Terms of Engagement: Marriage and Migration in India

Andrew Beauchamp*, Rossella Calvi†, and Scott Fulford‡

December 31, 2016

Abstract

Indian marriage markets are characterized by an enormous level of female migration, the presence of dowries, and by differing levels of participation in the decision by women. We formulate and estimate a dynamic, equilibrium, two-sided matching model which allows for estimation of separate preferences for men and women. We recover male and female preferences over partner characteristics, dowry, and migration costs in the presence of differing degrees of female independence and unobserved heterogeneity. In counter-factual simulations we focus on how likely changes in sex-ratios, female autonomy, and education affect equilibrium marriage matching and welfare. Our estimates suggest that men prefer less educated and less autonomous women, and so increases in female education and autonomy reduce the welfare of women in the marriage market, even if education and autonomy improve welfare outside of marriage. Declining sex-ratios improve welfare for some, but not all, women largely by increasing the value of marrying later.

Keywords: Empirical Matching; Marriage Markets; Migration; India

JEL Classification: J12 ; O15 ; D1

*Department of Economics, Wright State University 3640 Col Glenn Hwy Dayton OH 45435.

†Department of Economics, Rice University.

‡Economist, Consumer Financial Protection Bureau. The views expressed are the authors' and not necessarily those of the Consumer Financial Protection Bureau or the United States.

"I asked a family who came looking for a match for their son about what they wanted in a girl. They said they wanted a brand."
—Marriage broker for the Punjabi community in New Delhi.[1]

1 Introduction

With its large young population, more marriages take place annually in India than anywhere else in the world.[2] So in understanding how parents and spouses make decisions about marriage, India is not a special case, everywhere else is. This paper develops and estimates a dynamic, general equilibrium, two-sided matching model of the largest marriage market in the world. Our focus is on recovering different preferences for men and women, which allow us to understand how demographic trends, such as missing women or increased female education, affect male and female welfare differently.

The marriage market in India has a number of characteristics that make it particularly suitable for our analysis. Firstly, marriage is nearly universal and occurs early in the life-cycle, simplifying the dynamics of the problem. Secondly, we show that in practice India is divided into many distinct marriage markets by geography and caste. The distinct markets give rise to variation in the supply of available partners on both sides the marriage market, which is necessary to separately identify male and female preferences. Finally, in many ways the marriage market in India looks more like a market than it does in other parts of the world. Co-habitation before marriage is extremely rare and 68% of spouses first meet each other on the day of the wedding (Fulford, 2013, table 1). Marriage is frequently arranged by the parents: 60% of women report that their parents alone were responsible for the choice of spouse. So marriage takes place, for the most part, based on characteristics and terms that are, at least potentially, observable. The prevalence until recently of placing short matrimonial advertisements in newspapers among the middle class in India shows just how little information is necessary to form at least the initial pool of eligible spouses even among the educated elite (Banerjee et al., 2013). Marriage brokers have become increasingly professionalized, and their work is similar to the work of investment bankers or realtors in finding and connecting buyers and sellers (Trivedi, 2011).

Understanding the preferences that drive the Indian marriage market provides a useful insight both into how people make trade-offs among different traits they find desirable, but also into how the decision maker affects these preferences. Our model is built around equilibrium outcomes (terms) which include migration distance, dowry size and the degree of female independence, along with partner characteristics of age and education level. Many Indian women and, to a

[1]Quoted in Trivedi (2011).

[2]India is followed closely by China. Demographic calculations from the 2001 Indian census show that there were between 8.9 and 9.9 million women per year who married for the first time in India in 2001 and since then the population of women of marriageable age has increased. Remarriage after widowhood and divorce add to this number, although the census provides no easy way of calculating their totals. The United Nations Population Division lists the number of Chinese marriages in 2005 and 2006 as between 8.2 and 9.5 (see World Marriage Data, 2008, which omits India)

lesser extent, men, have little to no input into who they marry. Others are the primary decision maker. We find that this autonomy is valued heterogeneously by men and women, and that therefore women pay a price when exercising some degree of autonomy.

India is experiencing a growing shortage of women as a preference for sons and sex-selective abortions collide. The upcoming "marriage squeeze" (Guilmoto, 2012; Sautmann, 2011) is therefore of primary public policy interest since, at the very least, a large cohort of young men will find that there are no brides available. We examine the likely consequences for the marriage market in India. We find that decreases in the sex-ratio generate large changes in the age gap at marriage as men marry later while women marry younger. We also see substantial increases in involvement in the wedding decisions and decreases in marriage migration.

Imposing universal education and female participation in their marriage decisions (autonomy) has larger welfare effects than sex-ratio changes, and according to our estimates both policies universally reduce female welfare from marrying. For education, this loss arises because universal education forces women to compete in a fewer number of marriage markets, markets that men prefer not to match in. In equilibrium, this means women must compromise on other terms: they match with men from farther away and receive smaller dowries, which on net reduces welfare. Of course, this finding does not mean that universal autonomy or education are bad for welfare outside of the marriage market, but emphasizes the importance of modeling and general equilibrium; even well estimated partial equilibrium results are not a good guide for policy when competition is important.

Our modeling approach extends the work of Arcidiacono, Beauchamp and McElroy (2014) who build an equilibrium discrete types model in the spirit of Choo and Siow (2006). The principal difference from Choo and Siow (2006) is to pursue an identification strategy that exploits the observation of multiple matching markets, allowing us to estimate distinct male and female preferences. The approach also allows for preferences over relationship outcomes, which we call terms, in addition to partner characteristics. We extend the basic static equilibrium model of Arcidiacono, Beauchamp and McElroy (2014) to incorporate the full dynamic marriage problem on both sides of the market, endogenously modeling the distribution of singles and marriages over time. Because the original work of Arcidiacono, Beauchamp and McElroy (2014) and Choo and Siow (2006) builds up a discrete choice model of individual decision making, extending the model to include dynamics and unobserved heterogeneity is a straightforward application of the literature on dynamic discrete choice models.[3,4] We argue this is an important modeling contribution, since until now most two-sided marriage studies treat the problem as static in nature, necessarily abstracting away from inter-temporal substitution among partners which we find experiences dramatic changes in different counter-factual policy environments.

Our work fits into a large and growing literature examining the Indian marriage market (summarized in the next section). While there have been some attempts to build models to

[3]See Aguirregabiria and Mira (2010) for a detailed review of the recent literature.

[4]Salani and Galichon (2012) and Chiappori et al. (2012) discuss conditions for the identification of unobserved heterogeneity in a similar class of two-sided model.

explain various features of the Indian marriage market (Banerjee et al., 2013; Edlund, 2006; Fulford, 2013) and to take into account sex-ratios (Anderson, 2003, 2007; Rao, 1993; Sautmann, 2011), ours is the first to estimate the multiple ways that the general equilibrium can effect the market. The concept of equilibrium is central to this work since it would be easy when just looking at the decisions made by individuals to mistake indifference between choice sets for indifference over individual characteristics. For example, all women may prefer men with more education, but in equilibrium some women are going to have to marry men with less education. What do the women who marry the better educated give up to do so? In equilibrium, more women will search for better educated men until the marginal ones are just indifferent. That will harm the probability of matching and the marriage terms making educated men less valuable in other dimensions. Our framework can take into account these complex, multi-dimensional calculations and tradeoffs necessary in a two sided matching market with frictions.

Understanding the marriage market is also central to understanding both the causes and consequences of the growing scarcity of women in India (Guilmoto and Depledge, 2008). Indian women go "missing" (Sen, 1992) in two distinct ways: First, there is the growing prevalence of sex selective abortion in India. While the reasons are complex, at least one reason to abort is that women join their husband's family on marriage, and so investments in them are more difficult to justify (Bhat and Zavier, 2007); to use the common expression, why should one "water a neighbor's garden?" Since marriage is the intermediate event which imposes costs through dowries and moves the bride from her parent's household, understanding the preferences and equilibrium around marriage is crucial to addressing skewed sex-ratios in early life. The second wave of missing women occurs later in life, generally at post-reproductive ages (Anderson and Ray, 2010, 2012) in part because of declining bargaining power within the household (Calvi, 2016) and widowhood (Anderson and Ray, 2015). Since the terms set at marriage help determine womens' treatment within it, understanding how these terms are set might provide greater insight into outcomes later in life.[5]

2 Marriage in India

Marriage is close to universal in India. Figure 1 shows the marriage hazard rate—with no pun intended—of Indian men and women. Women start marrying early but the peak marriage years are between 15 and 19 with a significant fraction marrying between 20-24. More than 90% of women have married by age 25. Men start marrying later, but by the time they have reached their 30s, more than 90% of men have married at least once. Divorce is extremely uncommon,

[5]In this regard, the present paper marginally relates to recent works on how equilibrium in the marriage market determines intra-household allocation of resources. Chiappori et al. (2015), for instance, set out an equilibrium lifecycle model of education, marriage and labor supply and consumption in a transferable utility context. Anderson and Bidner (2015) develop a model of the marriage market with intra-household bargaining to shed light on the incentives for brides' parents to allocate the property rights over the dowry between their daughter and her groom. Cherchye et al. (forthcoming) combine the assumption of a stable marriage market with the collective model of the household, while Cherchye et al. (2016) apply this framework to investigate the economic gains to marriage and divorce in Malawi.

but does exist. Somewhat more common is widowhood. The average age of marriage for women is 17.3. A common practice in northern India separates the marriage ceremony from when the bride moves to join her husband's family. The average age of migration and consummation is 17.7.[6]

The practice of dowry, in which the bride's family gives a wide array of goods including gold to the groom and his family is a central focus of the literature studying the Indian marriage market.[7] More than 90% of marriages now include a dowry of some sort, and it can be several multiples of annual income for the bride's family (Anderson, 2007). Thought of as a negative price for brides, dowries appear to be puzzling because the evidence, slightly spotty since dowries are technically illegal in India, suggests they have increased in size and spread to new populations even as the poor and declining female-to-male sex-ratio suggests that women should be more in demand (Anderson, 2003, 2007; Rao, 1993). Dowries may be a form of conspicuous consumption (Bloch et al., 2004) or, as Edlund (2006) suggests, a form of early bequest that should rise as women become scarcer. Botticini and Siow (2003) examine dowries as bequests when daughters leave their natal family while sons stay. Then the "dowry contract" makes sense since it does not dilute the son's incentives to invest in family specific capital such as immovable, and often unsellable, land. The theory helps explain the disappearance of dowries in many parts of the world as such fixed capital has become less important, and their prevalence in patrilocal India with its poor capital markets. Sautmann (2011) introduces search frictions into the marriage market and shows that with frictions, rising populations can explain both a narrowing age gap and higher dowries. We add to this literature by modeling dowries as equilibrium objects, which men, women, and a woman's decision making family can all have different preferences over. We show that intra-household differences in preferences between the parties is yet another reason that dowries have not behaved in the manner economists would predict if it were a price.

