ISSN No. 2454 - 1427

CDE
January 2022

"Made in Heaven, Matched by Parents": Does Arranged Marriage Restrict Labour Market Autonomy and Participation of Women? Theory and Evidence from India

Sugata Bag
Email: sugata@econdse.org
Department of Economics
Delhi School of Economics

Anirban Kar
Email: anirban@econdse.org
Department of Economics
Delhi School of Economics

Working Paper No. 317

http://www.cdedse.org/pdf/work317.pdf

CENTRE FOR DEVELOPMENT ECONOMICS
DELHI SCHOOL OF ECONOMICS
DELHI 110007

"Made in Heaven, Matched by Parents": Does Arranged Marriage Restrict Labour Market Autonomy and Participation of Women? Theory and Evidence from India *

Sugata Bag† Delhi School of Economics

Anirban Kar‡ Delhi School of Economics

Abstract

Female labour force participation in India has stagnated despite gains in other aspects. Do Indian women prefer to stay out of labour market voluntarily or do social norms prevent their participation? We identify parental involvement in partner choice during marriage as an important bottleneck. We first find that women who had some degree of involvement in partner choice enjoy significantly more autonomy in post-marriage labour market choices than those whose marriages were arranged solely by parents. We use a marriage tradition instrument to estimate causal effects. Since autonomy and participation affect each other, next, we estimate simultaneous equations for autonomy and participation for only rural women. We use parental involvement in marriage and district-level share of drought-affected villages as two exogenous variables - the former for autonomy and the latter for participation. We find that autonomy significantly increases participation. We further explain the mechanism through a theoretical model. To distinguish between autonomy and participation; we introduce a new household delegation game. The main message of the theoretical model is that parental involvement in partner choice reduces women's ability to screen partners leading to relatively more mismatches, i.e., women who are inclined to work mismatched with men who prefer otherwise.

JEL Classification: J12, J16, D82

Key Words: Parental involvement, Partner selection, Female autonomy, Female labour market participation, Delegation game

*This research is supported by a Faculty Research Programme Grant under the IoE Scheme 2021, University of Delhi.

†sugata@econdse.org

‡anirban@econdse.org

1 Introduction

In her influential work on social-relational account of autonomy, philosopher Marina Oshana defines autonomy as a condition in which agents have ‘de facto power and authority over choices and actions significant to the direction of their lives’[35]. Any external condition that deprives or limits a person from fulfilling her aspirations is a restriction on her autonomy. In this paper we illustrate, both theoretically and empirically, that parental involvement in partner choice curbs female autonomy in post-marriage labour market choices. As far as we know, the existing literature on female autonomy has not investigated marriage practices at all; understandable given the lack of proximity between the two. First, there is a time lag between marriage and post-marriage labour choices. More importantly, the actors involved in these decisions are distinct. While marriage choices of a woman can be controlled by parents, they do not have any direct involvement in post-marriage labour market choices, which are typically controlled by the husband.

Economists have carefully documented the consequence of discrimination at natal home as well as husband’s control over occupation and fertility choices (among other aspects), but they have treated these as distinct spheres of women subjugation. There is a dearth of quantitative work to understand interaction and interrelation between different gender norms. We attempt to bridge the gap. Our theoretical argument shows that marriage practices affect post marriage labour market autonomy through screening of partner’s type and document its quantitatively large impact in India. We further show empirically that possession of (or lack of) autonomy translates into higher (or lower) female labour force participation thereby offering new insight into India’s dismal and stagnant female labour force participation. It turns out that despite the apparent separation in space and time mentioned above, combination of gender norms (at marriage and after) create distortions that cannot be understood if investigated separately. There is a small literature that has investigated similar questions. Huang et al.(2012) [25] shows that parental involvement in match-making adversely affects a couple’s marriage harmony and joint income, in China circa 1980s. In case of India, Jejeebhoy et al.(2013) [29] shows that lack of women’s involvement in their spouse selection leads to lower communication and interaction with their husbands, and exhibits lower agency and likely to experience marital violence. Chawla [12] finds evidence that women in parent-arranged marriages have lower bargaining power in household matters and enjoy lower couple interaction than those who had a say in choosing their spouse. While these papers focus on domestic matters; we investigate gender norms and labour market outcomes. Before we introduce the detail of our research, we first lay down the background of two key aspects of our thesis: labour market autonomy and parental involvement in marriage decisions.

Largely overlooked by economists, the role of marriage practices in the subjugation of women, however, has been emphasized by the feminists. Feminist literature recognizes marriage practices as a key site for the production of social hierarchies. For instance, Sherry Ortner persuasively argues that 'Control of the marriage system, always in the hands of men, transforms diffused authority ... into real power and controls' [34]. Uma Chakravarti (2002) [11] situates marriage restrictions at the heart of the Indian caste system. Since endogamous marriage is the backbone of the hierarchical caste system, she argues that it can only be maintained through constraining female sexuality. Consequentially 'marriage is ritualized as an act' through which parents can hand over their 'sexually pure' daughter to her husband (known as 'kanyadaan' in India). It is not surprising that parental involvement in daughter's partner search was a norm in all pre-industrial societies and continues to be the dominant mode of marriage in South Asia (Anukriti and Dasgupta, 2017 [4]). It is true that fully parent-arranged marriages, in which women have no choice at all, have declined to some extent over time. Instead, it is now quite common for the parents to short-list potential grooms, control the search process and restrict pre-marriage communication with potential matches (Desai and Andrist, 2010 [15]; Banerjee et al., 2013 [7]). Our paper is complementary to the work done by feminists. While feminists emphasize the arc of control through sexuality and purity, we document the disempowering aspect of marriage practices in purely economic terms.

The concept of labour market autonomy requires disambiguation. Economists have consistently argued that access to labour market brings greater autonomy for women. Anderson and Eswaran (2009) [1], Fletcher et al. (2017) [20] demonstrate that wage work contributes to women's autonomy in household decision making. Qian (2008) [38] find the same effect along with positive spillover on child schooling. Sivasankaran (2014) [41] observe that age at first childbirth increases with labour force participation, suggesting women have greater autonomy over fertility decisions. In these papers, autonomy in certain matters is the outcome that results from the control over labour market choices. But what determines autonomy over labour market choices in the first place? Literature on the determinants of female labour force participation exists (Goldin, 1995) [21]; Klasen, 2019 [30]). But we believe that participation and autonomy should be understood as two distinct concepts. Labour force participation is the eventual choice that a person makes from the set of available opportunities, whereas autonomy is the de facto power to make such a choice. Take an example of two women, living in the same village. Suppose that agricultural labour work is available in the village. The first woman does not face any restriction from her husband but finds the household care-giving more profitable than the labour market wage. In contrast, the second woman would have preferred to work as an agricultural labourer but is prohibited from working outside. Thus outcome wise these women are not different but the first woman still enjoys more autonomy than the latter. We can distinguish the concepts of outcome and autonomy in terms of decision theory. Suppose that a decision-maker is choosing from a choice set S. The element she chooses from S is the outcome. But availability in opportunities in S as compared to the universal choice set U, which gives us an idea about the external restriction that the decision-maker faces, can be considered as autonomy (or the lack of it).

Now, we are ready to provide an outline of our paper. Our theoretical model isolates a mechanism that channels the lack of female autonomy over partner selection into post-marriage labour market choices. Two types of mathematical models are quite common in 'gender' literature - an older generation of 'unitary models' (like, Becker, 1981 [9]) and relatively recent bargaining models (Lundberg and Pollak, 1996 [32]). In spirit, our model is closer to the latter, because we think household decisions arise from conflict and compromise. There are papers which also accommodate asymmetric information between partners (e.g. Anderson and Genicot, 2015 [2]). Our main point of departure from the bargaining models is the distinction between the concepts of autonomy and outcome outlined above. While bargaining models capture a negotiation process, its disagreement outcomes and compromise achieved, it is unsuitable for modelling opportunity sets and strategic choice of external constraints. Instead, we model the household negotiation as a delegation game.

Our basic theoretical argument is as follows. The game has two periods - marriage matching takes place in the first period followed by labour choice delegation game in the second period. As far as labour choice is concerned, a husband expects his wife to prioritize household care-work. To ensure it he may even explicitly constrains his wife's labour market choices. The wife, on the other hand, may or may not be inclined towards paid work but information on her 'type' is not available to the husband. If she is so inclined, then external restrictions lead to marital disharmony and some utility loss for the husband. Thus the husband has to balance his expectation for care-work and marital harmony. In our model, men also differ from each other in two aspects - earning capacity and outlook towards wife's claim on household resources, in short, 'patriarchy type'. We show that in equilibrium, the husband finds it optimal to leave the labour market choice to his wife but imposes an upper bound on what she can choose from. The greater the upper bound, the bigger is the opportunity set available to the wife and the higher is her autonomy in labour market. Further, the extent of autonomy offered by the husband depends on the husband's 'patriarchy type'. We find that irrespective of his earning capacity, men with a more egalitarian outlook allow higher labour market autonomy to their wives. To connect these observations with marriage practices, we go back to pre-marriage negotiations between a woman and her parents. Here the parents are primarily in charge. But parents only observe earning capacity of a potential match, not his patriarchy type. The woman, if allowed by her parents, can observe the patriarchy type as well, possibly through interactions before marriage. However, The decision to allow their daughter to interact with the potential match lies with the parents and parents may decide not to involve her when it is in their interest. The wedge between parents' and daughter's interests comes from the assumption that parents put more weight on the groom's earning capacity (symbolic of status) than the daughter does. This leads to another round of strategic decisions - parents can either approve/disapprove a match unilaterally or delegate the decision to their daughter.

Here delegation signifies autonomy in partner selection. We show that in equilibrium, irrespective of her type, a woman who gets to choose her partner also enjoys more autonomy in the labour market because she can screen a potential partner's patriarchy type which her parents cannot. Thus when parents control partner choice, all women are more likely to match with a restrictive partner. Does it have any impact on the outcome, that is actual female labour force participation? In our model, post-marriage labour force participation will be maximum when there is an assortative matching between work-inclined women and non-restrictive men. Instead, if a work-inclined woman marries a restrictive man then she cannot participate due to restriction. At the same time, a non-inclined woman prefers not to participate even when she is free to choose. Since parental involvement decreases autonomy for all types of women, it also reduces the overall labour market participation of women in society.

The prediction of our theoretical model leads to our first hypothesis that the different levels of autonomy enjoyed by a woman in partner choice at the time of marriage causally affect autonomy on her work choice after marriage. We test this empirically by using the second wave of the India Human Development Survey (2011- 12) [17]. This nationally representative database, as far as we know, is the only source that has information about labour market choices as well as marriage practices. The dataset consists of a sample of 39,523 women who were married between 1950 and 2012. The survey records female autonomy in work choice through a question that asks women who have the most say in their decisions to work. Roughly 46% women both in urban and rural areas claimed to have the most say on their work choice, while about 52% of women conceded that their husbands have the most say. As far as marriage is concerned, about 48% of the rural women and 31% of the urban women were married through the full parental arrangement. Roughly 48% of urban marriages and 68% urban marriages were jointly arranged and the residual, 4% to 5% were self-matched, respectively in rural and urban areas. The comparison across the three groups of matchmaking shows that compared to the women with parents as sole matchmakers, the other two groups, joint search and self-search, have relatively more autonomy in their work choice. The patterns are more or less the same across urban and rural areas. Based on the descriptive statistics, women who had self-searched marriages are about 9 percentage points more likely to have autonomy in their work choice vis-à-vis fully parents arranged marriages (see Table 2 and Figure 2).

The main challenge in estimating the causal effect is the endogenous choice of the match-making process. There could be unobserved attributes that are correlated to both marriage type and labour market autonomy. For example, it could be the case that more assertive women have autonomy in both partner choice and the job market, but it can not be controlled because assertiveness is not observed. It can also be the case that women with pre-marriage labour market experience, another aspect that we do not observe, enjoy more autonomy in both partner choice and post-marriage labour choice. To mitigate such selection problem and to isolate the true impact of marriage practices, we employ instrumental variable strategy. We build upon our instrumental variables on a longitudinal tradition of marriage pattern based on Huang et al.(2012) [25]. To instrument women's autonomy in partner selection (or, loosely speaking, marriage practices), we use state-wise (28 states and 5 Union territories) share of two marriage patterns (viz. jointly arranged and self-matched marriages) for an earlier birth cohort of women. The idea behind this is quite simple; parents are more likely to engage in partner selection for their off-springs in areas where there exists a social practice of parent-arranged marriages, both due to social learning and costly breaking of norms. However, these 'tradition' variables are independent of the individual characteristics and experience of the women and parents, and hence do not affect the post-marriage outcomes such as labour market autonomy at an individual level, satisfying the independence condition. Our IV estimates show that at the all-India level, women who had self-searched their spouses are 30 percentage point more likely to enjoy autonomy in their work choice vis-à-vis the ones who had fully parents arranged marriages, while likelihood is about 22.5 percentage points higher for the women had jointly arranged marriages. Although, the patterns are more or less same across urban and rural areas, as expected, the effects are even higher in rural areas than the all-India estimates. But in the urban area, the effects are much dampened, almost half of that of in rural areas, for both jointly arranged and self-matched marriage types. All these estimates are statistically significant and robust to alternative specifications. To put it in the perspective note that only 46% of women believed that they have the most say in labour market decisions. Thus parental involvement in marriage is a major determinant of future labour market prospects for women in India.

We next move to an important follow-up question: does autonomy translate into higher labour force participation for women? Female labour force participation has a prominent place in the Economics of Gender, as economists believe that paid work can bring financial independence to women resulting in stronger negotiating power in household matters. In India, the most baffling aspect has been abysmally low labour force participation of women. Using various survey datasets, Fletcher et al.(2017) [20] estimate that female labour force participation in 2011-12 was 27% compared to male labour force participation of 96%. There are indications that it has remained stagnant or fallen even further in recent years (see, Deloitte report, 2018 [14]). Several tentative explanations have been put forward to explain this phenomenon. It ranges from supply-side factors such as lack of suitable, remunerative jobs in the market, discriminatory hiring to demand-side factors like mobility restriction and gender norms. We provide robust evidence that lack of autonomy in labour market choice, i.e., external restrictions imposed upon women, significantly constrain women's participation.

The empirical challenge in estimating a causal relation comes from the bidirectional relation between autonomy and labour market outcome. Autonomy can potentially affect participation, but at the same time, labour market participation can enhance bargaining power and autonomy. We estimate a system of two simultaneous equations in which two outcome variables, viz. autonomy in work choice and labour market participation, influence each other. For the identification of our system in a Rural Indian set up, on one side we exploit the exogenous variation in labour market conditions due to shocks arising from droughts to determine woman's labour market participation, on the other side we use woman's actual marriage type in the past to determine her present-day labour market autonomy, along with usual controls. As expected, our regression shows that when women enjoy autonomy over work choice, they are likely to exhibit significantly 10 percentage point higher participation vis-à-vis those who do not possess the autonomy. On the flip side, when women had worked in the past one year (be it paid or unpaid), they tend to have on average 10 percentage point significantly higher probability of possessing the autonomy on their work choice vis-à-vis those who did not work.

Our main contribution is to uncover the role of marriage practices on post-marriage labour choices and outcomes. The interaction between marriage practices and post-marriage restrictions suggests that patriarchal norms and practices must be seen as a whole. It also points to potential intergenerational propagation of cultural norms, although Fernandez et al. (2004) [19] proposed a different channel. As far as policies are concerned, our results caution against 'quick-fix' responses. We illustrate that there could be inter-temporal, inter-sectoral links that can be as important as proximate causes. On the theoretical side, we introduce a new model in household negotiation. We show that delegation models rather than bargaining models capture the notion of autonomy more appropriately.

The paper is organized as follows. Section 2 lays down the theoretical foundation of our paper. Data description can be found in Section 3. Section 4 provides the empirical strategy, while Section 5 discusses our estimation results. The penultimate Section 6 deals with the empirical strategy and estimation results of our follow up question, whether autonomy translate into higher labour force participation. Finally, we draw some conclusions.

2 A Theoretical Model of Partner Selection and Post Marriage Household Bargaining

2.1 Strategic Interactions and Economic Environment

There are three decision making agents in our model - a bride/wife (F), a groom/husband (M), and bride's parents. The bride's parents are assumed to be homogenous and therefore they act like a single agent (P).

Let us start by describing the labour and consumption choices of a married couple, who will be referred to as a 'household'. In a household, we assume that there is a division of work - M works full time in labour market and F alone is involved in household public good production, that is care work. Although this is a simplifying assumption, it is not far from the reality in a patriarchal society like India. Male labour force participation is almost universal, so is their absence from performing care work.[4] Additionally, F

[4] 14% of Indian men took part in unpaid care-giving services for household members

can also work part-time in labour market. Allocation of her total work hours, normalized to 1, between care work and labour market is denoted by h and $(1-h)$ respectively. Care work production function is $v(h)$. We assume that v is twice continuously differentiable, strictly increasing, strictly concave and $v'(h) \to \infty$ as $h \to 0$. Cost of care work, σh, is solely borne by the wife, where σ is her intrinsic characteristic. It captures F's relative inclination towards care work against paid work. Higher the σ; greater the cost of care work; stronger the inclination towards labour market participation. Value of σ is F's private information but it is common knowledge that σ is drawn from a Uniform distribution over $[0, 1]$. Although the husband does not participate in care work, he still has to bear a cost $\theta\sigma h$, arising from marital conflict. This negative externality stems from misalignment of wife's inclination towards paid work and her care work duties. For instance if $\sigma = 0$, then the wife is fully inclined towards care work and there is no marital conflict. As σ increases, stronger misalignment arises which leads to greater marital conflict and higher negative externality on the husband. We take θ to be exogenously fixed and $0 < \theta < 1$.

