instrumental Variable (MIV)*

To further tighten the bounds of ATE, we make use of new information through the introduction of a MIV. As cautioned by Millimet and Roy (2015), a MIV should *not* be

viewed as a typical instrumental variable. The *only* condition that needs to be satisfied for an MIV to be valid is that potential outcomes must vary monotonically with the variable used as an MIV (Manski and Pepper, 2000). Following Kreider et al. (2012), the MIV assumption imposes

$$P[Y(1) = 1|v = u_1] \geq P[Y(1) = 1|v = u] \geq P[Y(1) = 1|v = u_2] \tag{11}$$

$$P[Y(0) = 1|v = u_1] \geq P[Y(0) = 1|v = u] \geq P[Y(0) = 1|v = u_2] \tag{12}$$

where v is the MIV and $u_1 > u > u_2$. In other words, higher values of v are associated with better potential outcomes.

Here, we use the literacy rate and gross state domestic product (GSDP) per capita (measured at constant prices) of the woman's state of residence as two alternative MIVs (MIV1 and MIV2 respectively). Higher overall levels of literacy and GSDP per capita are likely to reflect higher local economic development which by reducing patriarchal culture should increase the likelihood of women's participation in the labor force irrespective of whether or not they are in hypogamous marriage. We draw these variables from two different sources. The data for state level literacy rate comes from the Indian Census of 2011.[17] The data for GSDP per capita (for 2015-16 measured at 2011-12 constant prices) is compiled from the directorate of economics and statistics of respective state governments.[18]

Following Proposition 1 in Manski and Pepper (2000), the joint MTS-MIV assumption implies

$$\sup_{u_1 \geq u} \text{LB}(u_1) \leq P[Y(t) = 1|v = u] \leq \inf_{u_2 \leq u} \text{UB}(u_2), t = 0, 1. \tag{13}$$

where $UB(u)$ and $LB(u)$ denote the upper and lower bounds of the individual components

[17] The data are available at https://rbidocs.rbi.org.in/rdocs/Publications/PDFs/-6TABLE4134B659E3B243EE9CB292D36ABC281B.PDF

[18] The data are available at http://mospi.nic.in/download-tables-data.

of the ATE obtained under MTS assumption evaluated conditional on $v = u$.

Estimation and Inference Estimation of the bounds on the ATE is straightforward. For the worst case and MTS bounds, we just need to compute the empirical probabilities. For the MTS-MIV bounds, we use suprema and infima over the sample means in the subgroups defined by the MIVs. Since the MIV estimates suffers from a bias in the presence of finite sample analysis (Manski and Pepper, 2000), we use Kreider and Pepper's (2007) nonparametric finite sample bias-corrected MIV estimator. To address the uncertainty arising from sampling variability, along with the bounds, the Imbens and Manski (2004) 95% confidence intervals are reported (see Kreider et al. 2012).[19]

4 Results

4.1 Main Results

The main empirical results are presented in Figures 6 and 7. Figure 6 provides the results obtained using the NFHS sample while Figure 7 provides the results obtained using the IHDS sample. In Figure 6, the results corresponding to the four indicators of employment – currently employed in any work, employed in any work in the last twelve months, currently employed in paid work, employed in paid work in the last twelve months – are presented in four graphs. In Figure 7, the results corresponding to two indicators of employment – currently employed in paid work and employed in any work in the last twelve months – are presented in two different graphs (recall, IHDS provides information on only two employment indicators). In each graph of a given figure, we plot and report sharp bounds on the ATE and Imbens and Manski (2004) 95% confidence intervals under various assumptions regarding the selection process. Specifically, we report the ATE and confidence intervals under no assumption on selection (i.e., the worst case bounds), under the assumption of MTS, under

[19] We implement the bounds approach using codes written by McCarthy et al. (2015).

the assumption of MTS and MIV1 and under the assumption of MTS and MIV2.

Turning to the results, the following findings stand out. First, without imposing any assumptions concerning the selection process, the bounds are of width one and necessarily include zero as discussed in Section 3. Nonetheless, the bounds are useful in excluding possible values of the ATE. For example, Figure 6 shows that the bounds on the ATE of educational hypogamy on women's likelihood of being currently employed in paid work are $[-0.353, 0.647]$. Likewise, Figure 7 shows that the bounds on the ATE of educational hypogamy on women's likelihood of being currently in employed in paid work are $[-0.352, 0.648]$. Thus, a considerable range of values of the ATE, especially in the negative domain, is ruled out.

Second, the MTS assumption, is remarkably powerful in tightening the bounds. In particular, compared to the bounds obtained without any assumption concerning selection process, the bounds under MTS are significantly narrower. For example, Figure 6 reveals that, the imposition of MTS causes the bounds on the ATE of educational hypogamy on women's likelihood of being employed in any work during the last twelve months to shrink from $[-0.458, 0.542]$ to $[-0.458, -0.045]$. Likewise, Figure 7 reveals that, the imposition of MTS causes the bounds on the ATE of educational hypogamy on women's likelihood of being employed in any work during the last twelve months to shrink from $[-0.537, 0.463]$ to $[-0.537, -0.097]$. It is worth noting here that not only does the MTS assumption tighten the bounds, it also allows us to identify the sign of all the ATEs in Figures 6 and 7 as negative. Moreover, all the confidence intervals exclude zero. This indicates that, even without invoking further assumptions, we can claim, educational hypogamy has a significant negative effect on women's labor force participation.

Third, the MIV restrictions when imposed along with the MTS assumption leads to further tightening of the bounds. For example, under MIV1 (MIV2), Figure 6 shows that the bounds on the ATE educational hypogamy on women's likelihood of being currently employed in paid work and of being employed in paid work in the last twelve months

are $[-0.339, -0.045]$ ($[-0.314, -0.058]$) and $[-0.373, -0.063]$ ($[-0.353, -0.071]$) respectively. The corresponding bounds in absence of the MIV assumption are $[-0.353, -0.003]$ ($[-0.339, -0.045]$) and $[-0.392, -0.015]$ ($[-0.373, -0.063]$). Thus, the MIV restriction while is not necessary to sign the bounds, it is definitely useful to improve the inference about the true effect of educational hypogamy on women's labor force participation based on the estimated bounds.

Overall, our results clearly indicate that educational hypogamy leads to significant decline in women's overall employment as well as paid employment. Specifically, based on the NFHS sample, it can be concluded that educational hypogamy decreases the women's likelihood of being currently employed by at least 7-8% (5-6% if we consider only paid employment), and of being employed in the last twelve months by at least 8-10% (6-7% if we consider only paid employment). Based on the IHDS sample, we can conclude, educational hypogamy decreases the women's likelihood being currently employed in paid work by at least 5-6%, and of being employed in any work in the last twelve months by at least 16-31%.[20] These findings are in line with our hypothesis, and is also largely consistent with Pierce et al. (2013), Ratliff and Oishi (2013), Bertrand et al. (2015) and Roychowdhury and Dhamija (2022) which suggest that when women outcompete their husbands in terms of economic status, their lives are adversely affected in several ways.

Robustness Checks To examine the robustness of our results we carry out a battery of robustness checks. First, we use sample weights. Second, we varied the number of MIV cells. Third, we estimate the bounds for different NFHS rounds separately. The results of all the robustness checks can be found in the Appendix (Tables A2–A4). Thankfully, the results of the robustness checks are in line with our main results.

[20] The last figure must be interpreted keeping in mind that IHDS sample shows a much higher overall employment rate of women in the twelve months preceding the survey compared to NFHS sample (52% versus 40%).