Marriage is intimately linked with education in India, and for women the young age of marriage in particular is a constraint. Maertens (2013) presents evidence from several villages in central India that suggests that parents view the most desirable age of marriage as a constraint for their daughters, but not for their sons. Combined with a practice of marrying daughters in order of birth, Vogl (2013) shows that having a younger sister encourages parents to marry their older daughters faster, thus reducing their educational opportunities. On the other side, education seems to be a particularly desirable quality for grooms (Banerjee et al., 2013). For this reason our model incorporates the dynamic decision on the part of a woman or her family to either enter the marriage market with their teenage daughters, or to instead preserve the option value of future marriage.

One of the central characteristics of Indian marriages is that the woman leaves her birth village to join her husband's family. Patrilocal village exogamy where the woman moves out

[6] The practice is known as *gauna*. See table 1 in Fulford (2013). The age ranges are those reported in the census and match the approach taken in our estimation. Finer age ranges from surveys closely match the census.

[7] Anukriti et al. (2016) provide a fair description of the evolution of and the heterogeneity in dowry by caste, religion, and state in contemporary India.

of her village to join her husband's family is the practice throughout most of India, although very different practices prevail in the north-east. Table 1 shows the migration from place of birth by origin and reason of adults. Only 24% of women 25 and older still live where they were born, while 85% of men do. Women move almost entirely for marriage, and, to a much smaller extent, to accompany their husbands or parents. While migration is pervasive among Indian women, 73% of women stay within their birth district. Rosenzweig and Stark (1989) suggest one motive for female marriage migration may be consumption smoothing in the presence of geographically diverse shocks, a hypothesis Fulford (2013) dismisses since there is no evidence of transfers. Instead, Fulford (2013) suggests that a model of the marriage search is necessary to rationalize the regional differences in migration. We exploit the fact that women so frequently remain within the district to model the marriage markets across India at the district level.

3 Model

Following the work of Arcidiacono, Beauchamp and McElroy (2014) we formulate a discrete-types directed search model of Indian marriage markets. We treat the district as the primary matching market, motivated by empirical patterns in inter-district migration which show the majority of spouses come from the same district discussed in the previous section and evident in Table 1.

We categorize each male as a type m where $m \in \{1, 2, \ldots, M\}$. Similarly, each woman (or her household decision maker such as her parents) is given a type w where $w \in \{1, 2, \ldots, W\}$. An individual's type can denote some collection of observed characteristics such as age, caste, or education. For males (females) there are then W (M) types of mates. Let im indicate the i-th member of type m.

We index the possible terms of the relationship by $r \in \{1, ..., R\}$, and focus modeling on migration distance, dowry size and allowing women to participate in spousal selection as the terms of interest. Other terms could include social roles within the household, or even predetermined characteristics of the match such as how many children to have, whether the woman works outside the home, or how she will be treated within marriage. What distinguishes terms from characteristics is that an individual cannot change his or her own characteristics, but can look for better relationship terms, an action which is more likely when sex-ratios are more favorable. A term is a characteristic of a match rather than of either partner. Although it may seem at first glance that the distance moved is characteristic of a man, distance moved depends on the geographic position of both individuals, so we model it as a characteristic of the match.

Search within this framework is completely directed: men and women are able to target their search on both the characteristics of the partner as well as the terms of the relationship. Each woman (man) then makes a discrete choice to search in one of $M \times R$ $(W \times R)$ markets, resulting in $M \times W \times R$ types of matches (each element of which we denote as $\{m, w, r\}$). Thus, the different types of matches are segmented, consistent with needing to invest in particular networks or search channels to identify a partner within a given type-term combination (e.g. an older, educated man living nearby).[8]

Search is then modeled as a multi-stage matching game. Individuals who do not match today (provided they are young enough) can participate in future marriage markets. Individuals first decide in which market to search, precluding the option of not searching.[9] Since there are no direct search costs, but only opportunity costs, everyone finds it optimal to search in the current period. Following search, couples are joined with the probabilities of matching depending on the number of searchers on both sides of the market. Unmatched individuals who do not age out of the market participate in the market tomorrow, but as older agents whose matching prospects are different. Thus, the decision of where to search in a given time period endogenizes the uncertainty over whether one can match today, and the uncertainty over market prospects in the future, both of which are functions of the behavior of all types of men and women both today and in future periods.

3.1 Individuals

An individual's expected utility for searching in a particular market today depends upon four factors. For a w_t-type woman (whose type will change in subsequent periods so we denote it by subscript t), who matches with an m-type man on relationship terms r, during period t these four elements are:

1. the probability of matching in the relationship in the market today $P_t^{mr}(w_t)$,
2. a deterministic portion of utility conditional on matching given by $\mu^{mr}(w_t)$,
3. an individual-specific preference term $\epsilon_{it}^{mr}(w_t)$,
4. and the continuation value associated with participating in the market tomorrow, or, if old enough, exiting the market following a failure to match today given by: $V_{t+1}(w_t)$.[10]

Thus, the probability of matching and the observed utilities from matching vary only at the type-relationship level (here w_t, m, r), and not at the individual level i. This means the probability of matching is only affected by individual and partner types and relationship terms: all females of type w_t searching in the $\{m, r\}$ market have the same probability of matching. The only individual-specific element of expected utility is $\epsilon_{it}^{mr}(w_t)$. We assume, the $\epsilon_{it}^{mr}(w_t)$ are known to the individual before making their decision over where to search, so there is no match-specific component beyond what occurs through the observed characteristics of the partners (w_t and m) and the terms of the relationship (r). We treat the $\epsilon_{it}^{mr}(w_t)$'s as observed only to the individual:

[8]This feature of the model is consistent with newspaper advertising and match-maker services, forms of search specialization.

[9]Given the high marriage rates in India and the cultural preference for marriage, this assumption is not far from reality.

[10]The corresponding terms for men are $P_t^{wr}(m_t)$, $\mu^{wr}(m_t)$, $\epsilon_{it}^{wr}(m_t)$ and $V_{t+1}(m_t)$

only the distribution is known to the other participants in the market.[11] Continuation values similarly only differ based on an individuals' own type (w_t), and is not influenced by current decisions over where to search (m and r). In this framework, the sources of uncertainty from the individual's perspective are their probability of finding a match today $P_t^{mr}(w_t)$ and the continuation value, which can include participation in the marriage market tomorrow. Finally, we also assume $\mu^{mr}(w_t)$ is not a function of time, although age will be included in the state-space, so it does affect match utility.

The value from searching in a particular market takes an expected utility form: it is the probability of matching (or not) times the utility conditional on matching (or the continuation value). We specify the functional form of the utility such that the value function for a w_t-type woman's optimal search decision is given by:

$$V_{it}(w_t) = \max_{j \in \{MXR\}} P_t^j(w_t) \cdot \mu^j(w_t) + (1 - P_t^j(w_t)) \cdot E(V_{t+1}(w_{t+1})) + \epsilon_{it}^j(w_t), \tag{1}$$

where $E(V_{t+1}(w_{t+1}))$ is the unconditional expected-value function from the search problem in the following period, when agent i's type is w_{t+1}. Expressing the problem for each choice specific value function we have:

$$V_{it}^j(w_t) = P_t^j(w_t) \cdot \mu^j(w_t) + (1 - P_t^j(w_t)) \cdot E(V_{t+1}(w_{t+1})) + \epsilon_{it}^j(w_t). \tag{2}$$

The value function is expressed as a function of time varying types w_t (with analogous expressions for men as a function of m_t) only because individuals in our model age from one period to the next. Thus, the state transition probabilities in our model are degenerate and the rest of the type-space consists of permanent characteristics. For agents who are sufficiently old in t, we assume a terminal value function characterizes individuals utility. Since age is included in w_t an individual today who is sufficiently old will not participate in the matching market tomorrow, and instead receive a terminal value which we express as a function of their time-invariant type (w):

$$E(V_T(w_T)) = \beta X_w \cdot \tau_w. \tag{3}$$

where β is the discount factor. Here τ_w approximates the lifetime utility of reaching the post-marriage ages in T for women with time invariant-type w.

For agents who are younger and so face a matching market in subsequent periods the expected

[11] Rasul (2006) shows that when marriage markets are characterized by search, learning about marriage quality plays an important role determining the impact of divorce law changes. Brien, Lillard and Stern (2006) and Laufer and Gemici (2011) show that learning about partner quality is also important for rationalizing non-marital cohabitation. Svarer (2004) also presents evidence from Denmark in favor of learning through cohabitation. By contrast Indian marriage markets see virtually not pre-marital cohabitation and often spouses have no opportunity to learn about match quality prior to marriage. Across all of India 68% of women first meet their husbands on the wedding day, and another 9% have known him for less than a month Fulford (2013, Table 1).

value of the facing that market takes the following form:

$$E(V_{t+1}(w_{t+1})) = \beta \int_{\mathbf{P_{t+1}(w_{t+1})}} \int_{\varepsilon_{\mathbf{t+1}}(\mathbf{w_{t+1}})} V_{t+1}(w_{t+1})\, d\varepsilon dP, \tag{4}$$

where we omit i subscripts and expectations are both with respect to the vector of unobserved utility tomorrow ($\varepsilon_{\mathbf{t+1}}(\mathbf{w_{t+1}})$) and the vector matching probabilities across the $\{m, r\}$ types governing the market tomorrow for a type w_{t+1} woman, denoted $\mathbf{P_{t+1}(w_{t+1})}$.

We assume that the $\epsilon_{it}^{mr}(w_{t+1})$'s are i.i.d. Type-I Extreme Value errors and are unknown to the econometrician. This assumption allows for a closed-form representation of the expected value of facing the matching market tomorrow:[12]

$$E(V_{t+1}(w_{t+1})) = \beta \int_{\mathbf{P_{t+1}(w_{t+1})}} \left(\log \left(\sum_j e^{V_{t+1}^j(w_{t+1})} \right) + \gamma \right) dP, \tag{5}$$

where $V_{t+1}^j(w_{t+1})$ is the choice-specific value function under Equation (2) expressed in period $t+1$. We solve explicitly for the equilibrium probabilities of matching tomorrow for both sides of the market, obtaining $\mathbf{P_{t+1}(w_{t+1})}$ and $\mathbf{P_{t+1}(m_{t+1})}$, and impose rational expectations, replacing the integral in Equation (5) with the equilibrium probabilities which are consistent with decision making on both sides of the market at all time periods.

