

Revealed Beliefs and the Marriage Market Return to Education

Alison Andrew^{1,2} and Abi Adams-Prassl^{*2,3}

¹Department of Economics, University College London

²Institute for Fiscal Studies

³Department of Economics, University of Oxford

This Version: November 15, 2021

Most recent version: [Click here!](#)

Abstract

This paper explores how parents in Rajasthan, India make choices about their daughters' schooling and marriage. We develop a dynamic model in which parents take schooling and marriage decisions under uncertainty about the quality of future marriage offers. Parents' choices are thus partially driven by their beliefs about the likelihood of receiving high-quality marriage offers in the future and how this likelihood depends on a daughter's age and education. We identify this model by creating a novel hypothetical-choice tool. We prove that by varying the information about future marriage offers described in hypothetical scenarios, our tool enables us to identify both preferences and probabilistic beliefs without directly eliciting probabilities. We find that parents believe there to be a large marriage-market return to girls' education and that this drives much of their investment in their daughters' schooling. And while parents would prefer to delay a daughters' marriage until at least age 18, a belief that marriage-market prospects begin to deteriorate with age once a daughter is out of school creates an incentive for parents to accept early marriage offers.

*Email addresses: alison_a@ifs.org.uk and abi.adams@economics.ox.ac.uk. We thank Nava Ashraf, Orazio Attanasio, Oriana Bandiera, James Banks, Teodora Boneva, Rossella Calvi, Rachel Cassidy, Rachel Griffith, Seema Jayachandran, Willemien Kets, Sonya Krutikova, Hamish Low, Costas Meghir, Francisco Oteiza, Rohini Pande, Aureo de Paula, Imran Rasul, Gabriela Smarrelli, Anna Stansbury, Marcos Vera-Hernandez and Basit Zafar for helpful comments and feedback. We are enormously grateful to Abhishek Gautam, Hemlata Verma, Ronak Soni and Amit Kumar for invaluable support in developing and piloting this instrument and to Kuhika Seth for collecting insightful qualitative data. We would like to thank the Centre for Public Policy at the Institute for Fiscal Studies, the Children's Investment Fund Foundation and the John Fell Fund, University of Oxford for generous financial support.

1 Introduction

When to leave school? When to marry? Who to marry? These are some of life's most consequential decisions. Together, they profoundly shape economic and mental wellbeing. They are also fundamental in shaping gender inequalities. Early marriage and low education leave women vulnerable to poverty, poor mental and physical health, low decision-making power, and economic dependence on men (Duflo 2012; Jensen and Thornton 2003). In rural India, the context of this study, parents typically make schooling and marriage decisions for their daughters. Thus, understanding the drivers of parents' choices is crucial for understanding the processes that sustain India's exceptionally high levels of gender inequality across many dimensions of welfare.

This paper is the first to examine the role that beliefs about the marriage market play in shaping parents' decisions about whether to keep their daughters in education and when, and to whom, to get them married. Do parents invest in their daughter's education in the expectation of securing her a better groom? Does a belief that good marriage offers might not be forthcoming in the future prompt parents to accept offers that arrive early in a daughter's adolescence? If we removed uncertainty about a daughter's marriage market match, what would parents prefer regarding her education and age of marriage?

Identifying the role of beliefs about marriage market returns is challenging. With observational choice data, identifying beliefs separately from preferences is not possible without strong assumptions (Manski, 2004). To overcome these identification issues, we develop a novel set of strategic survey instruments motivated by a flexible dynamic model of parents' schooling and marriage decisions. In our model, marriage offers of varying qualities arrive stochastically with a likelihood that depends on the age and education of the daughter. Our identification approach is based on the novel insight that by varying the amount of information on future realizations of stochastic variables, discrete choice experiments can identify not only preferences, but also subjective beliefs. We formally prove this result and use it to design a set of novel discrete choice experiments that we use with a sample of 4,500 caregivers of adolescent girls living in the Dhaulpur district of Rajasthan. We use respondents' choices to these instruments to estimate parental beliefs and preferences using far fewer assumptions than would be required if estimating such a model from observational data.

More specifically, our novel discrete choice methodology uses respondents' choices to two sets of discrete choice experiments that present different hypothetical vignettes concerning a daughter's marriage and education. Crucially, these two sets of experiments provide different information about the realization of future marriage offers. This

allows us to identify parents’ preferences over daughters’ age of marriage, completed education, and marriage match quality *and* their subjective beliefs about how the quality of marriage offers varies with a young woman’s age and education. In the first set of experiments, which we call “ex-post” experiments, respondents were presented with vignettes describing a set of complete realized descriptions of a daughter’s age and education at marriage and groom characteristics. In these experiments, the marriage offers at future time points are known with certainty. Respondents are asked to choose which option is preferred from the set. Choices in the ex-post experiment identify parental preferences over girls’ age of marriage, education and match characteristics (Wiswall and Zafar 2018).¹

In the second set of experiments, which we call “ex-ante” experiments, we collect state-dependent stochastic choice data (Caplin 2016). We elicit respondents’ behavior in the case of vignettes that present take-it-or-leave-it marriage offers from different grooms, potentially when a daughter is still in school. In the ex-ante experiment, there is uncertainty over the realization of future marriage offers. Thus, the expected value of rejecting the marriage offer is shaped not only by parents’ preferences but also by their beliefs about the probability of receiving different qualities of marriage offers in the future (Sautmann 2017). We prove formally that choices to the ex-ante and ex-post experiments identify beliefs about how the quality of marriage offers varies with a daughter’s age and education. We term the beliefs that reconcile choices made with and without uncertainty over future marriage offers “revealed beliefs”.

Varying the information contained in vignettes provides an intuitive way of eliciting beliefs that does not require directly asking respondents for point probabilities. Recovering subjective beliefs without asking respondents directly about probabilities has many advantages. First, it has clear benefits when respondents are unfamiliar with probabilistic concepts or have low numeracy. Second, even respondents with a sophisticated theoretical understanding of probability make systematic errors when reporting probabilistic beliefs (Tversky and Kahneman 1974). Our approach is also well suited to contexts where individuals do not have easily retrievable beliefs over the realization of different states, but rather adopt a set of state-contingent decision rules that are “as if” they were acting on a well-defined set of beliefs.² Third, our approach works well when eliciting beliefs over constructs – such as groom quality – that are formed of multiple components, combined in a manner that is unknown to the researcher ex-ante.

We use our method with a sample of 4,500 caregivers in rural Rajasthan. We estimate parents’ preferences

¹See also Banerjee, Duflo, Ghatak, and Lafortune (2013) that asks families to rank responses to matrimonial adverts to estimate the strength of caste preferences in marriage in India.

²Such models are feature in both economics and psychology literatures on learning (Charness and Levin 2005; Bjornerstedt and Weibull 1994; Kahneman and Tversky 1973; Erev and Roth 1988). Nash proposed a model in his PhD thesis (Nash 1951). We provide a further discussion of the relation of our approach to these types of models in Section 3.

over their daughters' education, age of marriage and marriage match in addition to their subjective beliefs about the marriage offer distribution. In particular, we estimate beliefs over the likelihood of receiving marriage offers from grooms of different quality levels as a function of a daughter's age and education. We find that parents care a great deal about securing a high-quality marriage match for their daughter. They place particular weight on whether or not the groom has a government job, the most well-paid and secure form of employment in the area. This preference, in turn, motivates parents' pre-marital choices as they seek to increase the chance of their daughter receiving high-quality marriage offers.

We find that a substantial perceived marriage market return to education is a key driver of pre-marital investment in girls' education. This is especially the case for college education. We find that, conditional on a given marriage match, parents dislike college education. However, they perceive that college dramatically increases the likelihood of their daughter marrying a high-quality groom. For instance, parents believe that a 19 year old daughter with a college degree has a 26% chance of receiving a marriage offer from a high-quality groom, a 12% chance if she finished high school and a less than 1% chance if she only has primary education. The subjective belief estimates pass several validation checks: they are qualitatively consistent with preferences over brides' characteristics that we elicit on the groom's side of the marriage market; they exhibit qualitatively similar trends to explicitly eliciting respondents' expected marriage match; and they are consistent with patterns of assortative matching on education that we observe in national survey data.

Regarding age at marriage, we find that parents prefer to delay a daughter's marriage until the age of eighteen, the legal minimum, but have no preference for delaying further. In contrast, we find that parents believe a girl's marriage market prospects worsen with age immediately after leaving formal education. These decreasing marriage market returns to age incentivize parents to accept marriage offers earlier than they would otherwise prefer since they anticipate that they are unlikely to receive high-quality marriage offers in the future. This feature implies that shocks that push girls out of school can give rise to early marriage since families perceive that their daughters' marriage market prospects will worsen with every year they wait.

Finally, we allow parents' perceptions of the likelihood of receiving high-quality marriage offers to be affected by whether or not their daughter's behavior conforms to prevailing gender norms. We find that girls who "get home late" and are "friends with boys" are perceived to face worse marriage market prospects. This finding suggests that the marriage market plays an important role in sustaining such gendered norms around appropriate behavior for

young women.

These findings have implications for policy and program design. Many of the programs seeking to promote later marriage have primarily targeted adolescent girls and their families (Buchmann et al. 2017; Edmonds et al. 2021; Bergstrom and Özler 2021). Our results show that programs targeted at young men, and their families, to loosen the stringent gendered expectations to which prospective brides are held might also be fruitful in reducing rates of early marriage. Our results also highlight the protective value of education against early marriage. They also identify a group of young women at particular risk of early marriage: those who have recently dropped out of school and whose marriage prospects are believed to be sharply declining in age. Shocks that cause young women to drop out of school thus also create pressures for their marriage. Our results thereby complement work that shows the importance of economic shocks for early marriage in contexts where marriage payments are the norm (Corno, Hildebrandt, and Voena 2016).

Our findings contribute to three main literatures. First, we contribute to the literature on the drivers of education and marriage choices in contexts where traditional gender roles are the norm. Much of what we know about the drivers of these choices come from interpreting the effects of standalone programs or policies (see Bergstrom and Özler (2021) for a review). This literature has highlighted that, in addition to factors such as the availability, accessibility and cost of school (Muralidharan and Prakash 2017; Burde and Linden 2013), these choices are deeply rooted in stringent gender norms (Edmonds et al. 2021; Buchmann et al. 2017; Dhar et al. 2021). Our work complements this existing literature by directly estimating the underlying structural primitives in the parents' decision problem.

Second, this paper contributes to the growing literature on marriage market returns to education. Existing empirical work on marriage market returns to education has primarily focused on societies with high female labor force participation and typically finds large and positive marriage market returns to female education (Chiappori, Dias, and Meghir 2015; Attanasio and Kaufmann 2017; Lafortune 2013) although there are instances of negative returns to very high levels of education and professional ambition (Fisman, Iyengar, Kamenica, and Simonson 2006; Hitsch, Hortaçsu, and Ariely 2010).³ Our results suggest that a substantial marriage market return to women's education also exists in this very different context: one where women's labor force participation is very low and gender norms are much more conservative (Dhar, Jain, and Jayachandran 2021). Indeed, we find that marriage

³Bursztyn, Fujiwara, and Pallais (2017) show that women respond to such negative returns to career ambitions by avoiding signaling their career ambitions to potential partners.

market returns provide the primary motivation for investing in a daughter’s college education in our context.⁴

Finally, this paper contributes to a growing body of work that develops theory-driven *strategic survey instruments* (see, for example, Ameriks et al. (2020) and Attanasio (2021)) to identify structural primitives under far fewer assumptions than would be required for identification from observational data. Specifically, we design and use tools to identify both preferences and subjective beliefs within the context of a dynamic search model of the marriage market. Our key methodological innovation is to develop a new choice-based method for eliciting beliefs. Directly measuring subjective beliefs is increasingly popular. Recent work has, for example, directly measured beliefs over: the effects of contraception (Miller, de Paula, and Valente 2020), returns to education (Delavande and Zafar 2019; Wiswall and Zafar 2018; Boneva and Rauh 2018; Attanasio, Cunha, and Jervis 2019), the drivers of child development (Attanasio, Cunha, and Jervis 2019; Cunha, Elo, and Culhane 2013; Bhalotra, Delavande, Font Gilabert, and Maselko 2020), children’s ability (Dizon-Ross 2019; Bobba and Frisanchio 2020). We show, for the first time, that by adapting standard discrete choice experiments, which are increasingly popular in measuring preferences, it is also possible to identify subjective beliefs.⁵

In principle, our approach could be used to elicit individual-specific preferences and beliefs. This is unfortunately not possible in this paper given the cultural context of our application. Age of marriage is a sensitive topic in our study area, and thus misreporting of age of marriage is common (Borkotoky and Unisa 2014). Social desirability bias would be a major concern if we asked respondents about their own choices given legal reforms and widespread social and political campaigns targeting child marriage.⁶ To reduce systematic bias in reported choices and to avoid putting respondents in an uncomfortable position, we thus employ a hypothetical framing throughout the implementation of our approach: we ask respondents about the choices they think a hypothetical couple would make for their daughter.⁷ As a consequence of this hypothetical framing, we formally identify respondents’ perceptions of average preferences and average beliefs in the precise application in this paper.⁸ We provide a number of robustness

⁴We note that there are many reasons why the groom’s side of the marriage market might value educated brides. Some of these, such as valuing the income that she could contribute to the household suggest a break with traditional ideas of gender roles, while others, such as valuing her parenting skills, are perfectly compatible with traditional roles. Digging deeper into underlying reasons why the marriage market values women’s education is an important topic for future research.

⁵Delavande and Zafar (2018) and Arcidiacono, Hotz, and Kang (2012) combine directly elicited beliefs and choice data to estimate dynamic structural models but do not make use of comparisons of choice with and without uncertainty to identify beliefs themselves. Our approach is similar to work that identifies people’s subjective beliefs by the decisions that they make when faced with a choice between different state-contingent lotteries, conditional on knowledge of their preferences (Andersen, Fountain, Harrison, and Rutström 2014).

⁶The Prohibition of Child Marriage Act, 2006 increased mechanisms to prosecute adult men who marry children, those performing a child marriage and those promoting or permitting one, including parents and guardians.

⁷In addition to alleviating social desirability bias, the use of hypothetical vignettes also limits the role of unobserved characteristics in driving respondents’ decisions.

⁸While systematic bias in perceptions of average preferences can be an important issue for rare events (Bursztyn, González, and Yanagizawa-Drott 2018), it is unlikely to be a first-order concern in our study. In our setting, marriage is near universal (99.5% of women in rural Rajasthan are married by age 30 (NFHS-4).), marriage arrangements are public and the subject of much discussion

checks for systematic bias in these objects and find that partial knowledge of true average preferences and beliefs is not a first-order concern in our context. Thus, we consider that respondents' perceptions are unlikely to be systematically wrong and thus refer simply to "average preferences" and "average beliefs" throughout.

This paper proceeds as follows. In Section 2 we describe the key features of the marriage market, and the interaction between marriage and education decisions, in rural Rajasthan. We then develop a dynamic discrete choice model of parents' decisions about the education and marriage of their daughter that recognizes the uncertainty they face over future marriage offers. In Section 3, we discuss the challenges of identifying such a model from observational data and motivate our choice to use strategic survey instruments. In Section 4, we use our model to guide the design of the ex-post and ex-ante choice experiments. Section 5 shows how the combination of choice data from the ex-post and ex-ante experiment enables the identification of beliefs about the likelihood of marriage offers from high quality grooms as a function of the age and education of a young woman. Section 6.1 shows the key sources of variation present in choices to our instruments, while Section 6.2 gives our key structural results and shows that our revealed belief results and show that they are consistent with patterns from directly elicited expectations over average groom quality, groom-side preferences, and patterns of assortative matching in Rajasthan. Section 7 concludes.

2 Theoretical Framework

We begin by setting out a dynamic partial equilibrium model that captures the key contextual features of the marriage market. Parents' choice of whether to accept a marriage offer for their daughter and how much education she acquires are dynamic decisions made under uncertainty. The optimal choice depends not only on parents' preferences over their daughter's age and education at marriage but also on parental beliefs about how the quality of marriage offers depends on a daughter's age and education. In following sections, we will show how specifically-designed stated-choice experiments are able to identify these structural parameters of interest, both preferences and subjective beliefs.

within the community. We find no systematic differences in response patterns between respondents who have children who have already gone through the process of marriage for one of their children and those who haven't, nor by how similar a vignette is to respondents' own circumstances.

2.1 Context

There are four key contextual features of the marriage market that we embed in our model. First, we treat marriage and schooling as decisions made by parents rather than adolescent girls themselves. In our study communities, marriages are almost always arranged by parents and young women have little control over either the timing of their marriage or the choice of spouse.⁹ For example, only 12% of married adolescent girls in our survey communities had met and spoken to their husband alone before marriage, and only 13% of unmarried adolescent girls expected they would (Achyut et al. 2016).

Second, there is uncertainty about the quality of potential grooms. The study area is patrilocal, meaning that women join their husband’s natal community, and usually their natal home, upon marriage. The norm is for the husband’s natal community to be at least 10km away from the bride’s.¹⁰ Parents search for potential grooms through extended family and sub-caste networks.¹¹ The search process can be lengthy and these frictions leave a role for uncertainty over the quality of future matches.

Third, we do not explicitly model the relationship between a daughter and her parents after marriage. Although marital ties create extended family networks that can be important for risk-sharing, job search and migration (Rosenzweig and Stark 1989), married women’s primary economic unit is fundamentally that of the marital household. Married women have little autonomy to maintain independent economic connections with their natal family.¹² Indeed, it is rare for younger women in our study area to work outside the home or in family businesses. In our sample, only 37% of mothers had worked for cash in the past year (Table 1) and labor force participation is lower amongst younger women.¹³

Finally, marriage is a significant economic transaction. Despite having been illegal since 1961, dowry payment remains an important feature of most marriages in this part of India. Dowry is a transfer, typically made up of cash and gold or silver jewelry along with furniture, home appliances and sometimes a vehicle, from the bride’s

⁹In our paper, we do not make a distinction between marriage and gauna. Under the custom of gauna, a marriage is not consummated, and the bride does not move in with the husband, until several years after the official marriage. However, the custom is rapidly decreasing in prevalence across India. Indeed, only 6% of married women under the age of 25 in India report their gauna being performed after they were first married (IHDS); this is half the rate for married women aged 35-45.

¹⁰Focus group respondents mentioned this distance should be at least 10km (see notes from Caregivers’ Focus Group Discussion (FGD) 2, Online Appendix B).

¹¹Preference for within-caste marriage is very strong. However, previous work has suggested that because all castes share this preference, it has little impact on matching across other characteristics or matching efficiency (Banerjee, Duflo, Ghatak, and Lafortune 2013).

¹²In the marital household, women’s decision-making and mobility are often restricted. In our sample, 39% of mothers are not allowed to go to the market unaccompanied, while 92% do not own any asset they could dispose of at will. It is thus difficult for women to maintain an economic connection with their natal family that is not mediated through their marital household.

¹³See 2015/16 National Family Health Survey India (NFHD-IV).

family to the groom’s family at the time of marriage.¹⁴ Within our study communities, dowry is primarily viewed as a “groom price”.¹⁵ The value of dowry is substantial relative to household wealth in our study communities. Respondents in our confidential focus groups gave a monetary range corresponding to \$4,700 to \$15,600 or between 3 and 10 times current GDP per capita in Rajasthan (Researve Bank of India 2017), a ratio which is in line with previous estimates (Rao 1993; Bloch and Rao 2002).¹⁶ We do not directly model dowry bargaining but instead view dowry as a component of groom quality.

2.2 Model

States, Actions and Option Sets. We model a finite portion of the dynamic problem facing parents spanning a daughter’s adolescence. In practice, we consider the period from when a daughter is age 13 to when she is age 22,¹⁷ although we allow for her to remain unmarried at the end of this period. In periods before marriage, possible actions, d_t , that parents can take are:

$$d_t = \begin{cases} S & \text{if the daughter is kept in education,} \\ H & \text{if the daughter is kept at home,} \\ M & \text{if the daughter is married.} \end{cases} \quad (2.1)$$

The actions available in parents’ option sets, and their payoffs from these actions, depend on a vector of state variables $\omega_t = \{t, \mathcal{M}_t, Ed_t, Z_t, q_t, \varepsilon_t\}$. \mathcal{M}_t denotes whether the daughter is already married in period t ($\mathcal{M}_t = 1$) or not ($\mathcal{M}_t = 0$). Ed_t gives the daughter’s current level of education.¹⁸ q_t summarises information on the best marriage offer received in period t . If a marriage offer is received, $q_t \in \mathbf{Q} = \{Q^0, \dots, Q^K\}$, where Q^i represents finite groom quality types. If no marriage offer is received, then $q_t = N$. Z_t contains other factors, such as a daughter’s particular enjoyment of school or the household’s financial situation, that may influence the parents’ payoffs and that are constant over time and known to both parents and the econometrician. Finally, ε_t represents time-varying

¹⁴The 1849 respondents living in rural areas of Rajasthan in the 2011-12 Indian Health and Development Survey indicated that the following items were ‘usually’ included in dowries in their communities: gold (84% of respondents), silver (89%), cash (43%), TV (53%), furniture (86%), pressure cooker (66%), utensils (95%), bedding/mattress (76%), watch/clock (76%), sewing machine (43%) and scooter (15%).

¹⁵One focus group participant commented that: “Dowry is directly proportional to what kind of boy one is looking (for), if one is seeking a boy who is educated, has fields and a good house and family then the dowry is always higher” (FGD 3, Online Appendix B). This is consistent with literature showing the transition of dowry from a pre-mortem inheritance for the bride to a market-clearing price for grooms (Anderson and Bidner 2015).

¹⁶2018 price level. See FGD 3, Online Appendix B. Further, in addition to dowry transfers to the groom’s family the bride’s family also covers the cost of the wedding, which may be elaborate. Spending on expensive wedding celebrations may be used to signal the quality of a good match to neighbors to increase the social standing of the bride’s family (Bloch, Rao, and Desai 2004).

