Bounding the Causal Effect of Parental Educational Hypogamy on Child Health in India

Shreya Jain*[1], K Narayanan*[2], Tara Shankar Shaw*[3]

Abstract

This study examines the effect of parental educational hypogamy, which refers to a situation where the mother has higher educational attainment than the father, on child health outcomes in India using data from the National Family Health Survey-5 (2019-21). We employed recursive bivariate probit and nonparametric partial identification approach to estimate the relationship between parental educational hypogamy and three measures of child health, namely stunting, wasting, and underweight. We find that hypogamous parents have a lower probability of a child being both stunted and underweight. Our results suggest that parental educational hypogamy plays a protective role in improving early childhood health and nutrition outcomes. The changing patterns of partner selection, particularly in terms of educational hypogamy, may have implications for disparities in child health outcomes.

Keywords: Educational Hypogamy, Partial Identification, Child Health, India

JEL: C14, I15, I21, I25, J12

---

*Address: Department of Economics,
Indian Institute of Technology Bombay, India
Pincode- 400076

[1] Email id: jain.shreya7@iitb.ac.in (Corresponding author); Mobile no.- 8010244056

[2] knn@iitb.ac.in

[3] tara.shankar.shaw@gmail.com

1. Introduction

The structure of the marriage market and educational assortative mating[1] behaviours have undergone substantial changes as a result of the increased education of women in both developed and developing countries (De Hauw et al., 2017; Esteve et al., 2012, 2016; Kc et al., 2010; Lin et al., 2020; Van Bavel, 2012). There has been a decline in educational hypergamy (pertains to the situation where husbands are more educated than their wives)[2] and a rise in educational hypogamy (the pattern in which wives have more education than their husbands) in recent decades. There can be implications of growth of such hypogamic unions on child health. Existing scholarship has, for the most part, focused on the trends of educational homogamy[3] and its implication for various child health outcomes (Abufhele et al., 2022; Ariyo & Jiang, 2021; Pesando, 2022; Rauscher, 2020).

There is little evidence to date on the extent to which educational hypogamy is associated with child health, a question we investigate in the present study. Interest in this relationship stems from the well-established relationship between maternal education and child health (Abuya et al., 2012; Güneş, 2015). Higher education of women is associated with greater autonomy over financial decisions and household expenses pertaining to the health and nutrition of their children in the household (Maitra, 2004; Quisumbing & Maluccio, 2003; Thomas, 1994). Previous research also highlights that mothers with higher education levels have a better understanding of health-related behaviours, health risks and are more likely to seek out preventative care for their children making them better equipped to make informed decisions about their children's health (Mangrio et al., 2011).

According to family systems theory (Bowen et al., 1978; Kerr, 2000) mother's education relative to the father's may also have implications for infant health. Behrman (2020) discusses two possible explanations for the changing returns to a mother's relative educational status. First, there is differential selection into hypergamous and non-hypergamous unions over time, meaning that mothers with higher education levels are increasingly characterized by other characteristics such as older ages, more schooling, and better employment outcomes that

[1] Educational assortative mating is defined as the non-random matching of partners with respect to education.
[2] In India, we see a decline in educational homogamy as well (Lin et al., 2020; Sarkar, 2022). It refers to the situation where husband and wife have similar education attainments.
[3] Education homogamy and Parental education similarity have been used interchangeably.

enhance their intra-household bargaining. Second, non-hypergamous unions are becoming associated with more women's involvement in decision-making compared to hypergamous unions.

The impact of relative maternal education on child health is complex and multifaceted. Furthermore, factors like socioeconomic standing, healthcare availability, and environmental circumstances could wield notable influence. Moreover, the dynamics might differ based on the prevailing societal standards and cultural values within the specific country or area under consideration (Abufhele et al., 2022).

However, the status inconsistencies in education between husbands and wives where a wife has more education than their husband can sometimes lead to stress, tension, and intimate partner violence (IPV) as a form of backlash (Cools & Kotsadam, 2017; Hornung et al., 1981; Roychowdhury & Dhamija, 2022; Weitzman, 2014). Thus, a mother's higher relative educational status may have little or even negative effects on the child's nutrition and well-being. Improvements in mothers' relative educational status may do little to change mothers' bargaining power in the absence of better earnings prospects due to the poor quality of education and the limited returns to education in the labour market (Al-Samarrai & Bennell, 2007). Therefore, the effect of a mother's relative education on a child's well-being remains inconclusive.

Additionally, the persistently low female labour force participation rate in India, despite increased educational attainment and decreased fertility rates, can be ascribed to multiple contributing factors. One such factor is the potential trade-off that women face between investing in their education and participating in the labour force versus prioritizing home production and the marriage market. Research by Afridi et al. (2016), Behrman et al. (1999) and Lam & Duryea (1999) suggest that even if the returns to education in the labour market for women increase, they may not be sufficient to counteract the rise in returns to education in the marriage market and home production. The reasons for this can be multi-fold, with cultural and societal norms playing a significant role in shaping women's choices. The perceived role of women as primary caretakers of the home and family often conflicts with the notion of women participating in the labour force, leading to a higher valuation of education as a means of enhancing their position in the marriage market rather than the job market. Additionally, women who do participate in the labour force have often reported facing discrimination, harassment, and a lack of support in terms of family responsibilities, which further discourages labour force participation (Chatterjee et al., 2018). As a result, some women may prioritize investing in their homes and children's education and health rather than participating in the labour force even though their reservation wage rises.

Lin et al. (2020) demonstrated that women with extensive education often form marriages with men possessing lower levels of education, albeit from more advantaged backgrounds. This, in turn, could lead to a potential positive externality on household well-being, such as better spousal and child health, increased education, and the age of marriage of other unmarried household members. To better understand these complex dynamics and find some evidence in favour of the above arguments, this paper investigates the impact of parental educational hypogamy on child health outcomes in India.

This study goes beyond associational techniques to approach a causal estimate by employing both parametric as well as nonparametric estimation techniques. We are using observational data for the analysis so the non-random selection of women into hypogamous marriages becomes a paramount concern. Women entering hypogamous marriages are likely to differ from women in non-hypogamous marriages in ways that are not observed in the data like the level of patriarchy in their natal and marital home, and risk-prone behaviours. Women self-selecting into hypogamy based on attributes that are unobserved but correlated with the child health outcomes makes the relationship endogenous and estimation of causal effects requires more rigorous identification strategies. The typical strategy is to rely on an instrumental variable (IV). Given the binary nature of our outcome and treatment variable, we used a recursive bivariate probit model. This is a commonly used standard estimation technique in health economics literature to address endogeneity.

However, the validity (exogeneity) of an instrument is often under scrutiny. To overcome this challenge, we also used a nonparametric partial identification approach to address the identification problem (Manski, 1995; Manski & Pepper, 1998). A study by (Tamer, 2010) acknowledges that models that do not identify parameters of interest may still contain important information about these parameters. We get sharp bounds on the average treatment effect (ATE) of hypogamy on child health, even when the marriage type is not random. These bounds require weaker assumptions than traditional instrumental variable (IV)-based methods, but they result in bounds rather than precise estimates. Nevertheless, the bounds reveal what can be learned under different assumptions about the selection process. This approach favours the principle that empirical models with fewer assumptions are more robust and credible for making inferences and conclusions. Stronger assumptions may provide more information about a parameter, but the inferences may be less credible (Manski, 2003).

Using data from NFHS-5 (2019-2021) and employing the above methodologies we find a significant positive effect of parental educational hypogamy on child health outcomes in the Indian context. Children of hypogamous parents have a lower probability of being stunted and underweight. We also provide some suggestive evidence (potential mechanisms) that this result arises because women in hypogamous marriages have higher bargaining power than women in non-hypogamous marriages. They have a better prenatal care context and a higher say in decisions concerning their own and spousal health care in the household.

Geographical Context

India's unique cultural context offers a distinctive setting for investigating the impact of parental educational hypogamy on child health. One significant factor that sets India apart from other countries is the perception of women's education. In India, higher education for women is considered a cultural strategy to ensure proper socialization for children rather than a potential for higher family earnings. Despite the rise in educational attainment, female labour force participation has not increased proportionately. This cultural perspective, combined with the influence of arranged marriages and universal marriage practices, significantly impacts assortative mating in India. These practices may limit women's partner pools, leading to a greater emphasis on ascribed characteristics such as caste and kinship network rather than achieved characteristics like education (Lin et al. 2020).

Esteve et al. (2016) found that hypogamous women are more likely to be breadwinners in Europe. Due to India's strong hierarchical gender regime, the impact of women's education on educational hypogamy may differ from that of other countries. Rising hypogamy in India does not portend a dismantling of gender norms or a positive change and may have less social significance in traditional marriage power dynamics. The opposite, however, may be true (Lin et al. 2020); it may be a sign of other, more pervasive kinds of class and gender stratification in Indian culture. Given this particular Indian setting, one is intrigued by how such a rise in hypogamy may affect family-related dynamics. Greater investment in the welfare of children may be one of the results of more educated women as fewer of them enter the workforce. Therefore, it becomes critical to study the impact of increasing women's education and subsequent parental hypogamy on child health.

The subsequent sections of this paper are organized as follows. Section 2 provides related literature on the topic under investigation. The empirical strategy employed in this study is outlined in Section 3. Section 4 describes the dataset utilized for the analysis. The findings derived from the analysis are presented in Section 5. Finally, the paper concludes with a discussion section.

2. Related Literature

Scholarly research on assortative mating has traditionally concentrated on examining educational sorting tendencies and highlighting the mechanisms behind them. These are two very independent tasks: identifying patterns of spousal choice and analyzing their ramifications. In this regard, attempts to measure the contribution of educational homogamy to economic inequality had been the prime focus (Breen & Salazar, 2011; Eika et al., 2019; Pesando, 2021). Recently many attempts were made to study the relationship between parental educational homogamy and children's outcomes. Couples' educational homogamy is motivated by the potential for it to exacerbate gaps in families' capacity to invest in their offspring's growth, health, and well-being, or the potential to maintain societal inequities for generations (Fernández & Rogerson, 2001). Simply put, a society where high-educated people marry high-educated people and low-educated people marry low-educated people will be more unequal than a society where high-educated people marry low-educated people, assuming that both men and women have access to educational opportunities. This is rather obvious within generations societies become more unequal in terms of income and wealth, but it may also hold across cross-generations, parental resource heterogeneity translates into the heterogeneity of outcomes for children born to various couple types, shaping their later-life outcomes (Bingley et al., 2022; Holmlund, 2022).

Esteve et al. (2016) demonstrated a strong link between the reversal of the gender gap in education and the cessation of hypergamy, as well as the rise in hypogamy. They achieved this insight by analyzing census and survey microdata from 420 samples across 120 countries spanning the years 1960 to 2011. Han (2022) found that the degree of educational hypogamy is associated with the deficit in college-educated men in the marriage market and the economic empowerment of college-educated women. Lin et al. (2020) found that hypogamous marriages continue to increase in India, even after accounting for the rise in education levels. Women with higher levels of education often marry men with lower educational attainment but who come from more privileged backgrounds. Additionally, marriages involving blood relations are more inclined to exhibit hypogamy. The papers suggest that educational hypogamy is a growing trend in many countries. However, none of the papers directly address the effect of parental educational hypogamy on child health outcomes.