Imposing the logit-error structure and the probability of a w-type woman searching for a m-type man in an r-type relationship, $\phi_{w_t}^{mr}$ then follows a multinomial logit form:

$$\Pr(m, r | w_t) = \phi_{w_t}^{mr} = \frac{\exp\left(P_t^{mr}(w_t) \cdot \mu^{mr}(w_t) + (1 - P_t^{mr}(w_t)) \cdot E(V_{t+1}(w_{t+1}))\right)}{\sum_j^M \sum_k^R \exp\left(P_t^{jk}(w_t) \cdot \mu^{jk}(w_t) + (1 - P_t^{jk}(w_t)) \cdot E(V_{t+1}(w_{t+1}))\right)}, \tag{6}$$

which is a function of current and future probabilities of matching through the expected value terms.

3.2 Matching

We now specify the matching process within each period. The matching process is essentially a production function, taking as inputs the number searching men and women in each market and generating the number of matches in each market as an output.

We parameterize the number of matches, X_t, in market $\{m, w, r\}$ at time t as depending upon the number of single m-type men and w-type women searching in the market. Let N_t^m and N_t^w indicate the number of m-type men and w-type women respectively. Recall that $\phi_{m_t}^{wr}$ and $\phi_{w_t}^{mr}$ give the per-period probabilities, or shares, of m-type men and w-type women searching in market $\{m, w, r\}$. Thus, $\phi_{m_t}^{wr} N_t^m$ is the number of men of type m searching for women of type w on relationship terms r in time period t. The number of matches in market $\{m, w, r\}$ is then

[12]This requires the conditional independence assumption of Rust (1987) with respect to the distribution of unobserved utility and the expectations regarding prospects of matching in the future.

given by:[13]

$$
\begin{aligned}
X_t^{mwr} &= A^* \left[\frac{\left(\phi_{m_t}^{wr} N_t^m\right)^\rho}{2} + \frac{\left(\phi_{w_t}^{mr} N_t^w\right)^\rho}{2} \right]^{\frac{1}{\rho}} \\
&= A \left[\left(\phi_{m_t}^{wr} N_t^m\right)^\rho + \left(\phi_{w_t}^{mr} N_t^w\right)^\rho \right]^{\frac{1}{\rho}} \qquad (7)
\end{aligned}
$$

where ρ determines the elasticity of substitution $1/(1-\rho)$, and A measures search frictions. When $\rho \to 0$ the CES function becomes Cobb-Douglas, and as $\rho \to -\infty$ the CES function becomes Leontief. Note that $X_t^{mwr} = X_t^{wmr}$ for all m, w, and r.[14]

Under the assumption that all m-type men searching in the same market have the same probabilities of matching, $P_t^{mr}(w_t)$ is given by:

$$
\begin{aligned}
P_t^{mr}(w_t) &= \frac{X_t^{mwr}}{\phi_{w_t}^{mr} N_t^w} \\
&= \frac{A \left[\left(\phi_{m_t}^{wr} N_t^m\right)^\rho + \left(\phi_{w_t}^{mr} N_t^w\right)^\rho \right]^{\frac{1}{\rho}}}{\phi_{w_t}^{mr} N_t^w} \qquad (8) \\
&= A \left[\left(\frac{\phi_{m_t}^{wr} N_t^m}{\phi_{w_t}^{mr} N_t^w} \right)^\rho + 1 \right]^{\frac{1}{\rho}}.
\end{aligned}
$$

This term then enters into the multinomial logit probabilities of searching in particular markets and captures the influence of the sex-ratio on market search decisions today.

Given these representations of the probability of searching and matching for each type of agent, it is straightforward to express the flow of unmatched (young) individuals across periods. Namely at a particular point in time t we observe (N_t^m, N_t^m) (e.g. from the Indian census). In the subsequent period the number of agents participating in the market tomorrow will be:

$$
N_{t+1}^w = \sum_j^M \sum_k^R N_t^w (1 - P_t^{jk}(w)) \phi_{w_t}^{jk} + N_{t+1}^{w,n_g} \qquad (9)
$$

$$
N_{t+1}^m = \sum_j^W \sum_k^R N_t^m (1 - P_t^{jk}(m)) \phi_{m_t}^{jk} + N_{t+1}^{m,n_g} \qquad (10)
$$

$$
\qquad (11)
$$

[13] For ease of exposition we are assuming an interior solution such that the number of matches produced is less than both the number of men and the number of women in the $\{m, w, r\}$ market. In practice, we nest the CES matching function into a Leontief function to constrain the number of matches to be less than the number of searching men and women:

$$
X_t^{mwr} = \min \left\{ A \left[(\phi_{m_t}^{wr} N_t^m)^\rho + (\phi_{w_t}^{mr} N_t^w)^\rho \right]^{\frac{1}{\rho}}, \phi_{m_t}^{wr} N_t^m, \phi_{w_t}^{mr} N_t^w \right\}.
$$

[14] The interior share parameter is normalized to be one-half. Identification of ρ and A are discussed below. It is unclear which moments in the data would identify the share parameter in our framework, so we normalize it.

where $(N^{m,ng}_{t+1}, N^{w,ng}_{t+1})$ are the next generation of individuals who enter the market in the subsequent period, and the first terms correspond to the total unmatched individuals from the market at t.[15]

3.3 Equilibrium

The probabilities of searching in a particular market, the ϕ's, give the share of a particular set of individuals who will search in a particular market. These ϕ's also determine the probabilities of matching, the P's. We rewrite Equation (6) to make the dependence of $P^{mr}_t(w_t)$ on $\phi^{mr}_{w_t}$ and $\phi^{wr}_{m_t}$ explicit. In each period, except the agent's terminal market, the choice probability is a function of both current probabilities of matching $P^{mr}_t(w_t)$ and the vector of future matching probabilities, $\mathbf{P_{t+1}}$. This takes the following form:

$$\phi^{mr}_{w_t} = \frac{\exp\left(P^{mr}_t(w_t, \phi^{mr}_{w_t}, \phi^{wr}_{m_t}) \cdot \mu^{mr}(w_t) + (1 - P^{mr}_t(w_t, \phi^{mr}_{w_t}, \phi^{wr}_{m_t})) \cdot E(V_{t+1}(w_{t+1}), \mathbf{P_{t+1}}(\mathbf{\Phi_t}, \mathbf{P_t}, \mathbf{\Phi_{t+1}})\right)}{\sum_j^M \sum_k^R \exp\left(P^{jk}_t(w_t, \phi^{jk}_{w_t}, \phi^{wk}_{jt}) \cdot \mu^{jk}(w_t) + (1 - P^{jk}_t(w_t, \phi^{jk}_{w_t}, \phi^{wk}_{jt})) \cdot E(V_{t+1}(w_{t+1}), \mathbf{P_{t+1}}(\mathbf{\Phi_t}, \mathbf{P_t}, \mathbf{\Phi_{t+1}})\right)}. \tag{12}$$

Here the future value of facing the market tomorrow depends on the expected market conditions tomorrow $\mathbf{P_{t+1}}$. $\mathbf{P_{t+1}}$ is a function

1. current probabilities of searching: $\phi^{mr}_{w_t}, \phi^{wr}_{m_t}$,
2. the current probabilities of matching: $P^{mr}_t(w_t), P^{wr}_t(m_t)$,
3. next-period search probabilities: $\phi^{mr}_{w_{t+1}}, \phi^{wr}_{m_{t+1}}$.

Dependencies (1) and (2) occur through the flow conditions defined in (9), while point (3) comes from the equilibrium tomorrow. We collect these three terms into the vectors: $(\mathbf{\Phi_t}, \mathbf{P_t}, \mathbf{\Phi_{t+1}})$ whose elements respectively contain the elements of (1) through (3).

For the same cohort making decisions in their final matching market the choice probabilities generate another set of equations involving the elements of $\mathbf{P_{t+1}}$, namely:

$$\phi^{mr}_{w_t} = \frac{\exp\left(P^{mr}_t(w_t, \phi^{mr}_{w_t}, \phi^{wr}_{m_t}) \cdot \mu^{mr}(w_t) + (1 - P^{mr}_t(w_t, \phi^{mr}_{w_t}, \phi^{wr}_{m_t})) \cdot \beta X_w \tau_w\right)}{\sum_j^M \sum_k^R \exp\left(P^{jk}_t(w_t, \phi^{jk}_{w_t}, \phi^{wk}_{jt}) \cdot \mu^{jk}(w_t) + (1 - P^{jk}_t(w_t, \phi^{jk}_{w_t}, \phi^{wk}_{jt})) \cdot \beta X_w \tau_w\right)}. \tag{13}$$

Within each period the choice probabilities must sum to one for both men and women, so equilibrium in our model is characterized by stacking the $(W \times R - 1)$ and $(M \times R - 1)$ shares for each period and solving for the fixed point defined by the set of Equations (12) and (13). Since ϕ is a continuous mapping on a compact, convex space, Brouwer's fixed point theorem

[15] Note that with this formulation it is straightforward to adjust the stocks of agents participating in the market tomorrow (for instance because of out-migration or mortality), by simply adding another scaling of (N^w, N^m).

guarantees that an equilibrium exists.[16] We also note that because the decisions are based on expected utility and probabilistic search the stability of the resulting matching will not hold.