¹⁷91% of 24 year old women are married in rural Rajasthan according to the 2014/15 National Family Health Survey.

¹⁸The comparison of Ed_t with t means the state also contains whether or not the daughter is currently in school.

idiosyncratic shocks to the payoffs associated with taking different actions in period t .

As we discuss in the following subsections, the idiosyncratic preference shocks ε_t and offer qualities q_t are unknown in advance of period t . We abstract from all other sources of uncertainty. Thus the vector of deterministic state-variables, which are either time-invariant or update deterministically as a function of past choices, is therefore:

$$\bar{\omega}_t = \{t, \mathcal{M}_t, Ed_t, Z_t\}.$$
¹⁹

The options available to parents at t are given by the set $O_t(\omega_t)$.²⁰ A daughter can only be married if a marriage offer is received (i.e. if $q_t \neq N$). We assume that once a daughter is taken out of education, she cannot re-enter later and girls cannot continue in education once married.²¹ Furthermore, girls cannot continue in education after age 20, the typical for finishing college. We abstract from the possibility of divorce given its rarity in India.²² In what follows, we will keep the dependence of the option set on past actions implicit.

Preferences. Parents care about the education their daughter receives, the value of any home production she contributes to, when she gets married, and who she gets married to. These preferences will reflect a combination of altruism towards a daughter's future welfare within marriage (Jensen and Thornton 2003) or towards future grandchildren (Rosenzweig and Wolpin 1994; Carneiro, Meghir, and Parey 2013; Grépin and Bharadwaj 2015; Behrman, Foster, Rosenzweig, and Vashishtha 1999; Chari, Heath, Maertens, and Fatima 2017), social norms and psychological payoffs to one's daughter's education and marriage (Maertens 2013), economic costs of dowry and the costs of keeping a girl in school or at home (Corno and Voena 2016), and the economic value of a connection to the groom's family (Rosenzweig and Stark 1989).

Preferences are defined over the sequences of realised states and actions taken over a daughter's adolescence. Let a potential sequence, or path, through the model from period t onwards be given as $\Psi_t = \{\omega_\tau, d_\tau\}_{\tau=t, \dots, T}$. A path contains all information about every state reached and every action taken. Parents preference over a path is given by their utility function $U_t(\Psi_t)$. We place the following assumptions on the underlying structure of preferences:

¹⁹For example, education evolves according to:

$$Ed_{t+1} = Ed_t + \mathbf{1}(d_t = S)$$
 (2.2)

²⁰Formally, the choice set, O_t , available to parents at $t < T$ when is given as:

$$O_t = \begin{cases} \{S, H, M\} & \text{if } d_{t-1} = S \text{ \& } \mathcal{M}_t = 0 \text{ \& } q_t \neq N, \\ \{S, H\} & \text{if } d_{t-1} = S \text{ \& } \mathcal{M}_t = 0 \text{ \& } q_t = N, \\ \{H, M\} & \text{if } d_{t-1} = H \text{ \& } \mathcal{M}_t = 0 \text{ \& } q_t \neq N, \\ \{H\} & \text{if } d_{t-1} = H \text{ \& } \mathcal{M}_t = 0 \text{ \& } q_t = N, \\ \emptyset & \text{if } \mathcal{M}_t = 1. \end{cases}$$
 (2.3)

²¹Amongst married young women enrolled in the original study and married sisters of study participants, who were all aged between 14 and 24, less than 7% were continuing any form of education which includes those enrolled on distance learning courses.

²²Just 0.3% of ever-married women in India are currently divorced according to the 2011 census.

A1. Additive separability over time: We assume that preferences over complete paths can be represented as the (discounted) sum of payoffs accrued in different periods of the model with additive separability over time.

A2. Constant discount factor: We assume that the discount factor, β , used to discount future payoffs is constant across time.

A3. Irrelevance of rejected marriage offers: We assume that marriage offers only enter payoffs if they are accepted. This rules out features like regret, which would allow rejected offers to impact welfare.

Together, assumptions A1, A2, and A3 imply that $U_t(\Psi)$ can be expressed as the discounted sum of flow payoffs, $v(\omega_t, d_t)$:

$$U_t(\Psi_t) = \sum_{\tau=t}^T \beta^{\tau-t} v(\omega_\tau, d_\tau) \quad (2.4)$$

where $v(\omega_\tau, d_\tau)$ is the per-period payoff in period τ , which is a function only of that period's state variables and that period's actions.²³

Our object of interest is $U_t(\Psi_t)$ (and, in particular is $U_0(\Psi_0)$) rather than $v(\omega_\tau, d_\tau)$. Taking preferences over education, for example, we do not attempt to separate out whether preferences for education come from flow payoffs during years in which a daughter is in school or from preferences over the completed stock of education after her schooling is complete. Importantly, our identification results for subjective beliefs are not sensitive to what portion of preferences are accrued across different time periods, they only require that payoffs accrued in different periods are summed additively with a constant discount factor.

Idiosyncratic Preference Shocks The state variables ε_t describe idiosyncratic preference shocks faced by parents. We adopt a structure of additive idiosyncratic shocks that is common in the discrete dynamic choice literature. We assume that these shocks vary randomly across periods and assume that in period t , parents have no information about the value of their idiosyncratic preference shocks in $t+1$ onwards. We assume that flow payoffs $v(\omega_\tau, d_\tau)$ are additively composed of a non-random component, $\bar{v}(\bar{\omega}_\tau, q_\tau, d_\tau)$, which depends on the current deterministic state variables, current marriage offer, and current action, and a the idiosyncratic preference shock, $\varepsilon_\tau^{d_\tau}$, associated with

²³Assumption A3 implies that $v(Q^k, \bar{\omega}_\tau, H) = v(Q^{k'}, \bar{\omega}_\tau, H), \forall k \neq k'$ and $v(Q^k, \bar{\omega}_\tau, S) = v(Q^{k'}, \bar{\omega}_\tau, S), \forall k \neq k'$.

each of the three options:

$$U_t(\Psi_t) = \sum_{\tau=t}^T \beta^{\tau-t} [\bar{v}(\bar{\omega}_\tau, q_\tau, d_\tau) + \varepsilon_\tau^{d_\tau}] \quad (2.5)$$

$$= \bar{U}_t(\bar{\Psi}_t) + \sum_{\tau=t}^T \beta^{\tau-t} \varepsilon_\tau^{d_\tau} \quad (2.6)$$

where $\bar{\Psi}_t \equiv \{\{\bar{\omega}_t, q_t, d_t\}, \dots, \{\bar{\omega}_T, q_T, d_T\}\}$.

Marriage Offer Uncertainty. Parents do not know whether they will receive marriage offers in future periods nor the quality of any offers received. Parents hold well-defined expectations over the likelihood of receiving marriage offers from grooms of a given quality and how this varies with their daughter's age and education. We assume that, in each period, parents believe they will be in marriage offer state $q \in \mathbf{Q} = \{Q^0, \dots, Q^K, N\}$ with probability $\pi(\bar{\omega}_t, q)$. While we can allow for $\pi(\cdot)$ to depend on all deterministic state variables, in practise we will focus on age and education.²⁴ These offer probabilities, $\pi(\cdot)$, will depend on the preferences of grooms and the distribution of groom quality in the local marriage market. However, we do not put structure on the search and matching process to give micro-founded expressions for $\pi(\cdot)$ in terms of these structural primitives. This theoretical framework aims rather to provide a basis for the design of a set of experiments to identify subjective beliefs over the distribution of match quality, age, and education conditional on preferences.

Choice Parents act to maximize their discounted expected utility. Specifically, parents adopt optimal decision rules δ to solve:

$$\max_{\delta} \mathbb{E}[U_t(\Psi_t) | \omega_t] \quad (2.7)$$

where δ is the set of decision rules that define actions that parents will take at every possible state from t onwards that they may find themselves in. When the probability of receiving good quality marriage offers in the future depends on the education and age of their daughter, parents must account for the impact of their decisions today on their daughter's future marriage options; optimal choices will not simply reflect their own preferences.

²⁴Note that we assume that the probability of a particular quality offer is independent of the parents' past decisions, i.e. we assume that rejecting a marriage offer today has no direct impact on the likelihood of receiving offers from particular groom qualities in the future: $\pi(Q|t, E, d_{t-1}) = \pi(Q|t, E)$

3 Identification Challenges and Approaches to Measurement

It is challenging to use observed choice behavior to separately identify the role that parental beliefs over the marriage market return to youth and education play in driving these choices from parental preferences. This is for at least three main reasons. First, it is well-known that researchers face a fundamental identification problem in separately identifying preferences from beliefs from choice data (Manski 2004): without strong assumptions, one can usually rationalize behavior by multiple combinations of utility and beliefs. In this context, early marriage could be rationalized by parents intrinsically preferring to marry daughters at a young age, perhaps to comply with social norms, or by a belief that high-quality marriage offers are unattainable for older women. The fact that observational datasets typically record final marriage matches but not marriage offers further serves to break the link between observed choices and deep parameters of the model. Second, the marriage and education choices we see in observational data are affected by many factors that are often unobserved by the researcher. For example, in qualitative work for this project, several parents mentioned that their daughters had left school due to sexual harassment they had experienced at school, a factor that is rarely recorded in observational data. Therefore, working with observational data alone requires strong assumptions over the nature of these idiosyncratic factors and their relationship to the factors under investigation. Third, social desirability bias in our context often creates systematic mismeasurement in the observed age of marriage.

These challenges lead us to develop theory-driven strategic measurements to identify the structural primitives of our model under less restrictive assumptions than would be required for identification from observational data. This paper is part of a growing literature that takes such an approach.²⁵ The flexibility to ask respondents questions in a way that ensures a precise mapping from their answers to deep parameters has many advantages. A tightly-controlled environment can allow researchers to hone in on precise drivers of decision making, unconfounded by the multitude of unobserved (to the economist) idiosyncratic factors that drive real-life choices. To this end, we design measurement tools to identify our two objects of interest: preferences and beliefs over future marriage offers.

To identify preferences, our work builds on the stated choice methods, which use choices made in hypothetical scenarios to infer preferences (Ben-Akiva, McFadden, and Train 2019). Such methods have long been popular in marketing and environmental economics. And they have more recently been applied to a wider array of practical applications to measure preferences over, for example, job attributes (Mas and Pallais 2017; Wiswall and Zafar

²⁵See Orazio Attanasio's presidential address at the 2020 Econometric Society meetings for an overview of the history of this literature.

2018), healthcare plans (Kesternich, Heiss, McFadden, and Winter 2013; Ameriks, Briggs, Caplin, Shapiro, and Tonetti 2020), and the distribution of resources (Berry, Dizon-Ross, and Jagnani 2020; Alesina, Stantcheva, and Teso 2018; Almås, Cappelen, and Tungodden 2020).

Our main measurement innovation is to adapt stated-choice methods also to identify beliefs. By varying the amount of information about future marriage offers contained in the hypothetical scenarios, we can measure, for the first time, respondents' beliefs about the likelihood of future events (in our case marriage offers) *without* directly eliciting probabilities. In adapting a stated-preference experiment to identify objects other than preferences, our work relates to Delavande and Zafar (2019) who elicit students' stated choices over colleges with and without financial constraints, allowing them to identify the constraints.

Our work provides a new tool for the growing agenda around the measurement of subjective beliefs. Recent literature has, for example, measured beliefs over: the effects of contraception (Miller, de Paula, and Valente 2020), returns to education (Delavande and Zafar 2019; Wiswall and Zafar 2018; Boneva and Rauh 2018; Attanasio, Cunha, and Jervis 2019), the drivers of child development (Attanasio, Cunha, and Jervis 2019; Cunha, Elo, and Culhane 2013; Bhalotra, Delavande, Font Gilabert, and Maselko 2020), children's ability (Dizon-Ross 2019; Bobba and Frisancho 2020). Not directly asking respondents about probabilities, we contend, brings several advantages. Many people find it difficult to understand and respond to probabilistic statements. Although creative methods have been developed to help individuals respond to probabilistic statements (Manski 2004; Delavande, Giné, and McKenzie 2011), this remains a particular concern in contexts such as ours where respondents have low numeracy – our average respondent had one and a half years of schooling.

However, the concern is somewhat deeper. There is evidence that the subjective probabilities people report contain systematic inconsistencies even when they understand abstract probability concepts (Kahneman and Tversky 1973; Tversky and Kahneman 1974; Giustinelli, Manski, and Molinari 2020).²⁶ To the extent that reported probabilities, be they inconsistent or not, are accurate representations of the beliefs that causally shape agents' actual decisions, this is not a problem. In fact, a key advantage of directly measuring beliefs might be to capture how such inconsistencies affect decision making. However, if people's actions are more sophisticated than the subjective beliefs that they report, then taking self-reported beliefs at face value would lead us to mispredict actions;

²⁶For example, reported subjective beliefs can be insufficiently sensitive to the prior probability of different outcomes and based on immediately recallable reference groups (Kahneman and Tversky 1973; Tversky and Kahneman 1974). In our context, this could result in, for example, insufficient sensitivity of reported beliefs to the distribution of groom quality in the local marriage market or for stated beliefs to be based on realized marriages without any correction for the sample-selection problem caused by the refusal of some marriage offers.

Giustinelli, Manski, and Molinari (2020) provide a clear example of this.²⁷ Experimental studies on decision making under uncertainty highlight that people’s actions can take on board more information, and in a more sophisticated way, than what is contained within the subjective beliefs they report (Erev and Roth 1988; Charness and Levin 2005). Indeed, it is not even necessary to assume that agents have a well-defined and easily-retrievable “beliefs” to hand for their actions to be “as if” they were acting in an expected utility framework; instead, agents may learn a set of best-response actions that are consistent with such a belief.²⁸ In his PhD dissertation (Nash 1951) John Nash proposed that:

“It is unnecessary to assume that the participants have full knowledge of the total structure of the game, or the ability and inclination to go through any complex reasoning processes. But the participants are supposed to accumulate empirical information on the relative advantages of the various pure strategies at their disposal.”

In this light, an attractive feature of our choice-based approach is that it is not reliant on parents making decisions about their daughters’ futures based on beliefs that are easy for them to retrieve and report to an interviewer in the form of probabilities. Instead, our approach is also compatible with models of decision making in which, like Nash’s “mass action” interpretation, parents know a set of best-response actions for different scenarios that might arise, e.g. whether it would be wise to accept a given marriage offer at different points in time, but these actions aren’t directly based on beliefs that respondents would be able to report in a survey. In that case, our approach identifies a set of revealed beliefs that are “as if” parents’ were acting upon them when making their choices.

A final dimension of our approach which is appealing in our setting is its ability to measure beliefs over a one-dimensional match-quality index when the index is driven by multiple components of quality but the importance of the different components are not known ex-ante. In our case, we hypothesized that several features of marriage matches might be important, such as a groom’s age, education, occupation, and family wealth. However, we didn’t know ex-ante how parents would combine these different dimensions into a single index of groom quality. To collect data on beliefs probabilistically, therefore, we would have had to elicit respondents’ beliefs over the joint distribution of grooms’ age, education, occupation, and wealth. It is well known that eliciting beliefs about multidimensional objects is challenging and time-consuming. By contrast, our revealed belief approach neither requires ex-ante

²⁷Indeed, when laying the foundations of subjective expected utility, Ramsey (1926) argued that one should measure beliefs based on the actions that individuals take: “the degree of a belief is a causal property, which we can express vaguely as the extent to which we are prepared to act on it... Our judgment about the strength of our belief is really about how we should act in hypothetical circumstances.”

²⁸Such models often, but not always, converge to individuals taking the same actions as if they were Bayesian updaters despite not requiring individuals to have a set of beliefs to hand (Nash 1951; Bjornerstedt and Weibull 1994).

knowledge of preferences over the different dimensions nor eliciting a multidimensional probability distribution. Instead, our revealed belief approach just requires us to collect choice data capturing how respondents respond to take-it-or-leave-it marriage offers that vary across all the dimensions mentioned. We then use the preferences identified from the ex-post experiment to infer what the uni-dimensional quality of each groom we described in the ex-ante experiment was.²⁹

4 Methods and Instrument Design

We design two sets of strategic survey questions that vary the information that parents have about future marriage offers to disentangle preferences from beliefs: the “ex-post” and “ex-ante” experiments. In this section, we describe the design of these instruments in detail. In Section 5, we show formally how they facilitate the identification of the structural functions of interest.

4.1 Overview

Our instruments are based on hypothetical vignettes that describe choice scenarios faced by fictional families. The hypothetical framing alleviates social desirability bias and the role of unobserved characteristics (Hughes 1998; Finch 1987). Parents are given different amounts of information about the realization of marriage offers at different points in time in the two experiments. Comparing choices made in these different information environments provides a simple choice-based method to estimate beliefs about the distribution of marriage offers facing adolescent girls of a given age and education levels.

In the ex-post experiment, respondents are told the story of a hypothetical family who had a twelve-year-old daughter. Respondents are told to imagine that there are two options for when the daughter can leave school, and when, and to whom, she can get married. Respondents are then asked what option they think the hypothetical family would choose for their daughter. Formally, the choice scenario in the ex-post experiment corresponds to a decision over different discounted utility profiles $\bar{U}_0 (\bar{\Psi}_0)$.

In the ex-ante experiments, respondents are presented with scenarios in which hypothetical families receive a marriage offer when their daughter is a specific age. Respondents had to decide whether the family would accept the

²⁹We anticipate that such an advantage could be useful in the job search literature. Recent work has stressed that workers value multiple dimensions of jobs in addition to salary (Mas and Pallais, 2017; Wiswall and Zafar, 2018) but this brings the challenge of identifying multidimensional offer distributions, something that we think our approach could be useful for.

marriage offer or reject it in favor of keeping the daughter in school or having her help in the home. As the pattern of future marriage offers was left unspecified, choices reflect both preferences *and* beliefs about the likelihood of future marriage offers, $\pi(\cdot)$. Formally, the choice scenario corresponds to the decision facing a family in some state ω_t (Equation 2.7).

4.2 The Sample

We administered our survey instruments on a sample of 4,605 female caregivers of adolescent girls who live across 120 villages in the Dhaulpur district of Rajasthan.³⁰ The sample is representative of primary caregivers, almost always mothers, of unmarried adolescent girls in the study communities.³¹ Data collection was carried out between December 2017 and March 2018.

Table 1 reports descriptives for our analysis sample. It shows that this is an economically and socially disadvantaged population. 50.7% of respondents live in houses with a dirt floor, respondents have an average of just 1.5 years of education and only 10.4% can read a complete sentence. The mean respondent married when she was fifteen. While gender norms are conservative in the study area (Achyut et al. 2016), 94% of the sample reported that they had at least “some say” over when and to whom a child got married and when a child left school. More details about the study population can be found in the baseline report (Achyut et al. 2016).

Data collection took place in respondents’ homes and, whenever possible, in a quiet and private environment. All interviewers were female and had experience of working on large scale household surveys. Interviewers were given two days training specifically on this survey instrument in addition to training on the remainder of the survey and general interview skills. They carried out a further two days of field practice before the start of data collection. Each respondent was randomised into being asked about *either* about the ex-post or the ex-ante instrument given time constraints.

³⁰We ran our experiments as part of an endline data collection for a cluster randomized controlled trial of a community-based gender norms program for adolescent girls. We obtained written consent from all caregivers before proceeding to the survey and experiments. This study has received ethical approval from institutional review boards at the University of Oxford (R43389), the International Center for Research on Women (15-0001) and Sigma, New Delhi (10035/IRB/D/17-18). Since we aim to elicit perceptions of average preferences and beliefs in the population we did not anticipate treatment allocation would have any impact on how respondents answered the choice experiments. Appendix Tables A.3 and A.4 confirm that, at most, the effect of treatment is marginal. We thus ignore treatment allocation for the remainder of the paper.

³¹In 2015, a random sample of (then) unmarried adolescent girls aged 12-17 years was drawn from complete lists of all adolescent girls in the villages. At endline in 2017/18, we attempted to re-interview all adolescent girls and all primary caregivers. Of the 4,994 caregivers interviewed at baseline, we re-interviewed 93.4% and we have complete discrete choice experiments for 93.0% of the original sample. For more details on sampling see Achyut et al. (2016).

Table 1: Sample descriptives

	Mean	Standard Deviation	N
Age in years	41.92	8.365	4464
Own age at marriage in years*	15.57	3.361	4423
Years of school*	1.492	3.267	4605
Can read complete sentence (in Hindi)*	0.104	0.305	4353
Number of sons*	2.118	1.112	4343
Number of daughters*	2.447	1.320	4343
Owns asset that can dispose of at will	0.132	0.339	4604
Can go to market unaccompanied*	0.611	0.488	4463
At least some say over when child gets married	0.963	0.190	4536
At least some say over to whom child gets married	0.952	0.213	4532
At least some say over when child leaves school	0.942	0.235	4534
Has done any work (inc. on family farm) in last year	0.595	0.491	4604
Has worked for cash in last year	0.344	0.475	4604
Has child (male or female) who is married	0.364	0.481	4576
House has dirt floor*	0.507	0.500	4603
Scheduled caste or scheduled tribe*	0.352	0.478	4581
Other Backward Caste or Economically Backward Class*	0.451	0.498	4581
Hindu*	0.968	0.177	4602

Notes: Table reports descriptives for sample of 4,605 caregivers with complete data from the choice experiments. *Variable measured in baseline survey during 2016. All other variables collected in 2017/18 endline survey.

4.3 Introducing the Scenarios

We began by stressing that a respondent’s answers would not be used to make inferences about the choices they would make for their own children to limit social desirability bias. The interviewer introduced the scenarios with the following statement:

“We are going to tell you some stories about parents and marriage of their children. These stories are purely hypothetical. We will ask you some questions about how you think the parents in the story will take decisions based on the given options. There are no right and wrong answers. All your answers are confidential and you are free to stop at any time.”