Child health outcomes are a critical aspect of public health research, as they reflect an individual's well-being and future prospects. One factor that has gained attention in recent years is parental educational similarity and its association with child health outcomes. This literature review explores several studies that have investigated the relationship between parental educational similarity and child health outcomes, highlighting findings from different countries and considering variations in maternal education, labour force participation, and gender equality.

Rauscher (2020) constitutes a key point of reference for our work. Her study expanded the research on parental educational similarity and infant health by studying birth records from the United States. The study hypothesized that educational similarity affects infant health through its influence on maternal stress and characteristics of the prenatal context. The results supported the homogamy-benefit hypothesis, suggesting that parental educational homogamy is beneficial for infant health. However, the effects varied across birth cohorts and maternal education levels. To our knowledge, it's the only other study to use instrumental variables to address endogeneity and it failed to find significant estimates for educational hypogamy.

Following Rauscher's work, Abufhele et al. (2022) conducted a study in Chile using administrative data from births that occurred between 1990 and 2015. Furthermore, their findings imply that parents' educational similarity is linked to a lower risk of low birth weight and preterm birth, whereas educational hypogamy is harmful, highlighting the gender inequality of educationally hypogamous couples in Chile as well as any stigma that may be associated with them as a result of their nonnormative nature.

Focusing on contexts other than high-income countries to LMICs (Low- and middle-income countries), Pesando (2022) examined the association between parental educational similarity and child health outcomes using longitudinal data from Ethiopia, India, Peru, and Vietnam. The study found evidence in favour of the homogamy-benefit hypothesis but only in the more developed and less gender-unequal contexts, namely Peru and Vietnam. This suggests that the positive association between parental educational similarity and child health outcomes may be influenced by contextual factors. Behrman (2020) focused on a similar research question in Malawi and found results consistent with Pesando's findings in Ethiopia and India. The study suggested that children may fare better in non-homogamous unions, primarily hypergamous ones, indicating a different relationship between parental educational similarity and child health outcomes compared to the other countries studied.

Hahn et al. (2018) emphasized the potential influence of a woman's education on her child's health outcomes. The study acknowledged that highly educated women may be more effective in maintaining their children's health. However, it also noted the challenge of distinguishing between the direct effects of women's education and indirect effects resulting from changes in the husband's characteristics that could impact child health outcomes. The study found that educated women were more likely to marry more-educated men with better job prospects, indicating that education's influence on women's well-being and socioeconomic status was primarily through the marriage market rather than the labour market.

This study makes a significant contribution to the existing literature by addressing several limitations observed in previous studies. To the best of our knowledge, this study is among the first to specifically examine the effect of educational hypogamy on child health outcomes in India. Previous research has predominantly focused on educational homogamy, as highlighted in studies by Batyra et al. (2023) and Pesando (2022). However, our study goes beyond the association-based evidence provided in prior research. We utilize nationally representative data, distinguishing ourselves from Pesando (2022), who relied on the Young Lives (YL) dataset. Moreover, we go beyond associations by investigating the causal effects of educational hypogamy on child health outcomes. This research addresses these gaps in the literature, offering valuable insights into the relationship between educational hypogamy and child health in India. Examining the effect of educational hypogamy on child health in India is an important issue to study due to its potential implications for public health and social equity.

By addressing this specific aspect of parental educational dissimilarity, the study adds to the broader literature on the social determinants of child health and contributes to a deeper comprehensive understanding of the factors influencing child health outcomes.

3. Empirical Framework

3.1 Recursive Bivariate Probit Model

Recursive Bivariate Model

Following Li et al. (2019), the basic specification that we will consider here is a recursive bivariate model characterized by a structural equation determining a binary outcome (***C*** for child health) as a function of a binary treatment variable (***H*** for hypogamy) where the binary treatment, or dummy, variable is in turn governed by a reduced form equation:

$$C^* = X'\beta_C + H\,\alpha + \varepsilon_1, \quad C = \mathrm{I}\ (C^* > 0)$$
$$H^* = X'\beta_H + Z'\,\gamma + \varepsilon_2, \quad H = \mathrm{I}\ (H^* > 0) \tag{1}$$

where I (·) denotes the indicator function. In (1), X contains the common covariates and Z contains the instruments. The underlying continuous latent variables C^* and H^* are mapped into the observed outcome *C* and the observed (potentially endogenous) regressor *H* via threshold crossing conditions, and the joint distribution of *C* and *H* conditional on X and Z , $P(C = c, H = h|\ X = x\ , Z = z)$, which for notational convenience we abbreviate to P^{ch}, therefore has four elements:

$$P^{11} = P(\varepsilon_1 > -x'\beta_C - \alpha, \quad \varepsilon_2 > -x'\beta_H - z'\,\gamma)$$
$$P^{10} = P(\varepsilon_1 > -x'\beta_C, \quad \varepsilon_2 > -x'\beta_H - z'\,\gamma)$$
$$P^{01} = P(\varepsilon_1 < -x'\beta_C - \alpha, \quad \varepsilon_2 > -x'\beta_H - z'\,\gamma) \tag{2}$$
$$P^{00} = P(\varepsilon_1 < -x'\beta_C, \quad \varepsilon_2 < -x'\beta_H - z'\,\gamma)$$

The probabilities in (2) are fully determined once a joint distribution for ε_1 and ε_2 has been specified, and given data consisting of *N* observations (c_i, h_i, x_i', z_i') for *i* = 1, . . . , *N*, the log-likelihood function can then be calculated as

$$L(\theta) = \sum_{i=1}^{N} log\, P^{c_i h_i}\,(\theta) \tag{3}$$

where $P^{c_i h_i}\,(\theta)$ denotes the probabilities in (2) evaluated at the point (c_i, h_i, x_i', z_i') and emphasizes the dependence of the probabilities on the parameter θ, which contains the coefficients β_C, β_H , γ, and α, and other unknown parameters of the joint distribution of ε_1 and ε_2 that need to be estimated from the data.

Bivariate Probit Model

In the bivariate probit model, it is assumed that $(\varepsilon_1, \varepsilon_2)$ is drawn from a standard bivariate normal distribution with zero means, unit variances, and correlation coefficient ρ:

$$(\varepsilon_1, \varepsilon_2) \sim N_2\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}\right) \tag{4}$$

The specification in (1) and (2) together with the assumption in (4) is commonly referred to as the recursive bivariate probit (RBVP) model. The joint distribution function of ε_1 and $\varepsilon_2$in the RBVP model is therefore $\Phi_2(\varepsilon_1, \varepsilon_2;\ \rho)$ where $\Phi_2(\cdot,\cdot\ ;\ \rho)$ denotes the cumulative distribution function of bivariate standard normal distribution with coefficient of correlation ρ. In this case, the joint probability function of *C* and *H* can be expressed compactly as

$$P^{ch}(\theta)\ =\ \Phi_2(t_1, t_2;\ \rho^*) \tag{5}$$

where $t_1\ =\ (2c - 1)(x'\beta_C + h\ \alpha), t_2\ =\ (2h - 1)(x'\beta_H + z'\ \gamma\ ) and$

$$\rho^*\ =\ (2c\ -\ 1)(2h\ -\ 1)\rho\ ,$$

and the log-likelihood function for the RBVP model can be written as

$$log\ L(\theta) = \textstyle\sum_{i=1}^{N} \Phi_2(t_{1i}, t_{2i};\ \rho_i^*) \tag{6}$$

where for *i* = 1, . . . , *N*,

$$t_{1i}\ =\ (2c_i - 1)(x_i'\beta_C + h_i\ \alpha),$$
$$t_{2i}\ \ =\ (2h_i - 1)(x_i'\beta_H + z_i'\gamma\ )$$
$$\rho_i^*\ =\ (2c_i\ -\ 1)(2h_i\ -\ 1)\rho\ ,$$

with the subscript *i* indicating the *i*th unit observed in the sample.

Consistency and efficiency of the MLE is contingent on the presumption that the model fitted to the data coincides with the true data generating process (DGP). If the model is miss-specified optimality properties of the MLE can no longer be guaranteed. Nevertheless, the likelihood function of a model can still be used to construct an estimator, and the resulting quasi-maximum likelihood estimator (QMLE) will be consistent for the pseudo-true parameter and asymptotically normal, and certain optimality features can also be ascertained given suitable regularity.

Although the QMLE produces asymptotically biased estimates of the parameters of the true DGP (α, ρ), the QMLE estimates of parametric functions such as the predicted probabilities and ATE perform surprisingly well. That the predicted probabilities and ATE estimates are not sensitive to the miss-specification of the error distribution, and that the RBVP QMLE is able to reproduce the true probabilities and ATE with reasonable accuracy despite the model being misspecified[4].

Instruments: We use two instruments for this analysis namely level of educational hypogamy in women's district of residence and women's exposure to mass media.

[4] See Li et al. (2019) for details. We have derived our RBVP model empirical framework from their study.

District level hypogamy: The rationale behind the first IV is that if a significant proportion of women in a woman's district, are married to men with higher educational levels, it may influence her own choice of partner[5]. This influence can be driven by social norms prevailing in the district meaning that there is a generally accepted belief or expectation that women should marry men with higher educational levels or by the peer effect. The observation of others in similar circumstances opting for partners with higher education levels may shape her perception of what is desirable or acceptable in a partner. Consequently, these social dynamics and the influence of social networks (peer group) can play a role in shaping an individual's preferences and decisions regarding their choice of partner. Although a district level instrument used in IV analysis may satisfy the relevance criterion, they may violate the exclusion criterion[6]. We assume that the exogenous part of the original variable varies across district, whereas the endogenous part varies only within districts. If this is not the case then our IV fails to satisfy the exclusion restriction.

Exposure to mass media: Mass media primarily influences hypogamy rather than child health directly. Exposure to media can shape individuals' perceptions and preferences, affecting partner selection and the likelihood of hypogamy (Prakash & Singh, 2014). However, the impact of media on child health outcomes is limited as parents rely more on trusted sources like family doctors for health information or elders. While media can provide health knowledge, it requires behavioural changes to have a direct impact on child health. Therefore, the influence of mass media on child health is predominantly mediated through its effect on hypogamy and maternal health knowledge. Mass media has also been used as an IV in a study by Glewwe (1999).

3.2. Nonparametric Partial Identification Approach

The interest of this paper is to estimate the average causal effect of parental education hypogamy on child health. We use a nonparametric partial identification approach to identify average treatment effect (ATE). We have binary outcomes, so conditional ATE is given by

---

[5] Rauscher (2020) used the mean spousal similarity of education among married household heads in the mother's state and cohort as one of the IVs. Shajan & Sumalatha (2022) used the proportion of women who work in a district as an IV for maternal employment. District-level variables are commonly used an IV.

[6] Larcker and Rusticus (2016) demonstrated in financial literature that utilizing industry aggregates as instrumental variables does not typically resolve the issue of determining causality or account for omitted variables. This is because industry aggregates encompass both the exogenous and endogenous components of the original variables, as highlighted by Gormley and Matsa (2014).

$$ATE\ (1\ ,0|\ \mathrm{X}\in\Omega\ )\ =\ E[Y(1)=1|\mathrm{X}\in\Omega\ \ ]-\ E[\ Y(0)=1|\mathrm{X}\in\Omega\ \ ]$$

$$=\ P[Y(1)=1|\mathrm{X}\in\Omega\ \ ]-P\ [\ Y(0)=1|\mathrm{X}\in\Omega\ \ ] \qquad (7)$$

where Y is the realized health outcome of a child (which is binary), Y(1) denotes the potential health outcome of the child who has hypogamous mothers, Y(0) denotes the analogous outcome when a child has nonhypogamous mothers, and $\mathrm{X}\in\Omega$ denotes conditioning on observed covariates whose values lie in the set Ω. Thus, the average treatment effect reveals how the mean outcome would differ if all children had a hypogamous mothers versus the mean outcome if all children had a nonhypogamous mothers. In this analysis, Y=1 denotes good health outcome (i.e., child is not stunted, not wasted, not underweight) and Y=0 otherwise.