4.2 Differential Quality of Marriage

Recall that in Section 1 we had suggested that the marriage quality of women in hypogamous marriages is likely to be worse than marriage quality of women in non-hypogamous marriages. The reason is simple. Hypogamy undermines patriarchal norms and beliefs which is likely to cause the former (but not the latter) to face repercussions in form of marital conflict, stress and tension, and severely undermine the quality of marriage. Such conflicts, stress and tension could increase manifold and ultimately lead to marital destabilization if women work outside the home. Anticipation of this might cause most women in hypogamous women to simply not join the labor force or withdraw from the labor force since divorce is still a taboo in rural India and marital destabilization is extremely costly for rural Indian women. Of course, if marital quality is high (or initial conflicts and stress within marriage is low), married women working outside could still lead to some tensions, but that is unlikely to result in marital destabilization. Non-anticipation of marital destabilization might cause most women in non-hypogamous women to join the labor force or not quit from the labor force. In sum, therefore, as per our proposed explanation, the root cause of the difference in labor force participation of women in hypogamous marriages vis-a-vis women in non-hypogamous marriages is the difference in marital quality between the two groups of women.

Is the marital quality of women in hypogamous marriages really different from that of women in non-hypogamous marriages? More specifically, is the marital quality of women in hypogamous marriages worse than the marital quality of women in non-hypogamous marriages? For our proposed explanation to be true, the answer to this question must unequivocally be an "yes".

We examine this question, firstly, using information provided in the NFHS on women's exposure to factors that could potentially lead to severe emotional stress and tension within marriage. Specifically, NFHS asks married women the following questions: '*(i) Does/did your husband ever say or do something to humiliate you in front of others?, (ii) Does/did your husband ever threaten to hurt or harm you or someone close to you?,* and *(iii) Does/did*

your husband ever insult you or make you feel bad about yourself?'. Women could respond 'yes' or 'no' to each question. Based on the responses to these questions, we create a variable called emotional stress. This variable takes a value 1 if a woman answers 'yes' to at least one of the three questions, and 0 otherwise. Further, we use responses (which again are in form of 'yes'/'no') to the question *'Are you afraid of your husband?'*, and create another variable called anticipation of emotional stress. Both these variables, it can be argued, are likely to be good proxies of marital quality.

Using these variables as outcomes, we estimate bounds on the ATE of hypogamy under various assumptions regarding the selection process as before.[21] Figures 8-9 present the results. We find that the bounds on the ATE under the combined MTS-MIV assumptions are strictly positive and statistically significant for both emotional stress and anticipation of emotional stress. This indicates that women in hypogamous marriages are significantly more likely to be exposed to emotional stress as well as anticipate emotional stress within marriage than women in non-hypogamous marriages suggesting that quality of marriage of the former are likely to be worse than the quality of marriage of the latter. This finding considerably strengthens our proposed explanation of why women in hypogamous marriages are less likely to participate in the labor force than their counterparts.

5 Conclusion

Hypogamy has been increasing in rural India over the last two decades. In this paper, we examine whether this can explain rural India's declining FLFPR. This could be the case since theory suggests, owing to the possibility that marriage quality of women in hypogamous

[21] Note, the MTS assumption here is that women in hypogamous marriages, apriori, are less likely to be susceptible to marital stress than women in non-hypogamous marriages. The assumption can be justified as follows. As discussed previously, women in hypogamous marriages have higher levels of education (and perhaps ability also) than women in non-hypogamous marriages. Since women in hypogamous are relatively more educated, it is likely to cause the them to have higher options outside marriage, and therefore a higher likelihood of leaving an abusive relationship than their counterparts (Erten and Keskin, 2018). Given that women's options outside marriage and likelihood of leaving an abusive relationship is likely be negatively related to the stress that she is likely to experience within marriage, this is likely to make women in hypogamous marriages less exposed to emotional stress than their counterparts.

marriages is worse than women in non-hypogamous marriages, participation in labor force might be more costly for the former than the latter. This might make women in hypogamous marriages less likely to participate in the labor force than their counterparts. If this is indeed the case, a rise in hypogamy could be viewed as a factor driving the decline in FLFPR in rural India.

To examine the causal relationship between hypogamy and female labor force participation, we employ a nonparametric bounds approach. Relying on fairly weak and transparent assumptions, we find clear evidence that, compared to women in non-hypogamous marriage, women in hypogamous marriages are significantly less likely to participate in the labor force. Further we find that the marriage quality of women in hypogamous marriages is worse than women in non-hypogamous marriages which lends support to the underlying theory. Overall, our findings suggest that the rise in educational hypogamy is a possible explanation for the declining FLFPR in rural India.

It is important to reiterate that the marriage quality is likely determined by patriarchal gender norms. A non-hypogamous marriage is likely to be of better quality than a hypogamous marriage since the former occurs in accordance to the prescriptions of patriarchal gender norms while the latter violates the prescriptions of patriarchal gender norms. This suggests that if gender norms could be improved (i.e., if the society starts becoming more acceptive of hypogamous marriages), the difference in marital quality would disappear (or at least reduce). Clearly, this would translate into disappearance of the difference in the cost of participation in the labor market between women in hypogamous marriages and women in non-hypogamous marriages, in turn, implying that the former will no longer have higher disincentive to participate in the labor force than the latter.

In light of this and given that hypogamy is rising in rural India, it seems imperative to improve gender norms if the decline in FLFPR in rural India is to be arrested. While it is widely known that gender norms are sticky, recent research suggests policy measures and nudges that address gender biases at a young age, like engaging adolescent girls and boys

in classroom discussions about gender equality (Dhar et al., 2022) or making them interact with female role models (Kipchumba et al., 2021), could be highly effective. For empowering women and improving their lives, in addition to using conventional policies, it is high time that policymakers in India should start adopting such approaches.

References

Afridi, F., Dinkelman, T. and Mahajan, K. (2018). "Why are Fewer Married Women Joining the Work Force in Rural India? A Decomposition Analysis over Two Decades," *Journal of Population Economics*, 31, 783-818.

Afridi, F., Bishnu, M., & Mahajan, K. (2022). "What Determines Women's Labor Supply? The Role of Home Productivity and Social Norms," *Journal of Demographic Economics* (forthcoming).

Akerlof, G.A. & Kranton, R.E. (2000). "Economics and identity," *Quarterly Journal of Economics*, 115, 715-753.

Alesina, A., Giuliano, P. & Nunn, N. (2013). "On the Origins of Gender Roles: Women and the Plough," *Quarterly Journal of Economics*, 128: 469–530.

Alfano, M. (2017). "Daughters, dowries, deliveries: The effect of marital payments on fertility choices in India," *Journal of Development Economics*, 125, 89–104.

Allendorf, K. (2013). "Going nuclear? Family structure and young women's health in India, 1992–2006," *Demography*, 50, 853–880.

Allendorf, K., & Pandian, R. K. (2016). "The decline of arranged marriage? Marital change and continuity in India," *Population and Development Review*, 42, 435–464.

Anderson, S. and G. Genicot (2015). "Suicide and Property Rights in India," *Journal of Development Economics*, 114, 64-78.

Baland, J.M. & Ziparo, R. (2017). "Intra-household bargaining in poor countries," *Towards Gender Equity in Development*, 69.

Bernard, T., Doss, C., Hidrobo, M., Hoel, J., & Kieran, C. (2020). "Ask me why: Patterns of intrahousehold decision-making," *World Development*, 125, 104671.

Bertrand, M., Kamenica, E., & Pan, J. (2015). "Gender identity and relative income within households," *Quarterly Journal of Economics*, 130, 571-614.

Bertrand, M. (2019). "The Gender Socialization of Children Growing up in Non-Traditional Families," *AEA Papers and Proceedings*, 109, 115-21.