Labour market wage rate for women and men are denoted by w_F and w_M respectively. Irrespective of his inclination, the husband works full time in the labour market. Total household income, $w_F(1-h) + w_M$, is used for private good consumptions. Thus household budget is given by $c_F + c_M = w_F(1-h) + w_M$, where c_F and c_M are private good consumptions of F and M respectively, reached through negotiation. The married couple's negotiation over (c_F, c_M, h) will be described shortly. Aggregate consumption of F is $y_F = c_F + v(h)$ and that of M is $y_M = c_M + v(h)$. Utility of F is additively separable in domestic consumptions and cost, $u_F(c_F, c_M, h) = y_F - \sigma h$. On the other hand utility of M is $u_M(c_F, c_M, h) = (y_M - \theta\sigma h) - (y_M - \delta y_F)^2$. The first component of M's utility is his domestic consumptions and cost. The second is an inequality aversion factor, which decreases with the consumption gap between husband and wife. Therefore M has to balance between own consumption and that of his wife. For instance, if M was playing a dictator game then the inequality aversion factor would have stopped him from grabbing the entire pie. Note that wife's consumption is multiplied by a factor δ. This is M's intrinsic characteristic and represents the intensity of his patriarchal beliefs. We assume $\delta > 1$ because M believes 1 unit of y_M is comparable to more than one unit of y_F. Higher δ represents stronger patriarchal beliefs. To see this, consider two men with different δ values. Take an equal division of household consumptions (c, c, h). The higher δ male suffers from higher utility loss due to inequality aversion, while the first

[33].

component is the same for both. Thus higher δ male dislikes equal division more strongly than the lower δ male. Type δ is M's private information but it is common knowledge that δ is uniformly distributed over support $[\underline{\delta}, \bar{\delta}]$ and $\underline{\delta} > 1$.

To simplify, we assume that labour market wage w_F is the same for all women. However there is heterogeneity in w_M. The average male wage is denoted by $\bar{w_M}$. Further, w_M and δ are uncorrelated. Bride's parents, P, care about two factors in a potential match - status of the groom as reflected in w_M and welfare of their daughter after marriage, that is u_F. Relative weight attached to w_M is κ while relative weight on u_F is normalized to 1. Thus $u_P = u_F + \kappa w_M$. The parameter κ captures how closely bride's family follows patriarchal norms - higher κ implies stronger status concern; a patriarchal trait that puts family 'reputation' above daughter's welfare.

We now describe the game and its information structure. There are two periods in our model. Marriage matching takes place in the first period. The bride and her family, F and P, negotiate over groom selection at this stage. Once a matching is finalized, a married couple, F and M, take labour and consumption decisions in the second period. Payoffs are realized at the end of second period. Events of the first period are as follows. F receives a marriage proposal from a potential groom whose wage is drawn from the probability distribution of male wages. The wage is observed by both F and P. However, δ is not observable to P. As usual in a patriarchal society, parents have the primary control over the approval of a marriage proposal. But if it is beneficial, which it could be because δ is only observed by F, P can delegate the approval decision (accept or reject) to F. Delegation allows F to observe δ - for example through informal communications. If the proposal is accepted; either by P or by F following delegation; then the game moves to period 2. Otherwise if the proposal is rejected then a new proposal appears after a delay. This is the last match; hence is accepted under all circumstances. Subsequently, the game moves to period 2.[5] In the second period the matched pair M and F bargain on (c_F, c_M, h). Irrespective of how a match has occurred, δ becomes common knowledge at this stage but σ is still F's private information. The sequence of decisions are as follows. First the husband decides whether to delegate the choice of h to his wife. He can either choose a $h \in [0, 1]$ himself or let F choose it from a subset D of $[0, 1]$. If it is delegated then F chooses a h from D. Finally, after observing h, the husband chooses private good consumptions c_M and c_F. A natural but key feature of our set-up is that the husband cannot credibly commit to

[5]We can add an outside option of remaining unmatched, but it would not bring any additional insight.

contingent transfers before h is chosen.

Three patriarchal norms are in play in our model. First, within a marriage the power to make wife's labour market decision and household allocation lies with the husband. Similarly, parents have the primary control over mate selection of women. Second, discrimination within household is captured by husband's intrinsic characteristic δ. Third, bride's parents adherence to family pride is represented by κ. Autonomy of women appear through delegation of mate selection and labour market decisions. F has more autonomy if P delegates approval of marriage proposal to her. Similarly, between two labour choice sets D_1 and D_2, F has higher autonomy in D_1 when $D_2 \subseteq D_1$. In the following subsections, we show that an exogenous increase in autonomy in mate selection causes greater autonomy for women in the labour market.

2.2 Choice of Consumption and Work

We are looking for subgame perfect Nash equilibria of the game. This section contains the key arguments; formal proofs and derivations can be found in Appendix A. As usual, we use backward induction. The last decision in period 2 is made by M, who chooses private consumptions c_F and c_M after observing the extent of labour market participation of his wife, h. The choice of h may have revealed some information about wife's type σ. Let $r(\sigma)$ be the posterior belief of M on σ. M's optimization problem is to maximize $\int_\sigma u_M(c_F, c_M, h)r(\sigma)d\sigma$ subject to the household budget constraint $c_F + c_M = w_F(1-h) + w_M$. Optimal choice of c_F by M is given below.

$$c_F = \frac{1}{1+\delta}\big[w_M + w_F(1-h) + (1-\delta)v(h)\big] - \frac{1}{2(1+\delta)^2} \quad (1)$$

Note that optimal choice of c_F depends on h but is independent of posterior belief on σ. Therefore when F chooses h at an earlier stage, she simply chooses the best from available options. She has no incentive to strategically manipulate $r(\sigma)$ through her choice. After replacing Equation 1 in wife's utility function, her utility from choosing h becomes

$$u_F(h,\sigma) = \frac{1}{1+\delta}\big[w_M + w_F(1-h) + 2v(h) - (1+\delta)\sigma h\big] - \frac{1}{2(1+\delta)^2} \quad (2)$$

It is useful to look at the following special case as a benchmark. If M offers F an unrestricted choice of h (that is from the set $[0,1]$) then her optimal choice denoted by h^F should either satisfy the following or it is corner solution.

$$v'(h^F) = \frac{1}{2}(w_F + (1+\delta)\sigma) \quad (3)$$

Since $v' \to \infty$ as $h \to 0$, optimal choice for $\sigma = 1$ is an interior solution. On the other hand, we assume that $v'(1) > \frac{1}{2}w_F$, which means optimal choice for $\sigma = 0$ is corner solution $h^F = 1$ - even with unrestricted choice she prefers to do full-time care work rather than paid job.
Moreover, $u''_F(h) = \frac{2}{1+\delta}v''(h) < 0$; implying $u_F(h,\sigma)$ is single-peaked around h^F. Single-peakedness of u_F helps us to draw further predictions. Instead of the full choice set $[0,1]$, suppose a subset D is available to F. Then she chooses either her unrestricted best, if available in D, or one of the two options which are closest to h^F, if not. To mark the dependence of h^F on σ and choice set D more explicit, from now on, we shall use the notation $h^F(\sigma, D)$ to denote the optimal choice.
Since v is strictly concave, it follows from Equation 3 that $h^F(\sigma,[0,1])$ is a decreasing function of σ. This is intuitive - stronger the inclination towards paid job, smaller the choice of care work.

Proposition 1 *(i) For all σ, $u_F(h,\sigma)$ is a single-peaked function of h on the set $[0,1]$.*

(ii) Unrestricted optimal choice of F is given by: $v'(h^F(\sigma,[0,1])) = \frac{1}{2}(w_F + (1+\delta)\sigma)$ or it is a corner solution.

(iii) Range of unrestricted optimal choice is $[h^F(1,[0,1]),1]$.

(iv) $h^F(\sigma,[0,1])$ is a weakly decreasing function of σ.

Utility of M, when his wife has chosen h, can be obtained by replacing Equation 1 in $u_M(c_F, c_M, h)$.

$$u_M(h,\sigma) = \frac{\delta}{1+\delta}(w_M + w_F(1-h) + 2v(h)) - \theta\sigma h + \frac{1}{4(1+\delta)^2} \tag{4}$$

σ is his wife's type. Assuming that σ is known, let us find the optimal choice of $h \in [0,1]$ for M by maximizing Equation 4. The optimal choice is either a corner solution or it is given by the following first order condition:

$$v'(h^M(\sigma)) = \frac{1}{2}\left(w_F + \frac{\theta}{\delta}(1+\delta)\sigma\right) \tag{5}$$

Since σ is not known to M, for ex-ante optimal choice, M maximizes his expected utility. Thanks to linearity, ex-ante optimal h is the same as ex-post optimal choice of the average F type.

Proposition 2 *(i) For all σ, $u_M(h,\sigma)$ is a single-peaked function of h on $[0,1]$.*

(ii) *Optimal choice of* $h \in [0,1]$ *for* M*, facing* σ*, is either given by:* $v'(h^M(\sigma)) = \frac{1}{2}\left(w_F + \frac{\theta}{\delta}(1+\delta)\sigma\right)$ *or it is a corner solution.*

(iii) *Optimal ex-ante choice of* M *is denoted by* $h^M_{ex-ante}$*. Either* $v'(h^M_{ex-ante}) = \frac{1}{2}\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)$ *or* $h^M_{ex-ante} = 1$*. The latter holds when* $\frac{\theta(1+\delta)}{\delta}$ *is sufficiently small such that* $v'(1) > \frac{1}{2}\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)$*.*

Since v is concave, $\theta < 1$ and $\delta > 1$, comparing Equation 5 and Equation 3, we can conclude that $h^M(\sigma) \geq h^F(\sigma, [0,1])$ for all σ. This misalignment is central to the delegation problem that we study in the next subsection. If σ was known, M would have liked her wife to put in more hours of care work that she is willing to. Social norm gives the husband the right to set h but F has more information because she knows her type σ. As a consequence, M might be willing to delegate the labour market choice to his wife. Note that $\frac{\theta}{\delta}$ provides a measure of misalignment of interest between M and F.

Proposition 3 *(i)* *For all* σ*,* $h^M(\sigma) \geq h^F(\sigma, [0,1])$*.*

(ii) *Take* $h_1 > h_2$*. If* $u_F(h_1, \sigma) > u_F(h_2, \sigma)$ *for some* σ*, then* $u_M(h_1, \sigma) > u_M(h_2, \sigma)$*.*

The second part of the proposition shows that if certain type of F, given two choices of h, prefers the higher care work, them M, while facing that type, also prefers the same.

2.3 Delegation in labour choice

In this subsection we identify the equilibrium delegation set that a husband offers to his wife. When M offers choice set D to his wife, she would choose her restricted best within D, $h^F(\sigma, D)$, where her type is σ. Thus ex-ante utility of M, who does not know σ, from D is,

$$u_M(D, \delta) = \int_\sigma u_M(h^F(\sigma, D), \sigma))d\sigma \tag{6}$$

The equilibrium choice must maximize husband's ex-ante expected payoff among all possible subset of $[0,1]$. That is $D^*(\delta) = \arg\max_{D \subseteq [0,1]} u_M(D, \delta)$.

The following proposition provides a sufficient condition for delegation. If misalignment of interest between the husband and wife is not too severe, then M finds it beneficial to delegate. For the condition to hold $\frac{\theta(1+\delta)}{\delta}$ cannot be too small which implies that δ, the male patriarchy factor, cannot be too high.

Proposition 4 *If $h^M_{ex-ante} < 1$ then $D^*(\delta)$ cannot be singleton.*

$D^*(\delta)$ is singleton simply means that there is no delegation and M has picked his best ex-ante h; that is $D^*(\delta) = \{h^M_{ex-ante}\}$. If $h^M_{ex-ante} < 1$ then M can do better by offering the subset $\tilde{D} = [h^M_{ex-ante}, 1]$ instead of $\{h^M_{ex-ante}\}$. Those with unconstrained best $h^F(\sigma, [0,1]) < h^M_{ex-ante}$ continues to choose $h^M_{ex-ante}$ at $\tilde{D}$ due to single-peakedness of $u_F(h, \sigma)$ (see Proposition 1). The rest of σ would switch to their unconstrained best because those are now available in $\tilde{D}$. However such switch is beneficial for M because $h^M(\sigma) \geq h^F(\sigma, [0,1]) \geq h^M_{ex-ante}$ and $u^M(h, \sigma)$ in single-peaked in h (part (i) and (ii) of Proposition 2). For the rest of the section, we assume that the sufficient condition holds so that there is delegation in equilibrium.

The equilibrium delegation set must be a subset of the range of F's unrestricted optimal choice, the interval $[h^F(1, [0,1]), 1]$. If not then M can do weakly better by removing all options below $h^F(1, [0,1])$ and including $h^F(1, [0,1])$. If there was a σ choosing options below $h^F(1, [0,1])$, would now switch to $h^F(1, [0,1])$, because of single-peakedness of $u_F(h, \sigma)$. Such switch is also beneficial for the husband because from his perspective the peak is further to the right, $h^F(1, [0,1]) < h^F(\sigma, [0,1]) < h^M(\sigma)$.

Proposition 5 $D^*(\delta) \subseteq [h^F(1, [0,1]), 1]$.

We already know that $D^*(\delta)$ is not singleton - there are at least two choices in equilibrium. The next result shows that M can further gain from including more choices. If a delegation set contains two choices h_1 and h_3 but nothing in between then introduction of an additional intermediate choice h_2 is beneficial for M. The argument is as follows. After h_2 is introduced, some σ type switch from either h_1 or h_3 to the new option. Suppose $h_3 < h_2 < h_1$. From part (ii) of Proposition 3, we already know that a switch from h_3 to h_2 is also beneficial for M. On the other hand a switch from h_1 to h_2 has ambiguous effect but we show that the aggregate effect is always positive.

Proposition 6 *Suppose that $h_3, h_1 \in D$ such that $h_3 < h_1$ and the open interval (h_3, h_1) is not in D, that is $(h_3, h_1) \cap D = \emptyset$. Take any $h_2 \in (h_3, h_1)$. Let $\tilde{D} = D \cup \{h_2\}$. Then M prefers $\tilde{D}$ to D: for any δ, $u_M(\tilde{D}, \delta) > u_M(D, \delta)$.*

Now we can show that $D^*(\delta)$ is an interval. First note that $D^*(\delta)$ has to be closed. Otherwise if its an open set then there is at least one type σ who has no best choice in $D^*(\delta)$, which cannot happen in equilibrium. We also know from Proposition 4 that $D^*(\delta)$ has at least two elements. If $D^*(\delta)$ is not connected then there are two points z_1 and z_2 such that $z_1, z_2 \in D^*(\delta)$ but the open interval $(z_1, z_2) \cap D^*(\delta) = \emptyset$. By Proposition 6 this cannot be equilibrium as M can gain by adding another option between z_1 and z_2.

Proposition 7 *Equilibrium delegation set , $D^*(\delta)$ is an interval.*

As the husband finds it optimal to delegate job market decision to his wife, he needs to decide which decisions the wife should be allowed to make and which should be ruled out. It is intuitive that those choices for which misalignment between husband and wife are low are likely to be included in the delegation set. Our next result confirms this intuition.

Proposition 8 *The equilibrium delegation set is an interval which has upper bound* 1*.*

Therefore the optimal delegation set is of the form $[h, 1]$. Next, we derive the lower bound of the interval by maximizing $u_M([h, 1], \delta)$ with respect to h.

Proposition 9 *Lower bound of the equilibrium delegation set is the unconstrained optimum of $\sigma^* = \frac{1}{\frac{2\delta}{\theta}-1}$. That is, $D^*(\delta) = \left[h^F\left(\sigma^*, [0,1]\right), 1\right]$.*
$h^F\left(\sigma^*, [0,1]\right)$ acts as a measure of labour market autonomy of F; higher the $h^F\left(\sigma^*, [0,1]\right)$ lower the autonomy.

As patriarchal type of the husband, δ, increases; the misalignment of interest between the couple increases, which shrinks the delegation set. Formally, an increase in δ decreases σ^*, which in turn increases the lower bound $h^F(\sigma^*, [0, 1])$.

Proposition 10 *F's autonomy in job market is inversely related to M's patriarchy type.*

In equilibrium all type $\sigma > \sigma^*$, choose the least care work available in the delegation set $h^F\left(\sigma^*, [0,1]\right)$. This follows from single-peakedness of $u_F(h, \sigma)$. The rest choose their unrestricted best because these are included in $D^*(\delta)$.

$$h^F(\sigma, D^*(\delta)) = \begin{cases} h^F(\sigma^*, [0,1]) & \text{for } \sigma > \sigma^* \\ h^F(\sigma, [0,1]) & \text{for } \sigma \leq \sigma^* \end{cases}$$

Utility that F of type σ obtains from marrying M of type δ and wage w_M is denoted by $u_F(\sigma, \delta, w_M) = u_F(h^F(\sigma, D^*(\delta)), \sigma)$. Increase in patriarchy type of the husband, everything else remaining the same, adversely affect the wife through two channels. First, given a choice of h, it reduces transfer and hence $u_F(h, \sigma)$ (Equation 2); and second, it increases the lower bound of care work available in the delegation set.