Given that we ask respondents how they think the parents in the story would behave, formally, this method allows us to elicit respondents’ *beliefs* about the expected behavior of parents in the community under the scenario described. As marriage is a universal, public, and much discussed topic within the community, we do not consider the assumption that respondents have accurate perceptions of expected behavior to be an incredible one.³²

³²For rare events, it may be more likely for individuals to misperceive social norms. In Saudi Arabia, where female labor force participation is currently very low, Bursztyn, González, and Yanagizawa-Drott (2018) find that men systematically underestimate the acceptability of women working outside the home.

Table 2: Summary of Vignette Characteristics

Characteristic	Support
<i>Characteristics of the Bride and Bride's Family</i>	
Wealth	{Very Poor, Average, Very Wealthy}
Age of Daughter	Ex-Post: Always 12
	Ex-Ante: {13, 14, ..., 21, 22}
Currently in School	{In School, Out of School}
Likes School	{Like, Dislike}
School Costs	{Free, Costly}
Chores	{Relax, Help Needed}
Comply with Norms (Ex-ante only)	{Polite and Well Behaved, Friends with Boys}
<i>Age and Education at Marriage</i>	
Age at Marriage (girl)	{13, 14, ..., 21, 22}
School Grade at Marriage	{7, ..., 12, College}
<i>Characteristics of the Marriage Match</i>	
Wealth	{ Very Poor, Average, Very Wealthy}
Age at Marriage	{21, 22, ..., 29, 30}
School Grade at Marriage	{None, 1, 5, 7, ..., 12, College}
Occupation	{Government Job, No Gov Job}
Minimum Dowry Acceptable to Groom (lakh)	{0, 0.5, 1, ..., 7, 7.5}

4.4 Vignette Characteristics

The stories underpinning the ex-post and ex-ante experiments provided a rich description of the hypothetical family and their daughter, and prospective marriage market matches. This provided a rich set of state variables and determinants of groom quality for respondents to evaluate. The same set of core characteristics was enumerated in both the ex-post and ex-ante experiments. These characteristics are summarised in Table 2. Figures 2 and 3 give examples of the translated verbal descriptions used for each characteristic.

Drawing on the existing literature and insights from our focus group discussions,³³ we specified five attributes of the hypothetical family and their daughter: (i) Household wealth; (ii) Whether the mother needed extra assistance in the home; (iii) Whether the daughter was currently in school; (iv) Whether the daughter enjoyed school (for those still in school); (v) The cost of schooling (for daughters described as being currently enrolled in school).³⁴

Wealth was included given the large literature on positive assortative matching on wealth in marriage markets (Eika, Mogstad, and Zafar ; Greenwood, Guner, Kocharkov, and Santos 2014; Siow 2015; Pencavel 1998; Becker 1973). Descriptions of the different wealth levels were calibrated to the top, middle and bottom quintiles of an asset index estimated on data from a previous survey of this sample. For example, our description of the situation faced a family of “average wealth”, the middle quintile, was:

“Their house has two rooms with a dung floor. They own one bigha of land and two cows. They own

³³See Online Appendix B for details on the focus groups.

³⁴We ensured that the description of these attributes were meaningful to respondents through prior qualitative research and piloting.

an electric fan and a bicycle but not a TV.”

The cost of schooling is an important determinant of education decisions (Duflo, Dupas, and Kremer 2015), while a girl’s enjoyment of school was frequently mentioned in our qualitative work as a key factor explaining why she is still in school or had been allowed to drop out.³⁵ Whether the mother requires help at home was specified to create variation in the value of home production.

In the ex-ante experiments we additionally included an indicator of the extent to which the daughter’s behavior complied with gendered norms since focus group participants frequently brought up that girls whose behavior conflicted with norms should be married as soon as possible since they would not receive good offers in the future. Thus, we described some daughters as being *“polite and well behaved”* while others we described as being *“friends with some boys and sometimes stay[ing] out of the house until late”*.

When describing potential marriage options, the following five attributes were given: (i) Wealth of groom’s family; (ii) Minimum dowry (marriage gifts) acceptable to the groom’s family; (iii) Completed education of groom; (iv) Age of groom; (v) Whether groom has a government job (the most well paid and secure form of employment in the area). Wealth, education and occupation were included given the large literature on positive assortative matching discussed above and as a result of focus group interviews. Despite having been illegal since 1961, the payment of dowry remains an important feature of most marriages in this part of India.

While our vignettes painted a rich description of the hypothetical family and the options they faced, due to time and memory constraints we could not include all potentially relevant characteristics in the stories. For example, we did not specify a daughter’s physical beauty (Banerjee, Duflo, Ghatak, and Lafortune 2013) nor the presence of siblings (Singh and Vennam 2016). However, respondents were prompted to assume that any unspecified characteristics were the same across options were the same, to limit potential concerns about Bayesian updating on the basis of specified characteristics.

4.5 Varying Uncertainty about the Future

The ex-post and ex-ante experiments differ in terms of what is left unstated about the future. This is crucial for our method to separately identify preferences from beliefs. In particular, in the ex-post experiment, there is no uncertainty over the realization of marriage offers; respondents choose between two options that both specify all information relevant to the fictional parents’ utility (how long the daughter stays in education, when she marries,

³⁵See Online Appendix B for details on the focus groups.

and who she marries). By contrast, in the ex-ante experiment, respondents are asked to choose between accepting or rejecting a take-it-or-leave-it marriage offer in the current period. Importantly, the value of rejecting the marriage offer is uncertain as it depends on what future marriage offers materialize; the expected value of rejecting a marriage offer in hand depends on the perceived likelihood that a daughter will receive a good marriage offer in the future.

Ex-Post Experiment The ex-post experiment presented respondents with a choice between two options for the fictional daughter’s adolescence, both of which contain sufficient information for the fictional parents to evaluate, up to the value of the idiosyncratic preference shocks, the value of each option. Formally, each option corresponded to a specified path through the model, $\bar{\Psi}_0$ (starting at the beginning of adolescence). In practice, this meant that the ex-post vignettes always described a daughter as being twelve years old and each choice option fully specified the evolution of a daughter’s education and marriage until age 22.³⁶ The exercise was described to respondents as follows:

“[The parents] are considering when and to whom they will get [their daughter] married and until when they will keep her in education. Imagine there are two possible options, for when [their daughter] will leave education and when and to whom she will get married.”

The interviewer then went on to describe the options, each of which specified when the daughter would marry, to whom (including the characteristics of the groom and his family as well as the minimum dowry required), and how much education she would complete prior to marriage. In other words, the options specified everything (other than the idiosyncratic preference shocks) relevant to parents’ utility as defined by our model. Thus, there was no need for respondents to form expectations over the likelihood of receiving good marriage offers at different points in time.

After the interviewer described the scenario and two options, she asked the respondent:³⁷

“Which option do you think [the parents] will choose for their daughter?”

Figure 1 (a) shows the visual aid used in the ex-post experiment while Figure 2 gives the English translation of an example script for the experiment; the underlined elements are those that were randomly changed between rounds.

³⁶The start of model time in our setting is age 12 and the end is age 22.

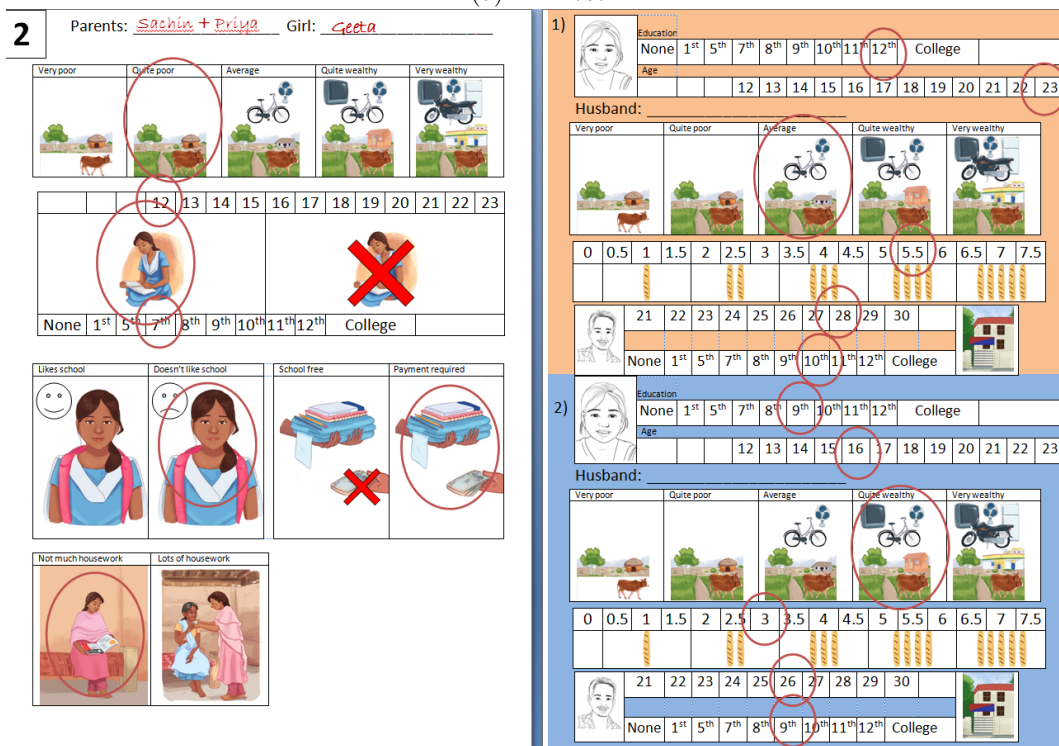
³⁷To ensure that respondents understood and took note of all relevant characteristics, each option was described twice and the relevant attributes were circled on the visual aids.

Ex-Ante Experiment In contrast to the ex-post experiment, the ex-ante experiments offered respondents a take-it-or-leave-it marriage offer that left respondents uncertain about what marriage offers might be received in the future. As the pattern of future marriage offers was left unspecified, the relative value of the various options reflects not only preferences but also beliefs about the likelihood of receiving good marriage offers in the future. Formally, the choices corresponded to a decision made in some state ω_t over the alternatives in the feasible choice set O_t ((Equation 2.7).

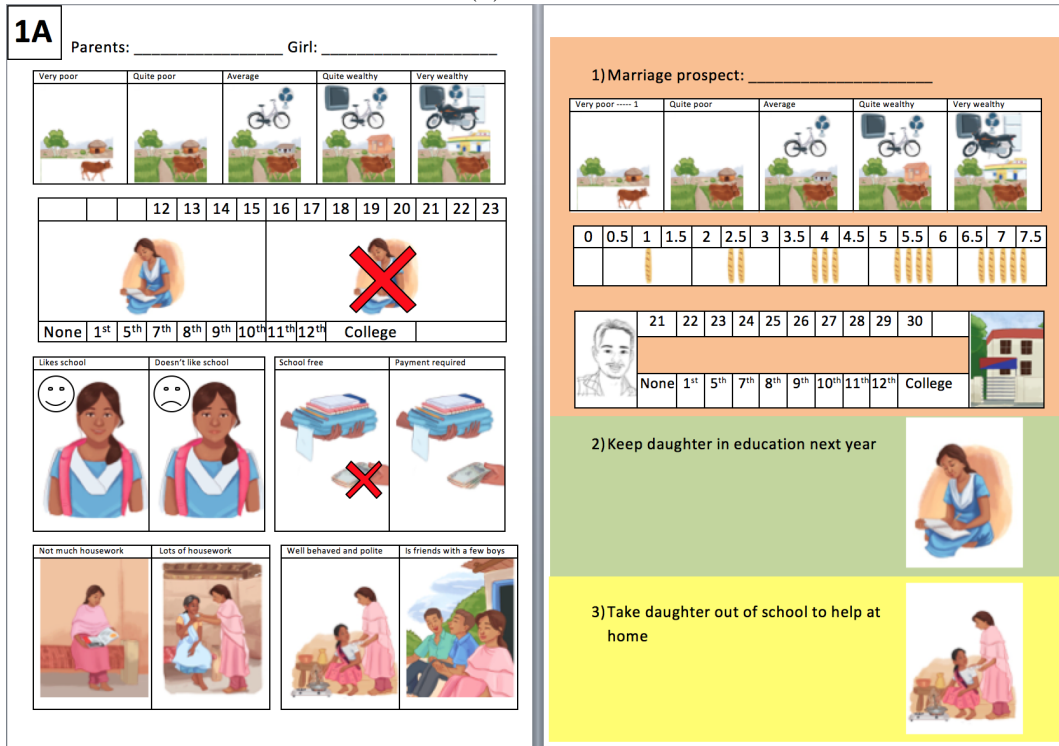
In the ex-ante experiment, we randomly varied the age of the fictional daughter between rounds, ranging from age 13 to age 22 (the end of model time). After the interviewer described the hypothetical family, respondents were told that the family had received a take-it-or-leave-it marriage offer for their daughter. The offer was described using the same characteristics used in the ex-post experiment, namely, it specified the groom's and his family's characteristics and the dowry. The interviewer told the respondent that the fictional parents needed to make a decision now: whether their daughter would get married in the next year to the groom described, whether the parents would keep her in school for another year, or whether their daughter would leave school to help at home. Respondents again had to decide which option they thought the family would choose.

Figure 1: Visual Aids

(a) Ex-Post



(b) Ex-Ante



Notes: Panel (a) shows a translated visual aid for the ex-post experiment. As the survey enumerator described the vignette to a respondent, they circled the relevant characteristics on the visual aid to help respondents keep track. The red circles give one example of a potential choice scenario. Panel (b) shows a translated visual aid for the ex-ante experiment. This is shown without any mark-up.

Figure 2: Ex Post Experiment: Example Script

Background

Please imagine a mother and father named Ramesh and Rita, they live in a village similar to your village. Ramesh and have a daughter named Sita.

Compared to other households in the village, Ramesh and Rita's household is of average wealth. Their house has two rooms with a dung floor. They own one bigha of land and two cows. They own an electric fan and a bicycle but not a TV.

Sita's age is 12. Sita's currently studying in 7th standard. Sita is getting average grades at school and really enjoys going. At Sita's school there is a scheme that covers all the costs of Sita's education, including stationary, uniforms and transport.

Sita's mother, Rita, really struggles to take care of all the work that needs doing at home. So Sita's help with cooking, cleaning, and taking care of elderly relatives is very useful.

Ramesh and Rita are considering when and to whom they will get Sita married and until when they will keep her in education. Imagine there are two possible options, for when Sita will leave education and when and to whom she will get married.

Option 1

Sita completes standard 9 at school. Sita marries at age 16.

She marries Rahul. Rahul is 24 years old. Rahul attended College and has a government job.

Compared to other household in the village, Rahul's household is very wealthy. Their house has three rooms with a cement floor. They own four bighas of land and two cows. As well as an electric fan and a TV they also own a refrigerator and a motorcycle.

Rahul's parents expect at least 6 lakhs in marriage gifts.

Option 2

Sita completes standard 12 at school. Sita marries at age 19.

She marries Bharat. Bharat is 22 years old. Bharat attended school until 10th standard.

Compared to other household in the village, Bharat's household is very poor. The whole household live in one room with a dung floor and they don't own any land. They have one cow. They own very other few assets, for example, they don't own an electric fan, a TV or a bicycle.

Bharat's parents expect at least 2.5 lakhs in marriage gifts.

Choice

Which option do you think Ramesh and Rita will choose for their daughter?

1. Keep Sita in education until she has finished standard 9 and then marry her to Rahul when she is age 16
2. Keep Sita in education until she has finished standard 12 and then marry her to Bharat when she is age 19

Figure 3: Ex Ante Experiment: Example Script

Background

Please imagine a mother and father named Raj Kumar and Aarti, they live in a village similar to your village. Raj Kumar and Aarti have a daughter named Jyoti.

Compared to other households in the village, Raj Kumar and Aarti's household is of average wealth. Their house has two rooms with a dung floor. They own one bigha of land and two cows. They own an electric fan and a bicycle but not a TV.

Jyoti's age is 15. She is currently studying in 10th standard.

Jyoti is getting average grades at school but does not enjoy school. Jyoti's parents have to pay for the full cost of Jyoti's education, including stationary, uniforms and transport.

Jyoti's family have no particular need for Jyoti to spend lots of time helping at home. When Jyoti is at home she spends lots of her time sitting and relaxing.

Raj Kumar and Aarti are worried about Jyoti as she is friends with some boys and sometimes stays out of the house until late.

Offer in Hand

Raj Kumar and Aarti are considering whether they will get Jyoti married in the next year, whether they will keep her in school for another year or whether Jyoti will leave school to help her parents at home.

Raj Kumar and Aarti know of a potential suitor for Jyoti, Amit.

Compared to other households in the village, Amit's household is very poor. The whole household live in one room with a dung floor and they don't own any land. They have one cow. They own very other few assets, for example, they don't own an electric fan, a TV or a bicycle.

Amit's parents expect at least 4 lakhs in marriage gifts.

Amit is 24 years old. Amit attended school until 11th standard.

Choice

Which option do you think Raj Kumar and Aarti will choose for their daughter?

1. Marry Jyoti to Amit this year
2. Keep Jyoti in school this year
3. Take Jyoti out of school so she can help in the home this year

4.6 Variation in Scenario Characteristics

We randomized at the individual level whether respondents participated in the ex-post experiment or the ex-ante experiment to prevent any confusion given the instruments' apparent similarity. Each respondent carried out three rounds of the experiment they were assigned to, with the characteristics of the hypothetical family, daughter, and marriage options randomly varying across rounds.³⁸ Scenario characteristics were drawn orthogonally across respondents, rounds, and options subject to the modifications described below. Online Appendix A provides full details of the option generation procedure and histograms of the distribution of all vignette characteristics and the correlations between them.

- **Infeasible/Perverse Options.** The completed education of a daughter was truncated at the maximum feasible for a described age of marriage, i.e. an option in which a daughter is described as being thirteen at marriage could not have higher than Grade 8 education. The minimum acceptable dowry was allowed to loosely correlate with the wealth of the groom's family: dowry was randomly drawn between: 0.5 and 3 lakh for "poor" groom families, 1.5 and 6 lakh for "medium" groom families, and 2.5 and 7.5 lakh for "rich" groom families. Note that the realism of the scenarios described is, however, irrelevant for identification purposes so long as respondents are able to rank the various options.
- **Sufficient Differences.** In the ex-post experiment, we redrew characteristics to ensure that the two options within the same round were sufficiently different from one another in terms of the daughter's age of marriage and education at marriage so that respondents found the choice meaningful. We also ensured that respondents were never choosing between two options in the same round in which the daughter married younger than fifteen in both cases because such scenarios sometimes made respondents uncomfortable in piloting.
- **Salient Vignette.** We randomly replaced the characteristics of the hypothetical family, and in the ex-post experiment, the description of the daughter at marriage in one option, in one round with the characteristics of the respondent's own household and daughter in order to test whether the similarity of the vignette to the respondent's own experience affected choices.³⁹

³⁸We ran the discrete choice experiments between December 2017 and March 2018. The experiments took place in respondents' homes and, whenever possible, in a quiet and private environment. The experiments were run by female interviewers who had experience of working on large scale household surveys but no particular experience of running lab-in-field experiments. Interviewers were given two days training on the experiments in addition to training on general interview skills. They carried out a further two days of field practice before the start of data collection.

³⁹In Appendix Table A.1 and Table A.2 we show that vignette salience has no impact.

5 Identification & Estimation

5.1 Identifying Preferences

In each round r of the ex-post experiment, respondents were offered a choice between two profiles, $\bar{\Psi}_{r1}$ and $\bar{\Psi}_{r2}$, that specified a daughter's age of marriage (A), the quality of the accepted marriage offer (q_A) and the value of all deterministic state variables at marriage ($\bar{\omega}_A$). Since the daughter was always aged 12 in the ex-post experiment, respondents were asked to compare these profiles from the start of model time, or time $t = 0$. Let $\bar{\Psi}^M = [\bar{\Psi}^0, \bar{\Psi}^1, \dots, \bar{\Psi}^M]$ give the set of all possible profiles of a daughter's adolescence starting from period 0 in which she marries at or before the final period T .

Let the probability that the first option is chosen in some round r be given as $P(\bar{\Psi}_{r1}, \bar{\Psi}_{r2})$. Since the descriptions in the ex-post experiment describe two alternative paths for the whole of a daughters' adolescence, we assume that in responding to it respondents draw two different sequences of idiosyncratic preference shocks $\{\varepsilon_{irj\tau}\}_{\tau=0}^T$ and attach one to each of the two alternatives j . ν_{irj} is the accumulated preference shock associated with taking path $\bar{\Psi}_{irj}$, and is thus defined as $\nu_{irj} \equiv \sum_{\tau=0}^T \beta^\tau \varepsilon_{irj\tau}^d$. It is thus with respect to this accumulated preference shock that respondents weigh up the two alternatives. As there is no uncertainty about the realisation of marriage offers in the ex-post experiment, respondents choose the profile that maximises discounted utility:

$$P(\bar{\Psi}_{r1}, \bar{\Psi}_{r2}) = Pr [\bar{U}_0(\bar{\Psi}_{r1}) + \nu_{ir1} > \bar{U}_0(\bar{\Psi}_{r2}) + \nu_{ir2}] \quad (5.1)$$

where $\bar{\Psi}_{r1}, \bar{\Psi}_{r2} \in \bar{\Psi}^M \times \bar{\Psi}^M$.

Theorem 1 provides a formal statement of identification of parents' utility function. For the purposes of identification, we treat choice probabilities as known at all possible combinations of adolescent profiles ending in marriage. This is a natural assumption in an experimental setting where we can generate sufficient variation in option characteristics. The second condition imposes a location and scale normalizations on utility as is required in any discrete choice setting. The third assumption gives the distributional assumption that we impose on the preference heterogeneity. The assumption of normality is not required for non-parametric identification of utility (Matzkin 1992), but is rather a convenient functional form assumption for estimation.