Bhalotra, S., Chakravarty, A. and Gulesci, S. (2020). "The Price of Gold: Dowry and Death in India," *Journal of Development Economics*, 143, 102413.

Bilodeau J., Marchand A., Demers A. (2020). Work, family, work–family conflict and psychological distress: A revisited look at the gendered vulnerability pathways. Stress and Health, 36, 75–87.

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.

Blundell, R., & MaCurdy, T. (1999). "Labor supply: A review of alternative approaches," (eds.) O.C. Ashenfelter & D.Card, *Handbook of Labor Economics*, 3, 1559-1695.

Boserup, E. (1970). *Woman's Role in Economic Development*. London: George Allen and Unwin Ltd.

Chakraborty, T. and Kim, S. (2010). "Kinship Institutions and Sex Ratios in India," *Demography*, 47, 989–1012.

Chakrabarti, S., & Biswas, C. S. (2012). "An exploratory analysis of women's empowerment in India: A structural equation modelling approach," *Journal of Development Studies*, 48, 164–180.

Chatterjee, U., Murgai, R. & Rama, M. (2014). "Job opportunities along the rural-urban gradation and female labor force participation in India," Mimeo, New York

Chowdhury, S. (2011). "Employment in India: What Does the Latest Data Show?" *Economic and Political Weekly*, 34, 32, 23-26.

Cygan-Rehm, K., Kuehnle, D., & Oberfichtner, M. (2017). "Bounding the causal effect of unemployment on mental health: Nonparametric evidence from four countries," *Health Economics*, 26, 1844-1861.

De Haan, M. (2011). "The effect of parents' schooling on child's schooling: a nonparametric bounds analysis," *Journal of Labor Economics*, 29, 859-892.

Deininger, K., Goyal, A. and Nagarajan, H. (2013). "Women's Inheritance Rights and Intergenerational Transmission of Resources in India," *Journal of Human Resources*, 48, 114–141.

Desai S, Vanneman R, and National Council of Applied Economic Research (N.C.A.E.R.). (2015). India Human Development Survey-II (IHDS-II), 2011–12. Inter-university Consortium for Political and Social Research [distributor]. Available at: https://doi.org/10.3886/ICPSR36151.v6

Desai, S. (2017). "Declining female labour force participation in rural India: The demand side," *Ideas for India* (March 7, 2017), Available at https://www.ideasforindia.in/topics/social-identity/declining-female-labour-force-participation-in-rural-india-the-demand-side.html

Dhar, D., Jain, T., & Jayachandran, S. (2022). "Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in India." *American Economic Review*, 112, 899-927.

Dhamija, G., & Roychowdhury, P. (2020). "Age at marriage and women's labour market outcomes in India," *Journal of International Development*, 32(3), 342-374.

Ebenstein, A. and S. Leung (2010): "Son Preference and Access to Social Insurance: Evidence From China's Rural Pension Program," Population and Development Review, 36, 47–70.

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.

Fletcher, J. M. (2013). "The effects of personality traits on adult labor market outcomes: Evidence from siblings," *Journal of Economic Behavior & Organization*, 89, 122-135.

Grzywacz J. G., Bass B. L. (2003). Work, family, and mental health: Testing different models of work-family fit. Journal of Marriage and Family, 65, 248–261.

Gundersen, C. & Kreider, B. (2008). "Food stamps and food insecurity: what can be learned in the presence of nonclassical measurement error?" *Journal of Human Resources*, 43, 352–382.

——— (2009). "Food stamps and food insecurity: what can be learned in the presence of nonclassical measurement error?" *Journal of Health Economics*, 28, 971–983.

Gundersen, C., Kreider, B., & Pepper, J. (2012). "The impact of the national school lunch program on child health: a nonparametric bounds analysis," *Journal of Econometrics*, 166, 79–91

Hakimi, S., Hejazi, E., & Lavasani, M. G. (2011). The relationships between personality traits and students' academic achievement. Procedia-Social and Behavioral Sciences, 29, 836-845.

Himanshu (2011). "Employment trends in India: a re-examination," *Economic & Political Weekly*, 46, 43–59

Imbens, G. W., & Manski, C. F. (2004). "Confidence intervals for partially identified parameters," *Econometrica*, 72, 1845-1857.

Jayachandran, S. (2015). "The roots of gender inequality in developing countries," *Annual Review of Economics*, 7, 63-88.

——— (2017). "Fertility Decline and Missing Women," *American Economic Journal: Applied Economics*, 9, 118-39.

——— (2021). "Social norms as a barrier to women's employment in developing countries," *IMF Economic Review*, 69(3), 576-595.

Kipchumba, E. K., Porter, C., Serra, D., & Sulaiman, M. (2021). "Infuencing youths' aspirations and gender attitudes through role models: Evidence from Somali schools," Working Paper. Available at https://pvsessions.tamu.edu/RePEc/SerraSomaliaRoleModel.pdf

Kishor, S., & Gupta, K. (2004). "Women's empowerment in India and its states: Evidence from the NFHS," *Economic and Political Weekly*, 39, 694–712.

Kreider, B. & Hill, S. (2009), "Partially Identifying Treatment Effects With an Application to Covering the Uninsured," *Journal of Human Resources*, 44, 409–449.

Kreider, B. & Pepper, J. (2007). "Disability and employment: reevaluating the evidence in light of reporting errors," *Journal of American Statistical Association*, 102, 432–441

——— (2008), "Inferring Disability Status From Corrupt Data," *Journal of Applied Econometrics*, 23, 329–349.

Kreider, B., Pepper, J. V., Gundersen, C., & Jolliffe, D. (2012). "Identifying the effects of SNAP (food stamps) on child health outcomes when participation is endogenous and misreported," *Journal of the American Statistical Association*, 107, 958-975.

Layard, R. (2017). "The economics of mental health," *IZA World of Labor*, 321 doi: 10.15185/izawol.321

Lin, Z., Desai, S., & Chen, F. (2020). "The emergence of educational hypogamy in India," *Demography*, 57, 1215-1240.

Macmillan, R. & Gartner, R. (1999). "When she brings home the bacon: Labor-force participation and the risk of spousal violence against women," *Journal of Marriage and the Family*, 947-958.

Manski, C. (1995). *Identification problems in the social sciences*. Harvard University Press, Cambridge

Manski, C. & Pepper, J. (2000). "Monotone instrumental variables: with an application to the returns to schooling," *Econometrica*, 68, 997–1010

Mazumdar, I., and Neetha, N. (2011). "Gender Dimensions: Employment Trends in India, 1993-94 to 2009-10," *Economic and Political Weekly*, 46, 43, 118-126.

McCarthy, I., Millimet, D. L., & Roy, M. (2015). "Bounding treatment effects: A command for the partial identification of the average treatment effect with endogenous and misreported treatment assignment," *Stata Journal*, 15, 411-436.

Mehrotra, S., & Parida, J. K. (2017). "Why is the labour force participation of women declining in India?," *World Development*, 98, 360-380.

Millimet, D.L. & Roy, M. (2015). "Partial identification of the long-run causal effect of food security on child health," *Empirical Economics*, 48, 83–141.

Molinari, F. (2008), "Partial Identification of Probability Distributions With Misclassified Data," *Journal of Econometrics*, 144, 81–117.

——— (2010), "Missing Treatments," *Journal of Business and Economic Statistics*, 28, 82–95.

Neff, D. F., Sen, K., & Kling, V. (2012). "The puzzling decline in rural women's labor force participation in India: a reexamination," GIGA Working Paper No. 196.