Proposition 11 *If $v(1) > \frac{1}{4}$ then for all σ, w_M, $u_F(\sigma, \delta, w_M)$ is a decreasing function of δ.*

2.4 Autonomy in match-making

Suppose that F has received an offer from M, whose wage is w_M and type is δ. If the offer is rejected then the second offer has expected wage $\bar{w}_M$. Note that distribution of w_M and δ are independent. If parents were taking the approval decision then the present offer will be accepted if

$$\kappa w_M + E_\delta u_F(\sigma, \delta, w_M) \geq \kappa \bar{w}_M + E_\delta u_F(\sigma, \delta, \bar{w}_M) \tag{7}$$

Since parents do not observe δ, their decision is independent of δ. That is there is no screening along M's type; only screening is along wage.
On the other hand if the approval decision is delegated to F then the proposal is accepted if $u_F(\sigma, \delta, w_M) \geq E_\delta u_F(\sigma, \delta, \bar{w}_M)$.

$$u_F(\sigma, \delta, w_M) = \frac{w_M}{1+\delta} + \frac{1}{1+\delta}\big[w_F(1-h) + 2v(h) - (1+\delta)\sigma h\big] - \frac{1}{2(1+\delta)^2}$$

where $h = h^F(\sigma, D^*(\delta))$. Note that apart from the first term, the rest is independent of w_M. We rewrite this as, $u_F(\sigma, \delta, w_M) = \frac{w_M}{1+\delta} + R(\sigma, \delta)$.

$$\begin{aligned}
&u_F(\sigma, \delta, w_M) \geq E_\delta u_F(\sigma, \delta, \bar{w}_M) \\
&\quad\Rightarrow \quad \frac{w_M}{1+\delta} + R(\sigma, \delta) \geq \frac{1}{\bar{\delta} - \underline{\delta}} \int_\delta \frac{\bar{w}_M}{1+\delta} d\delta + E_\delta R(\sigma, \delta) \\
&\quad\Rightarrow \quad R(\sigma, \delta) \geq \frac{\bar{w}_M}{\bar{\delta} - \underline{\delta}} \left[\ln(1+\bar{\delta}) - \ln(1+\underline{\delta})\right] + E_\delta R(\sigma, \delta) - \frac{w_M}{1+\delta}
\end{aligned}$$

The right hand side is an increasing function of δ. The left hand side, $R(\sigma, \delta)$ is a decreasing function of δ when $v(1) > \frac{1}{4}$. The proof is exactly the same as Proposition 11 and hence is skipped here. Therefore F accepts the proposal if δ is below a cut-off, called $\hat{\delta}$. This means F screens a proposal based on both w_M and δ. Moreover, high patriarchy types (above the cut-off) are rejected. In case, a proposal is rejected F gets a fresh draw of δ.
Therefore expected labour market autonomy measure under self-choice is

$$\int_{\underline{\delta}}^{\hat{\delta}} h^F\left(\sigma^*, [0,1]\right) d\delta + \int_{\hat{\delta}}^{\bar{\delta}} \left[\int_{\underline{\delta}}^{\bar{\delta}} h^F\left(\sigma^*, [0,1]\right) d\delta\right] d\delta \tag{8}$$

In case of parent's choice, expected autonomy measure is $\int_{\underline{\delta}}^{\bar{\delta}} h^F\left(\sigma^*, [0,1]\right) d\delta$. This is greater than Equation 8 because $h^F\left(\sigma^*, [0,1]\right)$ is an increasing function of δ.

Proposition 12 *Irrespective of her type σ, labour market autonomy for F is higher under self-choice compared to parent's choice.*

Our next result shows that an exogenous decrease in κ weakly increase delegation in match-making.

Proposition 13 *Everything else remaining the same, if parents delegate a matching decision to F then they continue to do so when κ falls.*

Proposition 13 and 12 together imply that an exogenous decrease in κ increases labour market autonomy of F.

3 Data Description

We draw data from the India Human Development Survey (IHDS) collected by the National Council of Applied Economic Research (NCAER) in New Delhi and the University of Maryland. The IHDS is a nationally representative household panel survey conducted in 2004–05 (IHDS-I) [16] and 2011–12 (IHDS-II) [17]. The IHDS has detailed questions on socio-economic characteristics of the households.[6, 7] Data relevant for our analysis comes from the module called 'Eligible woman' of the IHDS-II. It contains detailed information about health, education, fertility, beliefs, family planning, marriage, and gender relations in the household and community, for ever-married women in the age group of 15-49 years in each household.[8] A total of 39,253 women were interviewed privately, of which we consider women who are currently in their first wedlock, and responded about their marriage pattern, 'say in work choice' and 'family relations', resulting in a sample of about 30,000 women.

3.1 Key Variables - Match-Making Types, Women's Autonomy in Labour Market, Labour Market Participation

The key explanatory variable of our interest is women's autonomy or the say they have in selecting their partners (i.e. loosely speaking, type of mar-

[6]IHDS-I consists of 41,554 households in 1,503 rural villages and 971 urban neighbourhoods across India, whereas IHDS-II surveyed 42,152 households in 1,420 villages and 1,042 urban neighbourhoods. About 85% households from first round were surveyed again in second round.

[7]The IHDS adopts a sampling procedure in the survey to ensure a nationally representative sample. The districts were selected using stratified random sampling to represent a range of socio-economic conditions. Villages and urban centres and households were selected using a cluster sampling technique.

[8]The women interviewed in IHDS-I and above the age of 49 years in IHDS-II were also re-interviewed along with a new ever-married woman between 15-49 years in the same household for IHDS II.

riage). In order to account for the different extent of women's involvement in marriages in India, we construct a categorical variable with three marriage types. The survey has two questions regarding the marriage type (Questions no: 18.4A and 18.4B). First, ever-married women were asked "Who chose your husband?". The responses are divided into 4 categories: 1. the respondent herself; 2. the respondent and parents/ other relatives together; 3. Parents/other relatives alone arranged marriages; 4. a miscellaneous category of "other", which refers to cases where extended family members or members outside the family played a role in the choice of spouse. Moreover, for those who had chosen option (3) or (4) in the previous question, there was a follow up question for the women "Did you have any say in choosing him?". The responses were either "yes" or "no". Based on responses to these two questions, we construct a variable 'Marriage-Type' as follows. If the respondent chooses option (1) in the first question, indicating that she had full say in choosing her husband, this is coded as the marriage type 'Self-match'. While the response is option (2) in the first question, this is coded as 'Jointly-arranged'. Moreover, if she chooses either option (3) or (4) in the first question, and for the second question response is 'yes', this is also coded as marriage type 'Jointly-arranged', as this also accounts for the woman having some say in choosing her husband. Lastly, if the woman chooses option (3) or (4) in the first question, and answers 'no' to the follow-up question, this is coded as 'Parent-arranged'. The last type of matching is unilaterally arranged by others and the woman had no say at all in choosing her husband. We do not differentiate other relatives from parents, mainly because the relatives are an integral part of the parents' social networks.

Our main outcome variable is female autonomy in labour market. We construct a binary indicator variable for this based on the question "Who has the most say in decisions about your work?" (Question no: 17.47); for which only one response was recorded out of five options - Self, Husband, Senior male, Senior female and others. We create a binary variable – 'work-say' which takes the value 1 when response to the question is Self, that is, woman has the most say and 0 otherwise. This question has 32,823 responses and about 46% of them confirmed that they have say, while 52% said that their husbands have the most say.

Estimating labour participation, especially in the Indian context, is not an easy task. IHDS-II has separate modules for three different types of work - household farm, non-farm businesses and wage labour. The survey asks which household members participated in each type of work during the previous year for how many days and hours. There are questions on the type of occupation/business, number of days worked in the preceding year, and hours worked in a day in each occupation. Using this, total hours (or total days) worked across all categories in the preceding year is computed. It is to be noted that in many cases, individuals are, either periodically or simultaneously, employed in multiple works. We construct two definitions of workforce participation. In our first definition of workforce participation, which is the IHDS adopted official definition, an individual is considered to be employed if hours worked in the preceding year were at least 240 hours. For the second definition, an individual is considered to be employed if he/she worked for at least 180 days in the preceding year. This second is comparable to the 'usual principal status' definition used by the National Sample Survey (NSS) of India. We shall use both the definitions to demonstrate the robustness of our results. Naturally, both of these definitions of workforce participation exclude the domestic care-giving work i.e. work done by women for themselves and for other household members. Moreover, Indian women often remain engaged into caring for household animals, collection of firewood or other fuels, and fetching water from public sources. These household chores (following Psacharopoulos and Tzannatos [37]) were not included in our measures of workforce participation.

3.2 Summary Statistics - Patterns of Marriage and Women's Involvement in Partner Choice

In this section, we highlight the relevant aspects of summary statistics. One typical operational feature of the Indian match-making process is that the search process is generally initiated by either the boy's or the girl's family, when they consider the person is marriageable. However, for a small number of cases, the process is initiated even by the persons themselves. We however here focus on the girl's side, as our IHDS-II data does not permit us to explore the boy's side. We present various features of three different marriage practices using some graphs and tables, which are available in Appendix B.

Table 1 tells us about marriage-type across regions, as well as across urban and rural areas. Overall 42% of sampled women are in parent-arranged marriage, for 53% marriage is jointly arranged and only 4% is self-matched. However there is a clear distinction between urban and rural areas. Urban areas have significantly fewer parents arranged marriages; 31% compared to 48% in the rural areas. But this does not result in more self-matched marriages in urban areas - the gap is filled up by jointly arranged marriages. There is also a huge variation across region - parents arranged marriage is as high as 70% in the central India and as few as 13% in the Northeast region. Table 1 also highlights that women's autonomy in labour market varies a lot across region, but not so much over rural-urban gradation - women from Northeastern region enjoys the most autonomy, whereas women from Western India have the lest autonomy, closely followed by Northern region.

From Table 2 we can see that women for whom parents made the partner choice are typically less educated, marry at a younger age and have less educated parents compared to the other two types. Moreover, marriages are assortative in education. Husbands of women in parents arranged marriage type are also less educated compared to their counterparts in other two marriage types. On average, husbands of women in self-matched and jointly arranged types have about 2 years of more schooling than than the ones who had parent-arranged marriage. There is evidence of assortative matching in economic terms, as well in caste. More than 70% women in each marriage category expressed that their natal family and husband's family had similar economic status at the time of marriage. Interestingly the assortative matching, measured as above, is stronger for jointly arranged and self-match group and the difference between these groups with parents arranged is statistically significant. Although it is not included in the present version of this paper, our theoretical model predicts a similar result (available upon request). Further, marriages arranged either fully or jointly by parents of girls also has a strong preference over same caste marriage, compared to self-match marriages.

Many authors have noted that over the years match-making has undergone a huge transformation (Banerji et al., 2013 [8]; Chawla [12]; Jejeebhoy et al., 2013 [29]; Pesando and Abufhele, 2019 [36]). In the Figure 1 we capture the transitions of three match-making process during 70 years over different birth cohorts of women, separately for all-India, rural and urban sectors. Most strikingly, the self-match marriages, so called 'love' marriages, show a very slow uptick over the years. It accounts for meagrely 10% for the youngest birth cohort both in Rural and Urban India. The fully parent-arranged marriages have been gradually declining over time from its dominant position. However, there have been a compensatory increase in the jointly arranged marriages, both in rural and urban India, and this is dominant mode at present. The change over happened much earlier in Urban areas than the Rural.

Figure 2 tells how the women's autonomy is affected by their marriage patterns. Both in rural and urban areas, we notice that the women who had self-searched marriages enjoy higher autonomy. Figure 3 tells us how the labour force participation of women is linked to their autonomy in labour choice. For either definition of work participation, women who enjoy autonomy exhibit higher participation both in rural and urban sectors.

4 Empirical Strategy

The basic hypothesis that we seek to test is that how the different levels of autonomy enjoyed by a woman, as a potential bride in choosing her husband at the time of marriage, later on causally affect her autonomy within a conjugal relation on her work choice. The empirical model is of the following form.

$$Pr\,(WorkSay_{isrc} = 1) = \\ F\,(\gamma\; MarType_{isrc} + \alpha\; X_{isrc} + \beta\; H_{isr} + \psi_s + \delta_r) \quad (9)$$

The dependent variable is the probability of married woman i of birth cohort c residing in development region r,[9] state s having her autonomy in work choice ($WorkSay_{isrc} = 1$). On the right side, F is the standard normal cumulative distribution function. $MarType_{isrc} = 1$ is our main explanatory variable of interest capturing autonomy in partner choice - parents arranged (1), jointly arranged (2), and self matched (3). Alternatively, in place of three categories we can estimate the following model with two indicator variables related to jointly arranged and self-matched marriage, while keeping parents arranged marriage as the reference category -

$$Pr\,(WorkSay_{isrc} = 1) = \\ F\,(\gamma_2\; JointArranged_{isrc} + \gamma_3\; SelfMatched_{isrc} + \alpha\; X_{isrc} + \beta\; H_{isr} + \psi_s + \delta_r) \quad (10)$$

Further, on the right hand side, X_{isrc} is a vector of woman's individual-level controls that includes her along with her husband's characteristics, such as her educational attainments, dummy for having kids aged five years or below, age difference between the couple, difference in years of schooling between the couple. It is to be noted that instead of controlling for her age directly, we control for birth cohorts of women (we explain the construction of birth cohorts in a subsection describing IV later). This takes care of effects of time-varying changes in marriage patterns for women. The vector H_{isr} is a set of household-level controls for the husband's family, and these include — household head's religion and caste, household's main source of income. We shall provide detailed accounts of our set of controls later. ψ_s is the geographic regional/state fixed effect, whereas δ_r is fixed effect for development region of woman's current residing location. These two fixed

[9]IHDS-II identifies any locality in terms of level of development, there are four categories – less developed remote rural village, more developed rural village, metro urban and other urban.

effects respectively control for any local, state or region wide unobservable characteristics and cultural patterns or state policies that affect everyone uniformly.

Our interest lies on the vector of coefficients γ in Equation 9 (or equivalently, γ_2 and γ_3 in Equation 10) i.e. we are interested in measuring the causal effects of higher involvement of women in their match-making through jointly arranged or self-matched marriages vis-à-vis parents arranged marriages on the outcome variable. Had the marriage match between individuals been undertaken randomly, then a regression (a simple probit model or the Linear Probability Model (LPM) using ordinary least square (OLS)) using our either specification would result in unbiased and consistent estimates. However, woman's involvement, conversely her parental involvement, in matchmaking is not a matter of random occurrence, rather women (or their parents) often endogenously select themselves into marriage type as illustrated in our theoretical model. We have already discussed some patterns that can be seen in our data. Marriages are often assortative in family characteristics (such as caste, religion, and economic status) and individual characteristics (such as education levels). Moreover it may also be determined by state or regional culture as well as urban/rural location. Additionally, given our set up of cross-sectional analysis, there will be concerns about 'omitted variable bias' since it is impossible to measure and control all the relevant individual and parental characteristics (such as assertiveness of a woman) that may affect the labour market autonomy and matchmaking process. Therefore, our main explanatory variable is potentially endogenous. However, we can expressly discard the chances of reverse causation, as our outcome variable, i.e. woman's autonomy at present can no way affect her marriage-type which had happened in the past.

We try to mitigate this bias by instrumental variable approach. But we also expand our set of controls further. We include several variables that might have affected woman's involvement in her marriage. These are: difference in economic status of the two families at the time of marriage (if her natal family's economic status was better, worse or similar compared to potential husband's family), whether woman and her spouse grew up in same area/ locality, whether the woman had to migrate from natal village/town for her marriage, if the couple had blood relations, if the woman comes from a different caste than that of her husband (i.e. had inter-caste marriage), woman's age at first marriage. We also include factors that might potentially affect outcome variable such as a dummy for woman's poor-health status, several indicator variables for her memberships in women's group, self-help group, credit society or political party; moreover, a set of indicators referring whether woman's and her husband's parents are alive, who might influence couples' decisions in various ways; and lastly, total income of all other family members (excluding her income).

4.1 Instrumental Variable

To solve the potential endogeneity issues and to isolate the causal effect of woman's involvement in marriage on our outcome variable, we employ the instrumental variable (IV) procedure. We borrow the idea of IVs from Huang et al. (2012) [25] and build upon it.

In a related paper, in the context of marital harmony and marriage patterns in China before 1980s, Huang et al. (2012) [25] used the tradition of parental involvement in a location specific marriage market as the instrumental variable for an individual's choice of parental involvement. Idea behind this is simple; parents are more likely to engage in spouse selection for their off-springs in areas where there exists a (dominant) social practice of parent-arranged marriages, either due to the convenience of social learning or because it is costly breaking the prevalent social norms or both. Alternatively, during the matchmaking process parents would provide more freedom to or seek consent from their daughters when there exists a tradition or past precedences of doing so in their neighbourhood. The tradition variable is independent of the individual characteristics and experience of the women and parents, and hence do not affect the post-marriage outcomes such as labour market autonomy at an individual level. Specifically, Huang et al. (2012)[25] measured tradition by the 'share' of prevalence of parental involvement in the earlier marriage cohort in the same area, where an area can be a specific district or province, and be combined with its urban or rural nature.