To simplify our analysis, we do not consider the conditioning on subpopulations denoted by X. In the typical regression model, researchers strive to select control variables that satisfy the exogenous selection assumption. However, there is often controversy over whether important explanatory variables were overlooked. In contrast, our approach only uses conditioning on covariates to define subpopulations of interest and does not require satisfying regression orthogonality conditions since we are not estimating a regression model.

The main identification problem arises when assessing the effect of the parental educational hypogamy on children's health outcomes, we find that the outcome Y(1) is (unobserved) counterfactual for all children who had nonhypogamous mothers, and Y(0) is (unobserved) counterfactual for all children who had hypogamous mothers. In other words, for any given child, only one of two potential outcomes is observed. This is referred to as the selection problem. By Law of Total Probability, this identification problem can be written as follows:

$$P[Y(1)=1]=\ \ P[Y(1)=1|T=1]P[T=1]+\ \ P[Y(1)=1|T=0]P[T=0]$$

$$P[Y(0)=1]=\ \ P[Y(0)=1|T=1]P[T=1]+\ \ P[Y(0)=1|T=0]P[T=0] \qquad (8)$$

where, T=1 denotes that mother is in a hypogamous marriage, and T=0 otherwise. If we observe the actual marriage type, the sampling process identifies P[T=1] and P[T=0] and the expected potential outcome conditional on the outcome being observed, P[Y(1) =1|T=1] & P[Y(0)=1|T=0]. However, the sampling process cannot reveal the mean outcome for counterfactuals, P[Y(1)=1|T=0] & P[Y(0)=1|T=1] .Thus, P[Y(1)=1] & P[Y(0)=1] are not point-identified by the sampling process alone. This value could lie anywhere between 0 and 1in the absence of other information. Given the identification issue, under minimum and transparent assumptions, we derive bounds on the ATE. We use various assumptions concerning the nature of selection process to tighten the bounds which are discussed in detail below.

Worst-case bounds

When no assumptions are made to address the selection problem (Manski, 1995; Pepper, 2000) we end up with the 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|T=0] & P[Y(0)=1|T=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

$$P[Y=1|T=1]P[T=1] \leq P[Y(1)=1] \leq P[Y=1|T=1]P[T=1]+P[T=0]$$

$$P[Y=1|T=0]P[T=0] \leq P[Y(0)=1] \leq P[T=1]+P[Y=1|T=0]P[T=0] \quad (9)$$

The observed data identifies each term in these bounds. By subtracting the lower bound of P[Y(0)=1] from the upper bound of P[Y(1)=1], a precise upper limit on ATE can be obtained, and similarly, a precise lower limit can be obtained by subtracting the upper bound of P[Y(0)=1] from the lower bound of P[Y(1)=1].

$$UB_{ATE} = P[Y=1,T=1]-P[Y=1,T=0]+ P[T=0]$$

$$LB_{ATE} = P[Y=1,T=1]-P[Y=1,T=0]- P[T=1] \quad (10)$$

Nevertheless, these boundaries are clearly defined by a width of one, and they encompass zero. Consequently, determining the sign of ATE is not possible in this situation. In order to draw any meaningful inference about the ATE, the bounds must be narrowed by making certain assumptions about the relationship between child health and marriage type. To achieve this goal, we examine the identifying power of two monotonicity assumptions: monotone treatment selection (MTS) and a monotone instrumental variable (MIV) restriction[7].

[7] Generally, papers employ third assumption Montone Treatment Response (MTR) assumption to tighten the bounds. MTR (positive) states, ceteris paribus, that the outcome is a weakly increasing function of the treatment, such that $ATE_i \geq 0$ for every child. It assumes that individuals do not select into a treatment that would make them worse off in expectation.The assumption implies that there exist no negative

Monotone Treatment Selection

The first assumption introduced to tighten the bound is MTS assumption (Manski & Pepper, 2000) which supposes that sorting into treatments is not exogenous but expected potential outcomes move in a particular direction when individuals are compared across the treatment as well as the control group. We assume that children who had hypogamous mothers have a lower probability of good health than those children who had nonhypogamous mothers if they would have had a hypogamous mother. In other words, children of hypogamous mothers are more likely to have bad health outcome than children of nonhypogamous mothers conditional on treatment assignment[8]. Therefore, we assume children of hypogamous mothers to be negatively selected.

$$P[Y(1) = 1|T = 1] \leq P\,[\,Y(1) = 1|T = 0]$$

$$P[Y(0) = 1|T = 1] \leq P\,[\,Y(0) = 1|T = 0] \tag{11}$$

In our context, we argue that the MTSn[9] assumption is more plausible as negative selection bias in children of hypogamous mothers can occur due to various reasons, including the risk-prone nature of hypogamous mothers and the level of patriarchy in their natal and marital homes. Hypogamous mothers who self-select into these marriages may be more unconventional and risk-prone due to factors such as higher abilities reflected in their higher education and career aspirations, which may lead them to challenge traditional gender roles and norms. This can lead them to pursue non-traditional jobs that require longer working hours and are further away from home, potentially affecting their children's well-being as they may struggle to balance work and family responsibilities.

Moreover, the level of patriarchy in the natal and marital homes of hypogamous mothers can contribute to negative selection bias in various ways. Stigma and discrimination against hypogamous marriages in certain communities can limit resources available to women, while patriarchal and orthodox families may actively seek out hypogamous marriages as a point of pride, but also expecting women to suppress themselves to fit into specific gender roles.

impacts of hypogamy on child health. Arguably, the MTR (positive) assumption may be difficult to justify in the current application. So, we do not use it in our analysis.

[8] In our context, we assume that children of hypogamous mothers are potentially (or to start with) more likely to have worse health outcomes than children of nonhypogamous mothers.

[9] Negative Monotone Treatment Selection (MTSn): MTSn refers to the case of negative selection. In this case, individuals in the treatment group are more likely to experience a bad outcome conditional on treatment assignment.

These women may face increasing access to labour-market opportunities, but without more equitable opportunities within the household. This can lead to what is known as the "double burden," where women have to balance their work and family responsibilities (Abufhele et al., 2022). Such a burden can cause significant stress, intra-household conflict, relationship instability, and even severe domestic violence as a form of male backlash, affecting the health of the child (Roychowdhury & Dhamija, 2022). The probability of these situations faced by women in hypogamous marriages which affect child health negatively depends upon a class of socioeconomic variables such as religion, caste, wealth among others.

Further, the MTS assumption implies that children observed in hypogamy would not have better health than the children observed in nonhypogamy if both were in nonhypogamy.

To obtain the worst case and MTS bounds on ATE, it is sufficient to compute the empirical probabilities. The bounds on ATE under MTS assumption can be derived as derived in Manski and Pepper (2000) and Kreider et al. (2012). The bounds under MTS for Y(1) & Y(0) are as follows:

$$P[Y = 1|T = 1] \leq P[Y(1) = 1] \leq P[Y = 1|T = 1]P[T = 1] + P[T = 0]$$

$$P[Y(0) = 1|T = 0]P[T = 0] \leq P[Y(0) = 1] \leq P[Y = 1|T = 0] \tag{12}$$

Bounds on ATE under MTS is:

$$LB_{ATE} = P[Y = 1|T = 1] - P[Y = 1|T = 0]$$

$$\leq ATE \leq$$

$$P[Y = 1, T = 1] - P[Y = 1, T = 0] + P[T = 0] = UB_{ATE} \tag{13}$$

The upper bound under MTS is same as the worst case upper bound and the lower bound on ATE is associational difference.

Monotone Instrumental Variable

A second assumption to tighten the bounds on ATE is MIV. We make use of new information through the introduction of a MIV (*v*) (Manski and Pepper, 2000). With this additional variable *v*, it is possible to create subsamples for each value of *v* and then to obtain bounds on the mean potential outcomes within each of these subsamples (Manski & Pepper, 1998). A MIV should not be viewed as a typical instrumental variable as cautioned by Millimet & Roy (2015).

With a valid IV, we obtain a local average treatment effect (LATE) for a subpopulation of "compliers" whose treatment statuses are affected by the instrument (Imbens & Angrist, 1994). With a valid MIV, on the other hand, (partial) identification of the ATE for the full population of interest is obtained. For this reason, an MIV may allow conclusions of wider policy relevance in many cases. Another important distinction lies in the content of the assumptions themselves. In the LATE framework of Imbens and Angrist (1994), a valid instrument relies heavily on an exclusion restriction, which would replace the weak inequality in the MIV assumption with a strict equality. This conditional mean independence would yield the intuition that an IV affects the outcome only through its effect on the treatment. A valid MIV, on the other hand, needs to satisfy only one condition that potential outcomes must vary monotonically with the variable used as an MIV. Note that no restrictions are placed on the relationship between the MIV and the treatment variable. This means that MIVs are allowed to be endogenous with respect to treatment, or to have a non-monotonic impact on treatment. MIV weaker than IV assumptions but still a difficult statement to be tested. Following Kreider et al. (2012), the MIV assumption imposes:

$$P[\,Y(1) = 1|v = u_1] \leq \; P[\,Y(1) = 1|v = u] \leq \; P[\,Y(1) = 1|v = u_2]$$

$$P[\,Y(0) = 1|v = u_1] \leq \; P[\,Y(0) = 1|v = u] \leq \; P[\,Y(0) = 1|v = u_2] \qquad (14)$$

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

MIVs: A higher State Net Domestic Product (SNDP) per capita can lead to increased availability of resources for healthcare infrastructure, improved access to healthcare services, and better socioeconomic conditions, which can positively impact child health. It may result in better nutrition, higher immunization rates, improved sanitation facilities, and access to quality healthcare, all of which can contribute to healthier child development. Thus, our MIVs are captured by SNDP per capita, Health index by NITI Aayog, average parental education, more antenatal visits, district level female literacy rates, sanitation and exposure to media which are associated with weakly better potential child health outcomes, regardless of marriage type.

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

$$\sup_{u_1 \geq u} LB(u_1) \leq \mathrm{P}[\,\mathrm{Y(t)} = 1|\mathrm{v} = \mathrm{u}] \leq \inf_{u_2 \leq u} UB(u_2)\,, t = 0{,}1 \tag{15}$$

where, *UB(u)*and *LB(u)* denote the upper and lower bounds of the individual components of the ATE obtained under MTS assumption evaluated conditional on v=u.

To calculate these bounds in practice, the sample is divided into five roughly equally sized five cells (default ncells) based on the MIV values. We then calculate the MTS bounds for P[Y (1) = 1] and P[Y (0) = 1] for each cell. Weighted averages of the estimates of the UB and LB across the five cells yield joint MTS-MIV bounds on the individual components of the ATE (Corollary 1 of Proposition 1 in Manski and Pepper (2000)). Having obtained bounds for the individual components of ATE in this way, we can bound the ATE[10].