Parry, J. P. (1979). *Caste and Kinship in Kangra*. Routledge and Kegan Paul, London.

Pepper, J. (2000), "The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis," *The Review of Economics and Statistics*, 82, 472–488.

Pierce, L., Dahl, M.S., & Nielsen, J. (2013) "In Sickness and in Wealth: Psychological and Sexual Costs of Income Comparison in Marriage," *Personality and Social Psychology Bulletin*. 359–374.

Roychowdhury, P. & Dhamija, G. (2022). "Don't cross the line: Bounding the causal effect of hypergamy violation on domestic violence in India," *Journal of the Royal Statistical Society: Series A* (forthcoming)

Rangarajan, C., Iyer Kaul, P. and Seema (2011). "Where Is the Missing Labour Force?" *Economic and Political Weekly*, 46, 39, 68-72.

Ratliff, K.A. & Oishi, S. (2013). "Gender Differences In Implicit Self-Esteem Following a Romantic Partner's Success or Failure," *Journal of Personality and Social Psychology*. 688–702.

Sekhri, S. and Storeygard, A. (2014). "The Impact of Climate Variability on Crimes Against Women: Dowry Deaths in India," *Journal of Development Economics*, 111, 212–223.

Tamer, E. (2010). "Partial identification in econometrics," *Annual Review of Economics*, 2, 167-195.

Therborn, G. (2004). *Between sex and power: Family in the world 1900-2000*. Routledge.

Thorat, S., Madhewaran, S., and Vani, B.P. (2021). "Employment Discrimination and Its Impact on Poverty," Economic & Political Weekly, 56, 55-61.

Weitzman, A. (2014). "Women's and men's relative status and intimate partner violence in India," *Population and Development Review*, 40, 55–75.

World Bank (2010). *India's Employment Challenge. Creating Jobs, Helping Workers*. New Delhi: The World Bank, Oxford University Press.

Table 1. Summary statistics

	NFHS		IHDS	
	Mean	SD	Mean	SD
Panel A: Main outcomes				
Currently employed in any work	0.320	0.467		
Employed in last one year in any work	0.402	0.490	0.515	0.500
Currently employed in paid work	0.247	0.431	0.273	0.446
Employed in last one year in paid work	0.306	0.461		
Panel B: Covariate of interest/Treatment variable				
Educational hypogamy	0.208	0.406	0.170	0.375
Panel C: Demographic characteristics				
Women's age	32.682	8.210	35.387	9.873
Husband's age	37.352	8.654	40.404	10.825
Women's education	5.084	4.855	4.176	4.491
Husband's education	6.885	4.805	6.241	4.692
Women's height (in cm)	151.921	6.027	151.464	8.532
Religion				
Hindu	0.771	0.420	0.839	0.368
Muslim	0.103	0.304	0.101	0.301
Others	0.126	0.332	0.060	0.238
Don't know/Missing	0.000	0.015	0.000	0.000
Caste				
Schedule Caste (SC)	0.189	0.392	0.228	0.420
Schedule Tribe (ST)	0.214	0.410	0.105	0.306
Other Backward Caste (OBC)	0.372	0.483	0.405	0.491
Others	0.176	0.381	0.261	0.439
Don't know/Missing	0.049	0.216	0.001	0.031
Family size	5.686	2.533	5.628	2.517
Region of residence				
North/West/Central	0.608	0.488	0.655	0.475
South/East	0.392	0.488	0.345	0.475
N	108894		24120	

Notes: See text for definition of the outcome variables and variables of interest. NFHS includes the third, fourth and fifth round of data collected in 2005-06, 2015-16, and 2019-21 respectively. IHDS includes the second round of data collected in 2011-12. The North/West/Central region includes Bihar, Chandigarh, Chhattisgarh, Dadra and Nagar Haveli, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Madhya Pradesh, Maharashtra, Delhi, Punjab, Rajasthan, Uttar Pradesh and Uttarakhand. The South/East includes the Andaman and Nicobar islands, Andhra Pradesh, Arunachal Pradesh, Assam, Goa, Karnataka, Kerala, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Puducherry, Sikkim, Tamil Nadu, Tripura, Telangana and West Bengal. Information of paid work (currently as well as in last year) is missing for 6 observations in NFHS-3. Information of women's height is not available for 1559 (571) observations in the NFHS (IHDS) sample. Information of spousal age is not available for 4 observations respectively in the IHDS sample..

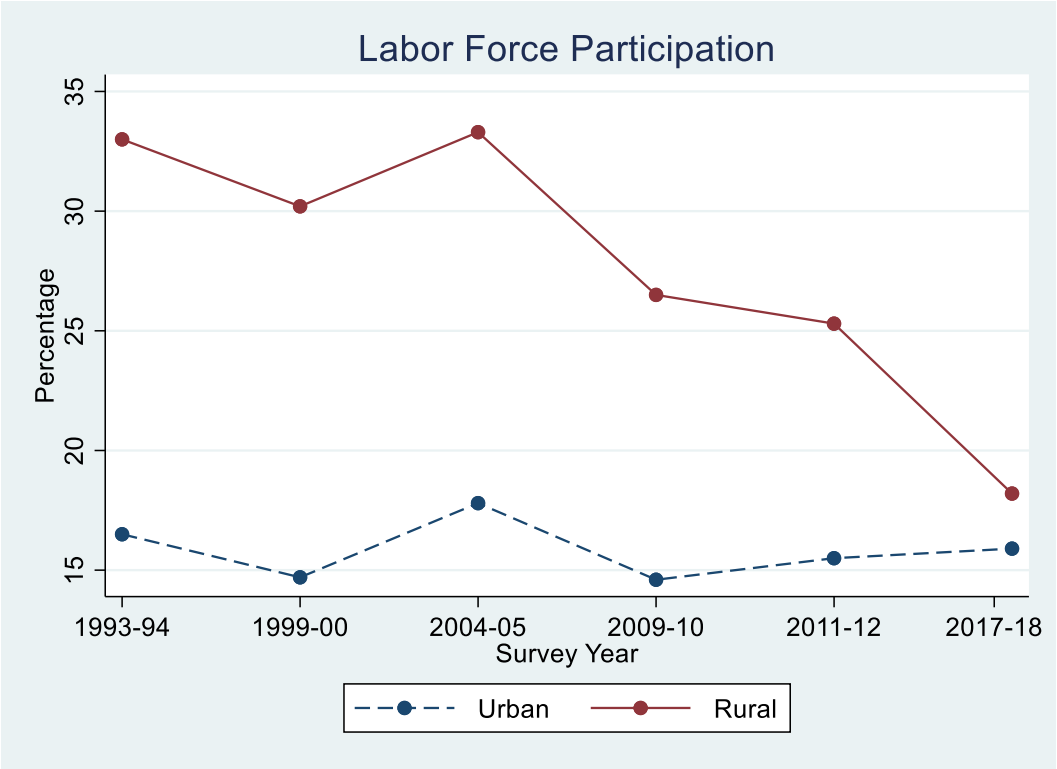

Figure 1. Female Labor Force Participation Rate
Data Source: National Sample Survey and Periodic Labor Force Survey