We adopt this strategy for constructing our IVs for two types of marriages – jointly arranged and self-matched marriages. However, adopting this method in Indian set up with the IHDS-II dataset has its own hurdles. First, unlike in China before 1980 (as in [25]), there have been considerable amounts of migration for Indian women owing to their marriages, both in rural and urban areas. Secondly, IHDS-II does not capture the information on natal residence of the woman (neither district nor state), but it has information on husband's family's current location and how far (in hours) is the natal place from husband's current house (where survey took place). From the data we have found that only 18% of women claimed to currently living in the same place of birth. About 42% of women live in a place (husband's house) which are 0-1 hour away from their natal place. For other 26% and 13%

women travel time are 2 and 3 hours to natal home, respectively.[10] Given these two aspects we can safely assume that most of the women are staying at least within the same state boundary as their natal home, whereas the same may not be true for districts. So, we construct our IVs as location-specific traditions at the 'state level' given the current location of husband's house, in the sense that every woman irrespective of their rural or urban natal residency faces the similar tradition across the state. Before we explain the method of construction of these IVs in detail, we want to highlight one fact that our approach imposes an obvious constraint that we can no longer use state level fixed effects in our model, therefore we will use fixed effects of geographic regions.[11]

We construct the traditions of different marriage type for every woman in a birth cohort i is measured by the prevalence of women's different levels of involvement, i.e. share of each type – parent-arranged, jointly-arranged and self-match, in an earlier cohort $(i + 2)$ in a state wide marriage market with state-urban or state-rural unit. We divide all women into 9 age cohorts in each of the 28 States and the 5 Union Territories marriage markets: the youngest cohort consists of women who were born after 1990, followed by cohorts where women born during 1986-90, 1981-85, 1976-80, 1971-75, 1966-70, 1961-65, 1955-60 and finally 1951-55. For instance, for the women in a particular state in cohort 'post-1990', the traditions of marriage types of joint-arranged marriages and self-matches will be measured by the shares of joint-arranged marriages and self-matches for all women born during 1981–85 in the same area, as our two IVs. Note that to avoid multi-collinearity we shall not use the share of parents-arranged marriage as another IV. Further, as there exist no corresponding measures of traditions for the two oldest cohorts in our sample, we will drop them from our regression analysis.

Following the literature, the validity of our instrumental variables approach

[10] To better understand migration patterns in India, we look at empirical evidence from earlier literature. Rosenzweig and Stark (1989) [40], Corno and Voena (2020) [13] note that in ICRISAT database, the average distance between a woman's current place of residence and her natal home is 30 km. They further note that in the NSS surveys of 1983 and 1987-88, only 6.1% of households are classified as "migrant households," defined as those for which the enumeration village differs from the respondent's last usual residence (Atkin, 2016[6]).

[11] We construct seven regions as follows: *Hills* – Jammu & Kashmir, Himachal Pradesh,Uttara Khand, Sikkim; *Northern* – Punjab, Chandigarh, Haryana, Delhi; *North East* – Arunachal Pradesh, Assam, Tripura, Meghalaya, Manipur, Mizoram, Nagaland; *Mid-Central* – Bihar, Madhya Pradesh, Chhattisgarh, Rajasthan, Uttar Pradesh; *Eastern* – West Bengal, Jharkhand, Odisha; *Western* – Gujarat, Daman & Diu, Dadra and Nagar Haveli, Maharashtra, Goa; *Southern* – Andhra Pradesh, Karnataka, Kerala, Tamil Nadu, Puducherry.

rests on three conditions. First, each IV must be meaningfully correlated with corresponding marriage type, even after including controls. Second, each IV must be uncorrelated with error term in equation (9 or, alternatively 10), also called 'orthogonality condition'. Third, they must be properly excluded from the model, so that their effect on the outcome variable is only indirect, via endogenous explanatory variable, known as the 'exclusion restriction'. We shall provide support for the first condition using first stage regressions, associations of each IV with different marriage types and first stage Cragg-Donald F-statistics across samples. And the second condition can be satisfied if our IVs are sufficiently random. Given the diversity in women's/parents' involvement in matchmaking, as noticed in Table 1 and Table 2, across the 33 state/UT-wide marriage markets and the 9 birth-cohort, our two IVs on traditions of jointly arranged and self-matched marriages actually vary a lot, almost randomly across areas and cohorts. Therefore, with our two IVs, we can achieve exact identification for 3-category endogenous variable Marriage-Type in Equation 9 (or equivalently, two dummies for marriage type in Equation 10). Further these tradition variables are independent of the individual characteristics of the women and parents, and hence deal with the endogenous selection into marriage type only and do not affect the post marriage outcomes such as her labour market autonomy, thereby this implies that the exclusion restriction is likely to hold.

Empirical Estimation Method:

As stated above, our IV model (as in Equation 9) involves a dichotomous depended variable, $WorkSay$, and a trichotomous endogenous explanatory variable, $MarType$. Therefore both of our first- and second- stage models are essentially inherently non-linear. Breaking the trichotomous explanatory variable in to two indicator variables (as in Equation 10), still retains the non-linear structure of the first stage model. Despite the non-linear structures, many authors still use LPM to estimate these kinds of models. However, (a) LPM does not estimate the structural parameters of a non-linear model (Horace and Oaxaca, 2006 [23]); (b) the LPM does not always produce consistent estimates of the marginal effects.[12]

Given the criticisms of LPM estimations, we use a special framework, CMP (Conditional Mixed Process).[13, 14] This allows to us to use the inherent non-

[12] See discussions by Giles, `http://davegiles.blogspot.co.uk/2012/06/yet-another-reason-for-avoiding-linear.html` and and Hausman et al., 1998[24]

[13] Many other studies in the similar context (e.g. Huang et al., 2012) often use the linear probability models (LPM). Because LPM requires less distributional assumptions and are often preferred than non-linear models when the main interest is to estimate the marginal effects of the explanatory factors (Angrist and Pischke, 2009)[3].

[14] But, for a non-linear setup like ours, where the dependent variable is binary and

linear structure of our model – a multinomial probit in the first stage for our trichotomous marriage-type variable and a probit in the second stage for dichotomous outcome variable. Additionally, it allows us to estimate average marginal effects (AME) more precisely for our non-linear model (See, Wooldridge (2010) [43] pp.599 for relevant discussion). We also use LPM for IV model corresponding to Equation 10, as this procedure allows us to produce conventional diagnostic tests for IV models, such as Cragg-Donald first-stage F test, Hausmann test for endogeneity, Anderson-Rubin t or chi2 test statistics. However, CMP procedure (corresponding to Eqn.9) has its own diagnostics test – significance of 'atanhrhoEx',[15] (which is provided in the bottom of Table 3). It is to be noted upfront that estimation results from LPM are consistent with the CMP ones presented in the next section (in Table 3). Additionally, we run a benchmark probit model and an IV model for all-India, as well as – rural and urban areas separately.

5 Estimation Results

In this section we shall discuss the causal effects of a woman's autonomy on marriage decision on her autonomy related to work choice related. First we shall present our findings, then discuss the robustness of our results.

5.1 Probit and IV Estimates - all-India, Rural and Urban Sectors

The estimation results of the models are presented in the Table 3. We present separate regression estimates for the All India, as well as rural and urban areas separately. Before we explain the results, let us describe the organisation of the table and the estimation methodology. Note that a similar Table 4 has been constructed using LPM.

On the top panel of Table 3, we have three columns for each sector presenting the results from the two sets of probit estimates and the estimates from

the suspected endogenous explanatory variable is trichotomous, the conventional STATA command "ivreg" is not suitable, moreover "biprobit" cannot handle our trichotomous endogenous variable. Therefore, we resort to 'cmp' command, developed by Roodman (2011)[39] following the idea of Seemingly Unrelated regression models using a recursive structure. The model is estimated by Maximum Likelihood Estimator (MLE).

[15] This test is based on the Wald test of correction of the error terms of two structural equations. High and significant Wald test statistic suggests the rejection of null hypothesis and that rejection means strong correlation between errors from first and second stages of regressions, which means that in such a situation two structural equations should be estimated together, further this can be interpreted as there exist endogeneity.

IV-models. The coefficients here represent average marginal effects (AME), showing the change in the probability of woman having the autonomy associated with a unit change in the main explanatory variable i.e. woman's involvement in marriage. For this three categorical variable, this is the difference with the reference category of parents arranged marriage. Figures in the parenthesis, across all panels, represent clustered robust standard errors at the sampling unit level. Two probit models in each sector differ from each other in terms of use of fixed effects and dropping sample birth cohorts. For probit-1 models (i.e. in columns (1), (4) and (7)), we include state fixed effect dummies (excluding Tamil Nadu, in the base) with full sample. However, for probit-2 models (i.e. in columns (2), (5) and (8)) we instead use region fixed effects (excluding southern region, in the base) and drop samples of last two birth cohorts, this done to achieve similar modelling strategy as in IV-models and to maintain the same sample across all regressions. As we move from column (1) to (2), we drop the last two birth cohorts of our sample women, majority of whom had arranged marriage by parents (during 1950s), the coefficients of jointly arranged and self-matched marriages appear more pronounced (vis-à-vis parents arranged marriage). Similar fact holds for column-pairs (4)-(5) and (7)-(8).

On the bottom panel, we present the associations of our two IVs for traditions (e.g. shares of jointly arranged marriage and self-matched marriages over different birth cohorts in rural or urban residencies of every state) with the two dummies of trichotomous marriage type variable from the first stage reduced forms (corresponding to our IV model for each sector, columns 3, 6 and 9). Note that coefficients presented for this first stage part are coming from multinomial probit estimations and are not AMEs. We find strong positive and significant correlations between our two IVs of traditions and the indicators of two marriage types (i.e. jointly arranged and self-matched). A similar and consistent result can be seen for the LPM models as well (Table 4) and let us interpret the coefficients from first stage estimates from this table. An increase of 1% in the tradition of 'self-match' type marriages 'a decade ago' tend to increase the probabilities of the woman having full say in partner choice by 0.596% and of woman's consented marriage by 0.214% at the pan India level. Similarly, an increase of 1% in the tradition of consented marriages a decade ago increases the probability that the woman will have some say (i.e. her consent will be asked by parents) in choosing her spouse along with her parents by 0.733%, however increases probability of woman's full say (i.e. self-match) only by 0.003% at the all India levels. However, in the subsectors of rural and urban areas, we see these effects vary in terms of magnitudes, but overall directions remain the same. Thus, in the first stage regressions, the instrumental variable is highly significant, and the F-statistics are reasonably high (i.e. 'arf' values are 33 for all-India, 28 for rural and 8 for urban in Table 4), which minimizes the concern of weak instruments. Moreover, from Table 3 for our non-linear models, we find that "atanhrhoEx" are also negative and significant for all India, rural as well as urban sectors, implying that our main explanatory variable is endogenous, and instruments are valid.

Let us now focus on the main results here (upper panel of Table 3). The impact of different level of woman's involvements in her partner selection on her post marriage autonomy in labour market choice are presented as average marginal effects (AME). All of these are positive and statistically significant, across the sectors and at all India level. Further, as expected from our theoretical predictions in earlier section, the magnitudes of AMEs of two types of woman's involvements (specifically of jointly arranged marriages and self-matched ones) vis-à-vis parents arranged marriages significantly differ and vary widely across all probit estimations for different sectors, and their scales in the IV estimations are also quite large, more pronounced, and significantly different. As a matter of fact, what is evident, even from probit results across sectors (see, columns 2, 4 and 6), the higher the autonomy a woman enjoys during her marriage, higher is the probability that she enjoys autonomy for her work decision. Needless to say, that a woman whose parents unilaterally fixed her marriage get the worst deal in terms of her work choice. Our IV results (columns 3, 6 and 9) not only support that, but also establish the causal effects and provide more precise estimates of the effects.

As per our IV estimates, at the all-India level the women who had jointly arranged marriage on average enjoys 22.6 percentage point (p.p., hereafter) higher probability of having autonomy in her labour market choice when compared to parents-arranged marriages. The corresponding figures in rural and urban areas are 26.1 and 10 p.p., respectively. Moreover, the women, who had chosen their spouse entirely by herself (i.e. self-matched), enjoy on average 30.1 p.p. higher probability of autonomy in her work choice; the corresponding figures in rural and urban areas are 34.2 and 15.3 p.p., respectively. Average marginal effects of IV estimates for each of two marriage types for all-India and rural area (columns 3 and 6) are in fact roughly four times higher that of corresponding probit estimates; and all-India results are mainly driven by rural areas. This seems to suggest that the probit estimates will underestimate the negative effects of parents involvement in marriage. We would also like to mention that similar models were run using LPM, results presented in Table 4, and we obtain consistently similar results as above, across all sectors. While the average partial effects of each OLS model there closely match with the AMEs of probit models in Table 3 above, but the IV estimates from LPM are way higher (especially for self-match marriage type) than that of Table 3, which obviously is a reason to rely more on our non-linear models here. However, on the whole, both types of estimation processes quite consistently conform and tell the similar story across sectors.

One additional important observation deserves our attention: we have included work status of woman in the past one year across the regressions above and it shows that the working status of women across the sectors consistently likely to increase their autonomy in labour market choice, and the effect is stronger in urban areas (although, from Table 2 panel B, it can be noticed that women's labour force participation rate is much lower in urban sector than rural as per both definitions of FLFP). We find support to this observation in the existing literature (e.g. Anderson and Eswaran (2009) [1] Fletcher et al. (2017) [20]) which states that access to labour market, particularly wage work, brings greater autonomy for women. It tells us that autonomy and labour market participation affect each other. We shall explore this in detail in next section.

5.2 Robustness Check

To check the robustness of our main results for all-India and rural/urban sectors presented in Table 3 in the previous subsection, we have further conducted various tests. We have adopted two-pronged strategy. In one case, we have tried an alternative way of constructing IVs of traditions of women's involvement in marriages. In other case, we reconstruct our trichotomous marriage type variable into a binary one (i.e. no involvement or some involvement in spouse selection), further we use an entirely new set of IVs for this. All other controls remain the same as before. We unambiguously confirm that our main results are robust to these changes.

Strategy-I:

As there is no one single definite way to measure the tradition of women's (or conversely, parental) involvement, the specific measure we used in our earlier estimation may appear a little arbitrary. To check whether our main results presented in Table 3 are affected by the choice or construction of IVs, we construct an alternative measure of the tradition as the new set of IVs, which are possibly more arbitrary and much less nuanced than the ones used before.

We keep the same birth cohorts as before, but now define the marriage markets at the state level but now include the rural/urban residency identity of the current location of husband's house. By including rural/urban residency identity, we are bringing in a further difference in the behaviour or culture of urban residents from rural habitants, even within a state. This rural-urban cultural difference could be, rather in fact are, significant in larger and predominantly agrarian states (e.g. Uttar Pradesh, Rajasthan). We could change the cohort distribution as well to bring more arbitrariness in the construction of traditions, but we do not.

The estimation results, using these new IVs, both from non-linear model (NLM) and LPM, are shown in Table 5, where the set of control variables are the same as in our main results in Table 3. Comparing the bottom panels of the first stage results across areas with that in Table 3, there is very little difference between the two especially for all-India and rural sector, not surprising though. However, these new IVs has slightly lower associations and thereby similarly lower explanation power than the old ones, particularly for the urban sector. Thereby, the results for the urban sector weakens further, we lose significance of 'atanhrhoEx' for our non-linear model for urban, and for LPM we notice that 'arf' is quite low. Albeit, the regression results, refer to the top panel now, are qualitatively the same and quantitatively similar as their counterparts in Table 3. It is also useful to note that our main results can be obtained by using a much more arbitrary measure of tradition as IV.

Strategy-II:

Following the main tenet of our causal result that the women whose parents unilaterally fixed their marriages get the worst deal in terms of her work choice, we reconstruct our trichotomous marriage type variable into a dichotomous one by clubbing the jointly arranged and self-match types into one category – so that it takes the value 0 when woman is not involved at all (parent-arranged marriages), and 1 when woman has some or full involvement. In this case, to come up with an exactly identified IV-model we need just one IV. However, we shall use two use instruments – 'number of sisters' and 'number of brothers' a woman had. Alternatively, clubbing them together, we could use the 'total number of siblings' as the only IV for our new binary endogenous marriage type variable.

The justification for these IVs comes from the cultural norms surrounding the order and timing of girls' marriage in Indian set up – when a potential bride has more siblings, especially more (younger) sisters, then often her parents are tempted to marry her off as quickly as possible and without even caring for her consent. The models of family resource dilution contend that family resources are finite and, thus, the sibship configurations—which includes not only the sibship size, but also their ordinal position of birth and sibship sex-composition—may shape the amount of family resources accessible for each child (Downey, 2001 [18]; Steelman et al., 2002 [42]). So long as family resources are relevant to parental and individual decisions about marriage—which is true in the presence of dowry practices in India—the resource dilution model should have implications for the associations between sibship size and/or birth order and the transition to marriage (Yu et al., 2012 [44]).[16] Unfortunately, we do not have information about her birth sequence among the siblings, else we could have used that instead or as well. In absence of that information, each of our proposed instruments crudely captures the scenario that given the limited economic resources within her natal household and more siblings, coupled with the prevalence of 'dowry system' that increases with bride's age, marrying off a daughter early, and without consulting her, is an attractive proposition. Thus, each of these proposed instruments is expected to meet all the three requirements of implementation of IV procedure – each of them is to be negatively correlated with our two-category marriage type; random enough across sampled women; and potentially not expected to affect our outcome variable. We go for an over-identified model using two instruments – numbers of sisters and brothers. We use the same set of controls and other fixed effects as before.