Finally, if one were to recursively substitute next-period equilibrium expressions of Equation (12) for a given cohort in the non-terminal period, the equilibrium governing the market tomorrow (through $\mathbf{P_{t+1}}$) will be a function of the choice probabilities of the next generation market participants. This is because despite having a non-stationary dynamic decision problem with terminal value functions, we also model equilibrium. Thus, for example, a cohort making search decisions at $T-2$ will be influenced by the future decisions (at $T-1$) of agents who have yet to enter the market at $T-2$. Those beginning the matching process in the subsequent period, whose populations were given by $N_{t+1}^{w,ng}, N_{t+1}^{m,ng}$, would have choice probabilities that are in-turn influenced by their expectations about decisions by the generation entering at T. In order to avoid this problem and solve the equilibrium we require a simplifying assumption, given that we cannot observe the entire history of matches (e.g. at some point we must stop this "overlapping" generations problem). We assume the market after the next-generation proceeds in a stationary manner, such that each element of $\mathbf{P_{t+2}} = \mathbf{P_{t+1}}$. Thus, when an agent at $T-2$ looks forward to equilibrium at $T-1$ they further assume that the match probabilities operating at $T-1$ will also govern equilibrium three periods from now at T. This approach allows us to not explicitly model the decisions of the next-generation, facilitating estimation with the available data, which is essentially a retrospective cross-section.[17]

4 Data and Descriptive Analysis

4.1 Data

For our empirical analysis, we combine 2001 Indian Census data of population by district, gender, age and caste with individual survey data from the 2005 India Human Development Survey (IHDS-I).[18] The IHDS-I is a nationally representative survey of 41,554 households (215,754 individuals) in 1,503 villages and 971 urban neighborhoods across India (Desai et al., 2008). The survey covers thirty-three states and union territories of India with the exception of small populations living in the island states (Andaman & Nicobar and Lakshadweep) and contains standard socio-economic and demographic information at the household and individual level. It includes a women's questionnaire asking ever-married women in the age group of 15-49 years

[16] As in macro models of the labor market, there is only one equilibrium where the search probabilities are positive in all markets in the static game. Diamond (1982) shows a necessary condition for multiple equilibria (with positive search probabilities) in a similar, but static, model (with endogenous search on both sides of the labor market) is increasing returns to scale in the matching technology. There are other equilibria of the static game that result from coordination failures where certain markets are empty. We focus on the static equilibrium where the the search probabilities are strictly positive in all markets. Given that no state variables other than age carry over into subsequent periods, we know for certain there exists a dynamic equilibrium which sees the interior static equilibrium in each period.

[17] The individual level data we use includes retrospective histories including marriage dates.

[18] District are the main administrative division of an Indian state or territory. In year 2001, India comprised 28 states and 7 union territories, which were divided into 593 districts.

a wide range of questions about education, health, income, and consumption patterns. Most importantly for the purpose of this paper, it also include information about gender relations, marriage practices, and marital history. Based on these questions, we construct measures of women's participation in their marriage decision, of their migration upon marriage, and a proxy for their dowry size. While the survey does not include direct questions about dowry payments at marriage, respondents are asked what is the amount of money usually spent by the girl's family and the boy's family at the time of marriage, in their community (jati) for a family like theirs.[19] In addition, we observe women's year of marriage and information on male and female education and caste.

We discretize our variables of interest as follows. $Involved$ is an indicator variable equal to 1 if the respondent answers "Respondent alone" or "Respondent and parents/other relatives together" to the question "Who choose your husband?"; Far is equal to 1 if the respondent reports having moved more than four hours away from her natal family at the time of her marriage; $HighDowry$ is equal to 1 if the bride's family's contribution at the time of marriage is at least 10 percent higher than that of the groom's family. Based on the schooling information, we create indicators for men and women for completing primary school ($Educ$). Table 2 contains some descriptive statistics.

According to the Census of India, the Indian population in 2001 was about 1.3 billion, with males outnumbering females by 35 million. 16 percent of the population belongs to Scheduled Castes (SCs) and almost 8 percent to Scheduled Tribes (STs). SCs generally consist of the *untouchables*, while the STs include a heterogeneous set of ethnic and tribal groups claimed to be the aboriginal population of India. For most states, the Census provides population data by gender, age group and marital status for the overall population, and for STs and SCs separately.[20] We therefore identify three main demographic groups, i.e., SC, ST, and not SC/ST. While this is far from fully capturing endogamous groups, it represents a first attempt to account for the fact that marriage markets are segregated both geographically and by caste.[21] For the reminder of the paper, we will use the terms caste or demographic group interchangeably.

The population breakdown by age, caste, gender, district and education level is not available from the 2001 Census of India. We achieve this higher level of disaggregation by estimating the probability of completing primary school for men and women conditional on their age, caste, marital status and district of residence using the IHDS-I data and a logit model.[22] Analogously,

[19]Given that dowries are illegal in India, the use of indirect questions aims at reducing problems related underreporting.

[20]Separate Census data for SCs and STs are not available for the following states: Haryana, Delhi, Punjab, Chandigarh, Nagaland, Sikkim, Mizoram, and Pondicherry.

[21]Indian men and women are normally bound to marry within their castes of birth. According to the IHDS-I, less than 6 percent of female respondents married men of different castes. In India, there are approximately 3,000 castes and 25,000 subcastes. The IHDS-I categorizes individuals in Brahmin, Other Backward Castes (OBCs), SCs, STs, and Other. To match the Census data, we combine Brahmin, OBCs, and Other in the not SC/ST category.

[22]Specifically, for individual i, of age a (5-year age groups), demographic group c (SC, ST and not SC/ST), of marital status m (currently married/not married), living in district d. we estimate the following logit model: $Educ_{iacmd} = \alpha_a + \alpha_c + \alpha_m + \alpha_{am} + \alpha_d + \epsilon_{iacmd}$. We estimate education probabilities for men and women,

we estimate the likelihood of women to be involved in their wedding decisions. This allows us to categorize women according to the cross-product of discrete measures of age, education, marital status, caste, and independence (i.e., participation in the choice of their husbands). For men, we use the cross-product of age, education, caste, and marital status. Table 3 highlights the actual distribution of men and women by age, caste, and marital status. Tables 4 and 5 report the estimated probabilities of completing primary school and of participating in the husband's choice by age and caste and, when available, by gender.

We match the survey respondents with the district-level census data. The women's questionnaire is administered to approximately 33,500 ever-married women in 373 districts.[23] The vast majority of these women (72 percent) are wives of heads of household.

4.2 Sex-Ratios and Marriage Terms: A Descriptive Analysis

Individuals may trade-off the probability of matching with partner's characteristics and the terms of the marriage. We here exploit the sex-ratio variation across marriage markets to shed light on these possible substitution patterns. Specifically, we investigate how the relationship terms of interest (i.e., migration distance, dowry size, and women's participation in spouse selection) relate to the sex-ratio (i.e., the number of women to men) in the relevant marriage market, which is defined by district and demographic group.

We estimate logit regressions for $Involved$, Far and $HighDowry$, conditional on the spouses' characteristics (i.e., demographic group, education, and age). To take into account the age gap at marriage between women and men (on average equal to 5 years in our sample) and to incorporate the potential dynamics of the decision process, we estimate the logit models using the female to male ratio in different age groups. We add state and state-demographic group fixed effects to control for potential unobserved heterogeneity. The logit models are estimated using data on married women aged 15 to 24 at the beginning of the year 2001 who got married between 2001 and 2005. Table 6 reports the average marginal effects of sex-ratios and spouses' characteristics on the terms of marriage.[24] Sex-ratios seem to matter for women's involvement in marriage decisions. The higher is the number of females to males the lower is women's probability of being involved in their spouse choice. While we cannot attach any causal interpretation to this result, it might indicate that the lower the probability of matching due to the higher competition, the higher is the inclination to arranged marriages. In addition, the higher is the relevant sex-ratio, the significantly lower is the probability of migrating at the time of marriage. By contrast, we find much weaker associations between dowry size and sex-ratios.

We hypothesize that future brides and their parents may have different substitution patterns and thus may respond differently to the marriage market conditions. We test this hypothesis

separately.

[23] 220 districts are not covered in the survey, while 2 districts (both in Delhi) are in the 2005 IHDS-I but not in the 2001 Census data. Thus, the match is successful for ever-married women in 371 districts.

[24] These findings are robust to using continuous variables for the distance from the natal family upon migration and for our proxy of dowry payments. Results are available upon request.

by allowing the relationship between Far (or $HighDowry$) and sex-ratios to differ depending on whether women are involved in the choice of their husbands. We present the results in Table 7. No significant difference exists when it comes to preferences regarding migration upon marriage. By contrast, when we look at dowry payments, women who participate in decision making regarding their marriages make significantly different tradeoffs than other women when faced with competitive marriage markets. Accordingly, in our model we allow preferences over partner and marriage characteristics to vary with women's participation in spouse selection.

5 Estimation

We begin the estimation section with a brief heuristic discussion of identification.[25] The parameters of the matching function we estimate, ρ and A, are identified by covariation in sex-ratios (N^w/N^w) and the discrete partner choices, and the overall match rates across districts respectively. The terminal value parameters (τ_w, τ_m) are identified by the willingness of different individuals to choose partners with low probabilities of matching in their final marriage market (e.g. the matching rates conditional $t = T - 1$ as distinct from $t < T - 1$). For instance we expect to see higher τ_m for men than for women, since male marriage rates are still relatively high past the age we stop modeling decisions (see Figure 1). Finally, in general the μ terms are identified as in a traditional discrete choice model, by covariation between the observed choices and individual and partner characteristics. To separately identify $\mu^{mr}(w_t)$ from $\mu^{wr}(m_t)$, in particular with respect to relationship terms r, requires variation in the choice set. This variation comes from estimating the model across the district-caste level in India, exploiting geographic and caste-divisions which generate different numbers of singles across markets. This approach allows us to see how type-w women or type-m men substitute toward different terms r and partner characteristics m and w when facing differing supplies of available partners.[26]

5.1 Estimation

[Note: so far, estimates obtained using not SC/ST only.] We estimate our model by combining the 2005 IHDS-I micro data and the aggregate information on the population of single men and women across districts in the 2001 Census of India. From IHDS-1 we observe whether individuals did not match in 2001 (that is, if they were un-married in the beginning of 2001 but between the ages of 15 and 24) as well as whether they matched by 2005. Importantly, we also observe male and female education, caste and the distance moved to reside with ones' spouse, along with whether the young woman had any input into the selection of her spouse. With these, we

[25]See Hsieh (2012) for a more detailed treatment identification of similar static models. Expanding the logit model to include dynamics does not appreciably change the identification discussion: the functions on the right-hand side of Equations (12) and (13) are now "more" non-linear and include more probabilities of matching (both current and future periods). In terms of the model primitives adding dynamics is the same as expanding the type-space (to include time matched) and changing the utility function to be more non-linear (it includes discounted future value terms which are functions of utility parameters as well).

[26]We set the discount factor at $\beta = .95$.

construct the type-space W which is the cross-product of binary measures of age, education, and independence (eight elements). For men, M is the cross-product of age and education (four elements). We allow R to take one of two values corresponding to whether the woman moved more or less than a four hour journey from her native village, so we classify matches as belonging to one of 64 groups within each period.