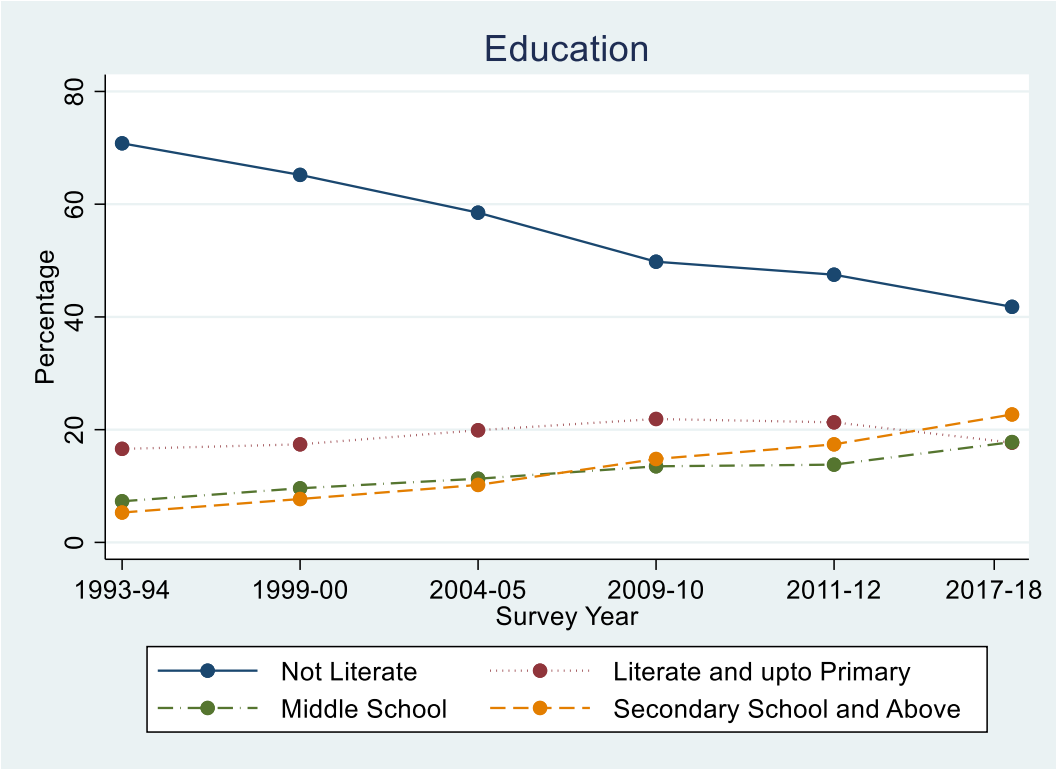

Figure 2. Educational Attainment of Females

Data Source: National Sample Survey and Periodic Labor Force Survey

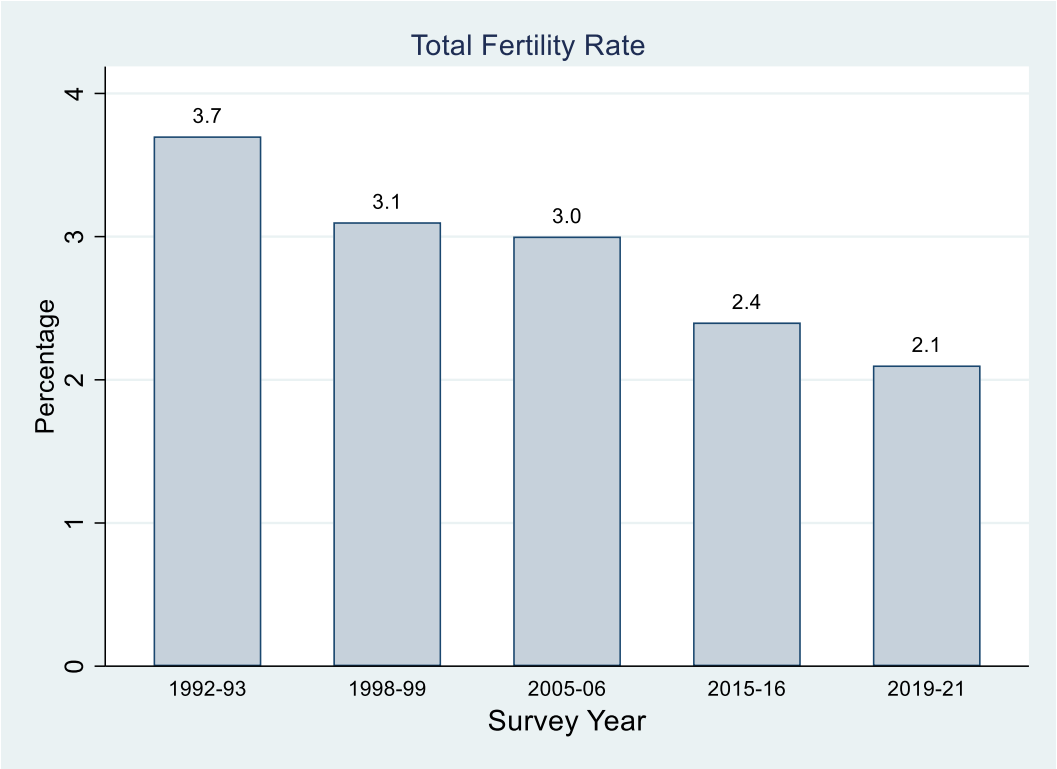

Figure 3. Fertility Rate
Data Source: National Family Health Surveys

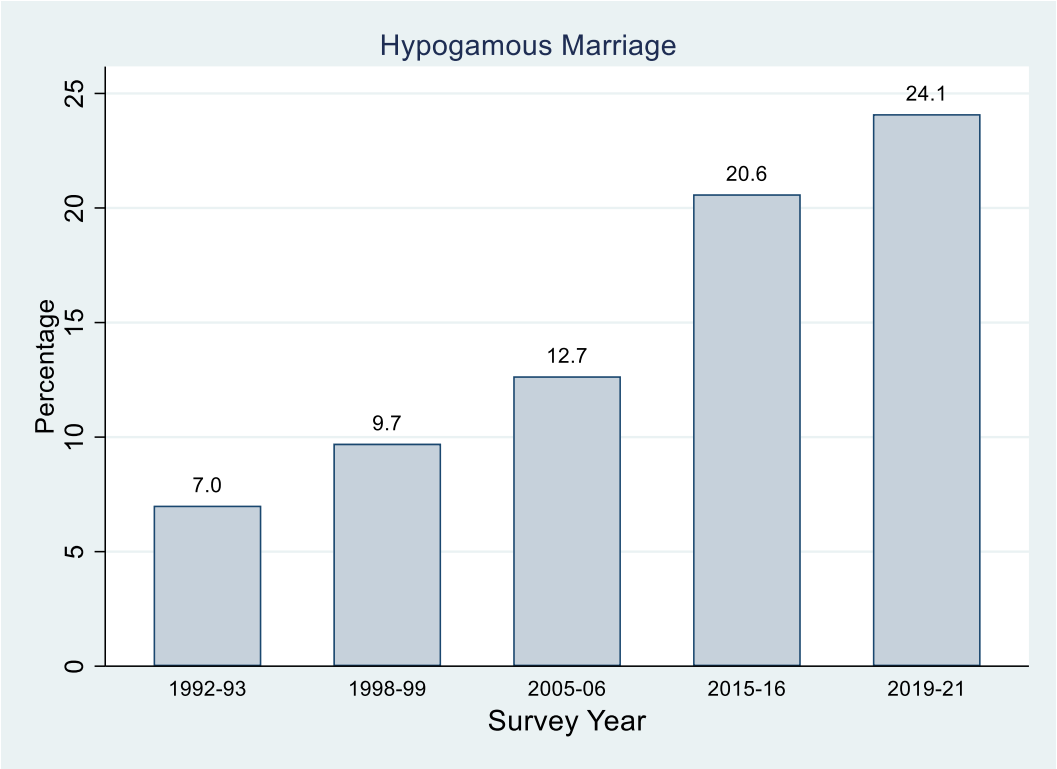

Figure 4. Hypogamous Marriages as a Percentage of all Marriages
Data Source: National Family Health Surveys