The estimation results for this model are provided in the Table 6. We provide estimates from both the non-linear model (NLM) and the LPM there, for all the sectors. At all-India level, according to IV-NLM model, we find that when women had the opportunity to engage into spouse selection process (i.e. either she chose her spouse or parents sought her consent), she enjoys on average 19 p.p. higher probability of possessing autonomy in work choice compared to the ones whose parents unilaterally arranged their marriages. However, for rural sector, the effect is even stronger, the probability of having autonomy in work choice stands at 30 p.p., but at the urban sector we do not see any significant impact. This is quite consistent with our main result.

6 Does Autonomy Cause Higher Participation

In this section we focus on our second hypothesis – post marriage autonomy in labour market causes higher participation. Female labour supply is an important outcome for determining gender justice and is hence deemed as one of the key markers for women empowerment. The emancipation of women through larger labour force participation is well documented in the literature (Anderson and Eswaran, 2009 [1], Atkin, 2009[5] Luke and Munshi, 2011

[16]See related discussions for India and other developing countries in Pesando and Abufhele (2019)[36].

[31]). We turn this question around and ask whether post marriage female autonomy in labour market actually causes greater participation in the labour force. The existing empirical evidence of this hypothesis is rather limited (see, Biljon et al., 2018 [10] for some evidence).

Female autonomy in labour market choice and actual labour force participation may be affected by the inter-linkages of these two aspects. We in fact argue that the relationship between woman's autonomy in work choice and female labour force participation is bidirectional and thereby endogenous. Autonomy allows a woman to participate in labour market, if she prefers so; on the other hand labour market participation makes her financially independent and more autonomous. Our summary statistics bring forth mixed evidence. For instance (see panel-A of Table 2),the women in the Northeast India, majority of whom, enjoy the autonomy in work choice, as a percentage of total sample which is the highest among all the geographic regions, their labour force participation (according to either definition) is among the lowest among all regions. On the other hand, in Western and Northern India women are comparable in terms of autonomy (which is lowest in entire India), but their participation rates differ a lot – it is relatively high in Western India but quite low in the North.

6.1 Empirical Strategy and Estimation Method

To tackle this issue of reverse-causality empirically, we resort to a system of two simultaneous equations (structural modelling) in which two outcome variables, viz. autonomy in labour market and labour market participation, are endogenous in the sense that they influence each other. We model the causal effect of women's autonomy related to their work choice on their likelihood to participate in the labour market, and the reverse effect, in a rural Indian set up.

Female labour participation is particularly sensitive to macroeconomic shocks. It is especially true in rural markets and the occurrence of drought is one such shock that significantly affect female LFP directly, but also indirectly through a strong interdependence with the employment status of other household members, particularly husband's. Therefore, the premise of our structural model rests on the assumptions that – on one side, women's greater involvement during their marriage determines their autonomy in labour market, which we already know; on the other, an external shock to labour market through spells of drought in rural areas can drastically affect employment opportunities for women. The dependence of each outcome variable on the other, as well as on other explanatory variables, is described with the following structural equations

$$\begin{aligned}
&Pr\left(WorkSay_{irs}\right) = \\
&G_1\left(\lambda_0\, Worked_{irs} + \lambda_1\, MarType_{irs} + \psi_s^1 + \delta_r^1 + \beta^1\, X_{irs}^1\right) && (11)\\
&Pr\left(Worked_{irs}\right) = \\
&G_2\left(\pi_0\, WorkSay_{irs} + \pi_1\, DSD_{rs}^{2011} + \pi_2\, DSD_{rs}^{2010} + \psi_s^2 + \delta_r^2 + \beta^2\, X_{irs}^2\right) && (12)
\end{aligned}$$

Here $WorkSay_{irs}$ and $Worked_{irs}$ are two interdependent, binary and observed outcome variables for woman i residing in development region r of the state s. G_1 and G_2 are the bivariate standard normal cumulative distribution functions. Further, as before, where ψ_s is the geographic regional/state fixed effect, whereas δ_r is fixed effect for development region of woman's current residing location. These two fixed effects respectively control for any local, state or region wide unobservable characteristics and cultural patterns or state policies and level of development that affect everyone uniformly. X_{isrc} is a vector of woman's individual-level controls that includes her, husband's, and household's characteristics. We shall provide the description of full set of controls in a while. Note that the superscripts for ψ_s, δ_r and X_{isrc} are equation specific.

We are particularly interested in the coefficients λ_0 and π_0, i.e. we want to estimate the effect of woman's work participation (in a span of year preceding the survey, using either definition of work status, discussed earlier) on her autonomy in labour market, and the reverse effect. Note that given the binary nature of our two endogenous outcome variables, nonlinearity is at work here. Multi-equation limited dependent variable models that are logically consistent often require fewer assumptions for formal 'identification' than classical linear ones. For the classical systems to be identified, a rank condition must be met. A common rule is the order condition: in each equation, at least one predetermined variable must be excluded for every endogenous dependent variable that is included (Greene, 2008, pp.449 [22]). Thus, we need at least one exogenous variable for each equation (and validly excluded from the other equation) that allows us to identify λ_0 and π_0. This is achieved by using – woman's marriage-type for post marriage labour market autonomy, and district level share of drought affected villages in 2010 and 2011, called DSD^{2010} and DSD^{2011} respectively, for labour force participation in a year preceding the survey.

The construction of these two drought variables, DSD^{2010} and DSD^{2011}, deserves some attention here. The IHDS-II collects information on occurrences of droughts (not severity) in almost all surveyed villages during the survey year and for several past years (calender year), and we see in some cases certain villages were affected by consecutive spells of droughts. We could have directly used this information and created village specific dummies for occurrences of droughts over some years, but that would not serve our purpose. We are of the view that effects/shocks of droughts are not necessarily localised and remain confined to the boundaries of villages, rather more wide spread (maybe at a district or a region wide or even bigger). Thus we expect the effect of drought on the labour market may not be limited to the village itself, rather the effect can be seen in many contiguous villages within district, so instead of using the drought affected status of each village, we construct these two variables as the share of drought affected villages of the total surveyed villages in a district for the two calender years 2011 and 2010. Note that these two years immediately precede the survey year (i.e. 2011-12) and the labour force participation was enumerated for a year just before the survey. So the more villages affected in a district by drought, the more acute the shock from droughts on labour market, and as a consequence people have to travel far and wide to search for a job.

For the estimation of our model, we again use the Seemingly Unrelated Regression (SUR) framework of CMP. However, we need to address two concerns. First, given our choice of instruments, in a recent paper by Corno and Voena (2020) [13] related to marriage patterns in India pointed out that occurrences and severity of droughts can significantly reduce the chances of child marriage. In other words, in our set up, droughts can alter the marriage pattern. Thereby, the instruments in our two structural equations become interlinked, especially for recent marriages during the years 2010-12. To avoid such a possibility when droughts are not validly excluded from first equation, we drop the samples of woman who got married after December 2009. Secondly, in our sample we have quite a few elderly women, as well as some younger ones below 18 years of age. However for the estimation of female labour force participation the convention is to restrict the analysis within working age population, therefore we only use the sample for the age group between 19-60 years, note that by the age of 19 above 95% women are married in rural India. Further, some more observations are dropped from those villages in certain districts where drought information for the concerned years is not available.

6.2 Estimation Results

We are not ready to present our estimation results for our simultaneous equation model for two endogenous dependent variables - women's autonomy and their labour force participation. In Table 7, we present the estimation results for two different models, corresponding to two different definitions of labour force participations (defined earlier). The coefficients here represent the system-wise average marginal effects for each variable. Each model has two equations - first one is for $SayWork$ i.e. women's autonomy, and the second one for work status. The dotted line in the middle segregates the equation specific IVs from the endogenous variables at the top.

Before we interpret the results, we must mention the set of controls used in the models for each equation. We control for state level fixed effects and development-region fixed effect for each model. We equation specific controls that are not used in other equation. Specific controls for equation 1 (Worked) include: Dummy for woman's migration from childhood residence, number of children under 5, number of female children aged 6-14, number of male children aged 6-14, household size, dummy for existence of debt, dummy for major expenditure-incurring incident in last 5 years, number of household members with major morbidity. Specific controls for equation 2 (WorkSay) include: Dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, marriage age, square of marriage age.

Besides the equation specific controls, we have a set of common controls for both the equations in each model include: dummies for religion and caste. The household specific controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the time of marriage, log of per capita income of other household members (except woman). The women specific controls are – dummy if her husband is away for work, a seven category variable for woman's education, woman's health poor dummy, set of dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, age of woman, absolute differences of age and years of schooling of woman with her husband, dummy for inter-caste marriage, dummy for pregnancy, her parents' education, husband's parents' education, dummy for practice of purdah in the household, dummies for English speaking ability, set of dummies if woman's parents and in-laws are alive, dummies for presence of elderly in the household ($>$ 65yr), dummies for husband's primary activity status.

Let us look at the results now. First, we focus on the lower part, there we see the significance of 'atanhrhoEx' for each model, meaning strong and significant correlation between the errors from two equations. Which further means that our two structural equations are linked and should be estimated together, i.e. two outcome variables are endogenous and significantly influencing each other. We also notice that each of our equation specific IVs is working as expected i.e. significantly correlated with the corresponding endogenous variable. Next, we focus the two of our endogenous outcomes - across the models each is positively and significantly affecting other; and the effect is slightly more pronounced in model-2 where we used the second definition of work status (240 hrs /yr). Therefore, when a woman possess the autonomy in labour market in the rural India, she is having on average 9.1 (resp. 10.6) percentage point higher probability of working vis-a-vis someone who does not have the autonomy, depending upon the definition of work we choose. While on the other side, when a woman did work in the past one year, she has 9.9 (resp. 10.6) percentage point higher probability of possessing autonomy in labour market compared someone who did not work.

7 Conclusion

Our paper identifies marriage practices, in particular, parental involvement in partner choice as a major cause behind the low labour force participation of women. The literature has missed it because these two events are not proximate enough to raise concern. Our theoretical model shows that parental involvement leads to mismatches. In a patriarchal society, women who are keen to work should find the right partner; a partner who would not be averse to her labour market participation. Parental involvement reduces women's ability to screen their partner and lead to costly mismatches for work-seeking women. It reduces their autonomy and hence participation. Empirically we do not observe whether a woman is a work-seeking type or not. However, we find that women who have some involvement in marriage are more likely to have the most say in their post-marriage labour market choice. The effect is strong in rural areas. By estimating simultaneous equations, we further show that gain in autonomy would lead to gain in participation. Thus marriage practices come out to be an important bottleneck for women's empowerment. Reform in cultural practices takes time and effort. But the first step is to recognize a problem. Women who choose their partner, particularly if the partner is from a different caste and religion, face many obstacles, harassment, threats from parents/relatives[26], from fascist vigilante groups [27] and increasingly from the state [28]. Unless this trend is broken, India may see further regression in marriage practices and a consequential decline in female labour market participation.

References

[1] Anderson, S. and M. Eswaran, 2009. What Determines Female Autonomy? Evidence from Bangladesh. Journal of Development Economics, 90, 179-191.

[2] Anderson, S. and Genicot, G., 2015. Suicide and property rights in India. Journal of Development Economics, 114, pp.64-78.

[3] Angrist, J. D. and Pischke, J.-S., 2009. Mostly Harmless Econometrics: an Empiricist's Companion. Princeton NJ: Princeton University Press.

[4] Anukriti, S. and S. Dasgupta, 2017. Marriage Markets in Developing Countries, in The Oxford Handbook of Women and the Economy, Edited by S. L. Averett, L. M. Argys, and S. D. Hoffman, Oxford University Press, UK.

[5] Atkin, D., 2009. Working for the Future: Female Factory Work and Child Health in Mexico. Technical report, Department of Economics, Princeton University.

[6] Atkin, D., 2016. The caloric costs of culture: Evidence from Indian migrants. American Economic Review, 106(4), 1144-81.

[7] Banerjee, A., E. Duflo, M. Ghatak and J. Lafortune, 2013. Marry for What? Caste and Mate Selection in Modern India. American Economic Journal: Microeconomics, 5, 33-72.

[8] Banerji, M., Martin, S. and Desai, S., 2013. Are the young and educated more likely to have 'love' than arranged marriage? A study of autonomy of partner choice in India. Working Paper Series (pp. 1Ā43). New Delhi: NCAER.

[9] Becker, G. S., 1981. A Treatise on the Family, Cambridge MA: Harvard University Press.

[10] Biljon, C., Fintel, D., and A. Pasha, 2018. Bargaining to work: the effect of female autonomy on female labour supply. Stellenbosch Working Paper Series No. WP04/2018. Accessed in January 2022 at https://resep.sun.ac.za/wp-content/uploads/2018/07/wp042018.pdf

[11] Chakravarti, U., 2002. Gendering Caste through a Feminist Lens, Stree, India.

[12] Chawla, A, 2020. Parental involvement in spouse choice and marriage outcomes: Evidence from India, Unpublished.

[13] Corno, L., Hildebrandt, N., and A. Voena, 2020. Age of marriage, weather shocks, and the direction of marriage payments. Econometrica, 88(3), 879-915.

[14] Deloitte, 2018. Opportunity or Challenge? Empowering Women and Girls in India for the Fourth Industrial Revolution, India.

[15] Desai, S. and L. Andrist, 2010. Gender Scripts and Age at Marriage in India. Demography, 47, 667-687.

[16] Desai, Sonalde, Vanneman, Reeve, and National Council of Applied Economic Research, New Delhi. India Human Development Survey (IHDS), 2005. Inter-university Consortium for Political and Social Research [distributor], 2018-08-08. https://doi.org/10.3886/ICPSR22626.v12

[17] Desai, Sonalde, Reeve Vanneman and National Council of Applied Economic Research. India Human Development Survey-II (IHDS-II), 2011-12. Inter-university Consortium for Political and Social Research [distributor], 2018-08-08. https://doi.org/10.3886/ICPSR36151.v6

[18] Downey, D. B., 2001. Number of siblings and intellectual development: The resource dilution explanation. American Psychologist 56(6/7): 497–504.

[19] Fernandez, R., A. Fogli, and C. Olivetti, 2004. Mothers and Sons: Preference Formation and Female Labor Force Dynamics, Quarterly Journal of Economics, 119, 1249–99.

[20] Fletcher, K. E., R. Pande and C.T. Moore, 2017. Women and Work in India: Descriptive Evidence and a Review of Potential Policies, HKS Working Paper No. RWP18-004, Harvard Kenedy School.

[21] Goldin, C., 1995. The U-shaped Female Labor Force Function in Economic Development and Economic History, in Investment in Women's Human Capital and Economic Development, Edited by In T.P. Schultz, University of Chicago Press, USA.

[22] Green, W., 2008. Econometric Analysis. Pearson.

[23] Horrace, W.C. and Oaxaca, R.L., 2006. Results on the bias and inconsistency of ordinary least squares for the linear probability model. Economics Letters, 90(3), pp.321-327.

[24] Hausman, J. A., J. Abrevaya, and F. M. Scott-Morton, 1998. Misclassification of the dependent variable in a discrete-response setting. Journal of Econometrics, 87, 239-269.

[25] Huang, F., G. Jin and L.C. Xu, 2012. Love and Money by Parental Match-Making: Evidence from Urban Couples in China, American Economic Review, 106, 555-560.

[26] Indian Express, 2020. https://indianexpress.com/article/india/tamil-nadu-dalit-killing-acquittal-life-term-6471698/, India.

[27] Indian Express, 2021. https://indianexpress.com/article/india/up-interfaith-marriage-stopped-woman-forced-to-leave-court-7429281/, India.

[28] Indian Express, 2021. https://indianexpress.com/article/cities/bangalore/karnataka-anti-coversion-bill-right-wing-love-jihad-7684737/, India

[29] Jejeebhoy, S. J., Santhya, K. G., Acharya, R., and R. Prakash, 2013. Marriage-related decision-making and young women's marital relations and agency: Evidence from India. Asian Population Studies, 9(1), 28-49.

[30] Klasen, S., 2019. What Explains Uneven Female Labor Force Participation Levels and Trends in Developing Countries?, The World Bank Research Observer, 34, 161–97.

[31] Luke, N., and K. Munshi, 2011. Women as agents of change: Female income and mobility in India. Journal of development economics, 94(1), 1-17.

[32] Lundberg, S. and R. A. Pollak, 1996. Bargaining and Distribution in Marriage, Journal of Economic Perspectives, 10, 139-158.

[33] National Statistical Office, 2019. NSS Report: Time Use in India-2019, India.

[34] Ortner, S. B., 1978. The Virgin and the State, Feminist Studies, 4, 19-35.

[35] Oshana, M., 2006. Personal Autonomy in Society, Aldershot: Ashgate Publishing, UK.

[36] Pesando, L. M. and A. Abufhele, 2019. Household Determinants of Teen Marriage: Sister Effects Across Four Low- and Middle-Income Countries. Studies in Family Planning, 50 (2): 113-136.

[37] Psacharopoulos, G., and Z. Tzannatos, 1989. Female Labor Force Participation: An International Perspective. World Bank Research Observer.

[38] Qian, N., 2008. Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance. Quarterly Journal of Economics, 123, 1251–1285.

[39] Roodman, D., 2011. Fitting fully observed recursive mixed-process models with cmp. The Stata Journal, 11(2), pp.159-206.

[40] Rosenzweig, M.R. and Stark, O., 1989. Consumption smoothing, migration, and marriage: Evidence from rural India. Journal of political Economy, 97(4), pp.905-926.

[41] Sivasankaran, A., 2014. Work and Women's Marriage, Fertility and Empowerment: Evidence from Textile Mill Employment, Working paper, Harvard University.