The major challenge to estimation is solving for the equilibrium probabilities of matching across time. We simplify the dynamic aspects of the model by assuming individuals only match during two periods which correspond to being ages 15 to 19 and 20 to 24 for women and 20 to 24 and 25 to 29 for men. This both corresponds to the Indian Census data, which only releases population estimates within those age bands, and allows for simplifying the number of future periods which must be explicitly solved for within each likelihood iteration. Given a sample of individuals who matched at younger and older ages in 2001 and older ages in 2005, we can express the likelihood contribution as a function of own and partner characteristics and the aggregate distributions of each type of single individual within each district in 2001 and 2005. The aggregate number of singles in 2005 will incorporate matching prior to 2005 and the maturation of the cohort who was aged 10 to 14 in 2001.

Given an initial vector of parameters, we construct the likelihood based on Equations (12) and (13). The likelihood contribution for an individual who matches in a given period is the product of the search and matching probabilities, while for an individual who did not match, we integrate out over the potential (unobserved) search decisions. Thus, in each period (for agents who have never successfully matched) we can express the log-likelihood as:

$$\begin{aligned} L_{itw}(\theta) \quad &= \quad \left[\prod_m \prod_r \left([\phi_w^{mr}(\theta)] \times [P_w^{mr}(\theta)]\right)^{I(d_{itw}=\{m,r\})}\right]^{I(y_{itw}=1)} \\ &\quad \times \left[\prod_m \prod_r \phi_w^{mr}(\theta) \times (1 - P_w^{mr}(\theta))\right]^{I(y_{itw}=0)} \end{aligned} \tag{14}$$

where y_{itw} is the binary indicator that a woman matched in period t and d_{itw} indicates which type of partner she matches with, and θ denotes the vector of matching function and utility parameters. The explicit functional form of ϕ_w^{mr} is drawn from either Equation (12) or Equation (13) and also depends on the year marital status was observed $t = 2001$ or $t = 2005$. Each likelihood iteration then involves specifying the initial vectors of $(\mathbf{P_t}, \mathbf{P_{t+1}})$ corresponding to $t = 2001$ and $t + 1 = 2005$ respectively. We then proceed by explicitly solving the equilibrium defined by equations (12) and (13), repeating until the probability vectors across time, type and district converge.

Finally, we additionally include unobserved heterogeneity governing female decision making, in which case the likelihood is written conditional on the unobserved state:

$$\mathcal{L}_w(\theta) = \prod_t \prod_i \Bigg(\sum_k p(k|x_i, \theta_k) L_{itw}(\theta_k) \Bigg). \tag{15}$$

A woman's right to inherit land and other property play a significant role in determining women's decision power. Inheritance rights in India differ by religion and since 1956 they are governed by the Hindu Succession Act (HSA).[27] The HSA only applied to Hindus, Buddhists, Sikhs, and Jains and established a law of succession whereby sons and daughters would enjoy (almost) equal inheritance rights. Gender inequalities, however, remained even after the introduction of the HSA. In case of a male dying intestate, i.e., without leaving a will, all his separate or self-acquired property, devolved equally upon sons, daughters, widow, and mother. By contrast, the deceased's daughters had no direct inheritance rights to joint family property. In the decades following the introduction of Hindu Succession Act, state governments passed amendments that equalized inheritance rights for daughters and sons (Kerala in 1976, Andhra Pradesh in 1986, Tamil Nadu in 1989, and Maharashtra and Karnataka in 1994). A national-level ratification of the amendments occurred in 2005. These amendments only applied to Hindu, Buddhist, Sikh or Jain women, who were not yet married at the time of the amendment. We use exposure to these reforms as a shifter x_i which enters the probability of being a different unobserved type but which is otherwise uncorrelated with decision making. We model the unobserved heterogeneity as being of two types, with the HSA amendments shifting the probability of being in type $k = 1$. In a similar vein, we use information on the height of the woman as an additional source of unobserved heterogeneity. Unfortunately, anthropometric indicators are missing for approximately half of the women in our sample.

6 Model Estimates and Fit

Estimates of the model parameters, including utility, matching function and unobserved heterogeneity parameters are presented in Table 8. We include three versions of the model, a baseline version without unobserved heterogeneity (listed under model (i)), which generally struggled to match the overall match rate, the fraction of matches involving less migration, and first-period matching statistics which we discuss in Table 9. With the inclusion of unobserved heterogeneity (listed under models (ii) and (iii)), fit improved substantially. Model (iii) is our preferred specification, and we can reject it relative to Model (ii) via a likelihood ratio test. Model (iii) includes a utility shifter which is the interaction between wife's age, whether the husband lives near, and the unobserved heterogeneity, which improved the areas mentioned above where fit was poor.

The upper panel includes the matching function and terminal period parameters. ρ captures the degree to which the sex-ratio is correlated with the decision to search for a particular partner, and the estimates reveal that sex-ratios do significantly effect partner sorting decision. The A-parameter captures the average match-rate, on a per-period basis, across districts. The estimates reveal both men and women prefer to avoid not-matching, though men have a slightly stronger preference for marriage, or at least marriage prior to age 30.[28] Especially important

[27] See Agarwal (1994) for more details.

[28] The female marriage window runs from age 15 to 24, while the male marriage window runs from age 20 to

is the difference between terminal values when including unobserved heterogeneity, which helps substantially in fitting both the first and second period match rates.

The middle panels show male and female preferences for partner characteristics and relationship terms. Accounting for unobserved heterogeneity in the preferences of women significantly alters the conclusion one draws from these preference estimates. For male preferences, the three models present very similar preference parameter estimates, with the exception for the value over a wife's education level and age. We estimate that men have a strong preference for less educated wives. Regardless of whether we control for unobserved heterogeneity, men are estimated to have stronger preferences for wives from nearby, and dislike both small dowries and female autonomy.

On the female side of the market, women value having a husband from nearby, and the unobserved heterogeneity reveals that this preference is especially strong among young women. We find that wives value larger dowries more than men. Accounting for unobserved heterogeneity among women shrinks the negative effect of female autonomy, a factor that may reflect differences in custom, given that controlling for Hindu Succession Act exposure (an important shifter in the unobserved heterogeneity) has a such a large effect on the female preferences autonomy. Finally, the female preference for husband age is substantially larger than male preferences. In terms of education, which we measure as primary school attendance, women on average put little value on the education of the husband once we account for heterogeneity.

Given the non-linear nature of our decision problem, not all of the estimates are directly interpretable. In Table 9 we compare the predicted matching probabilities from our model with those observed in our sample. The table reveals the baseline Model (i) struggles to hit the overall match rate and the probability of matching with a partner who lived near the woman's village, but this only holds in the first period. The amended model that includes unobserved heterogeneity does a much better job of matching the terms-distribution in both periods. Accounting for unobserved heterogeneity thus helps shrink the gap between the model predictions and observed sample means, especially in the first period. Fit is still less than ideal with respect to predicting the number of matches to nearby partners and the female-age profile. Nonetheless we turn to simulating a number of counter-factual environments using Model (iii).

7 Counterfactual Simulations

Given model parameters that fit the features of the terms' and characteristics' distributions fairly well, we turn to examining a series of counterfactual environments to see how female and male choices respond in equilibrium. First, we examine how a widening or shrinking sex-ratio affects equilibrium decision making and welfare, with attention to whether female autonomy, or lack thereof, mitigates some of the welfare gain. Second, we examine how universal schooling for young women and universal female autonomy in marital decision making alter the match

29, reflecting the ages where the vast majority of men and women married.

distribution.

Table 10 presents the probabilities of observing different relationships within counter-factual matching environments. This table presents the overall match probability which can change both because of partner substitution and because of differences in the populations of available partners or competitors. In column (2) we can see that increasing by one standard deviation the number of females within a district reduces the probability that a given woman matches overall, but the effects are disproportionately born in the first period, so one strategic response among women is to delay marriage.[29] In general, changes to the probability of matching on specific terms or with specific partner characteristics are smaller in the second period. This means, for instance, that women find it optimal when faced with more competition for husbands age 25 to 29 in period one, to wait until period two. In column three, we see the opposite pattern which flows from the same magnitude decrease in the percent female. We see substantially higher match rates in both periods. There are more matches from nearby, increased autonomy, older husbands and more educated spouses. To put these changes in a clearer context, we also simulate the model under the assumption that all districts faced the same sex-ratio as state with both a relatively low (Chahattisgarh) and a relatively high (Maharastra) sex-ratio.[30] The results are largely similar to our one standard deviation change in sex-ratios, but the changes are generally somewhat smaller overall. This is because sex-ratios at the district level are more variable than at the state level. An important point to note is that the increases or decreases in sex-ratios do not generate symmetric changes in the match distribution. For example, a one-standard deviation increase lowers the fraction matching educated husbands by four percentage points, while the same decrease only increases it by 1.5 percentage points.

In Table 11 we present the probability of searching in a given market conditional on having matched for the same counter-factual simulations, which helps eliminate the overall matching effects and allows one to focus on the substitution effects in each counter-factual. Again we note the asymmetry: increased sex-ratios lower local matching much more than a comparable decrease in sex-ratios increases near matching. This is capturing partner substitution. The largest changes from increased sex-ratios is female matching with younger men (reflected in the fraction matching older men (25 to 29) dropping by 9 percentage points). When the sex-ratio tilts towards women (decrease in percent female), one of the largest changes is also an increase in female autonomy in decision making, a point we highlighted above in the descriptive analysis.

We now turn to changing the characteristics distribution for women and altering the choice set. Table 12 shows the overall probability of matching impacts for universal female education (primary school completion) and allowing all women some say in marriage decisions. Both results show dramatically lower match rates in both the first and second period. In the baseline simulation roughly one third of all matches (0.135) involved female participation, column three

[29]This approach adds the standard deviation of the percent-female across districts in each age-education cell, while holding population constant.

[30]In the 2011 Indian Census Chahattisgarh ranked 27th for child sex-ratios and Maharastra ranked 4th. We choose these states because we have a largely representative sample of districts from these states.

requires female participation in all decisions. The net result is a decline in all types of marriages; similarly universal education lowers matching across terms and characteristics. These results reflect the strong male preferences against female autonomy and female education. To get a better sense of how terms changed for those who did match, we again examine the probability of searching in given market conditional on matching in Table 13. Here we see a number of effects of universal education: dowries fall, and female autonomy rises along with husband education in the first period and to a less extent in the second period. These come at a cost of women matching substantially younger men in the second period. The increase in female participation did not change the distribution of terms or characteristics much, with the exception of lowering the average husband's age in second period.