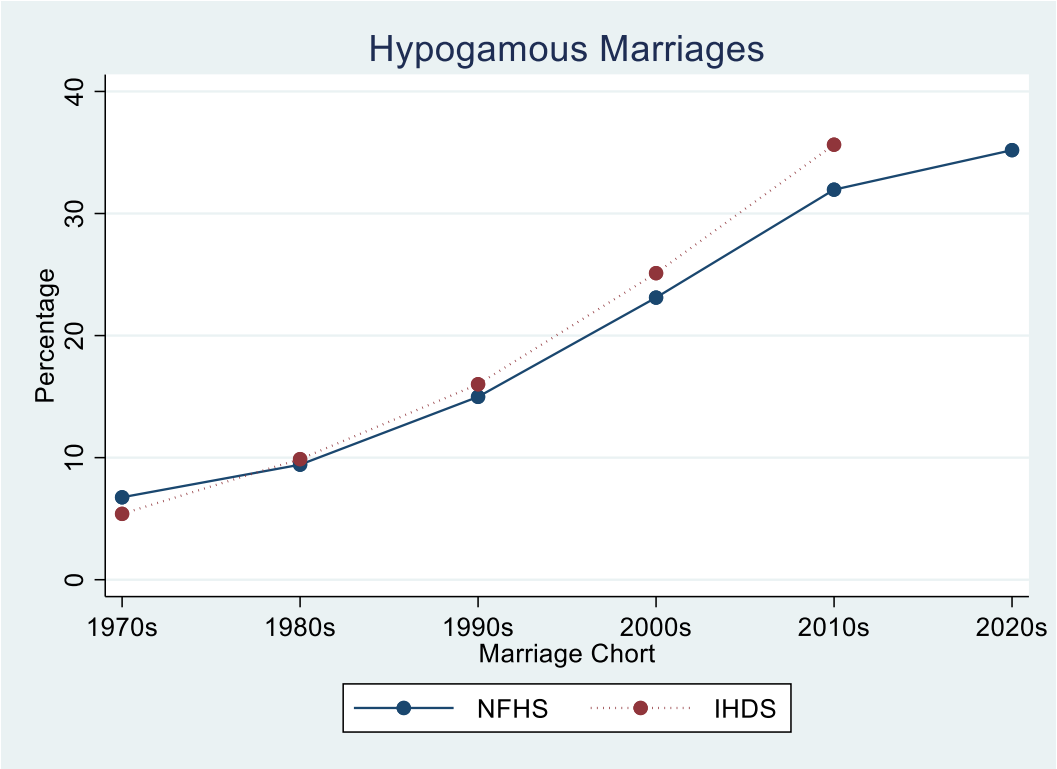

Figure 5. Hypogamous Marriages as a Percentage of all Marriages by Marriage Cohort

Data Source: National Family Health Surveys 2005-06, 2015-16, 2019-21 and Indian Human Development Survey 2011-12

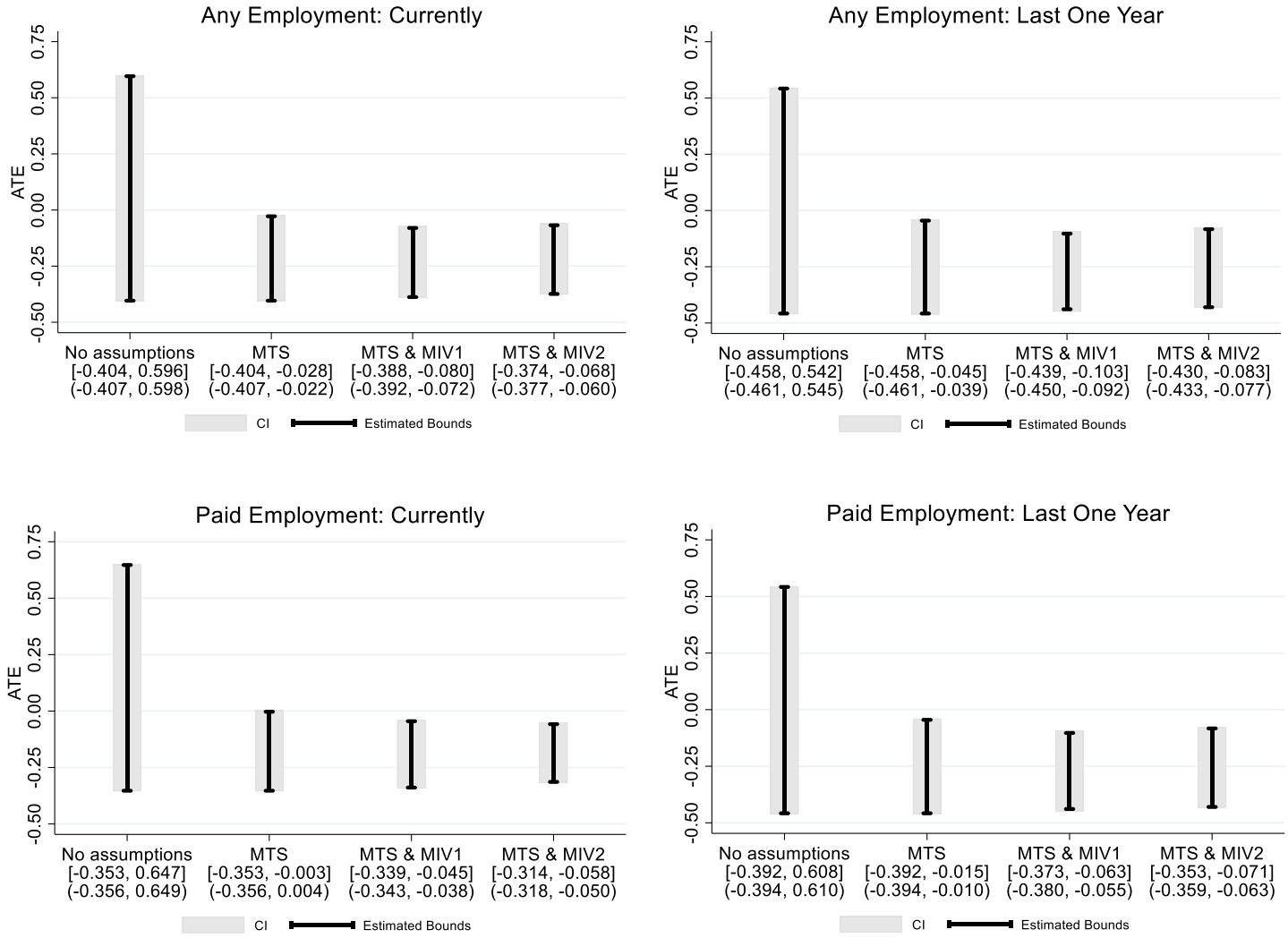

Figure 6. ATE of violation of educational hypogamy on women's labor force participation based on NFHS data.