[42] Steelman, L. C., Powell, B., Werum, R., and S. Carter, 2002. Reconsidering the effects of sibling configuration: Recent advances and challenges. Annual Review of Sociology 28: 243–269.

[43] Wooldridge, J.M., 2010. Econometric analysis of cross section and panel data. MIT press.

[44] Yu, W., Su, K., and C. T. Chiu, 2012. Sibship characteristics and transition to first marriage in Taiwan: Explaining gender asymmetries. Population Research and Policy Review 31(4): 609–636.

Appendix

A Proof of Theoretical Results

• Derivation of Equation 1:

M maximizes $\int_\sigma u_M(c_M, c_F, h)r(\sigma)d\sigma$ subject to household budget constraint $c_F + c_M = w_M + w_F(1-h)$. M chooses c_M and c_F given $r(\sigma)$ and h. From the budget constraint, we write $c_M = w_M + w_F(1-h) - c_F$ and replace c_M in the objective function. Hence M maximizes the following expression with respect to c_F.

$$\int_\sigma \Big[(w_M + w_F(1-h) - c_F + v(h) - \theta\sigma h) - \big[(w_M + w_F(1-h) - c_F + v(h)) - \delta(c_F + v(h))\big]^2\Big] r(\sigma)d\sigma$$

First order condition gives

$$\int_\sigma \big[-1 + 2(1+\delta)(w_M + w_F(1-h) - (1+\delta)c_F + (1-\delta)v(h)\big] r(\sigma)d\sigma = 0$$

$$\Rightarrow c_F = \frac{1}{1+\delta}\big[w_M + w_F(1-h) + (1-\delta)v(h)\big] - \frac{1}{2(1+\delta)^2}$$

• Proof of Proposition 1

We only need to show part (v). To see this, suppose that $h^F(\sigma, D)$ is not a decreasing function for some D. Then there are $\sigma_1 > \sigma_2$ such that $h^F(\sigma_1, D) > h^F(\sigma_2, D)$. By definition $h^F(\sigma_1, D)$

$$\frac{1}{1+\delta}\big[w_M + w_F(1 - h^F(\sigma_1, D)) + 2v(h^F(\sigma_1, D)) - (1+\delta)\sigma_1 h^F(\sigma_1, D)\big] - \frac{1}{2(1+\delta)^2}$$
$$> \frac{1}{1+\delta}\big[w_M + w_F(1 - h^F(\sigma_2, D)) + 2v(h^F(\sigma_2, D)) - (1+\delta)\sigma_1 h^F(\sigma_2, D)\big] - \frac{1}{2(1+\delta)^2}$$

Rearranging,

$$w_F\big(h^F(\sigma_2, D) - h^F(\sigma_1, D)\big) + 2\big[v(h^F(\sigma_1, D)) - v(h^F(\sigma_2, D))\big] > (1+\delta)\sigma_1[h^F(\sigma_1, D) - h^F(\sigma_2, D)]$$

Since $\sigma_1 > \sigma_2$ and $h^F(\sigma_1, D) > h^F(\sigma_2, D)$, we also have

$$w_F\big(h^F(\sigma_2, D) - h^F(\sigma_1, D)\big) + 2\big[v(h^F(\sigma_1, D)) - v(h^F(\sigma_2, D))\big] > (1+\delta)\sigma_2[h^F(\sigma_1, D) - h^F(\sigma_2, D)]$$

Rearranging again,

$$\frac{1}{1+\delta}\big[w_M + w_F(1 - h^F(\sigma_1, D)) + 2v(h^F(\sigma_1, D)) - (1+\delta)\sigma_2 h^F(\sigma_1, D)\big] - \frac{1}{2(1+\delta)^2}$$
$$> \frac{1}{1+\delta}\big[w_M + w_F(1 - h^F(\sigma_2, D)) + 2v(h^F(\sigma_2, D)) - (1+\delta)\sigma_2 h^F(\sigma_2, D)\big] - \frac{1}{2(1+\delta)^2}$$

This contradicts the definition of $h^F(\sigma_2, D)$

• Proof of Proposition 2

(i) u_M has unique maximum and it is concave because $u_M''(h) = \frac{2\delta}{1+\delta}v''(h) < 0$; hence it is single-peaked in h.
(iv) Since σ is uniformly distributed, the optimal ex-ante choice of M is obtained by maximizing $\int_\sigma u_M(h,\sigma)d\sigma$ with respect to h. Differentiating inside integration,

$$\int_\sigma \left[\frac{\delta}{1+\delta}(2v'(h) - w_F) - \theta\sigma\right] d\sigma = \frac{\delta}{1+\delta}\left[v'(h) - \frac{1}{2}\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)\right]$$

Here we have used the information that σ follows Uniform $[0,1]$. Since $\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)$ is smaller than $(w_F + (1+\delta))$, $h^M_{ex-ante}$ is strictly greater than $h^F(1,[0,1])$, which is the lower bound of unrestricted optimal choice of F. However, if $\frac{\theta(1+\delta)}{\delta}$ is sufficiently small then $v'(1)$ could be greater than $\frac{1}{2}\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)$ implying $h^M_{ex-ante} = 1$. Otherwise $v'(h^M_{ex-ante}) = \frac{1}{2}\left(w_F + \frac{\theta}{2\delta}(1+\delta)\right)$.

• Proof of Proposition 3

(ii) $u_F(h_1,\sigma) > u_F(h_2,\sigma)$ implies

$$\begin{aligned}
&\frac{1}{1+\delta}[w_M + w_F(1-h_1) + 2v(h_1) - (1+\delta)\sigma h_1] - \frac{1}{2(1+\delta)^2} \\
&\quad > \frac{1}{1+\delta}[w_M + w_F(1-h_2) + 2v(h_2) - (1+\delta)\sigma h_2] - \frac{1}{2(1+\delta)^2} \\
&\Rightarrow\ [w_M + w_F(1-h_1) + 2v(h_1)] - \sigma h_1 > [w_M + w_F(1-h_2) + 2v(h_2)] - \sigma h_2 \\
&\Rightarrow\ [w_M + w_F(1-h_1) + 2v(h_1)] - \frac{\theta}{\delta}\sigma h_1 > [w_M + w_F(1-h_2) + 2v(h_2)] - \frac{\theta}{\delta}\sigma h_2 \\
&\Rightarrow\ u_M(h_1,\sigma) > u_M(h_2,\sigma)
\end{aligned}$$

The third inequality follows from $h_1 > h_2$ and $\frac{\theta}{\delta} < 1$.

• We need the following lemmas to prove Proposition 6.

Lemma 1 *Take $h_2 < h_1$, both in the range $[h^F(1,[0,1]),1]$. Unrestricted optimum for σ_1 and σ_2 are h_1 and h_2 respectively; that is $h^F(\sigma_1,[0,1]) = h_1$ and $h^F(\sigma_2,[0,1]) = h_2$. Then we can find a cut-off on σ, called $\hat{\sigma}(h_2,h_1)$, such that those above it prefer h_2 over h_1 and the opposite is true for those below the cut-off. Moreover, $\sigma_1 < \hat{\sigma}(h_2,h_1) < \sigma_2$.*

Type σ prefers h_2 over h_1 if $u_F(h_2,\sigma) \geq u_F(h_1,\sigma)$. From Equation 2, we obtain an expression for the cut-off

$$\hat{\sigma}(h_2,h_1) = \frac{1}{1+\delta}\left[\frac{2(v(h_1) - v(h_2))}{(h_1 - h_2)} - w_F\right] \tag{13}$$

By mean value theorem, $\frac{v(h_1)-v(h_2)}{h_1-h_2} = v'(t)$ for some t in the open interval (h_2,h_1). Rearranging, we get, $v'(t) = \frac{1}{2}[w_F + (1+\delta)\hat{\sigma}(h_2,h_1)]$. From Equation 3, we obtain that t is the unrestricted best for $\hat{\sigma}(h_2,h_1)$. Since $h_2 < t < h_1$ and $h^F(\sigma,[0,1])$ is a decreasing function of σ (Proposition 1, part (iv)), we can conclude that $\sigma_1 < \hat{\sigma}(h_2,h_1) < \sigma_2$.

Lemma 2 *Take $h_3 < h_1$, both in the range $[h^F(1,[0,1]),1]$. Take $h_2 = \lambda h_3 + (1-\lambda)h_1$, where $0 < \lambda < 1$. Then*

$$(1-\lambda)\hat{\sigma}(h_3,h_2) + \lambda\hat{\sigma}(h_2,h_1) = \hat{\sigma}(h_3,h_1) \tag{14}$$

Further, $\hat{\sigma}(h_3,h_2) > \hat{\sigma}(h_3,h_1) > \hat{\sigma}(h_2,h_1)$

Since $0 < \lambda < 1$, h_2 is also in the range $[h^F(1,[0,1]),1]$ and $h_3 < h_2 < h_1$. We use definition of $\hat{\sigma}(h_3,h_2), \hat{\sigma}(h_2,h_1)$ and $\hat{\sigma}(h_3,h_1)$ (Equation 13) and $h_2 = \lambda h_3 + (1-\lambda)h_1$ to obtain Equation 14. We also know from Lemma 1 that $\hat{\sigma}(h_3,h_2)$ lies between σ_2 and σ_3 and $\hat{\sigma}(h_2,h_1)$ lies between σ_1 and σ_2, where $h^F(\sigma_k,[0,1]) = h_k$, for $k = 1,2,3$. Thus $\hat{\sigma}(h_2,h_1) < \sigma_2 < \hat{\sigma}(h_3,h_2)$. Equation 14 tells us that $\hat{\sigma}(h_3,h_1)$ is a convex combination of $\hat{\sigma}(h_2,h_1)$ and $\hat{\sigma}(h_3,h_2)$. Hence $\hat{\sigma}(h_3,h_2) > \hat{\sigma}(h_3,h_1) > \hat{\sigma}(h_2,h_1)$.

• Proof of Proposition 6

We can now proceed to prove Proposition 6. We have $h_3 < h_2 < h_1$; $h_3, h_1 \in D$ but $(h_3,h_1) \cap D = \emptyset$. Suppose $\tilde{D} = D \cup \{h_2\}$. We want to find $[u_M(\tilde{D},\delta) - u_M(D,\delta)]$.

Suppose σ_3 and σ_1 are such that their unconstrained best are h_3 and h_1 respectively. That is $h^F(\sigma_1,[0,1]) = h_1$ and $h^F(\sigma_3,[0,1]) = h_3$. Now, $h_3 < h_1$ implies $\sigma_3 > \sigma_1$ because $h^F(\sigma,[0,1])$ is decreasing in σ. From Lemma 14, we have $\sigma_3 > \hat{\sigma}(h_3,h_2) > \hat{\sigma}(h_3,h_1) > \hat{\sigma}(h_2,h_1) > \sigma_1$. We divide the σ space into following partitions:

$\sigma \leq \sigma_1$: These types have unconstrained best grater than h_1. Since $u_F(h,\sigma)$ is single-peaked in h and $h_1 \in D$ introduction of $h_2 < h_1$ in $\tilde{D}$ does not change their choice.
$\sigma_1 \leq \sigma < \hat{\sigma}(h_2,h_1)$: Since the interval (h_3,h_1) was not part of D, these types would have chosen h_1 at D. They continue to choose h_1 at $\tilde{D}$.
$\hat{\sigma}(h_2,h_1) \leq \sigma < \hat{\sigma}(h_3,h_1)$: These types switch from h_1 at D to h_2 at $\tilde{D}$.
$\hat{\sigma}(h_3,h_1) \leq \sigma < \hat{\sigma}(h_3,h_2)$: These types switch from h_3 at D to h_2 at $\tilde{D}$.
$\hat{\sigma}(h_3,h_2) \leq \sigma < \sigma_3$: Since the interval (h_3,h_1) was not part of D, these types would have chosen h_3 at D and there is no change at $\tilde{D}$.
$\sigma > \sigma_3$: These types have unconstrained best smaller than h_3. Since $u_F(h,\sigma)$ is single-peaked in h and $h_3 \in D$ introduction of $h_2 > h_3$ in $\tilde{D}$ does not change their choice.

$$\begin{aligned}
&u_M(\tilde{D}) - u_M(D,\delta)\\
&\qquad= \int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} [u_M(h_2,\sigma) - u_M(h_1,\sigma)]d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} [u_M(h_2,\sigma) - u_M(h_3,\sigma)]d\sigma\\
&\qquad= \frac{\delta}{1+\delta} w_F \left[\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} (h_1 - h_2)d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} (h_3 - h_2)d\sigma\right]\\
&\qquad+ \frac{2\delta}{1+\delta}\left[\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} (v(h_2) - v(h_1))d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} (v(h_2) - v(h_3))d\sigma\right]\\
&\qquad+ \theta\left[\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} \sigma(h_1 - h_2)d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} \sigma(h_3 - h_2)d\sigma\right]
\end{aligned}$$

Suppose $h_2 = \lambda h_3 + (1-\lambda)h_1$. The first term

$$\begin{aligned}
&\left[\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} (h_1 - h_2)d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} (h_3 - h_2)d\sigma\right] \\
=&(h_1 - h_2)(\hat{\sigma}(h_3,h_1) - \hat{\sigma}(h_2,h_1)) + (h_3 - h_2)(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_3,h_1)) \\
=&(h_1 - h_3)[\lambda(\hat{\sigma}(h_3,h_1) - \hat{\sigma}(h_2,h_1)) - (1-\lambda)(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_3,h_1)) \\
=&(h_1 - h_3)[\hat{\sigma}(h_3,h_1) - \lambda\hat{\sigma}(h_2,h_1) - (1-\lambda)\hat{\sigma}(h_3,h_2)] \\
=&0
\end{aligned}$$

The last equality follows from Equation 14. The second term,

$$\begin{aligned}
&\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} (v(h_2) - v(h_1))d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} (v(h_2) - v(h_3))d\sigma \\
=&(v(h_2) - v(h_1))(\hat{\sigma}(h_3,h_1) - \hat{\sigma}(h_2,h_1)) + (v(h_2) - v(h_3))(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_3,h_1)) \\
=&(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_2,h_1))[(1-\lambda)(v(h_2) - v(h_1)) + \lambda(v(h_2) - v(h_3))] \\
=&(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_2,h_1))[v(h_2) - (1-\lambda)v(h_1) - \lambda v(h_3)]
\end{aligned}$$

The second equality follows from Equation 14. Further, using Equation 13,

$$\begin{aligned}
&\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_2,h_1) \\
&\qquad =\frac{2}{1+\delta}\left[\frac{v(h_2) - v(h_3)}{h_2 - h_3} - \frac{v(h_1) - v(h_2)}{h_1 - h_2}\right] \\
&\qquad =\frac{2}{(1+\delta)(h_1 - h_3)}\left[\frac{v(h_2) - v(h_3)}{\lambda} - \frac{v(h_1) - v(h_2)}{1-\lambda}\right] \\
&\qquad =\frac{2}{(1+\delta)(h_1 - h_3)\lambda(1-\lambda)}\big[v(h_2) - (1-\lambda)v(h_1) - \lambda v(h_3)\big]
\end{aligned}$$

Thus the second term can be written as

$$\frac{(1+\delta)(h_1 - h_3)\lambda(1-\lambda)}{2}\big[\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_2,h_1)\big]^2$$

Finally, the third term is

$$\begin{aligned}
&\int_{\hat{\sigma}(h_2,h_1)}^{\hat{\sigma}(h_3,h_1)} \sigma(h_1 - h_2)d\sigma + \int_{\hat{\sigma}(h_3,h_1)}^{\hat{\sigma}(h_3,h_2)} \sigma(h_3 - h_2)d\sigma \\
=&\frac{(h_1 - h_2)}{2}\big[(\hat{\sigma}(h_3,h_1))^2 - (\hat{\sigma}(h_2,h_1))^2\big] + \frac{(h_3 - h_2)}{2}\big[(\hat{\sigma}(h_3,h_2))^2 - (\hat{\sigma}(h_3,h_1))^2\big] \\
=&\frac{1}{2}\big[(h_1 - h_2)(\hat{\sigma}(h_3,h_1) - \hat{\sigma}(h_2,h_1))(\hat{\sigma}(h_3,h_1) + \hat{\sigma}(h_2,h_1)) + \\
&\quad (h_3 - h_2)(\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_3,h_1))(\hat{\sigma}(h_3,h_2) + \hat{\sigma}(h_3,h_1))\big]
\end{aligned}$$

Using $h_2 = \lambda h_3 + (1-\lambda)h_1$ and Equation 14, the above expression can be rewritten as

$$-\frac{\lambda(1-\lambda)(h_1 - h_3)}{2}\big[\hat{\sigma}(h_3,h_2) - \hat{\sigma}(h_2,h_1)\big]^2$$

Combing all three terms we get

$$\begin{aligned}&u_M(\tilde{D},\delta) - u_M(D,\delta)\\ =&\lambda(1-\lambda)(h_1-h_3)\big[\hat{\sigma}(h_3,h_2)-\hat{\sigma}(h_2,h_1)\big]^2\left[\delta-\frac{\theta}{2}\right]\\ >&0\end{aligned}$$

• Proof of Proposition 8

We prove by contradiction. Suppose that $1 \notin D^*(\delta)$. Since $D^*(\delta)$ is an interval it must be $D^*(\delta) = [h_2, h_1]$ such that $h_1 < 1$. Consider an alternative delegation set $\tilde{D} = D^*(\delta) \cup \{1\}$. We want to compare $u_M(\tilde{D},\delta)$ and $u_M(D^*(\delta),\delta)$. Let σ_1 be the type for which unrestricted choice is h_1, that is $h^F(\sigma_1,[0,1]) = h_1$. The σ space can be partitioned as follows:

$\sigma > \sigma_1$: For these types the unrestricted choice is smaller than h_1. They are unaffected by inclusion of new choice $\{1\}$.
$\hat{\sigma}(h_1,1) < \sigma \leq \sigma_1$: Under both $D^(\delta)$ and $\tilde{D}$, these types choose h_1.
$\sigma \leq \hat{\sigma}(h_1,1)$: These types switch from h_1 at $D^*(\delta)$ to 1 at $\tilde{D}$.

$$u_M(\tilde{D},\delta) - u_M(D^*(\delta),\delta) = \int_0^{\hat{\sigma}(h_1,1)} [u_M(1,\sigma) - u_M(h_1,\sigma)]d\sigma$$

Since $\sigma \leq \hat{\sigma}(h_1,1)$ prefer higher level of care work than h_1, by part *(ii)* of Proposition 3, we know, $u_M(1,\sigma) > u_M(h_1,\sigma)$ for all such σ. Therefore $u_M(\tilde{D},\delta) > u_M(D^*(\delta),\delta)$.