In Table 14 we examine how the fraction unmatched and the age-gap evolve in each counterfactual. As hinted at above, autonomy and universal education raised the fraction who remain unmatched at age 25 substantially. Similarly, increasing the fraction female (either by a standard deviation, or by setting sex-ratios to Maharastra's levels) increases the number of women unmatched. The age gap is largely insensitive to sex-ratio changes, with the exception that an increase in the fraction female lowers the age gap by more than half a year. Under universal female participation and education, the second period age gap declines dramatically.

Finally incorporating all these changes into one metric, we calculate the ex-ante expected utility of facing the choice set as a 15 to 19 year old woman at the outset of the model. These are presented in Table 15, the first column of which presents dollar denominated utils using the convenient logit-consumer surplus functional form, and assuming the marginal utility of wealth is 1.44.[31] To measure how welfare changes across simulated environments we calculate the percentage point change in dollar denominated utils. In the second and third columns we calculate utility for the one-standard deviation sex-ratio changes. One key result from the simulations is that second period utility is generally more sensitive to policy changes: an increase in the sex-ratio raises utility more in the second period, and the decrease in the sex-ratio lowers it more in the second period relative to the first. This makes sense since gains and losses from a changed matching environment can be smoothed out across periods through marriage delay. A second point is that heterogeneity in preferences has an impact on whether women gain or lose from sex-ratio changes: uneducated women of unobserved-type $k = 1$ actually lose out from a fairly large global decrease in the sex-ratio. This is can happen because of the confluence of preference differences and competition for spouses: the decreased likelihood of matching overall (from increased competition) is overcome by an increased value in the case of a match for these women. The same pattern arises in the Chhattisgarh-Maharastra simulation. Imposing universal education and female participation has larger welfare effects than sex-ratio changes,

[31] Layard, Mayraz and Nickell (2008) present a tight range of estimates for this parameter, we use their estimate for singles. The logit form is:

$$E(CS_i) = \frac{1}{\alpha_i} log\left(\sum_j^J e^{V_i^j}\right) \tag{16}$$

where V_i^j is given in Equation (2), for period one.

and according to our estimates both policies universally reduce female welfare. For education this arises because universal education forces women to compete in a fewer number of marriage markets, markets that men prefer not to match in. As described above this means women must compromise on terms: they match with men from farther away and receive smaller dowries, which on net reduces welfare. Additionally heterogeneity matters here as well: welfare losses are almost entirely concentrated among women with $k = 0$. Finally, universal female autonomy is difficult to interpret here, since a woman's' household may be making these decisions on her behalf or jointly with her. Universal autonomy again reduces welfare much more in the second period than in the first, but again reflects the fact the women now must compete with one another in a market which men prefer not to match in.

8 Conclusion

We formulate and estimate a dynamic two-sided matching model which allowed us to uncover separate preferences for men and women. Our estimates reveal strong female preferences for living near ones native village, even more so than the male preference for proximity. We also find that female preferences for larger dowries is larger than males, casting doubt on the view of dowry as a bride price, and supporting the notion that dowry can be a form of transfer to a young woman. Importantly, once we include unobserved heterogeneity in our model, we find men have strong preferences against female autonomy and education. In our simulation, this translates directly into welfare losses among women from universal primary school and universal female participation in marriage decisions. Incorporating unobserved heterogeneity also shows that female participation in marriage decisions is an important driver in female preferences over their husbands age and education: after accounting for heterogeneity we find women prefer to marry older partners, but do not value education. Our counter-factual simulations reveal that the erosion of sex-ratios for women in India impacts women primarily through the availability of more partners and through marriage delay. Heterogeneity reveals that even increases in the sex-ratio are not universally welfare improving for women.

References

Agarwal, Bina, *A Field of One's Own*, Cambridge Books, 1994.

Aguirregabiria, Victor and Pedro Mira, "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, May 2010, *156* (1), 38–67.

Anderson, Siwan, "Why Dowry Payments Declined with Modernization in Europe but Are Rising in India," *Journal of Political Economy*, April 2003, *111* (2), pp. 269–310.

_ , "The Economics of Dowry and Brideprice," *The Journal of Economic Perspectives*, 2007, pp. 151–174.

_ **and Chris Bidner**, "Property Rights over Marital Transfers," *The Quarterly Journal of Economics*, 2015, *130* (3), 1421–1484.

_ **and Debraj Ray**, "Missing Women: Age and Disease," *The Review of Economic Studies*, 2010, *77* (4), 1262–1300.

_ **and** _ , "The age distribution of missing women in India," *Economic and Political Weekly*, 2012, *47* (47-48), 87–95.

_ **and** _ , "Missing Unmarried wWmen," 2015. Manuscript.

Anukriti, S, Sungoh Kwon, and Nishith Prakash, "Dowry: Household Responses to Expected Marriage Payments," 2016. Manuscript.

Arcidiacono, Peter, Andrew Beauchamp, and Marjorie McElroy, "Terms of Endearment: An Equilibrium Model of Sex and Matching," Boston College Working Paper 813 2014.

Banerjee, Abhijit, Esther Duflo, Maitreesh Ghatak, and Jeanne Lafortune, "Marry for What? Caste and Mate Selection in Modern India," *American Economic Journal: Microeconomics*, May 2013, *5* (2), 33–72.

Bhat, P. N. and A. J. Francis Zavier, "Factors Influencing the Use of Prenatal Diagnostic Techniques and Sex Ratio at Birth in India," *Watering the Neighbour's Garden: The Growing Demographic Female Deficit in Asia*, 2007, pp. 109–130.

Bloch, Francis, Vijayendra Rao, and Sonalde Desai, "Wedding Celebrations as Conspicuous Consumption: Signaling Social Status in Rural India," *The Journal of Human Resources*, Summer 2004, *39* (3), pp. 675–695.

Botticini, Maristella and Aloysius Siow, "Why dowries?," *American Economic Review*, 2003, pp. 1385–1398.

Brien, Michael J., Lee A. Lillard, and Steven Stern, "Cohabitation, Marriage, And Divorce In A Model Of Match Quality," *International Economic Review*, 05 2006, *47* (2), 451–494.

Calvi, Rossella, "Why Are Older Women Missing in India? The Age Profile of Bargaining Power and Poverty," 2016. Manuscript.

Cherchye, Laurens, Bram De Rock, Selma Telalagic-Walther, and Frederic Vermeulen, "Where Did It Go Wrong? Marriage and Divorce in Malawi," 2016. Manuscript.

_ , Thomas Demuynck, Bram De Rock, and Frederic Vermeulen, "Household Consumption When the Marriage is Stable," *American Economic Review*, forthcoming.

Chiappori, Pierre-André, Monica Costa Dias, and Costas Meghir, "The Marriage Market, Labor Supply and Education Choice," Technical Report, National Bureau of Economic Research 2015.

_ , Sonia Oreffice, and Climent Quintana-Domeque, "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market," *Journal of Political Economy*, 2012, *120* (4), 659 – 695.

Choo, Eugene and Aloysius Siow, "Who Marries Whom and Why," *Journal of Political Economy*, 2006, *114* (1), pp. 175–201.

Desai, Sonalde, Reeve Vanneman, and National Council of Applied Economic Research, "India Human Development Survey (IHDS) 2005," Inter-university Consortium for Political and Social Research (ICPSR) May 2008. Accessed 12 November 2011.

Diamond, Peter A., "Aggregate Demand Management in Search Equilibrium," *Journal of Political Economy*, 1982, *90* (5), pp. 881–894.

Edlund, Lena, "The Price of Marriage: Net vs. Gross Flows and the South Asian Dowry Debate," *Journal of the European Economic Association*, April-May 2006, *4* (2/3), pp. 542–551.

Fulford, Scott, "The Puzzle of Marriage Migration in India," Working Paper 820, Boston College 2013.

Guilmoto, Christophe Z., "Skewed Sex Ratios at Birth and Future Marriage Squeeze in China and India, 2005–2100," *Demography*, 2012, *49* (1), 77–100.

_ and Roger Depledge, "Economic, Social and Spatial Dimensions of India's Excess Child Masculinity," *Population (English Edition)*, 2008, *63* (1), pp. 91–117.

Hsieh, Yuwei, "Understanding Mate Preferences from Two-Sided Matching Markets: Identification, Estimation and Policy Analysis," Working Paper, University of Southern California August 2012.

Laufer, Steve and Ahu Gemici, "Marriage and Cohabitation," 2011 Meeting Papers 1152, Society for Economic Dynamics 2011.

Layard, R., G. Mayraz, and S. Nickell, "The marginal utility of income," *Journal of Public Economics*, August 2008, *92* (8-9), 1846–1857.

Maertens, Annemie, "Social Norms and Aspirations: Age of Marriage and Education in Rural India," *World Development*, 2013, *47*, 1–15.

Rao, Vijayendra, "The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India," *Journal of Political Economy*, August 1993, *101* (4), pp. 666–677.

Rasul, Imran, "Marriage Markets and Divorce Laws," *Journal of Law, Economics and Organization*, April 2006, *22* (1), 30–69.

Rosenzweig, Mark R. and Oded Stark, "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India," *Journal of Political Economy*, August 1989, *97* (4), pp. 905–926.

Rust, John, "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, September 1987, *55* (5), 999–1033.

Salani, Bernard and Alfred Galichon, "Cupid's Invisible Hand: Social Surplus and Identification in Matching Models," Sciences Po publications info:hdl:2441/5rkqqmvrn4t, Sciences Po September 2012.

Sautmann, Anja, "Partner Search and Demographics: The Marriage Squeeze in India," Working Papers 2011-12, Brown University, Department of Economics 2011.