THEOREM 1. *Identification of Preferences:* $\bar{U}_0(\bar{\Psi})$ is constructively identified at all $\bar{\Psi} \in \bar{\Psi}^M$ from ex-post choice probabilities under the following assumptions:

1. Ex-post choice probabilities, $P(\bar{\Psi}, \bar{\Psi}')$, are observed for all $(\bar{\Psi}, \bar{\Psi}') \in \bar{\Psi}^M \times \bar{\Psi}^M$.
2. Location and scale normalizations: $\bar{U}(\bar{\Psi}^0) = 0$ and $U(\bar{\Psi}^1) - \bar{U}(\bar{\Psi}^0) = 1$
3. The aggregated preference shock is distributed i.i.d normal, $\nu_{irj} \sim IN(0, \sigma_\nu^2)$, and is independent of $\bar{\Psi}_{rj}$ for all $\bar{\Psi}_{rj} \in \bar{\Psi}^M$.

Proof. See Appendix D.

We did not ask respondents to choose between options in which a daughter remained unmarried at T . Of course, if we had done this, identification of preferences over paths in which marriage does not occur at or before T would follow from the same argument as Theorem 1. We show in Appendix D that the valuation that parents associate with a daughter remaining unmarried at T is identified from choices in the ex-ante experiment.

5.2 Identifying Beliefs

In the ex-ante experiment, only information the current state of the world, in particular $\bar{\omega}_t$ and q_t , is specified and respondents make choices under uncertainty about future marriage offers. Before we proceed to show how choices made under uncertainty about future marriage offers can identify beliefs conditional on already-identified preferences, we introduce some new definitions and notation to aid our explanations.

An alternative notation for preferences To simplify notation, note that given Assumption A3, we can equivalently define the deterministic component of utility simply as a function of the deterministic state at marriage and the quality of the best offer received in the period of marriage for $\bar{\Psi} \in \bar{\Psi}^M$:

$$u(A, q_A, \bar{\omega}_A) \equiv \bar{U}_0(\bar{\Psi}_0) \tag{5.2}$$

where A is the age of marriage ($d_A = M$), q_A is the quality best marriage offer that period (i.e. the quality of the accepted offer), and $\bar{\omega}_A$ gives the deterministic state variables at marriage. $\bar{\omega}_A$ contains all payoff-relevant information, including information about the their daughter's final education.

Defining value functions At this stage, it is useful to define the value functions associated with the parents' problem. Somewhat unusually, we define all value functions as being evaluated from the point of view of period $t = 0$. We do this to avoid taking a stance on the allocation of payoffs across periods.⁴⁰

First, let's define the value function as the value that parents obtain from a daughter still being unmarried at period t with deterministic state variables $\bar{\omega}_t$ when they have received a marriage offer of quality q_t and a draw from the idiosyncratic preference shock distribution of ε_t . The period t value function is defined as the expected discounted value of total utility (summed across all model periods), viewed from the point of view of period 0, from behaving optimally in period t and in all periods $\tau > t + 1$ and conditional on the value of period t 's stochastic state variables, ε_t and q_t :

$$V(t, \bar{\omega}_t, q_t, \varepsilon_t) = \max_{\delta} \mathbb{E} [U_0(\Psi) | t, \bar{\omega}_t, q_t, \varepsilon_t] \quad (5.3)$$

$$= \max_{\delta} \mathbb{E} \left[\bar{U}_0(\bar{\Psi}) + \sum_{\tau=0}^T \beta^{\tau} \varepsilon_{\tau}^{d_{\tau}} | t, \bar{\omega}_t, q_t, \varepsilon_t \right] \quad (5.4)$$

where δ is the set of all possible decision rules defining actions at every possible state that parents may find themselves in.

Following Rust (1987), we next define the *expected value function*, which is the expected value that parents obtain from a daughter still being unmarried at period t with deterministic state variables $\bar{\omega}_t$. Again, we evaluate from the point of view of period 0. It is simply the expected value of the value function at states with deterministic state variables $\bar{\omega}_t$, with expectations taken over both stochastic state variables, the idiosyncratic preference shocks ε_t and the realisation of marriage offers q_t :

$$EV(t, \bar{\omega}_t) = \mathbb{E}_{q_t, \varepsilon_t} [V(t, \bar{\omega}_t, q_t, \varepsilon_t)] \quad (5.5)$$

Finally, we define choice-specific value functions. These describe the value obtained (evaluated from $t = 0$) conditional on being in state ω_t in period t , choosing action d_t and then choosing optimally in all periods thereafter.

The choice-specific value function associated with choosing marriage in period t is:

⁴⁰In words this is the current discounted value of currently being in period 0 but knowing that period t 's state variables will be ω_t . Importantly, knowledge of all information needed to form payoffs that occur in period between period 0 and $t - 1$ can be reconstructed from the period t deterministic state variables. There is thus a straightforward relationship between value functions defined from the point of view of period 0, $V(t, \omega_t)$ and those defined from the point of view of period t ($V_t(t, \omega_t)$). In particular: $V_0(t, \omega_t) = \sum_{\tau=0}^{t-1} \beta^{\tau} v(\omega_{\tau}, d_{\tau} | \omega_t) + \beta^t V_t(t, \omega_t)$ where $d_{\tau} | \omega_t$ is the action in time τ implied by ω_t .

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^M, d_t = M) = u(t, \bar{\omega}_t, q_t) + \beta^t \varepsilon_t^M \quad (5.6)$$

The choice specific value function associated with choosing either school or home is simply equal to the expected value of still being unmarried in period $t + 1$ with an education level updated to reflect period t 's schooling choice plus the value of the idiosyncratic preference shocks.

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^H, d_t = H) = EV(t, \bar{\omega}_{t+1}(\bar{\omega}_t, d_t = H)) + \beta^t \varepsilon_t^H \quad (5.7)$$

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^S, d_t = S) = EV(t, \bar{\omega}_{t+1}(\bar{\omega}_t, d_t = S)) + \beta^t \varepsilon_t^S \quad (5.8)$$

where $\bar{\omega}_{t+1}(\bar{\omega}_t, d_t = S)$ is the mapping of deterministic state variables in time t to those in time $t + 1$, conditional on period t 's actions.

Choice under uncertainty over future marriage offers Given this new notation, we can describe parents' choice behavior when acting under uncertainty over future marriage offers. The probability that parents choose each action $d_j \in O_t(\omega_r) \subseteq \{M, S, H\}$ is:

$$p_j(t, \bar{\omega}_t, q_t) = Pr \left[\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^j, d_t = j) > \tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^{j'}, d_t = j') \quad \forall j' \neq j \right] \quad (5.9)$$

Identification from ex-ante choice behavior Our model's predictions for how parents will behave at every possible action point in the model (Equation (5.9)), maps directly into the choice probabilities that we observe from the ex-ante choice experiment. We now show how using this relation allows us to identify expected value functions and beliefs. To gain insight into the intuition behind the identification, consider a simplified example where we assume that parents always receive at least one marriage offer. Further, we assume for this simplified example that there are only two types of grooms: high quality grooms ($q = 1$) and low quality grooms ($q = 0$). Let π_t be the probability that parents receive an offer from a high quality groom at time t and, thus, let $1 - \pi_t$ be the probability that they receive an offer only from a low quality groom.

Consider states where a daughter is already out of school and therefore parents are simply making a two-way choice, whether or not to accept the marriage offer of quality q that they have received. We suppress dependence on

all deterministic state variables (including education) other than age for ease of notation. Identification proceeds in two steps. In step one, we invert choice probabilities to identify the expected value of still being unmarried at every future unmarried state. In step two, we show how from the recursive relationship between expected value functions identifies the beliefs over the offer probabilities.

Step one. Identifying value functions from choice probabilities When respondent i in round r is described a hypothetical scenario with parents in period t trying to decide whether to accept a marriage offer of quality q , we assume that respondents act like parents in our model. That is, they draw a particular value of the idiosyncratic preference shocks at time t , ε_{irt} , from distribution $N(0, \sigma^2)$. Respondents then compare the value of parents keeping their daughter at home another year, $\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_{irt}^H, d_t = H)$, and the value of them accepting the marriage offer, $\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_{irt}^M, d_t = M)$. Under the assumption that idiosyncratic preference heterogeneity is normally distributed with mean 0 and standard deviation σ , we can form closed form expressions for the proportion of respondents, $p_M(t, q)$, who will chose marriage when at period t with an offer in hand of quality q :

$$p_M(t, q = 0) = Pr \left[\tilde{V}(t, q = 0, \varepsilon_{irt}^M, d_t = M) > \tilde{V}(t, q = 0, \varepsilon_{irt}^H, d_t = H) \right] \quad (5.10)$$

$$\begin{aligned} &= Pr \left[u(t, q = 0) + \beta^t \varepsilon_{irt}^M > EV(t + 1) + \beta^t \varepsilon_{irt}^H \right] \\ &= \Phi \left(\frac{u(t, q = 0) - EV(t + 1)}{\beta^t \sqrt{2}\sigma} \right) \end{aligned} \quad (5.11)$$

$$p_M(t, q = 1) = Pr \left[\tilde{V}(t, q = 1, \varepsilon_{irt}^M, d_t = M) > \tilde{V}(t, q = 1, \varepsilon_{irt}^H, d_t = H) \right] \quad (5.12)$$

$$\begin{aligned} &= Pr(u(t, q = 1) + \beta^t \varepsilon_{irt}^M > EV(t + 1) + \beta^t \varepsilon_{irt}^H) \\ &= \Phi \left(\frac{u(t, q = 0) - EV(t + 1)}{\beta^t \sqrt{2}\sigma} \right) \end{aligned} \quad (5.13)$$

We can invert and rearrange to identify the standard deviation of the preference shocks and the value associated with being unmarried at period $t + 1$:

$$EV(t + 1) = u(t, 0) + \sigma \beta^t \sqrt{2} \Phi^{-1}(p_M(t, q = 0)) \quad (5.14)$$

where:

$$\sigma = \frac{u(t, 1) - u(t, 0)}{\beta^t \sqrt{2} (\Phi^{-1}(p_M(t, q = 1)) - \Phi^{-1}(p_M(t, q = 0)))} \quad (5.15)$$

Having identified the value of being at each unmarried state and the variance of the additive preference shocks, we can identify the expected value of future preference shocks conditional on different choices being optimal. In particular, we can identify the expected value of ε_{irt}^M conditional on marriage being optimal in period t and the expected value of ε_{irt}^H conditional on home being optimal:⁴¹

$$E(\varepsilon_{irt}^M | d_t^* = M, t, q) = E(\varepsilon_{irt}^M | u_0(t, q) + \beta^t \varepsilon_{irt}^M \geq EV(t+1) + \beta^t \varepsilon_{irt}^H) \quad (5.16)$$

$$E(\varepsilon_{irt}^H | d_t^* = H, t, q) = E(\varepsilon_{irt}^H | u_0(t, q) + \beta^t \varepsilon_{irt}^M \leq EV(t+1) + \beta^t \varepsilon_{irt}^H) \quad (5.17)$$

Step two. Going from value functions to beliefs Having identified the expected value functions for the value of being unmarried at every state, we can use the recursive relationship between expected value functions to identify beliefs. If parents enter period t with an unmarried daughter, one of four outcomes might arise: (1) they get a low quality marriage offer and accept it; (2) they get a high quality marriage offer and accept it; (3) they get a low quality marriage offer and reject it; (4) they get a high quality marriage offer and reject it. Thus, the expected value of having an unmarried daughter at period t is equal to the mean of the expected value of each of these eventualities, weighted by the probability that that eventuality will occur:

$$\begin{aligned} EV(t) = & \underbrace{(1 - \pi_t) p_M(t, q = 0) [u(t, q = 0) + E(\varepsilon_t^M | d_t^* = M, t, q = 0)]}_{\text{low offer, accept}} \\ & + \underbrace{\pi_t p_M(t, q = 1) [u(t, q = 1) + E(\varepsilon_t^M | d_t^* = M, t, q = 1)]}_{\text{high offer, accept}} \\ & + \underbrace{(1 - \pi_t)(1 - p_M(t, q = 0)) [EV(t+1) + E(\varepsilon_t^H | d_t^* = H, t, q = 0)]}_{\text{low offer, reject}} \\ & + \underbrace{\pi_t(1 - p_M(t, q = 1)) [EV(t+1) + E(\varepsilon_t^H | d_t^* = H, t, q = 1)]}_{\text{high offer, reject}} \end{aligned} \quad (5.18)$$

Everything in this expression is identified other than π_t so we can simply rearrange to identify subjective beliefs.

The identification argument set out here holds true conditional on every level of education and thus can identify

⁴¹In Appendix D, we show that under normality this has a closed-form solution for two-way choices.

beliefs at every age-education combination when a daughter is already out of school. In Appendix D, we show that the same steps hold for all periods when a daughter is in school and thus parents are making a three-way choice; this identifies beliefs about offer probabilities at states where a daughter is still in school.

In Appendix D, we also show that identification of beliefs is possible when there are more than two groom types and when parents do not always receive a marriage offer in every period if there are “preference-shifting” instruments. These are characteristics that affect payoffs but are excluded from beliefs about offer probabilities. In practice, we use variation in whether or not the hypothetical daughter “likes school” to identify a model in which parents do not necessarily receive a marriage offer in every period.

THEOREM 2. Identification of Beliefs: $\pi(\cdot)$ and the variance of the idiosyncratic preference shocks, σ^2 , are constructively identified under the following assumptions:

1. Ex-ante choice probabilities, $p_j(\bar{\omega}, q)$ for $j \in \{M, S, H\}$, are observed for all $\bar{\omega} \in \Omega$ and $q \in Q$, with $p_j(\bar{\omega}, Q^k) \neq p_j(\bar{\omega}, Q^{k'})$ for $k \neq k'$.
2. The idiosyncratic preference shocks are distributed i.i.d normal, $\varepsilon_{irj} \sim IN(0, \sigma^2)$, and is independent of $(\bar{\omega}, q) \in \Omega \times Q$.
3. $\bar{U}_0(\bar{\Psi})$ is identified and satisfies Assumptions A1-A3.

Proof. See Appendix D.

5.3 Estimation

We estimate preferences and beliefs structurally through Maximum Likelihood. In doing so, follow a two-step procedure. First, we estimate the preference parameters indexing $\bar{U}_0(\Psi)$ for $\bar{\Psi} \in \bar{\Psi}^M$ from choice data from the ex-post experiment. Next, we estimate beliefs using choice data from the ex-ante experiment, taking preferences as given. We bootstrap the whole 2-step procedure 250 times to ensure that sampling error in the first step is accounted for in our estimated standard errors for our belief estimates. We resample at the respondent level to allow for correlation in how the same respondent answered different rounds of the experiment.

Preferences Recall the probability that the first option is chosen in some round r , $P(\bar{\Psi}_{r1}, \bar{\Psi}_{r2})$, is given as:

$$P(\bar{\Psi}_{r1}, \bar{\Psi}_{r2}) = Pr [\bar{U}_0(\bar{\Psi}_{r1}) + \nu_{ir1} > \bar{U}_0(\bar{\Psi}_{r2}) + \nu_{ir2}] \quad (5.19)$$

where $\nu_{irj} \sim N(0, \sigma_\nu^2)$.

We estimate the parameters of $\bar{U}_0(\cdot)$ by Maximum Likelihood, assuming a flexible functional forms for the dependence of utility on a daughter’s age of marriage, education, and match characteristics. We normalize the coefficient on “Government job” to one to facilitate comparison of the magnitude of effect sizes across our alternative specifications and estimate the variance of the idiosyncratic random error. To investigate the importance of observed preference heterogeneity, we analyse the impact of a rich set of respondent demographics on reduced form choice probabilities with a Lasso model (Section 6.1).

Beliefs We estimate subjective beliefs over marriage offer probabilities, in addition to the utility associated with having a daughter still unmarried at time $T + 1$ (i.e. $EV(T + 1)$) and the variance of the additive idiosyncratic preference heterogeneity by maximum likelihood, taking preferences as given. In this procedure, we also impose a constant discount rate of $\beta = 0.95$.

We place a set of functional form assumptions on the probability of receiving marriage offers to facilitate estimation. First, we restrict the number of marriage offer states to three: receive an offer from a high quality groom; receive an offer from a low quality groom; receive no marriage offer. We assume that the probability of receiving *no* marriage offer, $\pi(\bar{\omega}, N)$, is constant over time and across girls with different education levels. Second, we assume that, conditional on receiving an offer, the probability that it is from a high-quality groom is given by:

$$\pi(t, Ed_t) = \Phi(\mathbf{X}'\tau) \quad (5.20)$$

where \mathbf{X} contains a flexible set of age and education controls. Finally, we assume that the value of still being unmarried in $T + 1$ is constant across all education levels.

To find the belief parameters that maximize the likelihood of observing the actual choices that respondents made, we solve the whole model recursively in each iteration.

6 Results

6.1 Reduced Form Variation

We begin our results by simply describing response patterns across our two choice experiments. Figure 4 (a) shows how the probability that an option was chosen in the ex-post experiment varied with the age and education of a daughter, and with the characteristics of potential grooms. Figure 4 (b) shows how the probability that the marriage option was chosen in the ex-ante experiment varied with vignette characteristics.

In the ex-post experiment, respondents were more likely to choose options where the age of marriage was at least 18 but, beyond age 18, choice probabilities don't appear to increase further with age. In the ex-ante experiment, where respondents were left uncertain about the final marriage match if they rejected the offer in hand today, we see that the proportion of respondents accepting the marriage offer increases with age both up to and beyond age 18. The different pattern of choice probabilities with age of marriage across the two experiments *could* be driven by respondents believing that good marriage offers become less likely as girls get older. By bringing together the choice patterns in the two experiments within the structure of the model, we will shed light on whether this is in fact the case.

Turning to education, respondents in the ex-post experiment were more likely to choose options in which a daughter had more education at marriage. However, this doesn't necessarily imply a significant preference for high education due to the positive correlation between age of marriage and education in the options presented to respondents.⁴² Respondents to the ex-ante experiments were substantially more likely to accept a marriage offer if the daughter was already out of school compared to when she was still in school. Indeed, the impact of still being in school on the likelihood of selecting the marriage option is striking: respondents were as likely to accept the marriage offer for a 13 year old daughter who was already out of school as for a 19 year old daughter who was still in school. Respondents were twice as likely to accept the marriage offer when a 16 year old daughter was out of school compared to when she was in school. This is consistent with the views expressed in the qualitative focus groups in which it was said: "If a girl is not studying and is sitting at home, then conversations around her marriage begin."⁴³

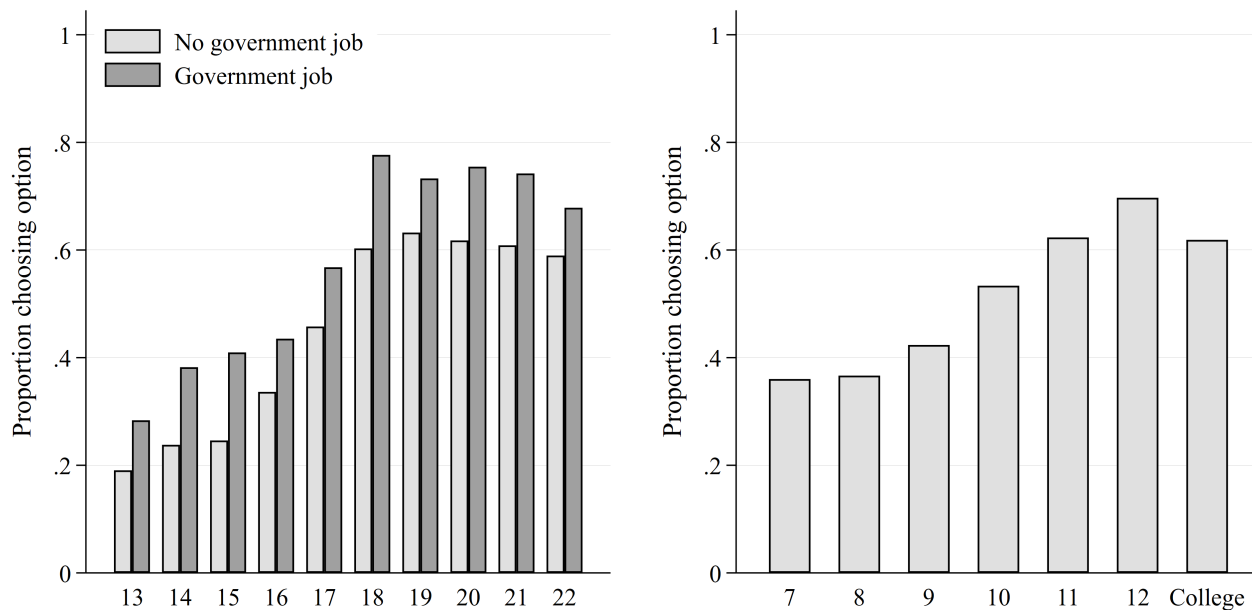
Finally, in both types of experiments, respondents were more likely to choose options associated with the daughter marrying a groom with a government job, the most well-paid and secure form of employment in the area.

⁴²Only older daughters can also acquire high levels of education.

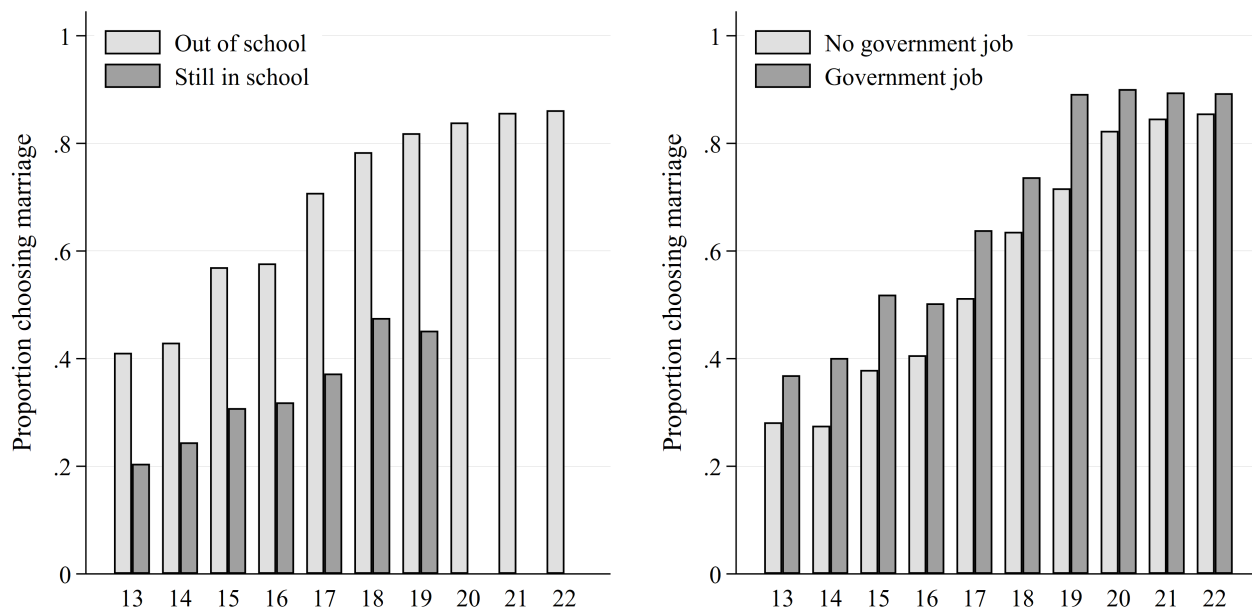
⁴³See Online Appendix B for our focus group discussions.