Estimating and making inferences on the bounds of the average treatment effect (ATE) is a straightforward process. To compute the worst case and MTS bounds, we simply calculate the empirical probabilities. For the MTS-MIV bounds, we use suprema and infima based on sample means within subgroups defined by the MIVs. To account for bias in the MIV estimates due to finite sample analysis, we employ the nonparametric finite sample bias-corrected MIV estimator proposed by Kreider and Pepper (2007). In addition to the bounds, we report Imbens and Manski's (2004) 95% confidence intervals to address the uncertainty resulting from sampling variability, following the approach outlined in Kreider et al. (2012).

4. Data and Measures

Data used in the study was taken from National Family Health Survey 2019-21 (NFHS-5), a comprehensive source offering insights into India's population, health, and nutrition at both national and regional levels, encompassing states, union territories, and 707 districts. The Ministry of Health and Family Welfare (MoHFW), Government of India, has overseen all five NFHS surveys. The International Institute for Population Sciences (IIPS), located in Mumbai, was designated by MoHFW as the principal entity responsible for coordinating all rounds of NFHS. Funding for NFHS-5 was provided by the MoHFW, Government of India. ICF, USA

[10] We implement the user-written *tebounds* command by McCarthy et al. (2015).

provided technical assistance through the Demographic and Health Surveys (DHS) Program, which is funded by USAID. Assistance for the Dried Blood Sample (DBS) component of the survey was provided by the Indian Council of Medical Research (ICMR) and the National AIDS Research Institute (NARI), Pune. NFHS-5 fieldwork for India was conducted in two phases— Phase-I from 17 June 2019 to 30 January 2020 covering 17 states and 5 UTs and Phase-II from 2 January 2020 to 30 April 2021 covering 11 states and 3 UTs — by 17 Field Agencies and gathered information from 636,699 (98%)[11] household, 724,115(97%) women aged 15-49 years, and 101,839 (92%) men aged 15-54 years.

Analytical sample

The final sample used in our analysis consists of 145,422 children aged between 0-59 months who reside with both of their parents. We specifically focused on children born to mothers who are currently married and have been married only once. Pregnant women and children from multiple births were excluded from the sample to ensure the accuracy and consistency of the data.

Outcome variables

1. Stunting: low height-for-age measurement, serves as an indicator of chronic undernutrition, indicating a prolonged lack of sufficient nutrition. Stunting can also be affected by recurrent and chronic illness.

2. Wasting: low weight-for-height, is a gauge of acute undernutrition, symbolizing insufficient nutrition intake shortly preceding the survey. Wasting can stem from inadequate food consumption or from a recent bout of illness that leads to weight loss.

3. Underweight: Weight-for-age is a composite index that takes into account both acute and chronic undernutrition.

We have defined our outcome variables as binary indicators. The variable "not stunted" takes a value of 1 if the child is not stunted, indicating normal growth. Similarly, for the variables "not wasted" and "not underweight," they take a value of 1 if the child is not wasted or underweight, indicating healthy nutritional status.

[11] Response rates in parentheses

Treatment/ Variable of interest

We measure parents' educational attainment using a seven-category variable based on parents' completed years of education: No education, below primary (1–4 years), primary (5–7 years), upper primary (8-9 years), secondary (10–11 years), senior secondary (12-14 years) and college graduate & above (15 years or more). Based on the education of both parents, we create indicator for whether the education level of child's mother is higher than father. Paternal educational hypogamy takes value 1 if the mother has higher education level than father and 0 if father has higher or equal educational level as the mother.

Controls

We included several variables as controls to account for factors that may be related to both parental educational hypogamy and child health outcomes. These control variables include the child's gender (with a value of 1 indicating male), the child's age in months, birth order (categorized), the average education level of both parents in years, the age difference between parents, the maternal age, the mother's height (with a value of 1 indicating a normal height of ≥ 145 cm), and the ratio of living children to total ever-born children. Additionally, we considered household characteristics such as religion, caste of the household head, wealth index, and residence (with a value of 1 indicating rural residence). We also accounted for the state of maternal residence to address any potential differences in child health outcomes that may be attributed to variations across states, such as differences in access to healthcare and prenatal care. Controlling for the state of residence allows us to capture any constant differences in child health outcomes between states.

Instrumental variables

We used two IVs namely, district level hypogamy and women's exposure to mass media.

Monotone Instrumental Variables

To tighten the bounds, we employed various alternative MIVs in our analysis. State Net Domestic Product per capita for the year 2019-20 was obtained from the Reserve Bank of India (RBI). We also utilized the Health Index provided by NITI Aayog, which incorporates indicators from three domains: health outcomes, governance and information, and key processes. These indicators were standardized on a scale of 0 to 100 and used to calculate composite index scores and rankings for the base year (2018-19) and the reference year (2019-20). Higher scores indicate better performance across these dimensions, reflecting improved healthcare services, accessibility, quality, and overall health system performance in a state or Union territory.

In addition, we considered average parental education and several district-level variables that represent favourable socioeconomic conditions. These variables include district-level sanitation, female literacy, and exposure to media. These factors have the potential to positively influence child health outcomes. Moreover, we incorporated the number of Antenatal Care (ANC) visits as a proxy for prenatal care, as it indicates better care during pregnancy, which can lead to improved child nutrition and overall health.

5. Results

5.1. Descriptive statistics

Table 1 provides a comprehensive overview of the key variables included in the regression models, disaggregated by mothers who are in a hypogamous and non-hypogamous union. The analysis reveals significant differences between these two groups across most variables, with the exception of not wasted and child sex.

Children born to hypogamous mothers exhibit favourable health outcomes compared to those with non-hypogamous mothers. Specifically, they have lower rates of stunting and underweight, indicating better nutritional status. Additionally, children from hypogamous mothers tend to have lower birth order, suggesting that they are more likely to be first or second-born rather than from higher-order births. Furthermore, these children have younger and more educated mothers, while their fathers tend to have lower levels of education. It is worth noting that the child survival ratio is higher among hypogamous mothers.

The prevalence of hypogamy is found to be lower in rural areas, among Hindu households, among scheduled tribes, and at both ends of the wealth spectrum. This implies that hypogamy is more prevalent in urban areas, among non-Hindu households, and among communities with higher socio-economic status.

These findings indicate that hypogamy is associated with certain disparities in child health outcomes and demographic characteristics. The results highlight the importance of considering the educational match between parents and its potential implications for child well-being.

Table A1[12] provides descriptive for the MIVs and also, we observe that hypogamous mothers have higher antenatal care visits which starts during first trimester of pregnancy itself and they have higher say in their spouse's health care.

Table 1. Descriptive statistics – analytical sample and by marriage type

Variable	Obs	Mean	Std. Dev.	Min	Max	Hypogamy	Non-hypogamy	Diff (sig)
Outcome variables								
Not Stunted (1=Yes; 0= No)	141,942	0.65	0.48	0	1	0.665	0.646	**
Not Wasted (1=Yes; 0= No)	138,781	0.81	0.39	0	1	0.816	0.812	ns
Not Underweight (1=Yes; 0= No)	144,633	0.69	0.46	0	1	0.704	0.689	**
Treatment Variable								
Mother educational hypogamy	145,422	0.25	0.43	0	1			
Child characteristics								
Child male	145,422	0.52	0.50	0	1	0.526	0.524	ns
Child age (in months)	145,422	29.86	17.47	0	59	29.26	30.05	**
Birth order	145,422							**
1		0.36	0.48	0	1	40.4	34.11	
2		0.34	0.47	0	1	37.1	33.43	
3		0.16	0.37	0	1	14.19	17.09	
4 or more		0.14	0.34	0	1	8.31	15.37	
Parental characteristics								
Mother's education (Years)	145,422	7.615	5.029	0	20	10.485	6.663	**
Father's education (Years)	145,422	8.516	4.703	0	20	6.447	9.203	**
Average Parental education (Years)	145,422	8.066	4.404	0	20	8.467	7.933	**
Parental age difference (M-F)	145,422	-4.398	3.589	-29	15	-4.665	-4.309	**
Mother's characteristics								
Mother's age	145,422	27.512	5.014	15	49	26.841	27.735	**
Mother's height (1 = Normal; 0 = H below 145cm)	144,781	0.89	0.32	0	1	0.884	0.888	*
Living to total ever born child ratio	145,422	0.97	0.09	0.17	1	0.979	0.973	**
Household characteristics								
Religion	145,422							**
Hindu		0.76	0.43	0	1	74.58	75.82	
Muslim		0.11	0.32	0	1	11.39	11.22	
Others		0.13	0.34	0	1	14.03	12.96	
Caste	145,422							**
None of them		0.17	0.38	0	1	17.45	17.39	
Scheduled caste		0.21	0.41	0	1	22	20.96	
Schedule tribe		0.23	0.42	0	1	19.89	23.58	
Other Backward Caste (OBC)		0.39	0.49	0	1	40.66	38.07	
Wealth index	145,422							**
Poorest		0.25	0.44	0	1	21.02	26.92	
Poor		0.23	0.42	0	1	22.88	22.76	
Middle		0.20	0.40	0	1	22.24	18.89	
Rich		0.18	0.38	0	1	19.73	17	
Richest		0.14	0.35	0	1	14.13	14.43	
Rural residence	145,422	0.78	0.41	0	1	75.16	79.2	**
Instrumental Variables								

[12] Table A1 is in Appendix

District hypogamy level	145,422	0.195	0.09	0.02	0.52	0.224	0.185	**
Exposure to mass media	145,422	0.579	0.493	0	1	66.43	55.02	**

Note: * $p < 0.05$, ** $p < 0.01$ (signifies significant differences between non-hypogamy and hypogamy based on chi-square test or t test)
ns- not significant

Figure 1 depicts the prevalence of educational hypogamy across states in India. The states with the highest percentages of hypogamous marriages, ranking in the top 5, are Kerala, Tamil Nadu, Tripura, Sikkim, and Meghalaya. On the other hand, the states with the lowest percentages, ranking in the bottom 5, are Jammu and Kashmir, Rajasthan, Bihar, Jharkhand, and Arunachal Pradesh. Hypogamous marriages are more prevalent in south and northeast India.

The results presented in Table A2 demonstrate changes in the prevalence of hypogamy, as well as the incidence of stunting, wasting, and underweight in India between NFHS-4 and NFHS-5 surveys. The prevalence of hypogamy has increased from 23% in NFHS-4 to 26% in NFHS-5. In contrast, the incidence of stunting has declined from 38% to 36%, wasting from 21% to 19%, and underweight from 36% to 32% over the same period. These findings indicate a shift in partner selection patterns and improvements in child nutritional indicators.

Moreover, the analysis reveals heterogeneity across both treatment variables (prevalence of hypogamy) and outcome variables (incidence of stunting, wasting, and underweight) at the state level. This suggests that the observed changes in partner selection and child health outcomes are not uniform across all states in India, and there may be varying factors contributing to these variations. These findings provide insights into the dynamic nature of hypogamy and child health outcomes in India.

Figure 2 illustrates the upward trend in average education levels and the narrowing gap between fathers and mothers' education across different cohorts of marriages.

Figure 3 presents the changes in educational assortative patterns over the past four decades among marriage cohorts. The prevalence of hypogamy has significantly increased from approximately 5% to nearly 35% during this period (Sarkar, 2022). Conversely, hypergamy and homogamy have consistently declined over the same timeframe.