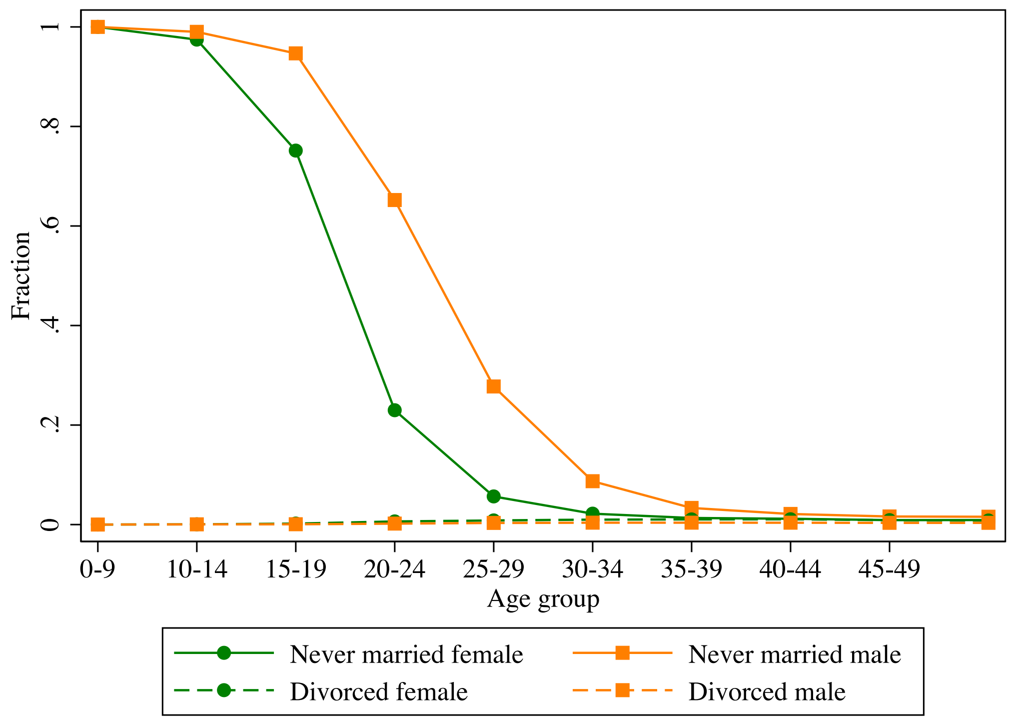
Sen, Amartya, "Missing Women: Social inequality outweighs women's survial advantage in Asia and north Africa," *British Medical Journal*, March 1992, *304*, 587–588.

Svarer, Michael, "Is Your Love in Vain? Another Look at Premarital Cohabitation and Divorce," *The Journal of Human Resources*, 2004, *39* (2), pp. 523–535.

Trivedi, Ira, "The (new) business of marriage," *Hindustan Times*, 2011, *October 28.*

Vogl, Tom S, "Marriage Institutions and Sibling Competition: Evidence from South Asia," *The Quarterly journal of economics*, 2013, *128* (3), 1017–1072.

Figure 1: Marriage and divorce by age in India



Notes: Statistics include both urban and rural. Source: Author calculations from Census of India 2001 table C-2 "Marital Status by Age and Sex".

Table 1: Adult migration in India by origin and reason

	Women	Men
Migration by origin (percent)		
Never moved	24.0	85.0
Same district	49.1	5.2
Same state (diff. dist.)	16.2	2.9
Another state/country	10.8	6.9
Living in same district	73.0	90.2
Reason for migration if migrated (percent)		
Employment	1.2	60.8
Education	0.1	1.8
Displaced	0.4	2.9
Marriage	87.3	5.8
Accompany parents/family	8.6	10.8
Other	2.4	18.0

Notes: Only for women and men 25 years and older to allow for marriage migration. Weighted to be nationally representative. Includes both urban and rural. Source: Author calculations from National Sample Survey 64th round 2008, Employment/Unemployment.

Table 2: Descriptive Statistics (IHDS-1)

	Obs.	Mean	Med.	St. Dev.
Ever married women 15-49				
Distance from natal family (hrs)	32,653	3.51	2.00	6.45
Relative contribution (bride's to groom's family)	32,740	1.88	1.50	2.43
Woman's age	33,003	33.00	33.00	8.00
Woman's education (years)	33,003	4.52	4.00	4.80
Husband's age	30,447	38.13	38.00	8.96
Husband's education (years)	30,325	6.71	7.00	4.87
$Involved$	31,078	0.40	0.00	0.49
Far	33,003	0.26	0.00	0.44
$HighDowry$	33,224	0.79	1.00	0.40
$Educ$	33,003	0.48	0.00	0.50
Husband's $Educ$	33,003	0.70	1.00	0.46
Ever married women 15-29				
Distance from natal family (hrs)	11,524	3.47	2.00	6.49
Relative contribution (bride's to groom's family)	11,549	1.86	1.50	2.78
Woman's age	11,654	24.25	25.00	3.20
Woman's education (years)	11,654	5.32	5.00	4.80
Husband's age	11,030	29.30	30.00	4.63
Husband's education (years)	10,979	7.14	8.00	4.68
$Involved$	10,901	0.38	0.00	0.49
Far	11,654	0.25	0.00	0.43
$HighDowry$	11,549	0.79	1.00	0.41
$Educ$	11,654	0.56	1.00	0.50
Husband's $Educ$	11,654	0.74	1.00	0.44

Source: Authors' calculations from IHDS-1.

Table 3: Pr($Married$ | Gender, Age, Caste)

	SC		ST		not SC/ST	
	Males	Females	Males	Females	Males	Females
Age group:						
15-19	0.0646	0.2939	0.0634	0.2654	0.0489	0.2254
20-24	0.4051	0.8175	0.4296	0.7842	0.3223	0.7397
25-29	0.7624	0.9533	0.7771	0.9289	0.6966	0.9348

Note: Average across districts. Source: Authors' calculations from Census of India 2001.

Table 4: Pr($Educ$ | Gender, Age, Caste)

	SC		ST		not SC/ST	
	Males	Females	Males	Females	Males	Females
Age group:						
15-19	0.7780	0.6254	0.7111	0.5394	0.8282	0.7259
20-24	0.7800	0.6398	0.7144	0.5538	0.8292	0.7403
25-29	0.7596	0.5995	0.6902	0.5128	0.8124	0.7033

Note: *Educ* means completed primary school. Average across districts and marital status. Underlying probabilities are estimated with data from IHDS-I and a logit model. Source: Authors' calculations from Census of India 2001.

Table 5: Pr($Involved$ | $Educ$, Age, Caste)

	SC		ST		not SC/ST	
	$Educ = 0$	$Educ = 1$	$Educ = 0$	$Educ = 1$	$Educ = 0$	$Educ = 1$
Age group:						
15-19	0.4324	0.4586	0.4422	0.4685	0.4187	0.4447
20-24	0.4199	0.4459	0.4296	0.4558	0.4063	0.4320
25-29	0.4171	0.4430	0.4267	0.4528	0.4035	0.4292

Note: *Involved* means woman's participation in husband's decision. Average across districts. Underlying probabilities are estimated with data from IHDS-I and a logit model (married women only). Source: Authors' calculations from Census of India 2001.

Table 6: Sex-Ratios and Marriage Terms

	Sex Ratio			
	F15-19/ M20-24	F20-24/ M25-29	F15-24/ M20-29	F15-29/ M15-29
Involved (N=2,638)				
SR	-0.159**	-0.144	-0.178**	-0.325**
	(0.031)	(0.115)	(0.038)	(0.020)
Woman's age	-0.00292	-0.00262	-0.00287	-0.00247
	(0.518)	(0.561)	(0.524)	(0.583)
Woman's education (years)	-0.000566	-0.000413	-0.000527	-0.000453
	(0.824)	(0.871)	(0.836)	(0.858)
Husband's age	0.00404	0.00384	0.00396	0.00401
	(0.178)	(0.200)	(0.187)	(0.182)
Husband's education (years)	0.00836***	0.00836***	0.00836***	0.00826***
	(0.001)	(0.001)	(0.001)	(0.001)
Far (N=2,872)				
SR	-0.153**	-0.368***	-0.260***	-0.331**
	(0.034)	(0.000)	(0.002)	(0.013)
Woman's age	0.00475	0.00421	0.00440	0.00505
	(0.256)	(0.311)	(0.292)	(0.225)
Woman's education (years)	0.00338	0.00331	0.00329	0.00348
	(0.163)	(0.170)	(0.173)	(0.151)
Husband's age	0.000263	0.000140	0.000258	0.000241
	(0.924)	(0.959)	(0.925)	(0.930)
Husband's education (years)	0.000959	0.000969	0.000952	0.000945
	(0.696)	(0.692)	(0.698)	(0.700)
HighDowry (N=2,814)				
SR	-0.0966	-0.0351	-0.0835	-0.200*
	(0.109)	(0.637)	(0.232)	(0.072)
Woman's age	-0.00586	-0.00545	-0.00571	-0.00570
	(0.150)	(0.181)	(0.161)	(0.161)
Woman's education (years)	-0.00164	-0.00146	-0.00156	-0.00157
	(0.440)	(0.492)	(0.462)	(0.459)
Husband's age	0.000308	0.000198	0.000257	0.000310
	(0.911)	(0.943)	(0.925)	(0.910)
Husband's education (years)	0.00260	0.00259	0.00260	0.00259
	(0.234)	(0.236)	(0.234)	(0.237)

Note: Logit model average marginal effects. Robust standard errors in parenthesis. Women aged 15-24 at the beginning of 2001 who got married between 2001 and 2005. All specifications include state, demographic group and state-demographic group fixed effects.

Table 7: Sex-Ratios, Marriage Terms and Involvement in Spouse Selection

	Sex Ratio			
	F15-19/ M20-24	F20-24/ M25-29	F15-24/ M20-29	F15-29/ M15-29
Far (N=2,654)				
SR	-0.204**	-0.406***	-0.314***	-0.374**
	(0.029)	(0.001)	(0.006)	(0.019)
SR×*Involved*	0.0541	0.000328	0.0463	-0.0210
	(0.673)	(0.998)	(0.762)	(0.916)
Involved	-0.0797	-0.0276	-0.0736	-0.00763
	(0.534)	(0.874)	(0.637)	(0.967)
HighDowry (N=2,567)				
SR	-0.161**	-0.0699	-0.143	-0.329**
	(0.033)	(0.450)	(0.101)	(0.013)
SR×*Involved*	0.240**	0.162	0.241*	0.560***
	(0.028)	(0.271)	(0.072)	(0.002)
Involved	-0.272**	-0.200	-0.277**	-0.546***
	(0.014)	(0.193)	(0.043)	(0.001)

Note: Logit model average marginal effects. Robust standard errors in parenthesis. Women aged 15-24 at the beginning of 2001 who got married between 2001 and 2005. All specifications include state, demographic group and state-demographic group fixed effects.