Figure 4: Ex-Post and Ex-Ante Descriptives

(a) Ex-Post Choices by Age of Marriage (left) and Education at Marriage (right)



(b) Ex-Ante Choices by Age, Schooling Status (left) and Match Characteristics (right)



Notes: Figure plots raw choice probabilities in both experiments. Panel (a) plots the proportion of respondents who chose a given option in the ex-post experiment by the age of marriage (left) and education at marriage (right) of the hypothetical daughter. The left hand graph also plots how this pattern varied by whether or not the groom was described as having a government job. Panel (b) plots the proportion choosing to accept the marriage offer in hand by age. The left graph splits by girls currently in education and those not. The right graph splits by whether the hypothetical groom has a government job.

Heterogeneity on Observables Before proceeding to our structural estimation, we first examine whether reduced-form response patterns differ by respondents’ observed characteristics in order to inform our decision about whether and how to allow structural parameters to differ by observed characteristics. It is important to note that as respondents were instructed to consider the behavior of hypothetical families, we did not anticipate substantial heterogeneity by the observed characteristics of the respondent. Nevertheless, respondents may have different reference points for the preferences and beliefs that they think a typical couple described in the vignette would have.

In practice, we look for heterogeneity by regressing (using a linear probability model) indicators of respondents’ choices on age of marriage, education at marriage, and all groom characteristics, while interacting all scenario characteristics with the set of eighteen respondent characteristics presented in descriptive Table 1. To prevent over-fitting, we employ a Lasso model to assess the most important predictors of heterogeneity. Lasso minimizes the mean-squared error of the prediction subject to a penalty on the absolute size of the estimated coefficients. Both the Extended Bayesian Information Criterion (EBIC) and Rigorous Lasso methods for selecting the optimal tuning parameter agree that no interaction terms should be included in the model and thus that there is no evidence of substantial heterogeneity in response patterns by observed characteristics. Appendix E gives full details of our approach.

We also test whether the similarity of a vignette to the respondent’s own experience affected choices. To do so, we randomly replaced the characteristics of the hypothetical family, and in the ex-post experiment, the description of the daughter at marriage in one option, in one round with the characteristics of the respondent’s own household and daughter (see Section 4). Appendix Tables A.1 and A.2 show the results of this exercise: we find no significant impact of vignette salience.

Given these findings, we will abstract from observed respondent heterogeneity in our structural estimation of parental preferences and beliefs.

6.2 Structural Results

Preferences Table 3 presents our structural preference parameters estimated under a number of different specifications. Across all specifications the coefficient on a groom having a government job is normalised to one to ease comparisons. Figures 5(a) and 5(b) graphically presents the coefficients (from the first specification) on the age of

marriage and education level of a daughter, along with the 95% confidence intervals.

Across all specifications, parents strongly prefer to delay a daughter's marriage up until she is eighteen years of age. Indeed, we estimate that before the age of 18, the value that parents place on delaying marriage by each additional year is between one half and one times the value they place on securing a groom with a government job. Beyond eighteen years of age, which is the legal minimum age of marriage for young women in India, parents show no preference for delaying marriage further. The discontinuity at age 18 is suggestive of parents incorporating the legal minimum age at marriage into their preferences despite evidence that the rules are often only laxly enforced (UNICEF 2011).

Preferences over education suggest that parents prefer a daughter to obtain more education up until the end of high school (12th Standard). Parents also prefer additional years in school if a daughter "likes school" (column 2 and 3).⁴⁴ Preferences for education appear, however, weaker than preferences over age at marriage. For example, we estimate that the value that parents place on each additional year of schooling is only around 14% of the value they attach to securing a groom with a government job. Many of the individual dummies are not statistically significantly different from zero. However, a linear trend in education (excluding college) is statistically significant (in column 3). That being said, conditional on a given marriage market match, parents strictly prefer a daughter to finish high school than to complete college. This might be driven by the higher financial cost of college compared to high school. Indeed, one participant in our focus group discussions remarked that "till class 12th, the expenditure is a little less [on education] but after that it is quite a lot". Note that, as the characteristics of the groom are given, this effect is not due to concerns about "over-education" and marriage market prospects.

Turning to "match utility", parents prefer more educated grooms, younger grooms, and grooms who don't come from a "low wealth" family. Although not statistically significant, grooms who were described as being associated with a lower dowry payment were more likely to be chosen. However, it is whether a groom has a government job that is the most important driver of groom quality. We estimate that parents value grooms having a government job four times as much as they value the groom coming from a medium or high wealth family (as opposed to a low wealth family). The groom having a government job increases the latent index by nearly 40% of a standard deviation of the 'unexplained' variation. The importance of government job we uncover here echoes the views of participants in focus group discussions, one of whom commented that "So many boys are sitting at home after

⁴⁴This corresponds closely to themes highlighted in our focus group discussions in which the importance of a daughter's own motivation in forming her parents' views of how long she should stay in school was stressed. For example, one participant mentioned: "If the child is hardworking and good, the parents will not have to ask him/her to study or to go to school, they themselves do".

doing BA and have nothing to do, there are no jobs or even if they do (get a job), they are not secure, in that case, having a government job really matters.”⁴⁵

In column (4), we impose a set of functional form restrictions to increase the ease with which we can use our preference results to estimate subjective beliefs. Given the coefficients on the age dummies in columns 1 and 2, we impose a piecewise linear specification for age, with a potential kink and discontinuity at age 18. We impose that preferences for education are linear, but allow preferences for college to differ through the inclusion of a dummy variable. Third, we remove the two shifters of preferences that were not significant.

⁴⁵Caregivers Focus Group 2, Online Appendix B.

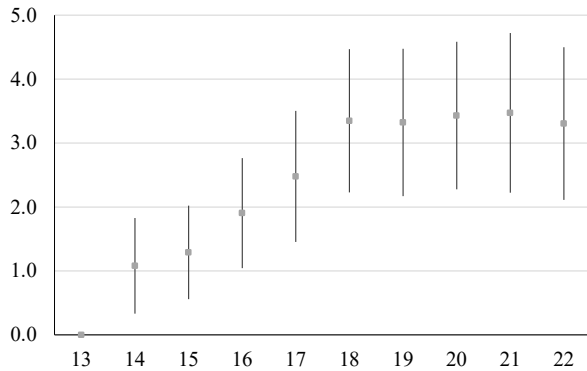
Table 3: Structural Preference Parameters

	(1)	(2)	(3)
Groom characteristics			
Groom has government job	1.000	1.000	1.000
Groom's age	-0.037 (.015)	-0.037 (.015)	-0.037 (.014)
Groom's education	0.061 (.019)	0.061 (.019)	0.062 (.018)
Groom college education	0.025 (.159)	0.025 (.159)	0.025 (.159)
Low wealth	0.000	0.000	0.000
Medium wealth	0.227 (.12)	0.220 (.121)	0.220 (.128)
High wealth	0.168 (.14)	0.165 (.145)	0.164 (.153)
Dowry (lakh)	-0.059 (.031)	-0.059 (.033)	-0.059 (.034)
Age of Marriage			
13	0.000	0.000	
14	1.079 (.382)	1.082 (.374)	
15	1.291 (.372)	1.295 (.363)	
16	1.904 (.439)	1.923 (.407)	
17	2.478 (.522)	2.505 (.48)	
18	3.348 (.572)	3.388 (.552)	
19	3.322 (.588)	3.368 (.566)	
20	3.428 (.588)	3.488 (.583)	
21	3.472 (.637)	3.540 (.58)	
22	3.305 (.609)	3.381 (.572)	
Age-13			0.511 (.094)
1(Age \geq 18)			2.969 (.517)
(Age-13) x 1(Age \geq 18)			-0.504 (.101)
Education at Marriage			
7th	0.000	0.000	
8th	0.136 (.148)	0.080 (.157)	
9th	0.265 (.202)	0.152 (.185)	
10th	0.450 (.184)	0.279 (.197)	
11th	0.546 (.184)	0.311 (.216)	
12th	0.906 (.175)	0.601 (.231)	
College	0.246 (.155)	-0.111 (.238)	-0.697 (.198)
Education (cont.)			0.132 (.038)
Preference shifters			
'Daughter likes school' x Years of schooling		0.074 (.04)	0.075 (.039)
'School is costly' x Years of schooling		0.047 (.04)	
'Help at home needed' x Years before marriage		-0.019 (.03)	
Preference shocks			
σ_ν	2.636 (.376)	2.636 (.376)	2.649 (.369)
Shifters		Y	Y
Number of experiments	6971	6971	6971
Number of respondents	2324	2324	2324

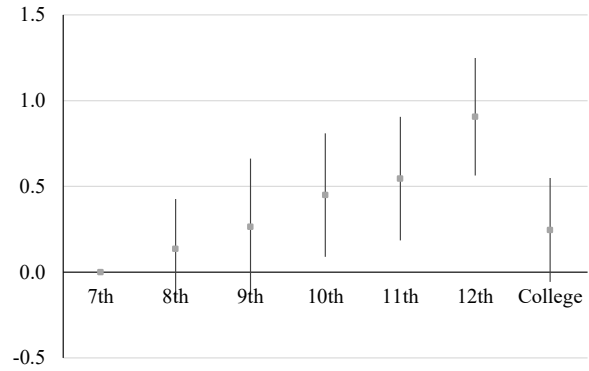
Notes: Table presents structural preference and standard errors (in parentheses). Standard errors calculated through bootstrap, re-sampling respondents with replacement, using 250 iterations. Value of a government job normalized to 1. Column (1) presents results from a model with only groom characteristics, age of marriage, and education at marriage. Column (2) additionally includes our three preference shifting instruments. Column (3) is the specification we carry forward to the belief estimation, adopting a piecemeal linear specification for age and education and dropping the insignificant instruments.

Figure 5: Preferences over a Daughter’s Age and Education at Marriage

(a) Age of Marriage



(b) Education at Marriage



Notes: Figure plots coefficients on age of marriage dummies (left) and education at marriage dummies (right) alongside 95% confidence intervals. Estimates taken from specification (1) in Table 3.

Subjective Beliefs To estimate revealed beliefs over the likelihood of getting marriage offers, we discretize groom quality into two levels on the basis of our estimated preference parameters. The value placed on a groom having a government job is so large relative to other groom characteristics that it creates a discontinuous support for the groom quality distribution amongst our fictional grooms (see Appendix Figure A.1). Thus, we split grooms into those with a government job (“high” quality grooms) and those without (“low” quality grooms). In practice, this split is at the 80th percentile of the fictional groom quality distribution. From both types of grooms, we use the average value of match utility associated with that type.

The probability of receiving an offer from a high type groom is allowed to depend flexibly on a daughter’s age, her education, and whether or not her behavior conforms to gendered norms. We assume a constant probability that the parents receive any marriage offer in a given period. As discussed in Section 5.2 and expanded on in Appendix D, identification of a model where offers are not necessarily received every period is facilitated by the inclusion of preference-shifting instruments, in our case an indicator of whether or not the daughter “likes school”.

Table 4 provides the structural belief coefficients and standard errors. Figure 6 plots the revealed beliefs over the likelihood of a daughter receiving a marriage offer from a high quality groom for a daughter whose behavior conforms to gendered norms.

Looking first at beliefs around how offer probabilities depend on a daughter’s age, we see a striking contrast from parents’ preferences. While parents prefer to delay a daughter’s marriage until the end of adolescence, parents

Figure 6: Revealed Beliefs

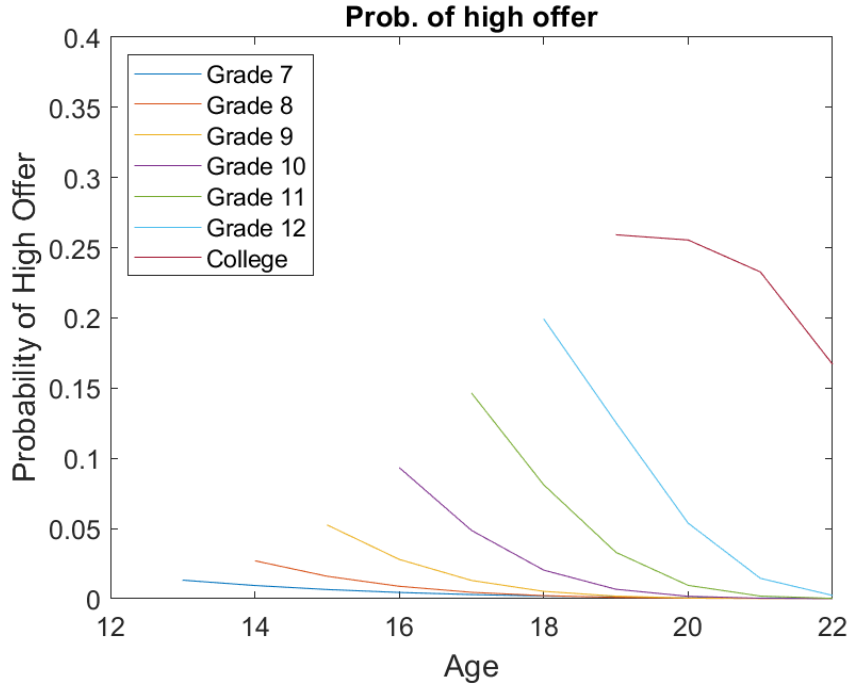


Figure shows the subjective probability of receiving an offer from a high quality groom at different ages and education levels of the daughter. Probabilities calculated for a daughter who conforms to gender norms.

Table 4: Structural Belief Parameters

(1)		
Age	-0.160	(.067)
Education	0.513	(.356)
Education ²	0.147	(.044)
College	1.708	(.335)
Age*Education	-0.123	(.028)
Daughter conforms to gender norms	0.306	(.137)
Constant	-1.475	(.423)
Offer Probability	0.260	(.026)
V(T+1)	-1.995	(1.035)
σ_ε	1.083	(.227)
Number of experiments	5524	
Number of respondents	2258	

Notes: Table provides the structural parameters and standard errors obtained by bootstrapping the whole 2-step estimation procedure 250 times.

believe that the likelihood of a marriage offer from a high quality groom is declining in her age once she is out of school. The coefficient on age is negative and highly significant. We estimate that this deterioration in offer prospects with age is particularly acute for daughters with more schooling, as seen through the significantly negative interaction between age and education. Take a daughter who has finished high school (grade 12) for instance. Figure 6 shows that we estimate that right after finishing school this daughter would have a 20% chance of receiving an offer from a groom with a government job but that this would fall to just 5% after two more years. These deteriorating marriage market returns to age are important because they imply that parents perceive that there is only a short window of opportunity for daughters to receive high quality marriage offers after they've left school. This creates incentives for parents to marry daughters as soon as possible after they leave education and thus contribute to instances of early marriage.

Turning next to education, we see that a daughter's level of schooling has a huge effect on how likely her parents think it is that they will receive high quality marriage offers for her. We recover a belief of a large marriage market return to education because respondents are less likely to accept marriage offers when a girl is in school in the ex-ante experiment than would be expected from the preference parameters alone. A strong expected marriage market return to education rationalizes these differences. Table 4 shows that both the linear and quadratic terms in education are positive and, on top of that, a dummy for college education is also large and highly significant. This suggests that there are particularly high returns to higher levels of education. We see this in Figure 6. If we consider 18 year old daughters, we see that parents perceive that they will have a 20% chance of an offer from a high quality groom if they've finished high school, an 8% chance if they've finished 11th standard and only a 2% chance if they've finished 10th standard.

Beliefs about the impact of college education on marriage offers differ markedly from parents' preferences for college conditional on a given marriage match. Recall that conditional on a given marriage match parents strictly preferred their daughter to leave education after high school rather than proceed to college. However, parents believe that college education hugely increases the probability of receiving good quality marriage offers.

A further feature of the beliefs we estimate is that parents perceive that daughters who's behavior breaks social norms, as described by the statement "she is friends with some boys and sometimes stays out of the house until late", will receive lower quality marriage offers than daughters who "are polite and well behaved". This suggests that stringent expectations regarding the proper behavior of young unmarried women can contribute to early marriage

in this context; when parents worry that their daughter might behave in a way that would damage her marriage prospects, they are inclined to secure her marriage rapidly.

6.3 Belief Validation

Given the novelty of our method, and the fact that our subjective expectations are structurally inferred rather than measured directly, we complement the results from our ex-ante experiment with two validation experiments.

- **Groom Side Preference Experiment** We also elicit “groom side” preferences over the characteristics of brides in a manner akin to our ex-post choice experiments. The likelihood of receiving a marriage offer from a high quality groom depends on the preferences of grooms over bridal characteristics. We presented respondents with a vignette of a hypothetical family who is looking to marry their *son*, specifying the same groom side attributes as discussed in Section 4 (see Figure A.3 for the visual aid). We presented two potential marriage options and ask respondents to select the one that they believe the hypothetical family would prefer. We included all verifiable bride side attributes (i.e. we do not describe whether the mother of the bride requires help at home or the attributes relating to the cost and enjoyment of schooling) and the highest dowry that the bridal family is willing to pay. This allows us to establish whether the patterns emerging from our revealed beliefs are consistent with groom side preferences over potential brides.
- **Expected Match Experiment** In a first validation exercise, we directly elicit how expected match quality varies with the characteristics of a potential bride. To do so, we provided respondents with information on the age, education, and wealth of a hypothetical bride and asked respondents to describe the attributes of the most likely groom they believed such a girl would marry if her parents wished to marry her at this point. Respondents identified the characteristics of the groom they thought most likely for a girl with the characteristics described to marry on a visual aid (see Figure A.2 for the visual aid). Delavande, Giné, and McKenzie (2011) find a strong correlation between answers to “what do you expect” questions and the mean of elicited subjective distributions.

Our revealed beliefs estimates are qualitatively consistent with the patterns revealed by our groom-side experiment and expected match experiment. Figure 7 (a) gives the key result from the groom-side experiment, showing the coefficients on a set of education dummies from a probit of option choice on vignette characteristics (see Table A.5 in Appendix A for all coefficients). We find a strong increasing preference for female education amongst

grooms with a government job and the coefficient on female college education is large and highly significant for these families. This is consistent with the patterns recovered in the revealed belief estimation, which also implied a much higher probability of matching with a groom with a government job if the daughter was educated to college level. Interestingly, we do not observe the same preference for female college education amongst families of grooms without government jobs. This might indicate that for these grooms the benefits of a highly educated wife might be outweighed by her negotiating a greater share of household resources once married (Anderson and Bidner 2015).

Likewise, we see that families of grooms dislike potential brides who have male friends (Table A.5). This is consistent with our revealed beliefs estimates that these young women are perceived to be less likely to receive high quality marriage offers.