Notes: Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

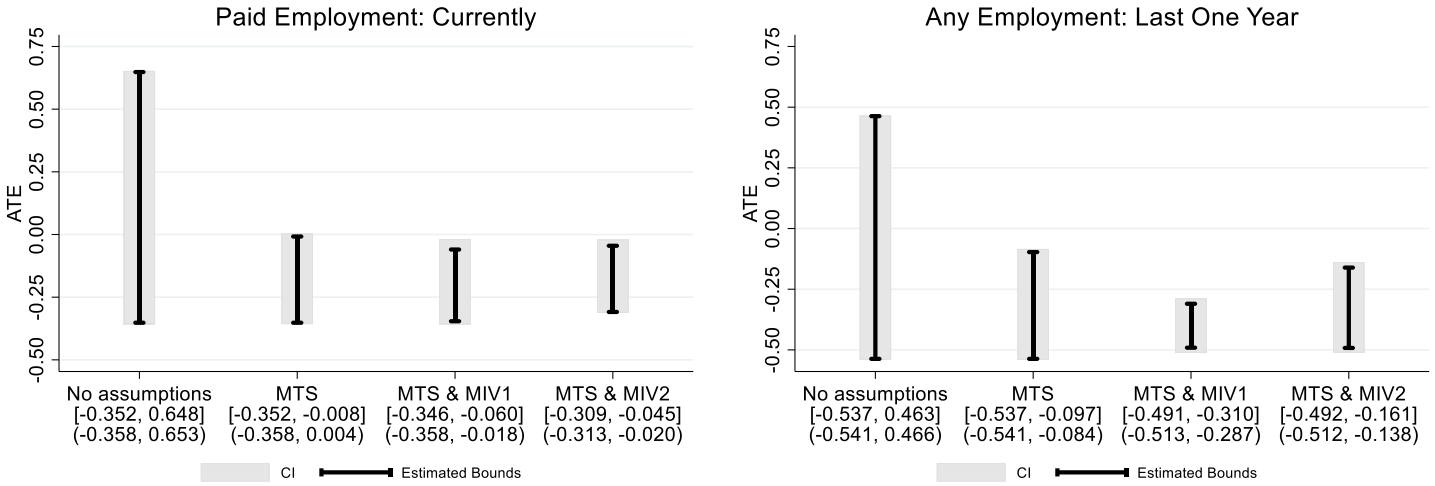

Figure 7. ATE of violation of educational hypogamy on women's labor force participation based on IHDS data.

Notes: Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

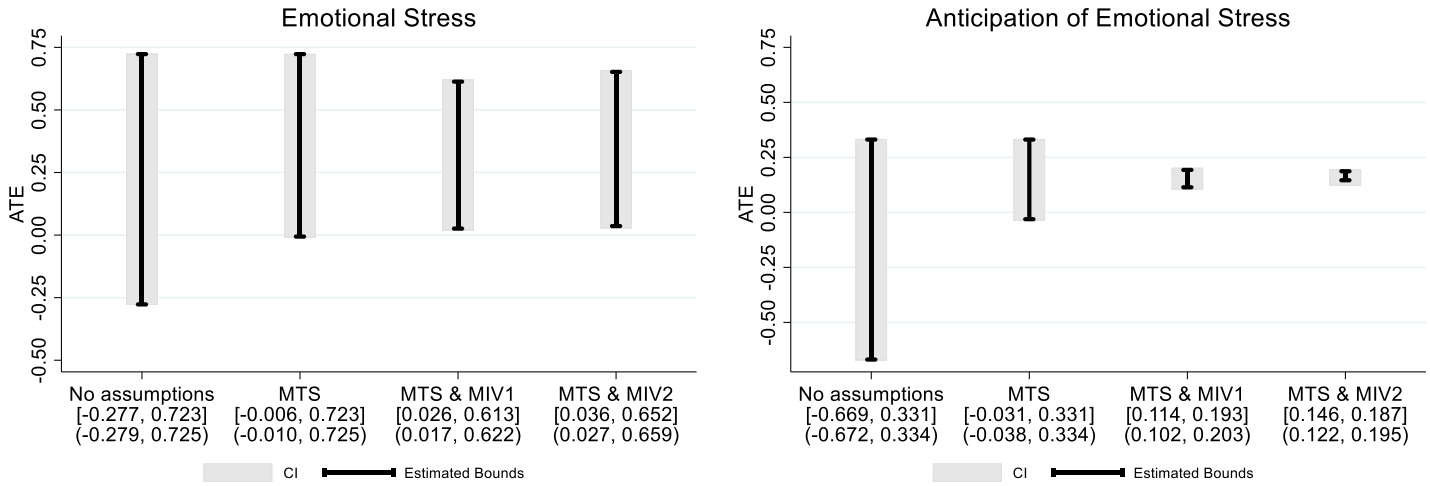

Figure 8. ATE of violation of educational hypogamy on emotional stress based on NFHS data.

Notes: Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

Appendix

Table A1. Background Characteristics, Treatment and Control Group

	NFHS				IHDS			
	Treated	Control	Difference		Treated	Control	Difference	
	Mean	Mean	Mean	SD	Mean	Mean	Mean	SD
Women's age	30.330	33.301	-2.970***	0.058	31.648	36.151	-4.503***	0.157
Husband's age	35.226	37.911	-2.685***	0.062	36.897	41.121	-4.224***	0.175
Women's education	9.226	3.994	5.233***	0.028	8.636	3.264	5.372***	0.059
Husband's education	5.581	7.228	-1.647***	0.032	5.152	6.463	-1.312***	0.070
Women's height (in cm)	151.993	151.903	0.091**	0.045	151.598	151.437	0.161	0.147
Religion								
Hindu	0.745	0.777	-0.032***	0.003	0.806	0.846	-0.040***	0.007
Muslim	0.102	0.103	-0.002	0.002	0.106	0.100	0.006	0.005
Others	0.153	0.119	0.034***	0.003	0.088	0.055	0.033***	0.005
Don't know/Missing	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Caste								
Schedule Caste (SC)	0.178	0.192	-0.014***	0.003	0.202	0.233	-0.031***	0.007
Schedule Tribe (ST)	0.197	0.218	-0.021***	0.003	0.083	0.109	-0.027***	0.005
Other Backward Caste (OBC)	0.366	0.373	-0.008**	0.004	0.410	0.404	0.006	0.008
Others	0.196	0.171	0.025***	0.003	0.304	0.252	0.052***	0.008
Don't know/Missing	0.062	0.046	0.017***	0.002	0.001	0.001	0.000	0.001
Family size	5.486	5.738	-0.252***	0.018	5.439	5.667	-0.227***	0.042
Region of residence								
North/West/Central	0.476	0.643	-0.168***	0.004	0.502	0.686	-0.184***	0.008
South/East	0.524	0.357	0.168***	0.004	0.498	0.314	0.184***	0.008
N	22687	86207			4093	20027		

Notes: See text for definition of the outcome variables and variables of interest. The North/West/Central region includes Bihar, Chandigarh, Chhattisgarh, Dadra and Nagar Haveli, Gujarat, Haryana, Himachal Pradesh, Jammu and Kashmir, Jharkhand, Madhya Pradesh, Maharashtra, Delhi, Punjab, Rajasthan, Uttar Pradesh and Uttarakhand. The South/East includes the Andaman and Nicobar islands, Andhra Pradesh, Arunachal Pradesh, Assam, Goa, Karnataka, Kerala, Manipur, Meghalaya, Mizoram, Nagaland, Odisha, Puducherry, Sikkim, Tamil Nadu, Tripura, Telangana anad West Bengal. In the treatment (control) group, information of women's height is not available for 284 (1275) and 108 (463) observations in the NFHS and IHDS sample respectively. In the treatment (control) group, information of spousal age and is not available for 0 (4) observations respectively in the IHDS sample.