5.2. Main Results

Probit and Biprobit results

Table 2 presents the findings of the probit and biprobit regressions examining the impact of parental educational hypogamy on child health outcomes, specifically the probabilities of not being stunted and not being underweight. Models 1 and 4 present the results of probit regressions without controls, assuming exogeneity of parental hypogamy. We report the marginal effects, which represent the change in the probability of not being stunted or not being underweight when the hypogamy dummy variable changes from 0 to 1. The marginal effect for not being stunted is 0.028, indicating that the propensity for not being stunted among hypogamous mothers is 2.8 percentage points higher than non-hypogamous mothers. Similarly, the corresponding propensity for not being underweight is 1.8 percentage points.

Models 2 and 5 introduce controls into the regression. While we still find significant results for not being stunted, the effect size decreases to 1%. However, we fail to find any significant results for not being underweight after accounting for covariates.

Models 3 and 6 report the marginal effects for the biprobit regressions. The direction and significance of the effects of educational hypogamy remain consistent compared to the standard probit models. Once again, a positive and significant effect is found regarding having hypogamous mothers on the probabilities of not being stunted and not being underweight. Having a hypogamous mother increases the probability of not being stunted and not being underweight by 15.3 and 7.4 percentage points, respectively.

To assess the validity of the instruments used, we calculate the first-stage F statistics and the Cragg-Donald Wald F statistic. The reported F statistics in Table 2 are significant and greater than 10, indicating that the instruments are not weakly correlated with the predictor variable. This is significantly higher than the critical statistic of 19.93 based on Stock and Yogo's 10 percent threshold, providing sufficient statistical evidence to reject the null hypothesis that the instruments are weak.

Furthermore, we test whether both instruments are jointly correlated with the outcome variable by estimating the Hansen J statistics. The insignificant result of this test suggests that both instruments are not jointly correlated with the outcome variable. This is the test of overidentifying restrictions. The joint null hypothesis is that the instruments are valid instruments, i.e., uncorrelated with the error term, and that the excluded instruments are correctly excluded from the estimated equation. Accepting the null hypothesis indicates that the instruments are valid and exogenous.

Finally, we examine whether biprobit models should be preferred over single-stage models by conducting the Wald test of rho. The significant result of this test confirms that biprobit should be preferred over single-stage probit due to the confounding bias present in the probit model. The results of all the tests are reported at the end of Table 2. Based on these various tests, we can conclude that both instruments are valid and provide consistent estimates.

Table 2: Effect of Parental education hypogamy on child health outcomes
(Probit & biprobit results: Marginal effects)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Not stunted			**Not underweight**		
	Probit	Probit	Biprobit	Probit	Probit	Biprobit
Edu Hypogamy	0.028**	0.010*	0.153**	0.018**	0.001	0.074*
	(0.004)	(0.004)	(0.030)	(0.004)	(0.004)	(0.033)
Child male		-0.021**	-0.020**		-0.024**	-0.024**
		(0.003)	(0.003)		(0.003)	(0.003)
Child age (in months)		-0.002**	-0.002**		-0.002**	-0.002**
		(0.000)	(0.000)		(0.000)	(0.000)
Birth order (Ref: 1)						
2		-0.030**	-0.029**		-0.022**	-0.021**
		(0.004)	(0.004)		(0.004)	(0.004)
3		-0.050**	-0.042**		-0.041**	-0.038**
		(0.006)	(0.006)		(0.006)	(0.006)
4 or more		-0.076**	-0.060**		-0.050**	-0.043**
		(0.007)	(0.008)		(0.007)	(0.008)
Average Parental education		0.008**	0.008**		0.006**	0.006**
		(0.001)	(0.000)		(0.000)	(0.000)
Parental age difference (M-F)		-0.002**	-0.002**		-0.002**	-0.002**
		(0.001)	(0.001)		(0.001)	(0.001)
Mother's age		0.005**	0.006**		0.002**	0.003**
		(0.000)	(0.000)		(0.000)	(0.000)
Mother's height		0.158**	0.155**		0.116**	0.116**
		(0.005)	(0.005)		(0.005)	(0.005)
Living to total ever born child ratio		-0.013	-0.012		-0.034*	-0.034^{+}
		(0.018)	(0.017)		(0.017)	(0.017)
Religion (Ref: Hindu)						
Muslim		-0.020**	-0.026**		-0.015*	-0.018**
		(0.006)	(0.006)		(0.006)	(0.006)
Others		0.009	0.006		-0.001	-0.002

		(0.010)	(0.010)		(0.010)	(0.010)
Caste (Ref: None of them)						
Scheduled caste		-0.048**	-0.046**		-0.045**	-0.045**
		(0.006)	(0.006)		(0.006)	(0.006)
Schedule tribe		-0.030**	-0.024**		-0.058**	-0.055**
		(0.007)	(0.007)		(0.007)	(0.007)
OBC		-0.029**	-0.029**		-0.031**	-0.031**
		(0.005)	(0.005)		(0.005)	(0.005)
Wealth index (Ref: Poorest)						
Poor		0.033**	0.029**		0.047**	0.045**
		(0.005)	(0.005)		(0.005)	(0.005)
Middle		0.061**	0.055**		0.078**	0.075**
		(0.006)	(0.006)		(0.006)	(0.006)
Rich		0.104**	0.099**		0.114**	0.112**
		(0.007)	(0.007)		(0.007)	(0.007)
Richest		0.129**	0.129**		0.140**	0.141**
		(0.008)	(0.008)		(0.008)	(0.008)
Rural		0.002	0.007		0.006	0.008
		(0.005)	(0.005)		(0.005)	(0.005)
State FE		Y	Y		Y	Y
N	141942	141322	141322	144633	144000	144000
Rho/ $\rho(\varepsilon_1, \varepsilon_2)$			-0.249			-0.129
Test of endogeneity						
Wald test of rho=0: Chi-sq(1)			20.235**			4.672*
Adequacy of instruments						
First-stage F-statistic			98.32**			100.33**
Underidentification test:						
Kleibergen-Paap rk LM statistic Chi-sq(2)			635.491**			650.150**
Weak identification test:						
Cragg-Donald Wald F statistic			705.501			721.970
Critical statistic by Stock-Yogo (10%)			19.93			19.93
Test of overidentifying restrictions:						
Hansen J statistic chi2(1)			0.001 (p = 0.973)			0.501(p=0.479)

Note: Figures are rounded up to 3 decimal points. Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$

In addition, we investigated the influence of Parental hypogamy on the probability of not being wasted. However, our analysis using both standard probit and biprobit models did not yield any statistically significant results for this outcome. The detailed table presenting these findings can be found in the appendix (table A3). Earlier research has also been unable to identify substantial findings regarding wasting. In the case of Cambodia, Miller and Rodgers (2009) discovered a connection between maternal education and stunting but did not observe a similar relationship with wasting. They hypothesized that the weaker association with wasting might be attributed to the limited effectiveness of maternal education in preventing illnesses like diarrhoea when there are widespread sources of infection.

The negative correlation coefficient (ρ) between the error terms $(\varepsilon_1, \varepsilon_2)$ indicates that unobservable or unmeasured factors that contribute to the likelihood of being in a hypogamous marriage are also associated with a reduced probability of not being stunted or not being underweight. This finding provides supporting evidence for assumption of negative selection bias under the Monotone treatment selection assumption. In other words, women who are more likely to enter hypogamous marriages are also more likely to have adverse child health outcomes, such as stunting and being underweight.

Bounds for Average Treatment Effect using Partial Identification Approach

We report sharp bounds on the ATE and Imbens and Manski (2004) 95% confidence intervals (using 30 bootstraps) 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 the assumption of MTS and multiple MIVs namely state net domestic product per capita (MIV1), health index (MIV2), district female literacy rate (MIV3), average parental education (MIV4), district level sanitation (MIV5), district level exposure to mass media (MIV6) and antenatal care (MIV7).

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. Nonetheless, the bounds are useful in excluding possible values of the ATE. For example, table 3 shows that the bounds on the ATE of parental educational hypogamy on child's likelihood of not being stunted are [-0.568, 0.432]. Corresponding values for not being wasted and underweight are [-0.655, 0.345] and [-0.591, 0.409]. Thus, a considerable range of values of the ATE, especially in the positive 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, table 3 shows that, the imposition of MTS causes the bounds on the ATE of parental educational hypogamy on child's likelihood of not being stunted shrinks from [-0.568, 0.432] to [0.020, 0.432]. Likewise, the imposition of MTS causes the bounds on the ATE of parental educational hypogamy on child's likelihood of not being wasted (not being underweight) shrinks from [-0.655, 0.345] to [0.004, 0.345] ([-0.591, 0.409] to [0.015, 0.409]). 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 as positive. Moreover, all the confidence intervals exclude zero. This indicates that, even without invoking further assumptions, we can claim, parental educational hypogamy has a significant positive effect on child's anthropometric indicators.

Third, the MIV restrictions when imposed along with the MTS assumption leads to further tightening of the bounds. For example, table 3 shows that under MIV1 the bounds on the ATE of parental educational hypogamy on child's likelihood of not being stunted are [0.034, 0.403]. Corresponding bounds on ATE of parental educational hypogamy on child's likelihood of not being wasted (not underweight) are [0.005, 0.313] ([0.022, 0.387]). 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 parental educational hypogamy on child health outcomes.

Overall, our results indicate that parental educational hypogamy leads to better child health and significant decline in the prevalence of stunting, wasting and underweight. Specifically, based on NFHS sample, it can be concluded that parental educational hypogamy reduces the child's likelihood of being stunted by at least 2- 3.4%, wasted by 0.4 - 1.9% and underweight by 1.5- 3.8%. These findings are in line with our results of biprobit for stunting and underweight. The ATE under biprobit fall within the bounds. We observe the overlapping of confidence intervals between point-estimated ATE and bound-estimated ATE.

Next, we explore some potential channels by which parental educational hypogamy can improve the child health.

Mechanism

One potential mechanism through which parental educational hypogamy may positively impact child health is by improving the prenatal and providing women greater decision-making power regarding healthcare within households. Hypogamous marriages symbolize higher bargaining power for women, leading to better prenatal care practices. Women in hypogamous marriages are more likely to adhere to the recommended number of antenatal visits, initiate care in the first trimester, and take necessary precautions and preventive measures during pregnancy. Despite facing similar societal constraints, mothers in hypogamous marriages tend to fare better in terms of their own health, spousal health, and consequently, the health of their children. This can be attributed to their higher awareness and better knowledge of health- related matters.