When asked to predict the characteristics of the best groom offer that girls of different characteristics would get, one observes a strong effect of education on the likelihood that a respondent predicts a match with a government job, while age controls (conditional on education) are insignificant. Figure 7 (b) shows the coefficient on daughter’s education for the least squares regression:

$$PredictGov_{ir} = X_r\beta + \epsilon_{ir} \tag{6.1}$$

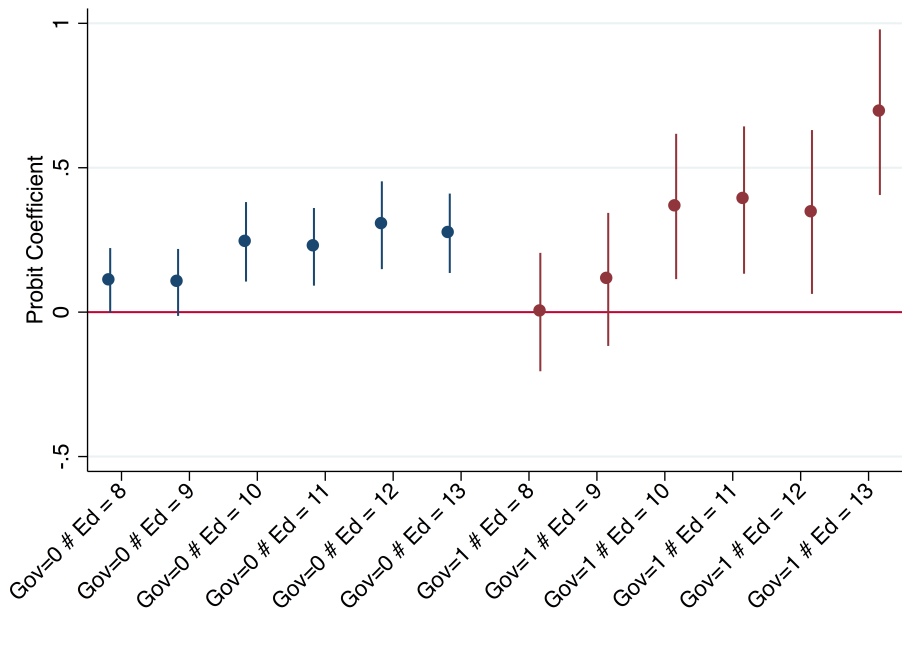
where $PredictGov_{ir} = 1$ if a respondent expects the daughter described in round r with characteristics X_r to receive an offer from a groom with a government job and $PredictGov_{ir} = 0$ otherwise (see Table A.6 in Appendix A for all coefficients). There is a linear relationship between a daughter’s education and the likelihood of predicting a match with a government job, which is broadly consistent with our revealed belief estimates.⁴⁶

Note that the probability levels implied by our revealed belief measures are more plausible than those arising from our direct expectations experiment. While the patterns from our direct expectations experiment also imply a belief in positive assortative matching, respondents in this experiment were much more likely to state that a daughter could expect a match with a government job than is plausible given the representative survey data: less than 10% of men aged between 21 and 30 have a professional job in Rajasthan but in our direct expectations experiment in 70% of rounds involving an 18 year old girl, respondents predicted a groom with a government

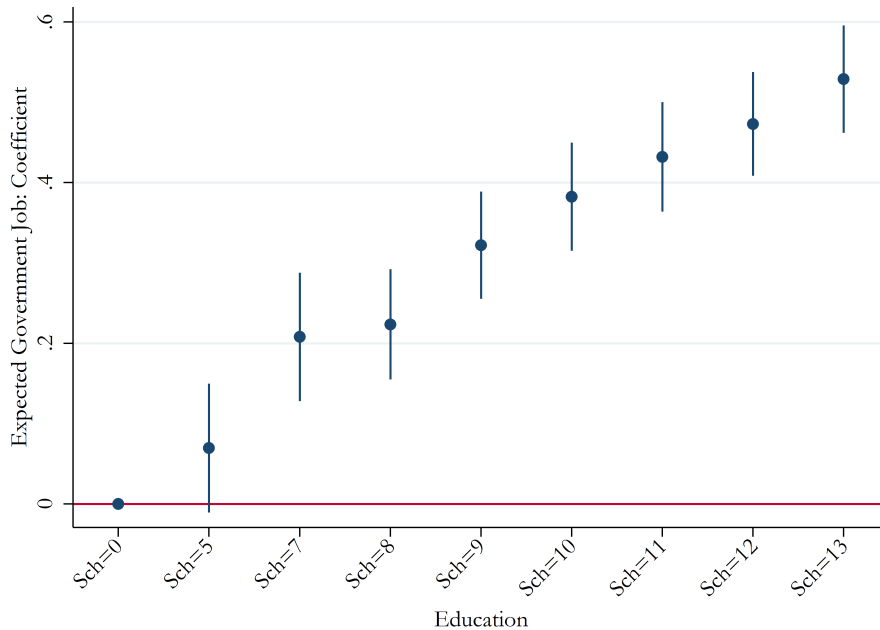
⁴⁶However, we note that the two sets of results are not directly comparable: while the revealed beliefs give the average conditional probability that respondents attach to a young woman getting an offer from a high quality groom, the expected match results provide information on how whether or not a respondent “expects” that the accepted match will have a government job varies with a young woman’s characteristics. We would anticipate that respondents answer that they “expect” the accepted match to have a government job if they perceive the probability to be greater than 50% although we are not aware of any research that explicitly assesses this assumption.

Figure 7: Validation Results

(a) Groom-Side Experiment
Groom Preferences over Age & Education



(b) Expected Match Experiment
Education & Groom with Government Job



Notes: Panel (a) plots coefficients on the education of a potential bride interacted with whether the groom has a government job and 95% confidence interval from a reduced form probit regression of whether the first option was chosen in the groom-side experiment on the characteristics of the hypothetical groom, groom's family and the potential bridal matches. Full set of coefficients given at Table A.5. Panel (b) plots coefficients on the education of a hypothetical daughter and 95% confidence interval from a linear regression of whether a match with a government job was predicted in the direct expectations experiment on the characteristics of the hypothetical daughter and her family. Full set of coefficients at Table A.6.

Table 5: Population Matching Patterns by Husband’s Occupation

All India (N=15,440) <i>Husband’s Occupation</i>	<i>Wife’s Education</i>			
	No Education	Primary	Secondary	Higher
Professional/Technical/Managerial/Clerical (7.51%)	5.78%	6.21%	49.01%	39.00%
Other (92.49%)	24.17%	14.93%	52.94%	7.96%

Rajasthan (N=1,091) <i>Husband’s Occupation</i>	<i>Wife’s Education</i>			
	No Education	Primary	Secondary	Higher
Professional/Technical/Managerial/Clerical (9.53%)	8.65%	14.42%	35.58%	41.35%
Other (90.47%)	36.27%	17.53%	35.97%	10.23%

Notes: Table presents the distribution of wives’ education by husband’s occupation type for couples where the husband is aged 21 to 30 in both the whole of India (top panel) and Rajasthan (bottom panel) from the NFHS-IV (2015/16).

job. This is consistent with findings in other settings in which the gradient of subjective expectations in different characteristics is more accurate than their level (Delavande and Rohwedder 2011).

Our revealed belief results are qualitatively consistent with representative data on matching patterns in the community and also with observations from our focus group discussions in which caregivers were optimistic that having a more educated daughter would mean that she married a more educated and better quality groom, one commenting that: “[if] a girl is educated she will go to a good house and will lead a good life”.⁴⁷ Other women commented directly on the assortative nature of matching in the area: “If our daughter is a graduate, we will look for a match for her who is a graduate as well.”⁴⁸ Table 5 gives the realized marriage matching patterns on education for India as a whole and Rajasthan specifically in 2015/16. We observe strong assortative matching on education with 41% of grooms with a professional job marrying a girl with higher education.

7 Conclusion

In this paper, we developed a novel methodology to identify parents’ subjective beliefs over the marriage market return to girls’ schooling and age and their preferences over female education and age at marriage. We administered our experiments with 4,605 female caregivers living in 120 villages in rural Rajasthan, a context with rigid patriarchal gender norms.

Drawing on the growing use of hypothetical choice experiments in economics to identify preferences, we offered respondents choices between randomly drawn options to identify average parental preferences over a daughter’s age of marriage, education and marriage match characteristics in the absence of uncertainty over future marriage

⁴⁷FGD 2, Online Appendix B

⁴⁸FGD 1, Online Appendix B

offers. We then extended the hypothetical choice approach by also offering respondents choices over whether or not to accept a given marriage offer in situations where there was uncertainty about which future offers would be received. Within a standard dynamic choice framework, these choices identify subjective beliefs over the joint distribution of age of marriage, education, and match quality. Our structural approach enables us to identify “revealed beliefs” that reconcile choices made with and without uncertainty, without directly eliciting probabilities from respondents.⁴⁹ This was an important benefit given our respondents had very low levels of literacy and numeracy and the experiments were carried out by non-specialist enumerators in a large field survey. Furthermore, the approach is well suited to elicit beliefs over continuous and multi-dimensional objects which is often challenging with existing approaches. Our revealed belief estimates are qualitatively consistent with estimates of groom-side preferences, directly elicited expected matches, and assortative matching patterns in observational data.

Our results suggest that the perceived marriage market return to education is the primary driver of female secondary schooling in our context. While parents place intrinsic value on a daughter’s education up until the end of high school, this preference is small in magnitude and decreasing in post-secondary education. However, our revealed belief estimates indicate that parents believe that the likelihood of a daughter receiving an offer from a high quality groom increases substantially with education, particularly so with college education. This creates a sizeable perceived marriage market return to a daughter’s education. This perceived marriage market return to education may be important in understanding why female education in India has continued to increase while female labor force participation has fallen further from an already low base (Fletcher, Pande, and Moore 2017).

Although parents have a strong distaste for marrying a daughter before age eighteen, a belief that a girl’s marriage prospects will start to deteriorate when she drops out of school can create a mechanism for early marriage. Indeed, our respondents were twice as likely to accept a marriage offer for a 16 year old daughter if the girl was not in school, and were equally likely to accept a given marriage offer made to a 13 year old girl out of school as for a 19 year old girl who was still in school. The beliefs we identify suggest that policies that ensure girls have safe and affordable access to high quality schooling and those that insulate adolescent girls from shocks that might otherwise cause premature school dropout will be fundamental in reducing rates of early marriage.

⁴⁹In future work we will experimentally validate our belief measures in a more simple and controlled environment in which we can compare our structurally derived measures with those gained through more standard elicitation methods.

References

- Achyut, P., A. Andrew, M. Das, A. Gautam, M. Huepe, S. Krutikova, S. Kumar, S. Sharma, R. Soni, H. Verma, and R. Verma (2016). Promoting Adolescent Engagement, Knowledge and Health Evaluation of PANKH : an adolescent girl intervention in Rajasthan, India. Baseline Report. Technical report.
- Ahrens, A., C. B. Hansen, and M. Schaffer (2019). LASSOPACK: Stata module for lasso, square-root lasso, elastic net, ridge, adaptive lasso estimation and cross-validation.
- Alesina, A., S. Stantcheva, and E. Teso (2018, feb). Intergenerational Mobility and Preferences for Redistribution. *American Economic Review* 108(2), 521–554.
- Almås, I., A. W. Cappelen, and B. Tungodden (2020). Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than Scandinavians? *Journal of Political Economy* 128(5), 1753–1788.
- Ameriks, J., J. Briggs, A. Caplin, M. D. Shapiro, and C. Tonetti (2020, jun). Long-Term-Care Utility and Late-in-Life Saving. *Journal of Political Economy* 128(6), 2375–2451.
- Andersen, S., J. Fountain, G. W. Harrison, and E. E. Rutström (2014). Estimating subjective probabilities. *Journal of Risk and Uncertainty* 48(3), 207–229.
- Anderson, S. and C. Bidner (2015, aug). Property Rights over Marital Transfers. *The Quarterly Journal of Economics* 130(3), 1421–1484.
- Arcidiacono, P., V. J. Hotz, and S. Kang (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics* 166(1), 3–16.
- Attanasio, O. (2021). Economics and Measurement: New Measures to Model Decision Making. In *Econometric Society Presidential Address*.
- Attanasio, O., F. Cunha, and P. Jervis (2019). Subjective Parental Beliefs. Their Measurement and Role. *NBER Working Paper*, No. 26516.
- Attanasio, O. P. and K. M. Kaufmann (2017, aug). Education choices and returns on the labor and marriage markets: Evidence from data on subjective expectations. *Journal of Economic Behavior & Organization* 140, 35–55.

- Banerjee, A., E. Duflo, M. Ghatak, and J. Laforune (2013, may). Marry for What? Caste and Mate Selection in Modern India. *American Economic Journal: Microeconomics* 5(2), 33–72.
- Becker, G. S. (1973, jul). A Theory of Marriage: Part I. *Journal of Political Economy* 81(4), 813–846.
- Behrman, J. R., A. D. Foster, M. R. Rosenzweig, and P. Vashishtha (1999). Women’s Schooling, Home Teaching, and Economic Growth. *Journal of Political Economy* 107(4), 682–714.
- Ben-Akiva, M., D. McFadden, and K. Train (2019). Foundations of Stated Preference Elicitation: Consumer Behavior and Choice-based Conjoint Analysis. *Foundations and Trends in Econometrics* 10(1-2), 1–144.
- Bergstrom, K. and B. Özler (2021). Improving the Well-Being of Adolescent Girls in Developing Countries. *World Bank Working Paper* (November).
- Berry, J., R. Dizon-Ross, and M. Jagnani (2020). Not Playing Favorites: An Experiment on Parental Fairness Preferences. *SSRN Electronic Journal* (June).
- Bhalotra, S., A. Delavande, P. Font Gilabert, and J. Maselko (2020). Maternal Investments in Children: The Role of Expected Effort and Returns. *IZA Discussion Paper* (13056).
- Bjornerstedt, J. and J. Weibull (1994). Nash Equilibrium and Evolution by Imitation. *IUI Working Paper* 407.
- Bloch, F. and V. Rao (2002, aug). Terror as a Bargaining Instrument: A Case Study of Dowry Violence in Rural India. *American Economic Review* 92(4), 1029–1043.
- Bloch, F., V. Rao, and S. Desai (2004). Wedding Celebrations as Conspicuous Consumption: Signaling Social Status in Rural India. *The Journal of Human Resources* 39(3), 675.
- Bobba, M. and V. Frisncho (2020). Self-perceptions about academic achievement: evidence from Mexico City. *Journal of Econometrics*.
- Boneva, T. and C. Rauh (2018). Parental Beliefs about Returns to Educational Investments—The Later the Better? *Journal of the European Economic Association* 16(6), 1669–1711.
- Borkotoky, K. and S. Unisa (2014, mar). Indicators to Examine Quality of Large Scale Survey Data: An Example through District Level Household and Facility Survey. *PLoS ONE* 9(3), e90113.
- Buchmann, N., E. Field, R. Glennerster, S. Nazneen, S. Pimkina, and I. Sen (2017). Power vs Money: Alternative Approaches to Reducing Child Marriage in Bangladesh, a Randomized Control Trial.

- Burde, D. and L. L. Linden (2013). Bringing education to afghan girls: A randomized controlled trial of village-based schools. *American Economic Journal: Applied Economics* 5(3), 27–40.
- Bursztyn, L., T. Fujiwara, and A. Pallais (2017, nov). 'Acting Wife': Marriage Market Incentives and Labor Market Investments. *American Economic Review* 107(11), 3288–3319.
- Bursztyn, L., A. L. González, and D. Yanagizawa-Drott (2018). Misperceived Social Norms: Female Labor Force Participation in Saudi Arabia. Working Paper 24736, National Bureau of Economic Research.
- Caplin, A. (2016). Measuring and modeling attention. *Annual Review of Economics* 8, 379–403.
- Carneiro, P., C. Meghir, and M. Parey (2013). Maternal Education, Home Environments and the Development of Children and Adolescents. *Journal of the European Economic Association* 11(1-4), 123–160.
- Chari, A., R. Heath, A. Maertens, and F. Fatima (2017, jul). The causal effect of maternal age at marriage on child wellbeing: Evidence from India. *Journal of Development Economics* 127, 42–55.
- Charness, G. and D. Levin (2005). When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect. *American Economic Review* 95(4), 1300–1309.
- Chiappori, P.-A., M. C. Dias, and C. Meghir (2015). The marriage market, labor supply and education choice. Technical report, National Bureau of Economic Research.
- Corno, L., N. Hildebrandt, and A. Voena (2016). Weather Shocks, Age of Marriage and the Direction of Marriage Payments. *Working Paper*.
- Corno, L. and A. Voena (2016). Selling Daughters: Age of Marriage, Income Shocks and the Bride Price Tradition. *Working Paper*, 1–67.
- Cunha, F., I. Elo, and J. Culhane (2013, jun). Eliciting Maternal Expectations about the Technology of Cognitive Skill Formation. Technical Report 9, National Bureau of Economic Research, Cambridge, MA.
- Delavande, A., X. Giné, and D. McKenzie (2011, mar). Measuring subjective expectations in developing countries: A critical review and new evidence. *Journal of Development Economics* 94(2), 151–163.
- Delavande, A. and S. Rohwedder (2011). Differential survival in Europe and the United States: Estimates based on subjective probabilities of survival. *Demography* 48(4), 1377–1400.
- Delavande, A. and B. Zafar (2018). University Choice: The Role of Expected Earnings, Non-pecuniary Outcomes and Financial Constraints. *Journal of Political Economy*.

- Delavande, A. and B. Zafar (2019). University choice: The role of expected earnings, nonpecuniary outcomes, and financial constraints. *Journal of Political Economy* 127(5), 2343–2393.
- Dhar, D., T. Jain, and S. Jayachandran (2021). Reshaping Adolescents’ Gender Attitudes: Evidence from a School-Based Experiment in India. *American Economic Review*.
- Dizon-Ross, R. (2019). Parents’ beliefs about their children’s academic ability: Implications for educational investments. *American Economic Review* 109(8), 2728–2765.
- Duflo, E. (2012). Women Empowerment and Economic Development. *Journal of Economic Literature* 50(4), 1051–1079.
- Duflo, E., P. Dupas, and M. Kremer (2015). Education, HIV, and early fertility: Experimental evidence from Kenya. *American Economic Review* 105(9), 2757–2797.
- Edmonds, E., B. Feigenberg, and J. Leight (2021, jul). Advancing the Agency of Adolescent Girls. *The Review of Economics and Statistics*, 1–46.
- Eika, L., M. Mogstad, and B. Zafar. Educational assortative mating and household income inequality.
- Erev, B. I. and A. E. Roth (1988). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review* 88(4), 848–881.
- Finch, J. (1987). The vignette technique in survey research. *Sociology* 21(1), 105–114.
- Fisman, R., S. S. Iyengar, E. Kamenica, and I. Simonson (2006). Gender differences in mate selection: Evidence from a speed dating experiment. *Quarterly Journal of Economics* 121(2), 673–697.
- Fletcher, E., R. Pande, and C. M. T. Moore (2017). Women and work in india: Descriptive evidence and a review of potential policies.
- Giustinelli, P., C. F. Manski, and F. Molinari (2020). Precise or Imprecise Probabilities? Evidence from Survey Response Related to Late-onset Dementia. *Journal of the European Economic Association*.
- Greenwood, J., N. Guner, G. Kocharkov, and C. Santos (2014). Marry your like: Assortative mating and income inequality. *American Economic Review* 104(5), 348–353.
- Grépin, K. A. and P. Bharadwaj (2015, dec). Maternal education and child mortality in Zimbabwe. *Journal of Health Economics* 44, 97–117.

- Hitsch, G. J., A. Hortaçsu, and D. Ariely (2010). Matching and sorting in online dating. *American Economic Review* 100(1), 130–163.
- Hughes, R. (1998). Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health & Illness* 20(3), 381–400.
- Jensen, R. and R. Thornton (2003). Early female marriage in the developing world. *Gender & Development* 11(2), 9–19.
- Kahneman, D. and A. Tversky (1973). On the psychology of prediction. *Psychological Review* 80(4), 237–251.
- Kesternich, I., F. Heiss, D. McFadden, and J. Winter (2013). Suit the action to the word, the word to the action: Hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics* 32(6), 1313–1324.
- Lafortune, J. (2013, jan). Making Yourself Attractive: Pre-Marital Investments and the Returns to Education in the Marriage Market. *American Economic Journal: Applied Economics* 5(2), 151–178.
- Maertens, A. (2013, jul). Social Norms and Aspirations: Age of Marriage and Education in Rural India. *World Development* 47, 1–15.
- Manski, C. F. (2004). Measuring expectations. *Econometrica* 72(5), 1329–1376.
- Mas, A. and A. Pallais (2017). Valuing alternative work arrangements. *American Economic Review* 107(12), 3722–3759.
- Matzkin, R. L. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica: Journal of the Econometric Society*, 239–270.
- Miller, G., A. de Paula, and C. Valente (2020). Subjective expectations and demand for contraception Subjective Expectations and Demand for Contraception.
- Muralidharan, K. and N. Prakash (2017). Cycling to School: Increasing Secondary School Enrollment for Girls in India. *American Economic Journal: Applied Economics* 9(3), 321–350.
- Nash, J. (1951). Non-cooperative games. *Annals of mathematics*, 286–295.
- Pencavel, J. (1998). Assortative Mating by Schooling and the Work Behavior of Wives and Husbands. *The American Economic Review* 88(2), 326–329.
- Ramsey, F. P. (1926). Truth and probability. Reprinted in. *Studies in subjective probability*, 61–92.

- Rao, V. (1993). The Rising Price of Husbands : A Hedonic Analysis of Dowry Increases in Rural India. *Journal of Political Economy* 101(4), 666–677.
- Researve Bank of India (2017). Handbook of Statistics on Indian Economy 2016-17. Technical report.
- Rosenzweig, M. R. and O. Stark (1989, aug). Consumption Smoothing, Migration, and Marriage: Evidence from Rural India. *Journal of Political Economy* 97(4), 905–926.
- Rosenzweig, M. R. and K. I. Wolpin (1994). Are There Increasing Returns to the Intergenerational Production of Human Capital? Maternal Schooling and Child Intellectual Achievement. *The Journal of Human Resources* 29(2), 670.
- Rust, J. B. Y. (1987). Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher. *Econometrica* 55(5), 999–1033.
- Sautmann, A. (2017). Age-dependent payoffs and assortative matching by age in a market with search. *American Economic Journal: Microeconomics* 9(2), 263–294.
- Singh, R. and U. Vennam (2016). Factors shaping trajectories to child and early marriage: Evidence from Young Lives in India.
- Siow, A. (2015, apr). Testing Becker’s Theory of Positive Assortative Matching. *Journal of Labor Economics* 33(2), 409–441.
- Tversky, A. and D. Kahneman (1974). Judgment under Uncertainty: Heuristics and Biases Amos. *Science* 185(4157), 1124–1131.
- UNICEF (2011). Desk Review of Child Marriage. Technical report.
- Wiswall, M. and B. Zafar (2018, feb). Preference for the Workplace, Investment in Human Capital, and Gender. *The Quarterly Journal of Economics* 133(1), 457–507.