Table 8: Parameter Estimates

		With Unobserved Heterogeneity	
	(i)	(ii)	(iii)
Match parameters ρ	-15.980	-10.360	-10.926
A	0.540	0.900	0.896
τ_m	-2.290	-0.249	-0.228
τ_w	-1.630	-0.184	-0.176
Male preferences μ_m			
Wife Near	6.550	4.165	4.132
Small Dowry	-4.420	-4.369	-4.479
Wife Autonomy	-3.250	-1.900	-1.954
Wife Age	0.730	1.666	1.650
Wife Ed	1.160	-5.331	-5.730
Female preferences μ_w			
Husband Near	3.390	4.774	4.805
Husband Near x 1{k=1}	.	-0.637	-0.689
Husband Near x 1{k=1} X Wife Young	.	.	3.977
Small Dowry	-6.040	-5.754	-5.852
Small Dowry x 1{k=1}	.	-1.703	-1.751
Own Autonomy	-2.040	-0.822	-0.850
Own Autonomy x 1{k=1}	.	-0.430	-0.467
Husband 25 to 29	1.930	3.156	3.206
Husband Ed	1.090	0.209	0.150
$P(k\|x_i)$			
Height	.	-0.945	-0.806
Height Missing		6.111	6.117
HSA Year of Repeal		4.078	4.096
HSA Religion Eligible		-0.229	-0.376
HSA Year X Eligibility		0.179	0.037
-log(like)	8620.9	7648.7	7581.0

Note: Hindu Succession Act-eligible religious groups include Hindu, Buddhist, Jain and Sikh. Estimates are for only castes not listed as scheduled castes or scheduled tribes.

Table 9: Model Fit, With and Without Unobserved Heterogeneity

	Baseline Model (i)			Model (iii)		
First Period	Model	Observed	Gap	Model	Observed	Gap
Overall	0.241	0.332	-0.091	0.316	0.332	-0.016
Village Near	0.152	0.281	-0.129	0.226	0.281	-0.055
Low Dowry	0.087	0.061	0.026	0.088	0.061	0.027
Female Autonomy	0.104	0.127	-0.024	0.135	0.127	0.008
Husband 25 to 29	0.105	0.129	-0.024	0.156	0.129	0.027
Husband Education	0.190	0.263	-0.073	0.245	0.263	-0.018
Female 20 to 24	0.148	0.232	-0.084	0.181	0.232	-0.050
Second Period						
Overall	0.125	0.123	0.002	0.138	0.123	0.015
Village Near	0.077	0.100	-0.023	0.096	0.100	-0.004
Low Dowry	0.044	0.022	0.022	0.038	0.022	0.016
Female Autonomy	0.054	0.051	0.003	0.060	0.051	0.009
Husband 25 to 29	0.059	0.086	-0.027	0.070	0.086	-0.016
Husband Education	0.098	0.100	-0.002	0.106	0.100	0.007

Note: Model Fit simulates behavior in all districts and compares the sample predicted fraction of matches with a given term or characteristic.

Table 10: Counter-factual sex Ratios, Match Probabilities

	Overall Match Probabilities				
	Model	1σ	1σ	(low sr)	(hi sr)
First Period	Prediction	Increase $\%F$	Decrease $\%F$	Chhattisgarh	Maharastra
Overall	0.316	0.258	0.351	0.288	0.329
Village Near	0.226	0.183	0.254	0.204	0.235
Low Dowry	0.088	0.073	0.097	0.081	0.091
Female Autonomy	0.135	0.110	0.151	0.123	0.141
Husband 25 to 29	0.156	0.110	0.165	0.144	0.166
Husband Education	0.245	0.205	0.260	0.228	0.251
Female 20 to 24	0.181	0.142	0.213	0.162	0.193
Second Period					
Overall	0.420	0.371	0.473	0.375	0.416
Village Near	0.291	0.257	0.328	0.260	0.288
Low Dowry	0.115	0.102	0.129	0.103	0.114
Female Autonomy	0.183	0.162	0.206	0.163	0.182
Husband 25 to 29	0.213	0.194	0.257	0.192	0.222
Husband Education	0.323	0.290	0.340	0.300	0.316

Note: Cells give the village weighted average probability of matching in a marriage with the given terms, own or partner characteristics. The simulation 1-σ increases the age-education specific sex-ratio (females over males) by one standard deviation within each district. Chhattisgarh and Maharastra simulations set the sex-ratio in all districts to the state-level sex-ratio in each of these two states. Second period match probabilities include only women in the cohort aged 15-19 at $t = 1$.

Table 11: Counter-factual sex Ratios, Search | Matching Probabilities

	Overall Match Probabilities				
	Model	1σ	1σ	(low sr)	(hi sr)
First Period	Prediction	Increase $\%F$	Decrease $\%F$	Chhattisgarh	Maharastra
Village Near	0.702	0.662	0.710	0.698	0.703
Low Dowry	0.290	0.276	0.289	0.291	0.289
Female Autonomy	0.419	0.396	0.421	0.419	0.419
Husband 25 to 29	0.445	0.355	0.422	0.454	0.462
Husband Education	0.821	0.789	0.790	0.834	0.811
Female 20 to 24	0.764	0.709	0.777	0.773	0.773
Second Period					
Village Near	0.686	0.647	0.687	0.686	0.686
Low Dowry	0.288	0.272	0.286	0.290	0.288
Female Autonomy	0.426	0.400	0.427	0.424	0.425
Husband 25 to 29	0.477	0.461	0.517	0.481	0.503
Husband Education	0.823	0.788	0.785	0.851	0.822

Note: Cells give the village weighted average probability of searching in a given marriage market, conditional on matching in any market. The simulation 1-σ increases the age-education specific sex-ratio (females over males) by one standard deviation within each district. Chhattisgarh and Maharastra simulations set the sex-ratio in all districts to the state-level sex-ratio in each of these two states. Second period match probabilities include only women in the cohort aged 15-19 at $t = 1$.

Table 12: Counter-factual Choice Sets, Match Probabilities

	Overall Match Probabilities		
	Model	All Women	All Women
First Period	Prediction	Educated	Autonomous
Overall	0.316	0.238	0.134
Village Near	0.226	0.155	0.095
Low Dowry	0.088	0.086	0.038
Female Autonomy	0.135	0.110	0.134
Husband 25 to 29	0.156	0.104	0.068
Husband Education	0.245	0.198	0.107
Female 20 to 24	0.181	0.129	0.072
Second Period			
Overall	0.420	0.266	0.210
Village Near	0.291	0.172	0.145
Low Dowry	0.115	0.089	0.057
Female Autonomy	0.183	0.121	0.210
Husband 25 to 29	0.213	0.078	0.086
Husband Education	0.323	0.211	0.164

Note: Cells give the village weighted average probability of matching in a marriage with the given terms, own or partner characteristics. The simulation 1-σ increases the age-education specific sex-ratio (females over males) by one standard deviation within each district. Chhattisgarh and Maharastra simulations set the sex-ratio in all districts to the state-level sex-ratio in each of these two states. Second period match probabilities include only women in the cohort aged 15-19 at $t = 1$.

Table 13: Counter-factual Choice Sets, Search | Matching Probabilities

	Overall Match Probabilities		
First Period	Model Prediction	All Women Educated	All Women Autonomous
Village Near	0.702	0.662	0.694
Low Dowry	0.290	0.341	0.294
Female Autonomy	0.419	0.440	1.000
Husband 25 to 29	0.445	0.436	0.457
Husband Education	0.821	0.853	0.842
Female 20 to 24	0.764	0.762	0.764
Second Period			
Overall			
Village Near	0.686	0.660	0.686
Low Dowry	0.288	0.321	0.279
Female Autonomy	0.426	0.431	1.000
Husband 25 to 29	0.477	0.293	0.353
Husband Education	0.823	0.835	0.824

Note: Cells give the village weighted average probability of searching in a given marriage market, conditional on matching in any market. The simulation 1-σ increases the age-education specific sex-ratio (females over males) by one standard deviation within each district. Chhattisgarh and Maharastra simulations set the sex-ratio in all districts to the state-level sex-ratio in each of these two states. Second period match probabilities include only women in the cohort aged 15-19 at $t = 1$.

Table 14: Counter-factual Unmatched and Age Gap

First Period	Model Prediction	1σ (+) $\%F$	1σ (—) $\%F$	(low sr) Chhattisgarh	(hi sr) Maharastra	All Women Educated	All Women Autonomous
P(Unmatched | 15-19 t=1)	0.399	0.487	0.314	0.463	0.392	0.605	0.717
Age Gap, t=1 | 15-19 t=1	5.43	4.78	5.40	5.56	5.51	5.415	5.487
Age Gap, t=2 | 20-24 t=2	2.38	2.31	2.59	2.51	2.41	1.464	1.765
Age Gap Overall | 15-19 t=1	3.86	3.51	3.97	4.03	3.92	3.415	3.200

Note: The probability of being unmatched is given only for the cohort of women age 15 to 19 at t=1, and is measured at t=3, when women have aged out of the model. The age-gap uses the mid-point of the observed age brackets to calculated the estimated difference in age between spouses (male age less female age) in years. Age gap ,$t = 1$ and $t = 2$ measure the age cap conditional on matching in the respective period, while Age Gap Overall presents the average overall age gap following all matching. Chhattisgarh and Maharastra simulations set the sex-ratio in all districts to the state-level sex-ratio in each of these two states. Second period match probabilities include only women in the cohort aged 15-19 at $t = 1$.

Table 15: Counter-factual Utility

		%Change in Female Utility					
	E(CS)	1σ	1σ	(low sr)	(hi sr)	All Women	All Women
Age 15-19	(utils)	(+) $\%F$	(—) $\%F$	Chhattisgarh	Maharastra	Educated	Autonomous
No Ed, k = 0	0.229	-1.142	0.597	-0.310	0.038	-	-2.598
With Ed, k = 0	0.869	-0.642	0.345	-0.474	0.050	-3.262	-1.666
No Ed, k = 1	0.127	-0.530	-0.665	0.305	-0.332	-	-0.671
With Ed, k = 1	1.339	-0.061	0.004	0.035	0.056	-0.596	-0.314
Age 20-24							
No Ed, k = 0	0.039	-0.910	0.464	-0.088	0.283	-	-6.413
With Ed, k = 0	0.282	-1.567	0.614	-0.279	0.410	-4.122	-4.884
No Ed, k = 1	0.036	-1.257	0.892	-0.772	0.093	-	-3.915
With Ed, k = 1	0.285	-0.791	0.316	-0.372	0.149	-0.047	-1.383

Note: Column one displays the logit-consumer surplus calculation for facing the choice at $t = 1$(using 1.44 as a the marginal utility of wealth, a value estimated for younger individuals by Layard, Mayraz and Nickell (2008)). Only the cohort aged 15 to 19 at $t = 1$ are included. k denotes the binary unobserved heterogeneity variable.