• Proof of Proposition 9

Suppose $D = [h,1]$ and $h^F(z,[0,1]) = h$. Then $\sigma < z$ chooses her unrestricted best because it is available in D, the rest choose the lower bound h.

$$\begin{aligned}u_M(D,\delta) &= \int_\sigma u_M(h^F(\sigma,D),\sigma)d\sigma\\ &= \int_0^z u_M(h^F(\sigma,[0,1]),\sigma)d\sigma + \int_z^1 u_M(h^F(z,[0,1]),\sigma)d\sigma\end{aligned}$$

Differentiating with respect to z, we obtain

$$\begin{aligned}&u_M(h^F(z,[0,1]),z) - u_M(h^F(z,[0,1]),z) + \int_z^1 \frac{d\left[u_M(h^F(z,[0,1]),\sigma)\right]}{dz}d\sigma\\ =&\int_z^1\left[-w_F + 2v'(h) - \frac{(1+\delta)\theta\sigma}{\delta}\right]\left(\frac{dh}{dz}\right)d\sigma\\ =&\frac{dh}{dz}\left[(-w_F+2v'(h))(1-z) - \frac{(1+\delta)\theta(1-z^2)}{2\delta}\right]\end{aligned}$$

Since $\frac{dh}{dz} < 0$ (Part *(iv)*, Proposition 1) and $z < 1$, F.O.C. implies

$$\begin{aligned}&-w_F + 2v'(h) - \frac{(1+\delta)\theta(1+z)}{2\delta} = 0\\ \Rightarrow\;&(1+\delta)\left[z - \frac{\theta(1+z)}{2\delta}\right] = 0\\ \Rightarrow\;&z = \frac{1}{\frac{2\delta}{\theta}-1}\end{aligned}$$

The equality in second line follows from Equation 3.

• Proof of Proposition 11

$$\begin{aligned}\frac{du_F(\sigma,\delta,w_M)}{d\delta} =& \frac{1}{(1+\delta)^3}\left[1-(1+\delta)(w_M+w_F(1-h^F(\sigma,D^*(\delta))+2v\big(h^F(\sigma,D^*(\delta))\big)\right]\\ &+\frac{d\left[h^F(\sigma,D^*(\delta))\right]}{d\delta}\left(\frac{1}{1+\delta}\right)\left[2v'\big(h^F(\sigma,D^*(\delta))\big)-w_F-(1+\delta)\sigma\right]\end{aligned}$$

For $\sigma<\sigma^*$, $h^F(\sigma,D^*(\delta))=h^F(\sigma,[0,1])$ implying $\left[2v\big(h^F(\sigma,D^*(\delta))\big)-w_F-(1+\delta)\sigma\ \right]=0$. On the other hand for $\sigma>\sigma^*$, $h^F(\sigma,D^*(\delta))=h^F(\sigma^*,[0,1])$. Thus $\frac{d\left[h^F(\sigma,D^*(\delta))\right]}{d\delta}=\frac{h^F(\sigma^*,[0,1])}{d\sigma^*}\frac{d\sigma^*}{d\delta}$. Both $\frac{h^F(\sigma^*,[0,1])}{d\sigma^*}$ and $\frac{d\sigma^*}{d\delta}$ are negative. Since σ chooses an h greater than her best (the best is not available in delegation set) thus $\left[2v\big(h^F(\sigma,D^*(\delta))\big)-w_F-(1+\delta)\sigma\ \right]<0$. Therefore the second term is either zero or negative, irrespective of σ. The first term,

$$\left[1-(1+\delta)(w_M+w_F(1-h^F(\sigma,D^*(\delta))+2v\big(h^F(\sigma,D^*(\delta))\big)\right]\leq[1-(1+\delta)(w_M+2v(1))]$$

This follows from $u_F(h^F(\sigma,D^*(\delta)),\sigma)\geq u_F(1,\sigma)$ because $u_F(h,\sigma)$, is single-peaked. Further, $[1-(1+\delta)(w_M+2v(1))]$ is negative because $2(1+\delta)v(1)>4v(1)>1$.

• Proof of Proposition 12

Difference in autonomy measure under self-choice and parent's choice is

$$\begin{aligned}T(\hat{\delta})&=\frac{1}{\bar{\delta}-\underline{\delta}}\left[\int_{\underline{\delta}}^{\hat{\delta}}h^F(\sigma^*,[0,1])\,d\delta+\int_{\hat{\delta}}^{\bar{\delta}}\left[\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\underline{\delta}}^{\bar{\delta}}h^F(\sigma^*,[0,1])\,d\delta\right]d\delta-\int_{\underline{\delta}}^{\bar{\delta}}h^F(\sigma^*,[0,1])\,d\delta\right]\\&=\frac{1}{\bar{\delta}-\underline{\delta}}\left[\int_{\underline{\delta}}^{\hat{\delta}}\left[h^F(\sigma^*,[0,1])-\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\underline{\delta}}^{\bar{\delta}}h^F(\sigma^*,[0,1])\,d\delta\right]d\delta\right]\end{aligned}$$

Then $T(\bar{\delta})=T(\underline{\delta})=0$. Moreover, $T''(\hat{\delta})=\frac{1}{\bar{\delta}-\underline{\delta}}\frac{d\left[h^F(\sigma^*,[0,1])\right]}{d\delta}>0$. Thus T is concave. Hence $T(\hat{\delta})<\lambda T(0)+(1-\lambda)T(1)=0$. Thus under self-choice, autonomy measure is smaller implying autonomy is higher.

• Proof of Proposition 13.

Parents' approval is dictated by Equation 7.

$$\begin{aligned}&\kappa w_M+E_\delta u_F(\sigma,\delta,w_M)\geq\kappa\bar{w}_M+E_\delta u_F(\sigma,\delta,\bar{w}_M)\\ \Leftrightarrow\ &\kappa w_M+\int_\delta\frac{w_M}{1+\delta}d\delta+E_\delta R(\sigma,\delta)\geq\kappa\bar{w}_M+\int_\delta\frac{\bar{w}_M}{1+\delta}d\delta+E_\delta R(\sigma,\delta)\\ \Leftrightarrow\ &w_M\geq\bar{w}_M\end{aligned}$$

If delegated F approves the proposal when $\delta\leq\hat{\delta}$. Otherwise the second proposal materializes. Parents' payoff from delegation is

$$\begin{aligned}&\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\underline{\delta}}^{\hat{\delta}}[\kappa w_M+u_F(\sigma,\delta,w_M)]\,d\delta+\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\hat{\delta}}^{\bar{\delta}}[\kappa\bar{w}_M+u_F(\sigma,\delta,\bar{w}_M)]\,d\delta\\ =\ &\kappa w_M\frac{\hat{\delta}-\underline{\delta}}{\bar{\delta}-\underline{\delta}}+\kappa\bar{w}_M\frac{\bar{\delta}-\hat{\delta}}{\bar{\delta}-\underline{\delta}}+\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\underline{\delta}}^{\hat{\delta}}u_F(\sigma,\delta,w_M)d\delta+\frac{1}{\bar{\delta}-\underline{\delta}}\int_{\hat{\delta}}^{\bar{\delta}}u_F(\sigma,\delta,\bar{w}_M)d\delta\end{aligned}$$

First take the case, $w_M > \bar{w}_M$. Then parents either approve the proposal or delegate. Suppose that the decision is delegated. Then,

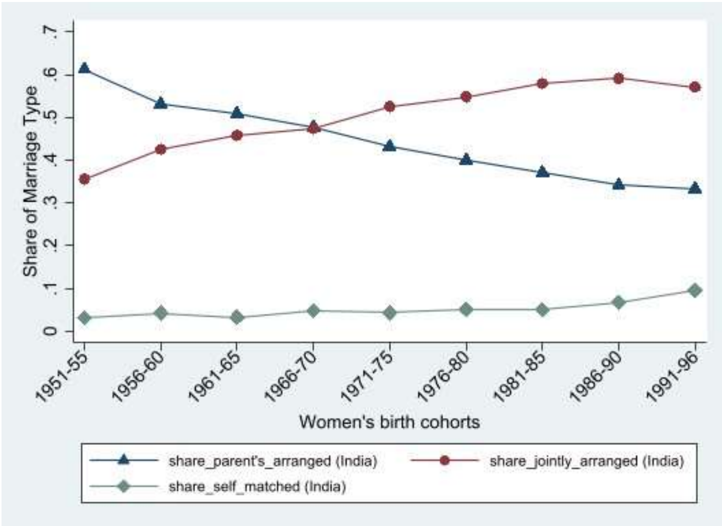
$$
\begin{aligned}
&\frac{1}{\bar{\delta} - \underline{\delta}} \left[\int_{\underline{\delta}}^{\hat{\delta}} u_F(\sigma, \delta, w_M) d\delta + \int_{\hat{\delta}}^{\bar{\delta}} u_F(\sigma, \delta, \bar{w}_M) d\delta \right] - E_{\delta} u_F(\sigma, \delta, w_M) \\
&\quad \geq \kappa w_M - \kappa w_M \frac{\hat{\delta} - \underline{\delta}}{\bar{\delta} - \underline{\delta}} - \kappa \bar{w}_M \frac{\bar{\delta} - \hat{\delta}}{\bar{\delta} - \underline{\delta}} \\
&\quad = \frac{\bar{\delta} - \hat{\delta}}{\bar{\delta} - \underline{\delta}} (w_M - \bar{w}_M) \kappa \\
&\quad > \frac{\bar{\delta} - \hat{\delta}}{\bar{\delta} - \underline{\delta}} (w_M - \bar{w}_M) \ \kappa'
\end{aligned}
$$

where $\kappa' < \kappa$. Hence delegation is also profitable at κ'. The proof is similar when $w_M \leq \bar{w}_M$

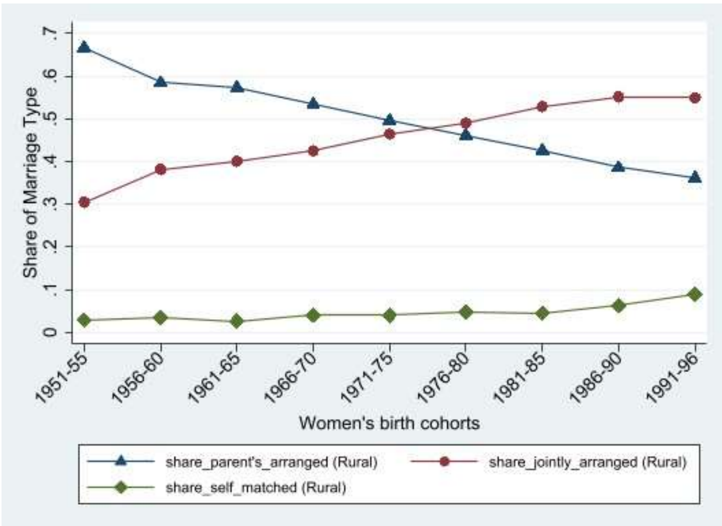
B Figures and Tables

Figure 1: Transition in Matchmaking Process over Different Birth Cohorts

All-India



Rural



Urban

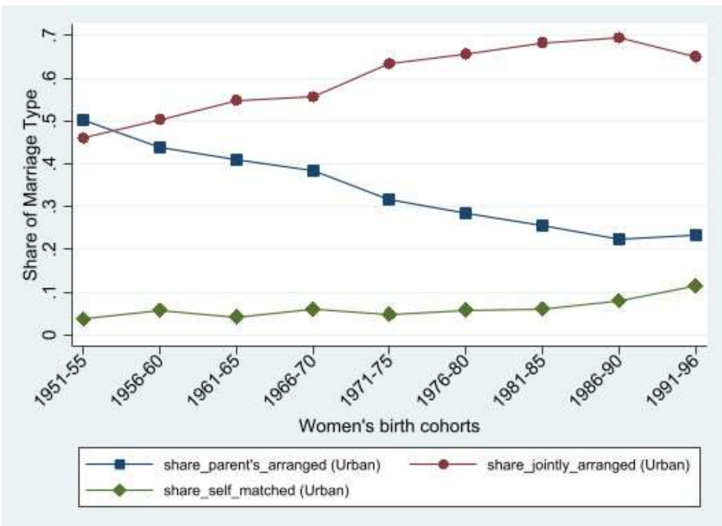


Figure 2: Women's Autonomy in Work and Marriage Type
in Rural and Urban Sectors

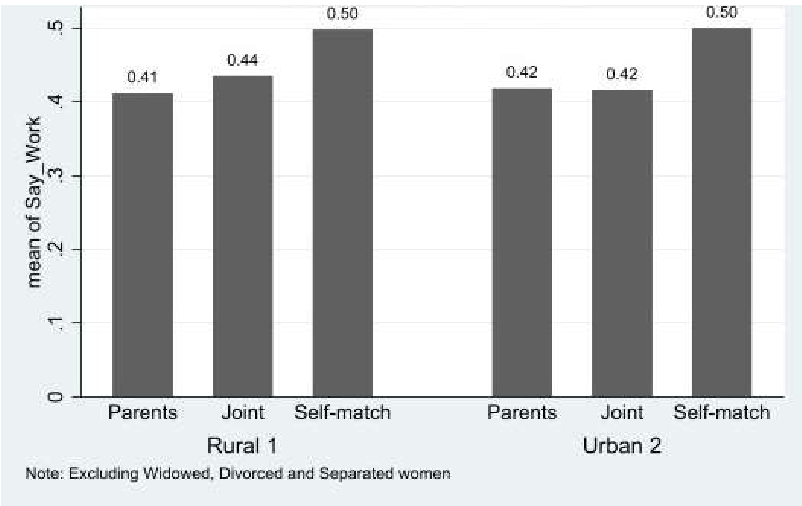


Figure 3: Women's Labour Force Participation Over Autonomy in Work
in Rural and Urban Sectors

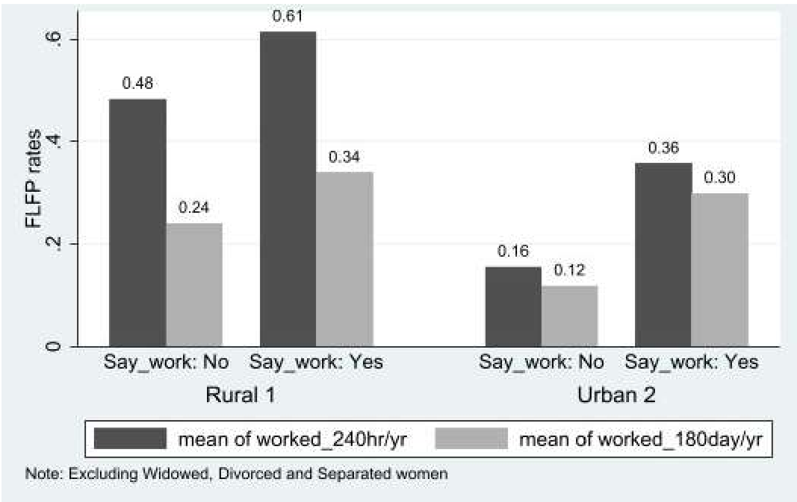


Table 1: Summary Statistics - Average Outcomes over Regions and Urban/Rural

Panel -A

	All	*Hills*	*Northern*	*North East*	*Mid Central*	*Eastern*	*Western*	*Southern*
Marriage Type								
…Parent	0.42	0.27	0.55	0.13	0.7	0.46	0.27	0.15
…Joint	0.53	0.66	0.42	0.53	0.28	0.46	0.69	0.81
…Self	0.04	0.08	0.02	0.35	0.02	0.08	0.03	0.05
Women's Autonomy on work	0.42	0.46	0.38	0.62	0.46	0.43	0.35	0.41
Women's FLFP								
…Worked (180d/yr)	0.26	0.35	0.18	0.15	0.24	0.15	0.29	0.33
…Worked (240h/yr)	0.45	0.54	0.29	0.31	0.49	0.31	0.47	0.5
work days /year	92.68	121.59	66.09	60.63	95.31	62.32	95.65	108.87
work hours /year	571.09	650.68	384.57	358.63	527.67	388.69	651.29	747.15
Observations	29857	2168	3053	503	8930	3902	4031	7270

Panel - B

	Rural	More Dev Vill	Less Dev Vill	**Urban**	Metro Urban	Other Urban
Marriage Type						
…Parent	0.48	0.40	0.55	0.31	0.26	0.32
…Joint	0.48	0.57	0.41	0.64	0.68	0.63
…Self	0.04	0.03	0.05	0.05	0.05	0.05
Women's Autonomy on work	0.43	0.42	0.43	0.42	0.32	0.45
Women's FLFP						
Worked (180d/yr)	0.29	0.29	0.28	0.20	0.16	0.21
Worked (240h/yr)	0.55	0.52	0.57	0.24	0.17	0.26
work days /year	106.63	104.02	109.01	63.91	50.43	67.29
work hours /year	646.36	660.55	633.45	415.81	336.5	435.71
Observations	20109	9584	10525	9748	1955	7793

Note: Covers the full sample of 29857 women who responded for 'say in work' and their marriage patterns. Women were born between 1951 and 1996. The values here represent the sample means.