A Additional Results

Table A.1: Importance of Vignette Salience in Ex-Post Experiment

	(1)		(2)	
Ex Post Choice				
Option is Like Respondent's Daughter	0.0716*	(0.0387)	0.0273	(0.0486)
Daughter's Age at Marriage=14	0.2298**	(0.1093)	0.2035*	(0.1225)
Daughter's Age at Marriage=15	0.3336***	(0.1077)	0.2827**	(0.1201)
Daughter's Age at Marriage=16	0.4992***	(0.1077)	0.5699***	(0.1339)
Daughter's Age at Marriage=17	0.6388***	(0.1104)	0.6900***	(0.1453)
Daughter's Age at Marriage=18	0.8870***	(0.1117)	0.9424***	(0.1370)
Daughter's Age at Marriage=19	0.8889***	(0.1101)	0.9389***	(0.1407)
Daughter's Age at Marriage=20	0.9785***	(0.1170)	0.9945***	(0.1754)
Daughter's Age at Marriage=21	0.9851***	(0.1176)	0.9139***	(0.1756)
Daughter's Age at Marriage=22	0.9885***	(0.1229)	1.0505***	(0.1977)
Daughter's Education =8	-0.0019	(0.0431)	-0.1363	(0.2211)
Daughter's Education =9	0.0024	(0.0475)	-0.0240	(0.2149)
Daughter's Education =10	0.0277	(0.0536)	0.1773	(0.2038)
Daughter's Education =11	0.0288	(0.0605)	-0.0455	(0.2035)
Daughter's Education =12	0.0982	(0.0671)	-0.0514	(0.2036)
Daughter's Education =13	-0.1248*	(0.0710)	-0.2225	(0.1890)
Daughter Likes School*Years in School	0.0206**	(0.0097)	0.0212**	(0.0098)
Cost of School Covered*Years in School	0.0123	(0.0096)	0.0115	(0.0096)
Help in Home*Years at Home	-0.0233**	(0.0103)	-0.0233**	(0.0104)
Groom's Age	-0.0095**	(0.0038)	-0.0095**	(0.0038)
Groom's Education	0.0158***	(0.0033)	0.0162***	(0.0033)
Government Job	0.2731***	(0.0335)	0.2725***	(0.0337)
Dowry	-0.0172*	(0.0090)	-0.0179**	(0.0090)
Groom's Wealth = Poor	-0.0456	(0.0324)	-0.0422	(0.0325)
Groom's Wealth = Wealthy	-0.0024	(0.0312)	0.0050	(0.0314)
Age=14 X Ed=8			0.2126	(0.2426)
Age=15 X Ed=8			0.1476	(0.2429)
Age=15 X Ed=9			0.1894	(0.2366)
Age=16 X Ed=8			0.0371	(0.2505)
Age=16 X Ed=9			0.0215	(0.2435)
Age=16 X Ed=10			-0.2113	(0.2301)
Age=17 X Ed=8			0.1905	(0.2597)
Age=17 X Ed=9			-0.2423	(0.2535)
Age=17 X Ed=10			-0.2403	(0.2411)
Age=17 X Ed=11			0.1260	(0.2348)
Age=18 X Ed=8			0.0738	(0.2611)
Age=18 X Ed=9			0.0493	(0.2577)
Age=18 X Ed=10			-0.3380	(0.2475)
Age=18 X Ed=11			-0.0843	(0.2402)
Age=18 X Ed=12			0.2080	(0.2280)
Age=19 X Ed=8			0.2299	(0.2601)
Age=19 X Ed=9			-0.1148	(0.2601)
Age=19 X Ed=10			-0.1021	(0.2513)
Age=19 X Ed=11			-0.0200	(0.2426)
Age=19 X Ed=12			0.1152	(0.2357)
Age=19 X Ed=13			0.0225	(0.2305)
Age=20 X Ed=8			0.0831	(0.2911)
Age=20 X Ed=9			-0.0755	(0.2787)
Age=20 X Ed=10			-0.2154	(0.2810)
Age=20 X Ed=11			0.1090	(0.2681)
Age=20 X Ed=12			0.0121	(0.2608)
Age=20 X Ed=13			0.1570	(0.2367)
Age=21 X Ed=8			0.1318	(0.2960)
Age=21 X Ed=9			0.1309	(0.2846)
Age=21 X Ed=10			0.1223	(0.2724)
Age=21 X Ed=11			0.1809	(0.2675)
Age=21 X Ed=12			0.2084	(0.2672)
Age=21 X Ed=13			0.1613	(0.2392)
Daughter Options	yes		yes	
Match Characteristics	yes		yes	
Interactions	no		no	
Number of Choice Experiments	6320		6320	

Notes: Table presents coefficients and standard errors for a probit regression of the following form: $Y_{ir} = 1(\lambda\Delta H_{ir} + \beta I_{ir} + \zeta_{ir} > 0)$ where Y_{ir} is equal to 1 if the respondent chose option 1 over option 2 in the ex-post experiment and where $\Delta H_{ir} = H_{ir1} - H_{ir2}$ and H_{irj} gives the characteristics of option $j = \{1, 2\}$ and $\zeta_{ir} \sim IN(0, 1)$. Included within H_{irj} is an indicator of whether the daughter described in that option is the same age and education (at marriage) as the respondents' daughter currently was. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.2: Importance of Vignette Salience in Ex-Ante Experiment

	(1)	(2)	(3)
Scenario Characteristics like Respondent's Daughter	0.0109	(0.0397)	0.0091 (0.0399)
Currently in School=0 × Daughter's Age=13	0.0000	(.)	0.0000 (.)
Currently in School=0 × Daughter's Age=14	0.0524	(0.1315)	0.0498 (0.1329)
Currently in School=0 × Daughter's Age=15	0.4044***	(0.1244)	0.4259*** (0.1260)
Currently in School=0 × Daughter's Age=16	0.4273***	(0.1248)	0.4363*** (0.1261)
Currently in School=0 × Daughter's Age=17	0.7782***	(0.1214)	0.8026*** (0.1226)
Currently in School=0 × Daughter's Age=18	1.0185***	(0.1217)	1.0404*** (0.1231)
Currently in School=0 × Daughter's Age=19	1.1429***	(0.1215)	1.1792*** (0.1222)
Currently in School=0 × Daughter's Age=20	1.2292***	(0.1245)	1.2513*** (0.1254)
Currently in School=0 × Daughter's Age=21	1.3034***	(0.1287)	1.3136*** (0.1300)
Currently in School=0 × Daughter's Age=22	1.3254***	(0.1275)	1.3611*** (0.1292)
Currently in School=1 × Daughter's Age=13	-0.5088***	(0.1544)	-0.4957*** (0.1558)
Currently in School=1 × Daughter's Age=14	-0.3958***	(0.1272)	-0.3792*** (0.1283)
Currently in School=1 × Daughter's Age=15	-0.1918	(0.1285)	-0.1862 (0.1304)
Currently in School=1 × Daughter's Age=16	-0.1772	(0.1288)	-0.1604 (0.1302)
Currently in School=1 × Daughter's Age=17	-0.0323	(0.1267)	-0.0251 (0.1279)
Currently in School=1 × Daughter's Age=18	0.2512*	(0.1304)	0.2728** (0.1319)
Currently in School=1 × Daughter's Age=19	0.1755	(0.1573)	0.1814 (0.1587)
Own Wealth = Poor	-0.0009	(0.0408)	-0.0076 (0.0412)
Own Wealth = Wealthy	-0.0640	(0.0408)	-0.0680* (0.0410)
Cost of School Covered	0.0406	(0.0503)	0.0431 (0.0504)
Daughter Likes School	-0.1986***	(0.0504)	-0.2043*** (0.0507)
Help in Home	0.0108	(0.0334)	0.0126 (0.0336)
Daughter has Male Friends	0.0995***	(0.0359)	0.0988*** (0.0362)
Groom's Age			-0.0174*** (0.0065)
Groom's Education			0.0208*** (0.0052)
Government Job			0.2773*** (0.0476)
Dowry			-0.0079 (0.0120)
Groom's Wealth = Poor			0.0053 (0.0439)
Groom's Wealth = Wealthy			-0.0147 (0.0437)
Wealth Disparity			-0.0538 (0.0432)
Groom Less Educated			-0.0573 (0.0653)
Constant	-0.2879**	(0.1126)	-0.0647 (0.2109)
Daughter Options	yes		yes
Match Characteristics	no		yes
Interactions	no		yes
Number of Choice Experiments	6836		6836

Notes: Table presents coefficients and standard errors for a probit regression of a binary indicator of whether the respondent choose the marriage option in the ex-ante experiment on characteristics of the hypothetical daughter, the hypothetical family, the hypothetical match and whether or not the option was based on the characteristics of the respondents' own daughter. All inference clustered at the respondent level. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Effect of RCT Treatment Status on Ex-Ante Experiment

	(1)	
Ex Post Choice		
Daughter's Age at Marriage=14	0.3285*	(0.1881)
Daughter's Age at Marriage=15	0.2964	(0.1833)
Daughter's Age at Marriage=16	0.4985***	(0.1833)
Daughter's Age at Marriage=17	0.5730***	(0.1883)
Daughter's Age at Marriage=18	0.9805***	(0.1892)
Daughter's Age at Marriage=19	0.9313***	(0.1866)
Daughter's Age at Marriage=20	0.9532***	(0.1971)
Daughter's Age at Marriage=21	1.1362***	(0.2012)
Daughter's Age at Marriage=22	1.0606***	(0.2090)
Daughter's Education =8	0.0092	(0.0724)
Daughter's Education =9	0.0716	(0.0789)
Daughter's Education =10	0.1197	(0.0890)
Daughter's Education =11	0.1572	(0.0990)
Daughter's Education =12	0.1398	(0.1117)
Daughter's Education =13	-0.0649	(0.1186)
Daughter Likes School*Years in School	0.0057	(0.0168)
Cost of School Covered*Years in School	0.0006	(0.0164)
Help in Home*Years at Home	-0.0104	(0.0176)
Government Job	0.4533***	(0.0516)
Girl Only × Daughter's Age at Marriage=14	-0.1436	(0.2587)
Integrated × Daughter's Age at Marriage=14	-0.0690	(0.2725)
Girl Only × Daughter's Age at Marriage=15	0.0523	(0.2559)
Integrated × Daughter's Age at Marriage=15	0.0856	(0.2678)
Girl Only × Daughter's Age at Marriage=16	-0.0349	(0.2563)
Integrated × Daughter's Age at Marriage=16	0.0464	(0.2682)
Girl Only × Daughter's Age at Marriage=17	0.0770	(0.2635)
Integrated × Daughter's Age at Marriage=17	0.1482	(0.2736)
Girl Only × Daughter's Age at Marriage=18	-0.1479	(0.2629)
Integrated × Daughter's Age at Marriage=18	-0.1396	(0.2793)
Girl Only × Daughter's Age at Marriage=19	0.0087	(0.2630)
Integrated × Daughter's Age at Marriage=19	-0.1517	(0.2723)
Girl Only × Daughter's Age at Marriage=20	0.1741	(0.2774)
Integrated × Daughter's Age at Marriage=20	-0.1475	(0.2881)
Girl Only × Daughter's Age at Marriage=21	-0.1691	(0.2809)
Integrated × Daughter's Age at Marriage=21	-0.3384	(0.2903)
Girl Only × Daughter's Age at Marriage=22	-0.0538	(0.2956)
Integrated × Daughter's Age at Marriage=22	-0.1734	(0.3022)
Girl Only × Daughter's Education =8	-0.0781	(0.1067)
Integrated × Daughter's Education =8	0.0552	(0.1021)
Girl Only × Daughter's Education =9	-0.1943*	(0.1175)
Integrated × Daughter's Education =9	-0.0147	(0.1112)
Girl Only × Daughter's Education =10	-0.2137	(0.1313)
Integrated × Daughter's Education =10	-0.0338	(0.1257)
Girl Only × Daughter's Education =11	-0.2835*	(0.1457)
Integrated × Daughter's Education =11	-0.0459	(0.1409)
Girl Only × Daughter's Education =12	-0.0536	(0.1631)
Integrated × Daughter's Education =12	-0.0200	(0.1568)
Girl Only × Daughter's Education =13	-0.2280	(0.1732)
Integrated × Daughter's Education =13	0.0869	(0.1677)
Girl Only × Daughter Likes School*Years in School	0.0318	(0.0242)
Integrated × Daughter Likes School*Years in School	0.0198	(0.0237)
Girl Only × Cost of School Covered*Years in School	0.0211	(0.0236)
Integrated × Cost of School Covered*Years in School	0.0186	(0.0232)
Girl Only × Help in Home*Years at Home	-0.0347	(0.0254)
Integrated × Help in Home*Years at Home	-0.0083	(0.0251)
Girl Only × Government Job	-0.1268*	(0.0737)
Integrated × Government Job	-0.1734**	(0.0731)
Daughter Options	yes	
Match Characteristics	yes	
Interactions	no	
Number of Choice Experiments	6320	

Notes: Table presents coefficients and standard errors for a probit regression of the following form: $Y_{ir} = 1(\lambda T_i \Delta H_{ir} + \beta I_{ir} + \zeta_{ir} > 0)$ where Y_{ir} is equal to 1 if the respondent chose option 1 over option 2 in the ex-post experiment and where $\Delta H_{ir} = H_{ir1} - H_{ir2}$ and H_{irj} gives the characteristics of option $j = \{1, 2\}$ and $\zeta_{ir} \sim IN(0, 1)$. T_i is a vector of treatment indicators. RCT treatment arms include: control, girl only (life skills and knowledge for adolescent girls) and integrated (with additional community component). Details of intervention given in (Achyut, Andrew, Das, Gautam, Huepe, Krutikova, Kumar, Sharma, Soni, Verma, and Verma 2016). Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.4: Effect of RCT Treatment Status on Ex-Ante Experiment

	(1)		(2)		(3)	
Control	0.0000	(.)	0.0000	(.)	0.0000	(.)
Girl Only	-0.0092	(0.0482)	-0.0075	(0.0479)	-0.0083	(0.0483)
Integrated	0.0976**	(0.0469)	0.0969**	(0.0465)	0.0983**	(0.0469)
Currently in School=0 × Daughter's Age=13	0.0000	(.)	0.0000	(.)	0.0000	(.)
Currently in School=0 × Daughter's Age=14	0.0472	(0.1330)	0.0490	(0.1318)	0.0457	(0.1329)
Currently in School=0 × Daughter's Age=15	0.4188***	(0.1260)	0.3969***	(0.1246)	0.4174***	(0.1259)
Currently in School=0 × Daughter's Age=16	0.4331***	(0.1261)	0.4244***	(0.1248)	0.4323***	(0.1260)
Currently in School=0 × Daughter's Age=17	0.7993***	(0.1224)	0.7751***	(0.1213)	0.7983***	(0.1223)
Currently in School=0 × Daughter's Age=18	1.0399***	(0.1225)	1.0177***	(0.1211)	1.0386***	(0.1224)
Currently in School=0 × Daughter's Age=19	1.1717***	(0.1223)	1.1368***	(0.1216)	1.1723***	(0.1222)
Currently in School=0 × Daughter's Age=20	1.2406***	(0.1257)	1.2213***	(0.1249)	1.2427***	(0.1256)
Currently in School=0 × Daughter's Age=21	1.3085***	(0.1301)	1.2995***	(0.1288)	1.3094***	(0.1300)
Currently in School=0 × Daughter's Age=22	1.3542***	(0.1293)	1.3207***	(0.1277)	1.3555***	(0.1292)
Currently in School=1 × Daughter's Age=13	-0.5498***	(0.1629)	-0.5237***	(0.1544)	-0.5123***	(0.1556)
Currently in School=1 × Daughter's Age=14	-0.4114***	(0.1300)	-0.3897***	(0.1222)	-0.3752***	(0.1232)
Currently in School=1 × Daughter's Age=15	-0.2199*	(0.1323)	-0.1915	(0.1255)	-0.1878	(0.1272)
Currently in School=1 × Daughter's Age=16	-0.1851	(0.1301)	-0.1767	(0.1254)	-0.1618	(0.1268)
Currently in School=1 × Daughter's Age=17	-0.0482	(0.1267)	-0.0349	(0.1241)	-0.0294	(0.1252)
Currently in School=1 × Daughter's Age=18	0.2610**	(0.1293)	0.2458*	(0.1277)	0.2653**	(0.1291)
Currently in School=1 × Daughter's Age=19	0.1710	(0.1520)	0.1742	(0.1507)	0.1780	(0.1519)
Cost of School Covered	0.0431	(0.0505)	0.0403	(0.0504)	0.0429	(0.0505)
Daughter Likes School	-0.2045***	(0.0506)	-0.1992***	(0.0504)	-0.2047***	(0.0507)
Help in Home	0.0114	(0.0336)	0.0094	(0.0334)	0.0112	(0.0336)
Daughter has Male Friends	0.1011***	(0.0361)	0.1005***	(0.0359)	0.0997***	(0.0362)
Groom's Age	-0.0176**	(0.0065)			-0.0177***	(0.0065)
Groom's Education	0.0190***	(0.0057)			0.0209***	(0.0052)
Government Job	0.2730***	(0.0480)			0.2770***	(0.0476)
Dowry	-0.0105	(0.0093)			-0.0081	(0.0120)
Wealth Disparity	-0.0629	(0.0400)				
Groom Less Educated	-0.0527	(0.0651)				
Groom's Wealth = Poor					0.0051	(0.0438)
Groom's Wealth = Wealthy					-0.0114	(0.0437)
Constant	-0.0155	(0.2264)	-0.3306***	(0.1133)	-0.1034	(0.2108)
Daughter Options	yes		yes		yes	
Match Characteristics						
Interactions						
Number of Choice Experiments	6836		6836		6836	

Notes: Table presents coefficients and standard errors for a probit regression of a binary indicator of whether the respondent choose the marriage option in the ex-ante experiment on characteristics of the hypothetical daughter, the hypothetical family, the hypothetical match and on indicators of RCT treatment status. RCT treatment arms include: control, girl only (life skills and knowledge for adolescent girls) and integrated (with additional community component). Details of intervention given in (Achyut, Andrew, Das, Gautam, Huepe, Krutikova, Kumar, Sharma, Soni, Verma, and Verma 2016) Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Determinants of Groom Choices: Reduced Form Probit

	(1)	(2)	(3)
Bride's Age at Marriage=14	0.1021 (0.0765)	0.1173 (0.0769)	0.1429* (0.0866)
Bride's Age at Marriage=15	0.2030*** (0.0710)	0.2191*** (0.0715)	0.2410*** (0.0813)
Bride's Age at Marriage=16	0.2563*** (0.0717)	0.2755*** (0.0722)	0.3329*** (0.0815)
Bride's Age at Marriage=17	0.2581*** (0.0716)	0.2678*** (0.0721)	0.3117*** (0.0810)
Bride's Age at Marriage=18	0.5207*** (0.0700)	0.5345*** (0.0704)	0.5976*** (0.0793)
Bride's Age at Marriage=19	0.5871*** (0.0750)	0.6062*** (0.0753)	0.6620*** (0.0851)
Bride's Age at Marriage=20	0.5850*** (0.0726)	0.6042*** (0.0731)	0.6790*** (0.0823)
Bride's Age at Marriage=21	0.5789*** (0.0697)	0.5933*** (0.0702)	0.6685*** (0.0793)
Bride's Age at Marriage=22	0.6506*** (0.0754)	0.6625*** (0.0759)	0.7612*** (0.0863)
Bride's Education =8	0.0854* (0.0491)	0.0917* (0.0499)	0.1098* (0.0572)
Bride's Education =9	0.1023** (0.0501)	0.1165** (0.0521)	0.1028* (0.0591)
Bride's Education =10	0.2611*** (0.0579)	0.2829*** (0.0607)	0.2435*** (0.0702)
Bride's Education =11	0.2456*** (0.0556)	0.2715*** (0.0596)	0.2261*** (0.0686)
Bride's Education =12	0.2876*** (0.0622)	0.3168*** (0.0678)	0.3009*** (0.0776)
Bride's Education =13	0.3117*** (0.0514)	0.3532*** (0.0621)	0.2731*** (0.0701)
Bride has Male Friends	-0.3283*** (0.0306)	-0.3248*** (0.0307)	-0.3343*** (0.0344)
Dowry	0.0598*** (0.0103)	0.0591*** (0.0103)	0.0552*** (0.0117)
Bride's Wealth = Poor	0.0071 (0.0374)		
Bride's Wealth = Wealthy	-0.0192 (0.0365)		
Bride's Wealth = Poor, Groom's Wealth = Poor		-0.0361 (0.0634)	0.0177 (0.0716)
Bride's Wealth = Wealthy, Groom's Wealth = Poor		-0.2571*** (0.0602)	-0.2509*** (0.0682)
Bride's Wealth = Poor, Groom's Wealth = Average		-0.0209 (0.0600)	-0.0138 (0.0663)
Bride's Wealth = Wealthy, Groom's Wealth = Average		0.0101 (0.0609)	0.0342 (0.0671)
Bride's Wealth = Poor, Groom's Wealth = Wealthy		0.0875 (0.0601)	0.0531 (0.0682)
Bride's Wealth = Wealthy, Groom's Wealth = Wealthy		0.2262*** (0.0624)	0.2246*** (0.0713)
Bride More Educated		-0.0483 (0.0428)	-0.0229 (0.0477)
Gov = 1 & Bride's Age = 14			-0.1068 (0.1914)
Gov = 1 & Bride's Age = 15			-0.0608 (0.1738)
Gov = 1 & Bride's Age = 16			-0.2081 (0.1798)
Gov = 1 & Bride's Age = 17			-0.1923 (0.1810)
Gov = 1 & Bride's Age = 18			-0.2279 (0.1763)
Gov = 1 & Bride's Age = 19			-0.2037 (0.1865)
Gov = 1 & Bride's Age = 20			-0.3089* (0.1841)
Gov = 1 & Bride's Age = 21			-0.3137* (0.1745)
Gov = 1 & Bride's Age = 22			-0.4170** (0.1852)
Gov = 1 & Bride's Ed = 8			-0.1095 (0.1192)
Gov = 1 & Bride's Ed = 9			0.0105 (0.1316)
Gov = 1 & Bride's Ed = 10			0.1224 (0.1463)
Gov = 1 & Bride's Ed = 11			0.1619 (0.1471)
Gov = 1 & Bride's Ed = 12			0.0459 (0.1642)
Gov = 1 & Bride's Ed = 13			0.4192*** (0.1622)
Gov = 1 & Bride has Male Friends			0.0486 (0.0773)
Gov = 1 & Dowry			0.0195 (0.0252)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Poor			-0.2537 (0.1584)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Poor			-0.0231 (0.1473)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Average			-0.0616 (0.1585)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Average			-0.0917 (0.1625)
Gov = 1 & Bride's Wealth = Poor, Groom's Wealth = Wealthy			0.1683 (0.1465)
Gov = 1 & Bride's Wealth = Wealthy, Groom's Wealth = Wealthy			0.0181 (0.1493)
Gov = 1 & Bride More Educated			-0.2075 (0.1351)
Daughter Options	yes	yes	yes
Match Characteristics	no	yes	yes
Interactions	no	no	yes
Number of Choice Experiments	4596	4596	4596