Table 3. ATE Estimates of parental educational hypogamy on child health based on NFHS-5 data

	Estimates [LB, UB]	**Confidence intervals** [LB, UB]
Stunting [Y= 1 , if child is not stunted]		
No assumption (worst case bounds)	[-0.568, 0.432]	[-0.570, 0.433]
MTS	[0.020, 0.432]	[0.014, 0.433]
MTS + MIV1 (SNDP_pc)	[0.034, 0.403]	[0.028, 0.408]
MTS + MIV2 (Healthidx)	[0.027, 0.362]	[0.023, 0.365]
MTS + MIV3 (dist_Flit)	[0.020, 0.424]	[0.014, 0.425]
MTS + MIV4 (avg_parentaledu)	[0.020, 0.428]	[0.019, 0.431]
MTS + MIV5 (dist_sanitation)	[0.020, 0.431]	[0.014, 0.433]
MTS + MIV6 (dist_expo)	[0.020, 0.409]	[0.017, 0.414]
MTS + MIV7 (ANC)	[0.020, 0.319]	[0.016, 0.321]
Wasting [Y= 1 , if child is not wasted]		
No assumption (worst case bounds)	[-0.655, 0.345]	[-0.657, 0.346]
MTS	[0.004, 0.345]	[0.002, 0.346]
MTS + MIV1 (SNDP_pc)	[0.005, 0.313]	[0.003, 0.319]
MTS + MIV2 (Healthidx)	[0.019, 0.267]	[0.017, 0.268]
MTS + MIV3 (dist_Flit)	[0.004, 0.294]	[0.002, 0.295]
MTS + MIV4 (avg_parentaledu)	[0.004, 0.298]	[0.002, 0.301]
MTS + MIV5 (dist_sanitation)	[0.004, 0.333]	[0.001, 0.339]
MTS + MIV6 (dist_expo)	[0.004, 0.311]	[0.000, 0.316]
MTS + MIV7 (ANC)	[0.004, 0.233]	[0.003, 0.235]
Underweight [Y= 1 , if child is not underweight]		
No assumption (worst case bounds)	[-0.591, 0.409]	[-0.593, 0.411]
MTS	[0.015, 0.409]	[0.010, 0.411]
MTS + MIV1 (SNDP_pc)	[0.022, 0.387]	[0.017, 0.390]
MTS + MIV2 (Healthidx)	[0.038, 0.327]	[0.028, 0.332]
MTS + MIV3 (dist_Flit)	[0.015, 0.406]	[0.007, 0.407]
MTS + MIV4 (avg_parentaledu)	[0.015, 0.392]	[0.015, 0.396]
MTS + MIV5 (dist_sanitation)	[0.015, 0.409]	[0.012, 0.411]
MTS + MIV6 (dist_expo)	[0.015, 0.387]	[0.013, 0.388]
MTS + MIV7 (ANC)	[0.015, 0.292]	[0.013, 0.306]

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.

The literature on bargaining dynamics has given limited attention to the implications of significant improvements in women's educational attainment observed in recent decades. Notably, educational hypergamy has declined globally across high-, middle-, and low-income countries (Esteve et al., 2012, 2016). This decline may have important implications for child health, nutrition, and overall well-being. Research has shown that when mothers have higher educational status relative to fathers, it is associated with improved child nutrition, reduced morbidity, and lower child mortality rates in regions such as Africa, Latin America, and Asia. This association could be attributed to several factors, including better labor market outcomes for women, higher economic status within the family, and increased control over decisions related to child health and expenditures (Maitra, 2004; Quisumbing & Maluccio, 2003; Thomas, 1994).

We examine this potential channel by estimating the ATE of parental educational hypogamy on prenatal care, first time visit to antenatal care, say in own and spousal healthcare under various assumptions regarding the selection process as before. Note, the MTS assumption here is that children having hypogamous mothers apriori, are less likely to have prenatal care than women in non-hypogamous marriages. The MTS assumption that children with hypogamous mothers are less likely to receive prenatal care can be further supported by considering the additional challenges that women in hypogamous marriages face when accessing healthcare. In particular, women in such marriages may have partners with lower incomes, which could limit their ability to afford healthcare costs, including prenatal care. Additionally, partners with lower levels of education may be less knowledgeable about the importance of prenatal care, which could make them less likely to encourage their partners to seek it.

Moreover, women in hypogamous marriages may face the double burden of balancing work and family responsibilities. This can cause significant stress, intra-household conflict, relationship instability, and even severe domestic violence as a form of male backlash. Such challenges can further reduce the likelihood that these women will seek prenatal care, leading to a decrease in the number of prenatal visits. Furthermore, women in hypogamous marriages may start antenatal care later than their non-hypogamous counterparts. This delay in seeking care can result in complications during pregnancy and childbirth, leading to adverse health outcomes for both mother and child.

Additionally, women in hypogamous marriages may have lower say in decisions regarding their own and their spouse's healthcare due to the inherent patriarchy in Indian households. In such cases, women may be considered as undermining their husband's authority if they put forward their opinions and, which can hurt their ego. This can result in women being less likely to have control over their own healthcare decisions, including the decision to seek prenatal care.

Table 4. ATE Estimates of parental educational hypogamy on various mechanism based on NFHS-5		
	Estimates [LB, UB]	**Confidence intervals** [LB, UB]
Antenatal visits		
No assumption (worst case bounds)	[-0.528, 0.472]	[-0.529, 0.473]
MTS	[0.058, 0.472]	[0.052, 0.473]
MTS + MIV (dist_flit)	[0.058, 0.462]	[0.052, 0.464]
First antenatal visit		
No assumption (worst case bounds)	[-0.624, 0.376]	[-0.625, 0.377]
MTS	[0.008, 0.376]	[0.006, 0.377]
MTS + MIV (dist_flit)	[0.008, 0.341]	[0.006, 0.347]
Say in decision about their own healthcare		
No assumption (worst case bounds)	[-0.648, 0.352]	[-0.653, 0.357]
MTS	[0.008, 0.352]	[0.005, 0.357]
MTS + MIV (dist_flit)	[0.008, 0.302]	[0.005, 0.315]
Say in decision about husband's healthcare		
No assumption (worst case bounds)	[-0.591, 0.409]	[-0.594, 0.411]
MTS	[0.014, 0.409]	[0.003, 0.411]
MTS + MIV (dist_flit)	[0.014, 0.371]	[0.003, 0.391]

We present results in table 4. We find that, for all the outcomes, the bounds on the ATE under the combined MTS-MIV assumptions are strictly positive and statistically significant. This indicates that hypogamy increases the likelihood of women to have antenatal care, visit of antenatal care in first trimester itself. They have higher say in their own and their husband's health care. For example, under MTS and MIV (district female literacy) assumption, the bounds on the ATE reveal that hypogamy increases the likelihood of women to have antenatal care by at least 5.8%, first antenatal care visit in first trimester and say in own healthcare decisions by at least by at least 0.8%, and say in spousal's healthcare by at least 1.4%.

These findings indicate that parental educational hypogamy leads to an increase in the likelihood of women to have 4+ antenatal care visits, visit of antenatal care in first trimester itself, higher say in their own and their husband's health care. Because we also find evidence that parental educational hypogamy decreases the likelihood of having stunted, wasted and underweight children, this result could be interpreted as providing suggestive evidence that a channel through which parental educational hypogamy impacts child health is by improving the prenatal context and higher say in decisions related to healthcare.

6. Discussion

In this study, we investigate whether there is a causal relationship between parental educational hypogamy and a child's health, specifically focusing on indicators such as stunting, wasting, and underweight. Given India's unique social and economic landscape, an increase in educational hypogamy might not necessarily lead to greater bargaining power for women. We hypothesized that women who are self-selecting in hypogamous relationships, compared to those in non-hypogamous relationships, are more likely to have children who experience stunting, wasting, or being underweight due to the violation of patriarchal gender beliefs and norms regarding gender roles. A discrepancy in educational status between husbands and wives can result in stress, tension, and intimate partner violence (IPV) as a form of backlash. Consequently, a mother's higher relative educational status may have minimal or even adverse effects on her child's nutrition and well-being. Even if mothers' relative educational status improves, it may not enhance their bargaining power if they face limited earning prospects due to the poor quality of education and low returns in the Indian labour market.

We employed data from the NFHS-5 and utilized recursive bivariate probit analysis to examine the direction of selection bias, supporting the assumptions in our nonparametric bound analysis. Our results reveal a positive and significant effect of hypogamy on the probability of a child not experiencing stunting or being underweight. This suggests that women in hypogamous marriages possess greater bargaining power compared to those in nonhypogamous marriages. Our findings indicate that a mother's higher relative educational status significantly influences child health.

These findings underscore the importance of both the overall education level of parents and the relative education of mothers in relation to fathers' for a child's health. The increasing prevalence of educational hypogamy, where women have higher levels of education than their husbands, should not be mistaken as a sign of a significant breakdown in gender barriers or a promising shift in gender norms within Indian society. On the contrary, it might reflect the deep-rooted social and gender stratification that persists despite educational advancements. While expanding women's education is undoubtedly a step in the right direction, it should be recognized as an initial milestone rather than a comprehensive solution for achieving gender equality. The fact that educational hypogamy is on the rise suggests that other societal factors, such as traditional gender roles and expectations, continue to shape and limit women's opportunities and choices. Educational hypogamy can be seen as a manifestation of the unequal power dynamics and entrenched gender norms that persist in Indian society. It may indicate that even as more women access education, traditional notions of male dominance and female subordination are still deeply ingrained. This can hinder the progress toward true gender equality and limit the transformative potential of women's education.

To address these underlying social and gender stratification, it is crucial to go beyond educational advancements alone. Efforts should be directed toward challenging and transforming the existing social norms, dismantling systemic barriers, and promoting women's empowerment in all aspects of life. This includes enhancing women's economic opportunities, promoting their leadership roles, ensuring equal access to resources and decision-making, and fostering a supportive environment that encourages gender equality.

Therefore, while acknowledging the importance of expanding women's education and its positive effect on child's health through more women entering into hypogamous marriages, it is imperative to recognize that it is just the beginning of a larger and more complex journey toward achieving gender equality. It requires concerted efforts across various domains to challenge and change the deep-rooted social and gender stratification in Indian society. We can only truly create a more equal and inclusive society for all by addressing these underlying factors.

References

Abufhele, A., Pesando, L. M., & Castro T., A. F. (2022). Parental educational similarity and inequality implications for infant health in Chile: Evidence from administrative records, 1990–2015. *Research in Social Stratification and Mobility*, *82*, 100736.

Abuya, B. A., Ciera, J., & Kimani-Murage, E. (2012). Effect of mother's education on child's nutritional status in the slums of Nairobi. *BMC Pediatrics*, *12*(1), 80.

Afridi, F., Dinkelman, T., & Mahajan, K. (2016). Why are Fewer Married Women Joining the Work Force in India? A Decomposition Analysis Over Two Decades. *SSRN Electronic Journal*.

Al-Samarrai, S., & Bennell, P. (2007). Where has all the education gone in sub-Saharan Africa? Employment and other outcomes among secondary school and university leavers. *The Journal of Development Studies*, *43*(7), 1270–1300.

Ariyo, T., & Jiang, Q. (2021). Trends in the association between educational assortative mating, infant and child mortality in Nigeria. *BMC Public Health*, *21*(1), 1493.

Behrman, J. A. (2020). Mother's Relative Educational Status and Early Childhood Height-for-Age z Scores: A Decomposition of Change Over Time. *Population Research and Policy Review*, *39*(1), 147–173.

Behrman, J. R., Foster, A. D., Rosenzweig, M. R., & Vashishtha, P. (1999). Women's Schooling, Home Teaching, and Economic Growth. *Journal of Political Economy*, *107*(4), 682–714.

Bingley, P., Cappellari, L., & Tatsiramos, K. (2022). Parental assortative mating and the intergenerational transmission of human capital. *Labour Economics*, *77*, 102047.

Bowen, M., Dysinger, R., Brodey, W., & Basmania, B. (1978). In M. Bowen. *Family Therapy in Clinical Practice. New York: Jason Aronson*, 3–15.