Table A2. Robustness Check, Using Sampling Weights

	NFHS				IHDS	
	Currently employed in any work	Employed in last one year in any work	Currently employed in paid work	Employed in last one year in paid work	Currently employed in paid work	Employed in last one year in any work
	(1)	(2)	(3)	(4)	(5)	(6)
MTS & MIV1	[-0.375, -0.136]	[-0.407, -0.188]	[-0.331, -0.101]	[-0.357, -0.136]	[-0.302, -0.105]	[-0.468, -0.200]
	(-0.383, -0.124)	(-0.416, -0.181)	(-0.337, -0.092)	(-0.361, -0.116)	(-0.318, -0.082)	(-0.506, -0.127)
MTS & MIV2	[-0.412, -0.062]	[-0.460, -0.084]	[-0.353, -0.051]	[-0.388, -0.071]	[-0.311, -0.087]	[-0.503, -0.131]
	(-0.418, -0.053)	(-0.463, -0.072)	(-0.359, -0.041)	(-0.393, -0.066)	(-0.323, -0.075)	(-0.508, -0.107)

Notes: Sampling weights provided in the NFHS used. The treatment group includes women whose educational attainment is at least as high as that of their husbands'. Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

Table A3. Robustness Check, Alternate Number of MIV Cells

	NFHS				IHDS	
	Currently employed in any work	Employed in last one year in any work	Currently employed in paid work	Employed in last one year in paid work	Currently employed in paid work	Employed in last one year in any work
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: MIV Cells = 3						
MTS & MIV1	[-0.404, -0.042]	[-0.458, -0.058]	[-0.353, -0.014]	[-0.392, -0.030]	[-0.352, -0.013]	[-0.537, -0.114]
	(-0.407, -0.038)	(-0.461, -0.055)	(-0.355, -0.009)	(-0.395, -0.023)	(-0.358, -0.002)	(-0.540, -0.103)
MTS & MIV2	[-0.404, -0.041]	[-0.458, -0.057]	[-0.353, -0.014]	[-0.392, -0.026]	[-0.352, -0.019]	[-0.533, -0.130]
	(-0.406, -0.037)	(-0.460, -0.052)	(-0.356, -0.009)	(-0.394, -0.022)	(-0.359, -0.000)	(-0.538, -0.112)
Panel B: MIV Cells = 10						
MTS & MIV1	[-0.350, -0.208]	[-0.389, -0.248]	[-0.311, -0.171]	[-0.341, -0.209]	[-0.318, -0.126]	[-0.430, -0.474]
	(-0.378, -0.182)	(-0.403, -0.231)	(-0.328, -0.152)	(-0.355, -0.186)	(-0.344, -0.072)	(-0.448, -0.438)
MTS & MIV2	[-0.308, -0.174]	[-0.331, -0.234]	[-0.258, -0.133]	[-0.275, -0.180]	[-0.224, -0.180]	[-0.386, -0.330]
	(-0.315, -0.159)	(-0.345, -0.219)	(-0.271, -0.121)	(-0.281, -0.165)	(-0.237, -0.161)	(-0.400, -0.312)

Notes: Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

Table A4. Subsasmple Analysis, ATE of Education Hypogamy on Female Employment in Three Rounds of NFHS

	Currently employed in any work	Employed in last one year in any work	Currently employed in paid work	Employed in last one year in paid work
	(1)	(2)	(3)	(4)
Panel A: NFHS 2005-06				
MTS & MIV1	[-0.466, -0.168]	[-0.511, -0.193]	[-0.372, -0.120]	[-0.406, -0.111]
	(-0.473, -0.148)	(-0.520, -0.156)	(-0.380, -0.100)	(-0.413, -0.079)
MTS & MIV2	[-0.478, -0.155]	[-0.527, -0.168]	[-0.370, -0.101]	[-0.401, -0.114]
	(-0.485, -0.123)	(-0.535, -0.134)	(-0.381, -0.066)	(-0.409, -0.079)
Panel B: NFHS 2015-16				
MTS & MIV1	[-0.363, -0.061]	[-0.421, -0.080]	[-0.321, -0.029]	[-0.363, -0.040]
	(-0.368, -0.052)	(-0.426, -0.064)	(-0.324, -0.021)	(-0.368, -0.031)
MTS & MIV2	[-0.329, -0.060]	[-0.391, -0.079]	[-0.289, -0.043]	[-0.333, -0.057]
	(-0.334, -0.050)	(-0.396, -0.067)	(-0.295, -0.032)	(-0.338, -0.043)
Panel C: NFHS 2019-21				
MTS & MIV1	[-0.386, -0.068]	[-0.427, -0.085]	[-0.352, -0.030]	[-0.383, -0.052]
	(-0.390, -0.060)	(-0.433, -0.079)	(-0.358, -0.018)	(-0.389, -0.038)
MTS & MIV2	[-0.376, -0.034]	[-0.426, -0.059]	[-0.326, -0.038]	[-0.361, -0.050]
	(-0.384, -0.027)	(-0.434, -0.051)	(-0.331, -0.027)	(-0.366, -0.039)

Notes: Point estimates of LB and UB around the unknown parameter Ψ in brackets; 95% Imbens-Manski confidence intervals calculated using bootstrap method in parentheses. See text for further details.

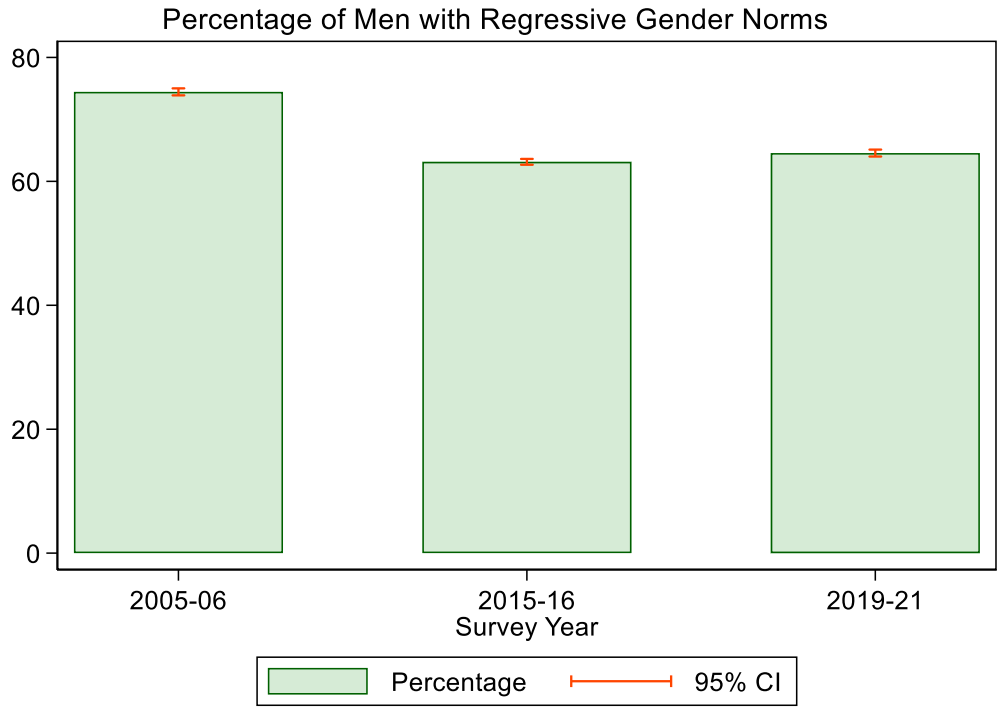

Figure A1. Percentage of Men with Regressive Gender Norms.

Notes: We have classified a man as a ‘man with regressive gender norms’ if he justifies violence against his wife for any of the reasons (such as going out without informing, neglects house or the children, argues with the husband, refuses to have sex with the husband, doesn't cook food properly, being unfaithful, shows disrespect for in-laws) and/or believes husband should be the sole decision maker in any of the household decisions including major household purchases, purchases for daily household needs, visits to the wife's family or relatives, use of money earned by wife, or number of children to have.