Table 2: Summary Statistics - Average Outcomes over Marriage Types

	Parents		Joint		Self		Mean Differences		
	Mean	SE	Mean	SE	Mean	SE	(1-2)	(1-3)	(2-3)
	(1)		(2)		(3)		(4)	(5)	(6)
Women's autonomy in work	0.41	0.00	0.43	0.00	0.50	0.00	-0.01*	-0.09***	-0.07***
Women's Characteristics									
Schooling yrs	1.35	0.01	2.68	0.02	2.78	0.02	-1.33***	-1.43***	-0.1
Marriage Age	16.77	0.02	18.82	0.03	19.02	0.03	-2.05***	-2.25***	-0.2
Mother's schooling yr	0.74	0.02	2.04	0.02	2.31	0.03	-1.30***	-1.57***	-0.27*
Father's schooling yr	2.73	0.03	3.88	0.04	3.77	0.04	-1.16***	-1.05***	0.11
Inter-caste Marriage	0.04	0.00	0.04	0.00	0.17	0.00	0	-0.13***	-0.13***
Spouse's Characteristics									
(Husbands) Schooling yrs	5.94	0.03	7.74	0.04	7.87	0.04	-1.79***	-1.93***	-0.14
Mother's schooling yr	0.49	0.02	1.55	0.02	1.94	0.02	-1.06***	-1.46***	-0.40***
Father's schooling yr	2.23	0.02	3.41	0.03	3.48	0.04	-1.18***	-1.24***	-0.07
Econ status - Natal vs Spouse Families									
Same	0.72	0.00	0.76	0.00	0.76	0.00	-0.04***	-0.04**	0
Natal Better	0.17	0.00	0.15	0.00	0.16	0.00	0.02***	0.02	-0.01
Natal Worse	0.11	0.00	0.08	0.00	0.08	0.00	0.02***	0.03***	0
Women's FLFP									
Worked (180d/yr)	0.25	0.00	0.26	0.00	0.26	0.00	-0.02**	-0.01	0.01
Worked (240h/yr)	0.48	0.00	0.42	0.00	0.42	0.00	0.05***	0.06***	0.01
work days /year	94.3	0.68	91.31	1.00	94.05	0.95	2.98*	0.25	-2.74
work hours /year	552.56	4.73	583.32	6.78	598.03	6.75	-30.75**	-45.47	-14.71
Observations	12575		15934		1319		28509	13894	17253

Note: Sample restrictions: IHDS-II interviewed a total of 39,253 women privately, we restrict our sample to those women who are currently in their first wedlock, and responded about their marriage patterns, and 'say in work choice', resulting in a sample of about 29,857 women.
Last three columns represent the differences of mean values and their significance between columns: (1)-(2) [i.e. Parent vs. Joint arranged marriages], (1)-(3) [i.e. Parent vs Self-match marriages] and (2)-(3) [i.e. Joint vs Self-match marriages].

Table 3: Estimation Results of Woman's Autonomy in Work (Average Marginal Effects from Non-linear Models)

	All India			Rural			Urban		
Dep Var: Work_Say	Probit	Probit	IV	Probit	Probit	IV	Probit	Probit	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Marriage Type (*Ref: Parents arranged*)									
…Jointly arranged	0.0289^{***}	0.0538^{***}	0.225^{***}	0.0294^{**}	0.0600^{***}	0.256^{***}	0.0205	0.0334^{**}	0.133^{**}
	(0.0104)	(0.0110)	(0.0246)	(0.0133)	(0.0142)	(0.0277)	(0.0152)	(0.0161)	(0.0581)
…Self-match	0.0473^{**}	0.0945^{***}	0.299^{***}	0.0365	0.0912^{***}	0.334^{***}	0.0424	0.0791^{**}	0.186^{**}
	(0.0213)	(0.0217)	(0.0363)	(0.0281)	(0.0284)	(0.0430)	(0.0303)	(0.0315)	(0.0801)
First Stage Estimates (Multinomial Probit)									
Mar_type;_Jointly arranged									
IV1: share Joint arrgd (cohort+State)			3.684^{***}			4.277^{***}			2.603^{***}
			(0.209)			(0.256)			(0.315)
IV2: share Self-match (cohort+State)			2.213^{***}			2.757^{***}			1.587^{**}
			(0.589)			(0.779)			(0.793)
Mar_type: Self-match									
IV1: share Joint Mrrg (cohort+State)			1.391^{***}			1.369^{***}			1.480^{***}
			(0.191)			(0.250)			(0.327)
IV2: share Self-match (cohort+State)			5.567^{***}			5.649^{***}			5.787^{***}
			(0.626)			(0.828)			(0.877)
atanhrhoEx			-0.294^{***}			-0.344^{***}			-0.185^{**}
(Constant)			(0.0392)			(0.0444)			(0.0941)
Observations	29144	27187	27187	19683	18387	18387	9453	8800	8800
Pseudo R^2	0.068	0.055		0.063	0.049		0.110	0.089	
State FE	Y	N	N	Y	N	N	Y	N	N
Region FE	N	Y	Y	N	Y	Y	N	Y	Y
Cohort 8,9 Drop	N	Y	Y	N	Y	Y	N	Y	Y

Notes: All regressions include dummies for religion and caste, as well as birth cohort dummies, development region dummies.

The other controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the of marriage, log of per capita income of other household members (except woman), dummies for if woman's parents and in-laws are alive, dummies for woman's education categories, woman's marriage age categories, woman's health poor dummy, dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, absolute differences of age and years of schooling with husband.

Clustered (at PSU level) robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 4: Estimation Results of Woman's Autonomy in Work – Linear Probability Models

	All India			Rural			Urban		
Dep Var: Work_Say	OLS-1	OLS-2	IV-LPM	OLS-1	OLS-2	IV-LPM	OLS-1	OLS-2	IV-LPM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Marriage Type (*Ref: Parents arranged*)									
…Jointly arranged	0.029***	0.055***	0.356***	0.029**	0.061***	0.372***	0.022	0.034**	0.261*
	(0.010)	(0.011)	(0.059)	(0.013)	(0.014)	(0.065)	(0.016)	(0.016)	(0.134)
…Self-match	0.047**	0.095***	0.769***	0.035	0.090***	0.837***	0.045	0.080**	0.708***
	(0.021)	(0.022)	(0.140)	(0.028)	(0.028)	(0.185)	(0.031)	(0.031)	(0.213)
First Stage Estimates (OLS for each of two marriage type indicators)									
Mar_type;_Jointly arranged									
IV1: share Joint arrgd (cohort+State)			0.839***			0.991***			0.553***
			(0.044)			(0.055)			(0.064)
IV2: share Self-match (cohort+State)			-0.019			0.160			-0.230*
			(0.093)			(0.128)			(0.136)
Mar_type: Self-match									
IV1: share Joint Mrrg (cohort+State)			0.015			0.005			0.021
			(0.017)			(0.021)			(0.027)
IV2: share Self-match (cohort+State)			0.775***			0.771***			0.742***
			(0.074)			(0.097)			(0.115)
First Stage F (Cragg Donald)			52.049			30.735			24.215
archi2			65.872			56.452			15.995
arf			32.841			28.106			7.930
Observations	29144	27187	27187	19684	18387	18387	9460	8800	8800
R^2	0.089	0.073	-0.024	0.083	0.065	-0.048	0.141	0.116	0.048
F	21.656	20.740	19.582	.	11.700	11.556	.	15.907	14.139
State FE	Y	N	N	Y	N	N	Y	N	N
Region FE	N	Y	Y	N	Y	Y	N	Y	Y
Cohort 8,9 Drop	N	Y	Y	N	Y	Y	N	Y	Y

Notes: All regressions include dummies for religion and caste, as well as birth cohort dummies, development region dummies.
The other controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the of marriage, log of per capita income of other household members (except woman), dummies for if woman's parents and in-laws are alive, dummies for woman's education categories, woman's marriage age categories, woman's health poor dummy, dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, absolute differences of age and years of schooling of woman with her husband, and her working status (240 hrs/yr) in the past year.
Clustered (at PSU level) robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 5: Estimation Results of Woman's Autonomy in work – (Strategy-1: Alternative IVs)

Dep Var: Work_say	All India		Rural		Urban	
	IV-NLM#	IV-LPM	IV-NLM#	IV-LPM	IV-NLM#	IV-LPM
	(1)	(2)	(3)	(4)	(5)	(6)
Marriage Type (*Ref: Parents arranged*)						
…Jointly arranged	0.225***	0.315***	0.261***	0.373***	0.101*	0.180
	(0.0252)	(0.059)	(0.0270)	(0.065)	(0.0577)	(0.144)
…Self-match	0.301***	0.964***	0.342***	0.820***	0.154**	0.901***
	(0.0371)	(0.167)	(0.0426)	(0.203)	(0.0749)	(0.260)
Worked_240hrs/yr	0.141***	0.145***	0.106***	0.113***	0.223***	0.221***
	(0.00895)	(0.010)	(0.0106)	(0.012)	(0.0141)	(0.017)
First Stage Estimates						
Mar_type:_Jointly arranged -						
IV1: share Joint	3.218***	0.736***	3.894***	0.876***	1.971***	0.422***
(cohort+state+UrRu)	(0.169)	(0.038)	(0.218)	(0.047)	(0.257)	(0.057)
IV2: share Self-match	2.723***	0.224***	2.261***	0.227**	3.024***	0.080
(cohort+state+UrRu)	(0.498)	(0.086)	(0.551)	(0.110)	(0.779)	(0.136)
Mar_type: Self-match -						
IV1: share Joint	1.289***	0.032**	1.442***	0.032*	1.212***	0.044**
(cohort+state+UrRu)	(0.171)	(0.014)	(0.215)	(0.019)	(0.290)	(0.021)
IV2: share Self-match	4.413***	0.597***	3.979***	0.571***	5.464***	0.598***
(cohort+state+UrRu)	(0.608)	(0.066)	(0.633)	(0.095)	(0.807)	(0.092)
atanhrhoEx	-0.295***		-0.356***		-0.111	
(Constant)	(0.0403)		(0.0440)		(0.0896)	
First Stage F (Cragg Donald)		36.250		16.038		19.503
archi2		73.581		56.023		16.215
arf		36.685		27.893		8.040
Observations	27187	27187	18387	18387	8800	8800
Pseudo R^2						
State FE	N	N	N	N	N	N
Region FE	Y	Y	Y	Y	Y	Y
Cohort 8,9 Drop	Y	Y	Y	Y	Y	Y

Notes: # NLM in the top line stands for non-linear model using Probit for second stage and Multinomial probit for first stage in CMP framework; whereas LPM stands for linear probability model and uses OLS for both first and second stages, using ivreg of STATA.
Further, for the NLM models the coefficients represent the AME.
All regressions include dummies for religion and caste, as well as birth cohort dummies, development region dummies. The other controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the of marriage, log of per capita income of other household members (except woman), dummies for if woman's parents and in-laws are alive, dummies for woman's education categories, woman's marriage age categories, woman's health poor dummy, dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, absolute differences of age and years of schooling of woman with her husband, and her working status (240 hrs/yr) in the past year.
Clustered (at PSU level) robust standard errors in parenthesis. $^{*} p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$

Table 6: Estimation Results of Woman's Autonomy in Work (Strategy-2: Alternative specification of marriage type and IVs**)**

Dep Var: Work_say	All India			Rural			Urban		
	Probit #	IV-NLM#	IV-LPM	Probit #	IV-NLM#	IV-LPM	Probit #	IV-NLM#	IV-LPM
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Marriage Type (*Ref: Parents arranged*)									
…Joint/Self-match	0.0542***	0.193***	1.044**	0.0607***	0.295***	1.014**	0.0312**	0.00979	1.166
	(0.0107)	(0.0628)	(0.419)	(0.0137)	(0.0900)	(0.404)	(0.0157)	(0.0804)	(1.178)
Worked (240hr/yr)	0.149***	0.142***	0.136***	0.114***	0.102***	0.108***	0.223***	0.224***	0.198***
	(0.00888)	(0.00959)	(0.0129)	(0.0110)	(0.0124)	(0.0134)	(0.0131)	(0.0133)	(0.0440)
First Stage Estimates									
Marriage_type_2cat									
IV1: No. of sisters		-0.0221***	-0.005***		-0.0236***	-0.006***		-0.0172*	-0.004
		(0.00610)	(0.002)		(0.00731)	(0.002)		(0.0104)	(0.003)
IV2: No. of brothers		-0.0195***	-0.005**		-0.0280***	-0.007**		-0.00322	0.001
		(0.00742)	(0.002)		(0.00900)	(0.003)		(0.0125)	(0.004)
atanhrhoEx		-0.236**			-0.429**			0.0353	
(Constant)		(0.113)			(0.203)			(0.131)	
First Stage F (Cragg Donald)			6.105			6.110			0.972
archi2			10.91			10.60			2.097
arf			5.442			5.279			1.040
Observations	29049	29049	29049	19619	19619	19619	9430	9430	9430
Pseudo R^2									
State FE	N	N	N	N	N	N	N	N	N
Region FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Cohort 8,9 Drop	N	N	N	N	N	N	N	N	N

Notes: NLM in the top line stands for non-linear model using Probit for both first and second stages, using in Biprobit od STATA; whereas LPM stands for linear probability model and uses OLS for both first and second stages, using ivreg of STATA.
Further, for the Probit and IV-NLM models the coefficients represent the AME.
All regressions include dummies for religion and caste, as well as birth cohort dummies, development region dummies. The other controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the of marriage, log of per capita income of other household members (except woman), dummies for if woman's parents and in-laws are alive, dummies for woman's education categories, woman's marriage age categories, woman's health poor dummy, dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, absolute differences of age and years of schooling with husband.
Clustered (at PSU level) robust standard errors in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Autonomy in Work Choice and Work Force Participation in Rural India
(Average marginal effects from Simultaneous Equations Model)

	Model 1 for Work Definition 1- 180d/yr		**Model 2** for Work Definition 2- 240hr/yr	
	(Eqn.1) Say_work	(Eqn.2) Worked	(Eqn.1) Say_work	(Eqn.2) Worked
Worked 180 days/yr	0.099^{***}			
	(0.027)			
Worked 240 hr/yr			0.106^{***}	
			(0.029)	
Say_Work		0.091^{***}		0.106^{***}
		(0.009)		(0.009)
Jointly arranged	0.058^{***}		0.060^{***}	
	(0.015)		(0.015)	
Self-matched	0.081^{***}		0.080^{***}	
	(0.030)		(0.031)	
Dist % drought vill in 2011		-0.031^{***}		0.003
		(0.004)		(0.013)
Dist % drought vill in 2010		0.026^{***}		0.037^{***}
		(0.004)		(0.005)
Observations	17227	17227	17076	17076
atanhrhoEx	-7.084^{***}		-7.232^{***}	
	(0.0882)		(0.0913)	

Note: Our model of system of two simultaneous equations has been estimated using SUR framework of CMP in STATA. Both outcome variables are binary and endogenous and influence each other. (Eqn 1: probit and Eqn2: probit). *Samples dropped*: Women married during 2011 and 2012 were excluded from estimation process, as we expect the prevalence of drought in a district may adversely affect the marriage pattern of the women. Further some observations are dropped from those villages in certain district where drought information for the concerned years is not available.

We control for state level fixed effects and development-region fixed effect for each model.

Common controls for both the equations in each model: include dummies for religion and caste. The household specific controls include household's perception on neighbourhood eve teasing, household's main income group dummies, dummies for relative economic status of two families at the time of marriage, log of per capita income of other household members (except woman). The women specific controls are – dummy if her husband is away for work, a seven category variable for woman's education, woman's health poor dummy, set of dummies for woman's membership in Female Group, SHG, Credit Saving Group, Political Orgnisation, age of woman, absolute differences of age and years of schooling of woman with her husband, dummy for inter-caste marriage, dummy for pregnancy, her parents' education, husband's parents' education, dummy for practice of purdah in the household, dummies for English speaking ability, set of dummies if woman's parents and in-laws are alive, dummies for presence of elderly in the household (>65), dummies for husband's primary activity status.

Specific controls for equation 1 (Worked): Dummy for woman's migration from childhood residence, number of children under 5, number of female children aged 6-14, number of male children aged 6-14, household size, dummy for existence of debt, dummy for major expenditure-incurring incident in last 5 years, number of household members with major morbidity

Specific controls for equation 2 (Work_say): Dummy for her marriage with blood relative, dummy for woman grow up in the same neighbourhood as husband, marriage age, square of marriage age.

Clustered (at village level) robust standard errors in parenthesis. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$