Breen, R., & Salazar, L. (2011). Educational Assortative Mating and Earnings Inequality in the United States. *American Journal of Sociology*, *117*(3), 808–843.

Chatterjee, E., Desai, S., & Vanneman, R. (2018). Indian Paradox: Rising Education, Declining Womens' Employment. *Demographic Research*, *38*, 855–878.

Cools, S., & Kotsadam, A. (2017). Resources and Intimate Partner Violence in Sub-Saharan Africa. *World Development*, *95*, 211–230.

De Hauw, Y., Grow, A., & Van Bavel, J. (2017). The Reversed Gender Gap in Education and Assortative Mating in Europe. *European Journal of Population*, *33*(4), 445–474.

Eika, L., Mogstad, M., & Zafar, B. (2019). Educational Assortative Mating and Household Income Inequality. *Journal of Political Economy*, *127*(6), 2795–2835.

Esteve, A., García-Román, J., & Permanyer, I. (2012). The Gender-Gap Reversal in Education and Its Effect on Union Formation: The End of Hypergamy? *Population and Development Review*, *38*(3), 535–546.

Esteve, A., Schwartz, C. R., Van Bavel, J., Permanyer, I., Klesment, M., & Garcia, J. (2016). The End of Hypergamy: Global Trends and Implications. *Population and Development Review*, *42*(4), 615–625.

Fernández, R., & Rogerson, R. (2001). Sorting and Long-Run Inequality. *The Quarterly Journal of Economics*, *116*(4), 1305–1341.

Glewwe, P. (1999). Why Does Mother's Schooling Raise Child Health in Developing Countries? Evidence from Morocco. *The Journal of Human Resources*, *34*(1), 124–159.

Güneş, P. M. (2015). The role of maternal education in child health: Evidence from a compulsory schooling law. *Economics of Education Review*, *47*, 1–16.

Hahn, Y., Nuzhat, K., & Yang, H.-S. (2018). The effect of female education on marital matches and child health in Bangladesh. *Journal of Population Economics*, *31*(3), 915–936.

Han, S. W. (2022). Is It Only a Numbers Game? A Macro-Level Study of Educational Hypogamy. *Demography*, *59*(4), 1571–1593.

Holmlund, H. (2022). How Much Does Marital Sorting Contribute to Intergenerational Socioeconomic Persistence? *Journal of Human Resources*, *57*(2), 372–399.

Hornung, C. A., McCullough, B. C., & Sugimoto, T. (1981). Status Relationships in Marriage: Risk Factors in Spouse Abuse. *Journal of Marriage and Family*, *43*(3), 675–692.

Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, *62*(2), 467–475.

Kc, S., Barakat, B., Goujon, A., Skirbekk, V., Sanderson, W., & Lutz, W. (2010). Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005-2050. *Demographic Research*, *22*, 383–472.

Kerr, M. E. (2000). *One family's story: A primer on Bowen Theory. The Bowen Center for the Study of the Family.*

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*(499), 958–975.

Lam, D., & Duryea, S. (1999). Effects of Schooling on Fertility, Labor Supply, and Investments in Children, with Evidence from Brazil. *The Journal of Human Resources*, *34*(1), 160–192.

Li, C., Poskitt, D. S., & Zhao, X. (2019). The bivariate probit model, maximum likelihood estimation, pseudo true parameters and partial identification. *Journal of Econometrics*, *209*(1), 94–113.

Lin, Z., Desai, S., & Chen, F. (2020). The Emergence of Educational Hypogamy in India. *Demography*, *57*(4), 1215–1240.

Maitra, P. (2004). Parental bargaining, health inputs and child mortality in India. *Journal of Health Economics*, *23*(2), 259–291.

Mangrio, E., Hansen, K., Lindström, M., Köhler, M., & Rosvall, M. (2011). Maternal educational level, parental preventive behavior, risk behavior, social support and medical care consumption in 8-month-old children in Malmö, Sweden. *BMC Public Health*, *11*, 891.

Manski, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard University Press.

Manski, C. F. (2003). Identification Problems in the Social Sciences and Everyday Life. *Southern Economic Journal*, *70*(1), 11–21.

Manski, C. F., & Pepper, J. V. (1998). *Monotone Instrumental Variables with an Application to the Returns to Schooling* (Working Paper No. 224). National Bureau of Economic Research.

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*(2), 411–436.

Millimet, D. L., & Roy, M. (2015). Partial identification of the long-run causal effect of food security on child health. *Empirical Economics*, *48*(1), 83–141.

Miller, J. E., & Rodgers, Y. V. (2009). Mother's education and children's nutritional status: new evidence from Cambodia. *Asian Development Review*, *26*(1), 131-165.

Pepper, J. V. (2000). The Intergenerational Transmission of Welfare Receipt: A Nonparametric Bounds Analysis. *The Review of Economics and Statistics*, *82*(3), 472–488.

Pesando, L. M. (2021). Educational Assortative Mating in Sub-Saharan Africa: Compositional Changes and Implications for Household Wealth Inequality. *Demography*, *58*(2), 571–602.

Pesando, L. M. (2022). A Four-Country Study on the Relationship Between Parental Educational Homogamy and Children's Health from Infancy to Adolescence. *Population Research and Policy Review*, *41*(1), 251–284.

Prakash, R., & Singh, A. (2014). Who marries whom? Changing mate selection preferences in urban India and emerging implications on social institutions. *Population Research and Policy Review*, *33*, 205-227.

Quisumbing, A. R., & Maluccio, J. A. (2003). Resources at Marriage and Intrahousehold Allocation: Evidence from Bangladesh, Ethiopia, Indonesia, and South Africa. *Oxford Bulletin of Economics and Statistics*, *65*(3), 283–327.

Rauscher, E. (2020). Why Who Marries Whom Matters: Effects of Educational Assortative Mating on Infant Health in the United States 1969–1994. *Social Forces*, *98*(3), 1143–1173.

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: Statistics in Society*, *185*(4), 1952–1978.

Sarkar, K. (2022). Can status exchanges explain educational hypogamy in India? *Demographic Research*, *46*, 809–848.

Shajan, J., & Sumalatha, B. S. (2022). Maternal employment and children's health in India: An exploratory analysis. *Journal of Public Affairs*, *22*(3), e2580.

Tamer, E. (2010). Partial Identification in Econometrics. *Annual Review of Economics*, *2*(1), 167–195.

Thomas, D. (1994). Like Father, like Son; Like Mother, like Daughter: Parental Resources and Child Height. *The Journal of Human Resources*, *29*(4), 950–988.

Van Bavel, J. (2012). The reversal of gender inequality in education, union formation and fertility in Europe. *Vienna Yearbook of Population Research*, *10*, 127–154.

Weitzman, A. (2014). Women's and Men's Relative Status and Intimate Partner Violence in India. *Population and Development Review*, *40*(1), 55–75.

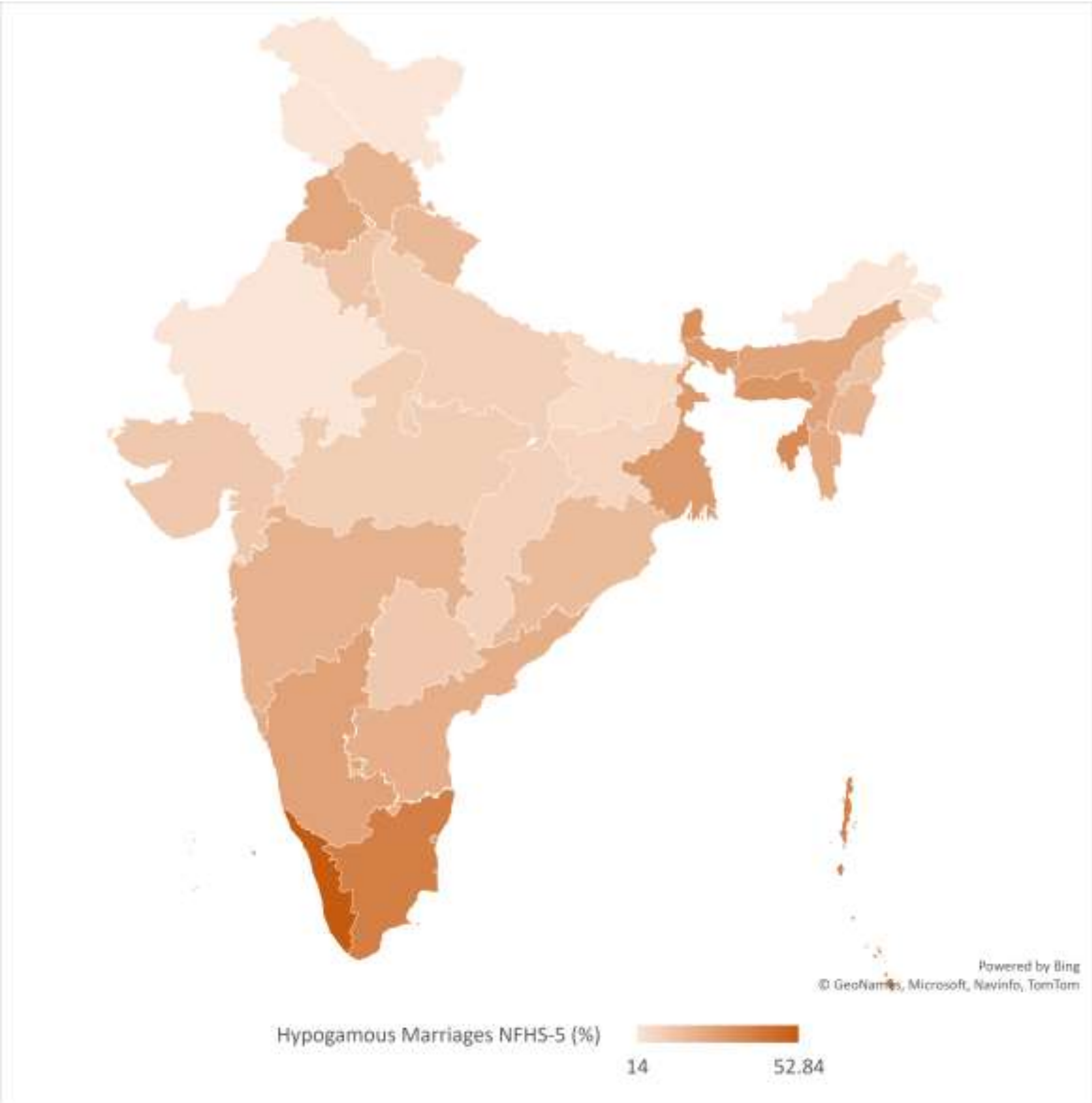

Figure 1. State wise prevalence of Educational Hypogamy in India

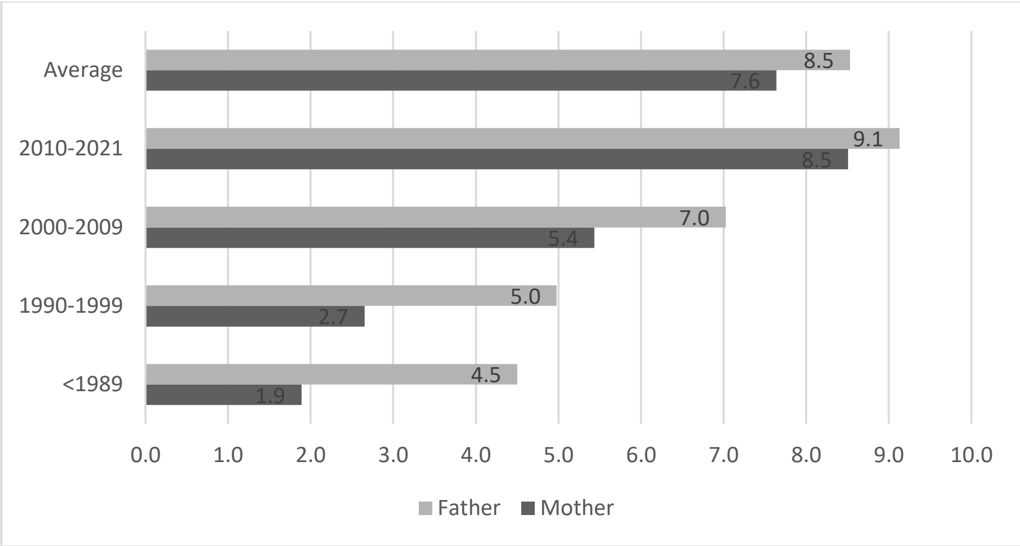

Figure 2. Mean years of education by parent gender across marriage cohorts

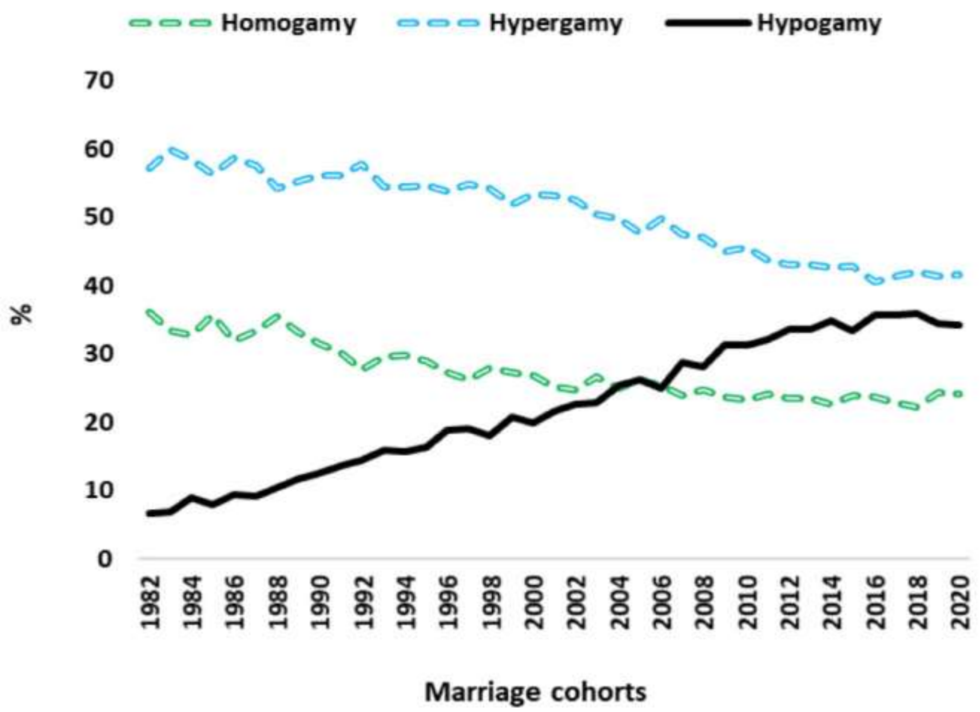

Figure 3. Educational assortative mating patterns over marriage cohorts in India

Source: Sarkar (2022) based on NFHS-5 (2019-2021)

Appendix

Table A1. Descritive Statistics for MIVs and Other Outcome Variables								
Variable	**Obs**	**Mean/ proportions**	**Std. Dev.**	**Min**	**Max**	**Hypogamy**	**Non-hypogamy**	**Diff(sig)**
Montone Instrumental Variables								
State net domestic product per capita (2019-20)	144,494	95226.77	64144.3	29794	313973			
Health index 2019-20(by NITI Aayog)	145,118	47.43	14.24	27	82.2			
District female literacy rate	145,422	57.49	13.21	23.86	98.29			
District level exposure to mass media	145,422	52.30	19.02	12.52	95.04			
District level sanitation	145,422	76.93	15.60	31.88	100			
Average Parental Education	145,422	8.07	4.40	0	20			
Antenatal care visits								
No anc	110,840	0.060	0.238	0	1			
1 visit	110,840	0.059	0.236	0	1			
2-3 visits	110,840	0.280	0.449	0	1			
4+ visits	110,840	0.601	0.490	0	1			
Mechanism								
Antenatal care visits (1= +4 visits; 0= 3 or less)	110,840	0.60	0.49	0	1	0.644	0.585	**
First anc visit in first trimester (1=yes; 0= No)	105,345	0.76	0.43	0	1	0.765	0.757	**
Say in their own healthcare (1= alone/Jointly;0= husband/others)	21,975	0.80	0.40	0	1	0.806	0.798	ns
Say in their spouse's healthcare (1= wife alone/Jointly;0= husband/others)	18,783	0.69	0.46	0	1	0.703	0.689	+
Note: + p< 0.1 * p < 0.05, ** p < 0.01								
ns- not significant								

Table A2. State wise Prevalence of Parental Educational Hypogamy and Stunting, Wasting, Underweight for NFHS-4 & NFHS-5

Region	State	Hypogamous Marriages NFHS-5 (%)	Hypogamous Marriages NFHS-4 (%)	Stunting NFHS5 (%)	Stunting NFHS4 (%)	Wasting NFHS5 (%)	Wasting NFHS4 (%)	Underweight NFHS5 (%)	Underweight NFHS4 (%)
North	Chandigarh	34.35	27.2	25.3	28.7	8.4	10.9	20.6	24.5
	Delhi	22.89	23.07	30.9	31.9	11.2	15.9	21.8	27.0
	Haryana	23.01	17.04	27.5	34.0	11.5	21.2	21.5	29.4
	Himachal Pradesh	27.1	22.15	30.8	26.3	17.4	13.7	25.5	21.2
	Jammu & Kashmir	14	17.44	26.9	27.4	19	12.1	21	16.6
	Punjab	31.02	27.35	24.5	25.7	10.6	15.6	16.9	21.6
	Rajasthan	14.02	11.63	31.8	39.1	16.8	23.0	27.6	36.7
	Uttarakhand	26.22	20.77	27	33.5	13.2	19.5	21	26.6
East	Bihar	16.92	12.68	34.6	37.6	18.9	23.1	31.3	37.7
	Jharkhand	18.53	16.94	35.7	42.0	18.9	25.8	33	42.8
	Odisha	25.49	22.04	39.7	46.2	17.3	17.9	32.1	39.5
	West Bengal	34.6	29.52	42.9	48.3	22.9	20.8	41	43.9
Central	Chhattisgarh	19.52	19.15	39.6	45.3	22.4	29.0	39.4	47.8
	Madhya Pradesh	20.46	16.28	31	34.1	18.1	20.4	29.7	34.4
	Uttar Pradesh	19.94	15.82	33.8	32.5	20.3	20.3	32.2	31.5
North-East	Arunachal Pradesh	18.79	16.16	28	29.3	13.1	17.3	15.4	19.4
	Assam	31.88	28.23	35.3	36.4	21.7	17.0	32.8	29.8
	Manipur	27.47	24.15	23.4	28.9	9.9	6.8	13.3	13.8
	Meghalaya	36.34	34.3	46.5	43.8	12.1	15.3	26.6	28.9
	Mizoram	30.23	29.56	28.9	28.1	9.8	6.1	12.7	12.0
	Nagaland	24.09	26.41	32.7	28.6	19.1	11.3	26.9	16.7
	Sikkim	37.31	30.9	22.3	29.6	13.6	14.2	13.1	14.2
	Tripura	38.76	31.02	32.3	24.3	18.2	16.8	25.6	24.1
West	Dadra & Nagar Haveli	24.72	20.29	39.4	41.7	21.6	27.6	38.7	38.8
	Goa	29.68	28.96	25.8	20.1	19.1	21.9	24	23.8
	Gujarat	21.97	20.31	39	38.5	25.1	26.4	39.7	39.3
	Maharashtra	28.25	25.08	35.2	34.4	25.6	25.6	36.1	36.0
South	Andaman & Nicobar Islands	42.27	36.33	22.5	23.3	16	18.9	23.6	21.5
	Andhra Pradesh	29.01	24.12	31.2	31.4	16.1	17.2	29.6	31.9
	Karnataka	32.76	28.7	35.4	36.2	19.5	26.1	32.9	35.2
	Kerala	52.84	50.12	23.4	19.7	15.8	15.7	19.7	16.1
	Lakshadweep	40.45	39.64	32	26.8	17.4	13.7	25.8	23.6
	Puducherry	43.88	34.7	20	23.7	12.4	23.6	15.3	22.0
	Tamil Nadu	42.38	35.06	25	27.1	14.6	19.7	22	23.8
	Telangana	21.99	20.25	33.1	28.0	21.7	18.0	31.8	28.3
All India		25.98	22.95	35.5	38.4	19.3	21.0	32.1	35.7

All India averages are inclusive of Daman and Diu & Ladakh.
Survey weights have been applied.
Hypogamy here is measured by difference in years of education between parents not by education level.

Table A3: Probit and biprobit results for wasting (Marginal effects)

	Not wasted		
	Probit	Probit	Biprobit
Edu Hypogamy	0.005 (0.004)	0.002 (0.004)	-0.024 (0.029)
Child male		-0.016** (0.003)	-0.016** (0.003)
Child age (in months)		0.002** (0.000)	0.002** (0.000)
Birth order (Ref: 1)			
2		-0.007+ (0.004)	-0.007+ (0.004)
3		-0.005 (0.005)	-0.007 (0.005)
4 or more		0.004 (0.006)	0.002 (0.007)
Average Parental education		0.002** (0.000)	0.002** (0.000)
Parental age difference (M-F)		-0.001+ (0.000)	-0.001+ (0.000)
Mother's age		-0.001+ (0.000)	-0.001+ (0.000)
Mother's height		0.007 (0.005)	0.007 (0.005)
Living to total ever born child ratio		-0.026+ (0.016)	-0.027+ (0.016)
Religion (Ref: Hindu)			
Muslim		-0.005 (0.005)	-0.004 (0.005)
Others		-0.002 (0.009)	-0.002 (0.009)
Caste (Ref: None of them)			
Scheduled caste		-0.021** (0.005)	-0.021** (0.005)
Schedule tribe		-0.043** (0.006)	-0.044** (0.006)
OBC		-0.022** (0.005)	-0.022** (0.005)

Wealth index (Ref: Poorest)			
Poor		0.017**	0.017**
		(0.004)	(0.004)
Middle		0.034**	0.035**
		(0.005)	(0.005)
Rich		0.035**	0.035**
		(0.006)	(0.006)
Richest		0.037**	0.036**
		(0.007)	(0.007)
Rural		0.009*	0.009^{+}
		(0.005)	(0.005)
State FE		Y	Y
N	138781	138176	138176
Rho			0.056
Wald test of rho=0: Chi-sq(1)			.803

Note: Figures are rounded up to 3 decimal points
Standard errors in parentheses
$^{+}$ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$