Notes: Table presents coefficients and standard errors for a probit regression of the following form: $Y_{ir} = 1(\lambda\Delta H_{ir} + \zeta_{ir} > 0)$ where Y_{ir} is equal to 1 if the respondent chose option 1 over option 2 in the groom's side experiment and where $\Delta H_{ir} = H_{ir1} - H_{ir2}$ and H_{irj} gives the characteristics of option $j = \{1, 2\}$ and $\zeta_{ir} \sim IN(0, 1)$. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Determinants of Expected Match

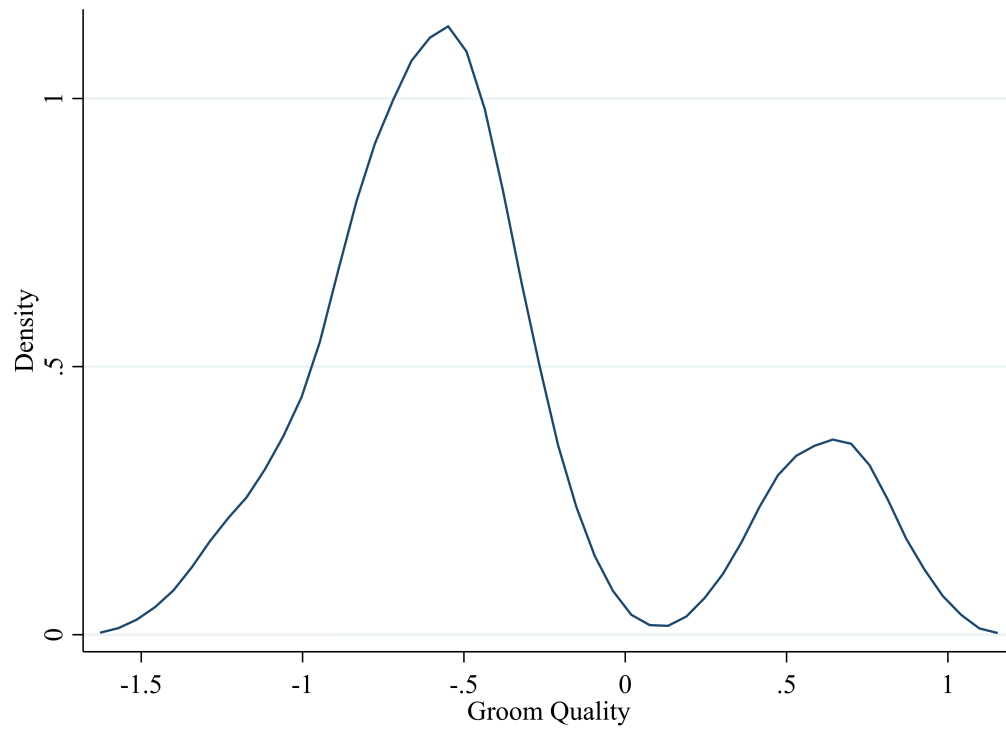
	(1) Education	(2) Government Job	(3) Wealth	(4) Dowry	(5) Dowry	(6) Quality
Age=13	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)
Age=14	0.3276 (0.3188)	0.0586 (0.0724)	0.0379 (0.0728)	-0.3232 (0.2188)	-0.3900* (0.2013)	0.0274 (0.0221)
Age=15	0.1562 (0.2912)	-0.0435 (0.0622)	0.0302 (0.0663)	-0.1468 (0.1939)	-0.1670 (0.1788)	-0.0075 (0.0190)
Age=16	0.3952 (0.2838)	-0.0134 (0.0611)	0.0053 (0.0655)	-0.1423 (0.1906)	-0.1693 (0.1754)	0.0059 (0.0187)
Age=17	0.3439 (0.2838)	0.0469 (0.0594)	-0.0001 (0.0640)	-0.0449 (0.1851)	-0.0931 (0.1700)	0.0183 (0.0182)
Age=18	0.3358 (0.2822)	0.0575 (0.0583)	0.0156 (0.0627)	-0.0190 (0.1830)	-0.1010 (0.1678)	0.0168 (0.0179)
Age=19	0.2987 (0.2827)	0.0034 (0.0586)	0.0074 (0.0629)	0.0717 (0.1831)	-0.0082 (0.1682)	-0.0043 (0.0180)
Age=20	0.4383 (0.2779)	0.0636 (0.0579)	0.0324 (0.0624)	-0.0165 (0.1827)	-0.1755 (0.1693)	0.0090 (0.0178)
Age=21	0.4442 (0.2801)	0.0626 (0.0578)	0.0508 (0.0627)	-0.0641 (0.1824)	-0.2701 (0.1727)	0.0029 (0.0178)
Age=22	0.4067 (0.2792)	0.0641 (0.0579)	0.0253 (0.0627)	0.0326 (0.1815)	-0.1837 (0.1726)	-0.0042 (0.0178)
Sch=0	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)
Sch=5	1.1546*** (0.2305)	0.0696* (0.0409)	0.0317 (0.0419)	-0.0099 (0.1262)	-0.1286 (0.1242)	0.0417*** (0.0135)
Sch=7	1.7528*** (0.2235)	0.2080*** (0.0408)	0.0799* (0.0436)	0.2600** (0.1262)	0.0295 (0.1263)	0.0870*** (0.0132)
Sch=8	1.9991*** (0.2174)	0.2235*** (0.0350)	0.0582 (0.0367)	0.1408 (0.1055)	-0.1032 (0.1085)	0.0972*** (0.0117)
Sch=9	2.4154*** (0.2164)	0.3220*** (0.0341)	0.1136*** (0.0362)	0.3333*** (0.1060)	-0.0011 (0.1117)	0.1296*** (0.0114)
Sch=10	2.7225*** (0.2167)	0.3823*** (0.0343)	0.1112*** (0.0376)	0.4146*** (0.1095)	0.0390 (0.1178)	0.1501*** (0.0115)
Sch=11	3.0420*** (0.2160)	0.4320*** (0.0347)	0.1183*** (0.0395)	0.5162*** (0.1144)	0.0918 (0.1239)	0.1679*** (0.0116)
Sch=12	3.2783*** (0.2150)	0.4730*** (0.0330)	0.1435*** (0.0374)	0.5835*** (0.1087)	0.1117 (0.1202)	0.1817*** (0.0112)
Sch=13	3.3834*** (0.2156)	0.5287*** (0.0342)	0.1797*** (0.0423)	0.7177*** (0.1278)	0.2034 (0.1382)	0.2005*** (0.0115)
Wealth of Bride's family=1	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)	0.0000 (.)
Wealth of Bride's family=2	0.0148 (0.0488)	0.0716*** (0.0162)	0.2441*** (0.0185)	1.1221*** (0.0515)	0.9465*** (0.0514)	0.0011 (0.0048)
Wealth of Bride's family=3	0.1573*** (0.0478)	0.1571*** (0.0155)	0.6446*** (0.0194)	2.5414*** (0.0570)	2.0778*** (0.0645)	0.0028 (0.0047)
Daughter Has Male Friends	0.0113 (0.0416)	0.0119 (0.0138)	0.0095 (0.0162)	-0.0381 (0.0504)	-0.0532 (0.0488)	0.0032 (0.0041)
Expect Match with Gov Job					0.2849*** (0.0504)	
Expected Age of Match					0.0491*** (0.0147)	
Expected Education of Match					0.0723*** (0.0172)	
Expected Wealth of Match					0.6260*** (0.0474)	
Constant	8.9729*** (0.3287)	0.2369*** (0.0579)	1.8411*** (0.0638)	1.3460*** (0.1784)	-1.5747*** (0.3817)	0.0074 (0.0183)
Observations	4599	4599	4599	4599	4599	4599

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Table presents coefficients and standard errors for OLS regressions of characteristics of the 'expected match' given by respondent on characteristics of hypothetical daughter. Standard errors in parentheses. Significance of coefficients indicated by: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Figure A.1: Kernel density plot of groom quality



Notes: Figure plots kernel density of estimated groom quality index, $\mathbf{X}_{irj\alpha}$, over the potential grooms that we presented to respondents in the ex-post experiments.

B Visual Aids for Validation Experiments

Figure A.2: Validation Experiments: Expected Match Visual Aid

Parents: _____ Girl: _____

Very poor	Quite poor	Average	Quite wealthy	Very wealthy

	12	13	14	15	16	17	18	19	20	21	22	23
None	1 st	5 th	7 th	8 th	9 th	10 th	11 th	12 th	College			

Well behaved and polite

Is friends with a few boys

If _____ got married this year she probably marry someone like this...

21	22	23	24	25	26	27	28	29	30
----	----	----	----	----	----	----	----	----	----

None	1 st	5 th	7 th	8 th	9 th	10 th	11 th	12 th	College
------	-----------------	-----------------	-----------------	-----------------	-----------------	------------------	------------------	------------------	---------

Government Job

No Government Job

Very poor

Average

Very wealthy

0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5

Figure A.3: Validation Experiments: Groom Experiment Visual Aid

3 Parents: _____ Son: _____

Very poor	Quite poor	Average	Quite wealthy	Very wealthy

	Age									Government Job
	21	22	23	24	25	26	27	28	29	
	Education									
None	1 st	5 th	7 th	8 th	9 th	10 th	11 th	12 th	College	

Marriage prospect 1:

	12	13	14	15	16	17	18	19	20	21	22	23
None	1 st	5 th	7 th	8 th	9 th	10 th	11 th	12 th	College			

Very poor	Quite poor	Average	Quite wealthy	Very wealthy

0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5

Well behaved and polite

Is friends with a few boys

1) Marriage prospect 2:

	12	13	14	15	16	17	18	19	20	21	22	23
None	1 st	5 th	7 th	8 th	9 th	10 th	11 th	12 th	College			

Very poor	Quite poor	Average	Quite wealthy	Very wealthy

0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5

Well behaved and polite

Is friends with a few boys

C Identification of Preferences

THEOREM 1. Identification of Preferences: $\bar{U}_0(\bar{\Psi})$ is constructively identified at all $\bar{\Psi} \in \bar{\Psi}^M$ from ex-post choice probabilities under the following assumptions:

1. Ex-post choice probabilities, $P(\bar{\Psi}, \bar{\Psi}')$, are observed for all $(\bar{\Psi}, \bar{\Psi}') \in \bar{\Psi}^M \times \bar{\Psi}^M$.
2. Location and scale normalizations: $\bar{U}(\bar{\Psi}^0) = 0$ and $U(\bar{\Psi}^1) - \bar{U}(\bar{\Psi}^0) = 1$
3. The aggregated preference shock is distributed i.i.d normal, $\nu_{irj} \sim IN(0, \sigma_\nu^2)$, and is independent of $\bar{\Psi}_{rj}$ for all $\bar{\Psi}_{rj} \in \bar{\Psi}^M$.

Proof From the point of view of period 0, parents have preferences over complete paths, $\bar{\Psi}$, that a daughter's adolescence might take. The non-stochastic component of these preferences are described by utility function $\bar{U}_0(\bar{\Psi})$. Let \mathbf{P} give the $M(M-1)/2$ -vector of choice probabilities (that the first option is chosen) at the possible combinations of paths presented in the ex-post experiment. Inverting the expression for ex-post choice probabilities (Equation 5.1) and leveraging our assumption on the error term gives:

$$\sqrt{2}\Phi^{-1}(\mathbf{P}) = BU_0 \tag{C.1}$$

where

$$\Phi^{-1}(\mathbf{P}) = \begin{bmatrix} \Phi^{-1}(P(\bar{\Psi}^0, \bar{\Psi}^1)) \\ \Phi^{-1}(P(\bar{\Psi}^0, \bar{\Psi}^2)) \\ \vdots \\ \Phi^{-1}(P(\bar{\Psi}^{M-1}, \bar{\Psi}^M)) \end{bmatrix}, \quad \mathbf{U}_0 = \begin{bmatrix} 1/\sigma_\nu \\ \bar{U}_0(\bar{\Psi}^2)/\sigma_\nu \\ \vdots \\ \bar{U}_0(\bar{\Psi}^M)/\sigma_\nu \end{bmatrix}$$

and

$$B = \begin{bmatrix} -1 & 0 & \cdots & 0 \\ 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & -1 \end{bmatrix}$$

That is, $\Phi^{-1}(\mathbf{P})$ is an $M(M - 1)/2$ -vector of the quantiles of the normal distribution at the observed choice probabilities, U_0 is the $M - 1$ vector of utility values (relative to the path $\bar{\Psi}^0$ and scaled by the standard deviation of the experimental error), and B is the design matrix that gives every unique combination of utility differences. Pre-multiplying both sides of Equation C.1 by B' and then inverting we get:

$$U_0 = \sqrt{2}(B'B)^{-1}B'\Phi^{-1}(\mathbf{P}) \tag{C.2}$$

Thus, utility relative to path $\bar{\Psi}^0$ can be identified from the observed ex-post choice probabilities.

D Identification of Beliefs

We structure our discussion of the identification of beliefs in three parts. First, we consider the case when marriage offers are received every period (i.e. $\pi(\bar{\omega}, N) = 0$ for all $\bar{\omega}$). Proving identification of subjective beliefs in this case is straightforward and will often be the most relevant in alternative settings where this methodology is applied.

Second, we analyse the case where we allow for no marriage offers to be received, i.e. $\pi(\bar{\omega}, N) \geq 0$ and utility is identified at all $\bar{\Psi} \in \bar{\Psi}$. Finally, we analyse the case where we allow for no marriage offers to be received, i.e. $\pi(\bar{\omega}, N) \geq 0$ and utility is only identified for paths that end in marriage, i.e. $\Psi \in \Psi^M$. This final case will not be generally relevant for researchers hoping to apply our methodology and is only included because we did not elicit choice behaviour over options in which a daughter remained unmarried at 23 in the ex-post experiment.

We first show how beliefs are identified with with two marriage qualities. We next show how the presence of instruments which affect flow payoffs but not beliefs permit the identification of more types. Finally, we discuss the

special case where we allow for the case that an offer may not be received at all.

D.0.1 Setup and notation

As in the main text we define the non-stochastic component of utility simply as a function of the state at marriage for $\Psi \in \Psi^M$:

$$u_t(A, q_A, \bar{\omega}_A) \equiv \bar{U}_t(\bar{\Psi}_t) \quad (\text{D.1})$$

Value functions, $V(\cdot)$, expected value functions $EV(\cdot)$, and choice-specific value functions, $\tilde{V}(\cdot)$, are also defined as in the main text:

$$V(t, \bar{\omega}_t, q_t, \varepsilon_t) = \max_{\delta} \mathbb{E} \left[\bar{U}_0(\bar{\Psi}) + \sum_{\tau=0}^T \beta^\tau \varepsilon_\tau^{d_\tau} | t, \bar{\omega}_t, q_t, \varepsilon_t \right] \quad (\text{D.2})$$

$$EV(t, \bar{\omega}_t) = \mathbb{E}_{q_t, \varepsilon_t} [V(t, \bar{\omega}_t, q_t, \varepsilon_t)] \quad (\text{D.3})$$

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^M, d_t = M) = u(t, \bar{\omega}_t, q_t) + \beta^t \varepsilon_t^M \quad (\text{D.4})$$

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^H, d_t = H) = EV(t, \bar{\omega}_{t+1}(\bar{\omega}_t, d_t = H)) + \beta^t \varepsilon_t^H \quad (\text{D.5})$$

$$\tilde{V}(t, \bar{\omega}_t, q_t, \varepsilon_t^S, d_t = S) = EV(t, \bar{\omega}_{t+1}(\bar{\omega}_t, d_t = S)) + \beta^t \varepsilon_t^S \quad (\text{D.6})$$

For ease of notation, we drop dependence on deterministic state variables other than education.

D.1 Case 1: $\pi(\bar{\omega}, N) = 0$ and $K = 2$

Step 1: Identifying Expected Value Functions and Distribution of Preference Shocks Consider any generic period t in which parents are making a two-way choice between keeping a daughter at home another year or marrying her (i.e. where she is already out of school). Under the assumption $\varepsilon_t^d \sim N(0, \sigma^2)$, acceptance probabilities are described by:

$$\begin{aligned}
p_M(t, Ed, outschool, q = 0) &= Pr(u(t, Ed, q = 0) + \beta^t \varepsilon_t^M > EV(t + 1, Ed, outschool) + \beta^t \varepsilon_t^H) \\
&= \Phi \left(\frac{u(t, Ed, q = 0) - V(t + 1, Ed, outschool)}{\beta^t \sqrt{2}\sigma} \right)
\end{aligned}$$

$$\begin{aligned}
p_M(t, Ed, outschool, q = 1) &= Pr(u(t, Ed, q = 1) + \beta^t \varepsilon_t^M > EV(t + 1, Ed, outschool) + \beta^t \varepsilon_t^H) \\
&= \Phi \left(\frac{u(t, Ed, q = 0) - V(t + 1, Ed, outschool)}{\beta^t \sqrt{2}\sigma} \right)
\end{aligned}$$

Inverting identifies σ and $EV(t, Ed, outschool)$, $\forall t = 2, \dots, T + 1$, for all states where the daughter has been out of school for at least one period already.

For daughters still in school, marriage probabilities take the form:

$$p_M(t, Ed, inschool, q) = Pr \left[\begin{array}{l} u_0(t, Ed, q) + \beta^t \varepsilon_M > EV(t + 1, Ed, outschool) + \beta^t \varepsilon_H, \\ u_0(t, Ed, q) + \beta^t \varepsilon_M > EV(t + 1, Ed + 1, inschool) + \beta^t \varepsilon_S \end{array} \right] \quad (D.7)$$

$$p_S(t, Ed, inschool, q) = Pr \left[\begin{array}{l} EV(t + 1, Ed + 1, inschool) + \beta^t \varepsilon_S > u_0(t, Ed, q) + \beta^t \varepsilon_M, \\ EV(t + 1, Ed + 1, inschool) + \beta^t \varepsilon_S > EV(t + 1, Ed, outschool) + \beta^t \varepsilon_H \end{array} \right] \quad (D.8)$$

Although closed form solutions do not exist, these expressions allow us to identify $EV(t, Ed, \cdot)$ for all states when a daughter is still in school or states one period after a daughter has left school.

Having identified all value functions and the variance of the idiosyncratic preference heterogeneity, we can form expressions for the expected value of idiosyncratic preference shocks conditional on optimal actions ($E(\varepsilon_t^d | d^*, \cdot)_i$). In two-way choices, these expressions have closed form solutions, for example:

$$E(\varepsilon_t^M | d_t^* = M, t, Ed, outschool, q) = E(\varepsilon_t^M | u_0(t, Ed, q) + \beta^t \varepsilon_t^M \geq EV(t + 1, Ed, outschool) + \beta^t \varepsilon_t^H) \quad (D.9)$$

$$= E \left(\varepsilon_t^M | \varepsilon_t^M - \varepsilon_t^H \geq \frac{EV(t + 1, Ed, outschool) - u_0(t, Ed, q)}{\beta^t} \right) \quad (D.10)$$

$$= \frac{1}{\sqrt{2}} \sigma_\varepsilon \frac{\phi \left(\frac{EV(t+1, Ed, outschool) - u_0(t, Ed, q)}{\beta^{T-1} \sqrt{2}\sigma_\varepsilon^2} \right)}{1 - \Phi \left(\frac{EV(t+1, Ed, outschool) - u_0(t, Ed, q)}{\beta^{T-1} \sqrt{2}\sigma_\varepsilon^2} \right)} \quad (D.11)$$

where we use the fact that if ε^M and ε^H are independent and follow identical normal distributions. These are similarly defined in the three-way choices although no closed forms exist.

Step 2: Going from Value Functions to Beliefs After having identified the value functions for the value of being unmarried at every state, we can use the recursive relationship between expected value functions to identify beliefs. In particular, the value of being unmarried at period t is equal to:

$$\begin{aligned}
& EV(t, Ed, outschool) \\
&= \underbrace{(1 - \pi(t, Ed))p_M(t, Ed, q = 0) [u(t, Ed, q = 0) + E(\varepsilon_t^M | d_t^* = M, t, Ed, outschool, q = 0)]}_{\text{low offer, accept}} \\
&+ \underbrace{\pi(t, Ed)p_M(t, Ed, q = 1) [u(t, Ed, q = 1) + E(\varepsilon_t^M | d_t^* = M, t, Ed, outschool, q = 1)]}_{\text{high offer, accept}} \\
&+ \underbrace{(1 - \pi(t, Ed))(1 - p_M(t, Ed, q = 0)) [EV(t + 1, Ed, outschool) + E(\varepsilon_t^H | d_t^* = H, t, Ed, outschool, q = 0)]}_{\text{low offer, reject}} \\
&+ \underbrace{\pi(t, Ed)(1 - p_M(t, Ed, q = 1)) [EV(t + 1, Ed, outschool) + E(\varepsilon_t^H | d_t^* = H, t, Ed, outschool, q = 1)]}_{\text{high offer, reject}}
\end{aligned}$$

Since, $\pi(t, Ed)$ is the only unknown in this relation, we can rearrange to identify $\pi(t, Ed)$ at every state where a daughter is out of school. .

Likewise, in states where a three-way choice exists, we have:

$EV(t, Ed, inschool)$

$$\begin{aligned}
&= \underbrace{(1 - \pi(t, Ed))p_M(t, Ed, q = 0) [u(t, Ed, q = 0) + E(\varepsilon_t^M | d^* = M, t, Ed, inschool, q = 0)]}_{\text{low offer, accept}} \\
&+ \underbrace{\pi(t, Ed)p_M(t, Ed, q = 1) [u(t, Ed, q = 1) + E(\varepsilon_t^M | d^* = M, t, Ed, inschool, q = 1)]}_{\text{high offer, accept}} \\
&+ \underbrace{(1 - \pi(t, Ed))p_S(t, Ed, q = 0) [EV(t + 1, Ed + 1, inschool) + E(\varepsilon_t^H | d^* = S, t, Ed, inschool, q = 0)]}_{\text{low offer, school}} \\
&+ \underbrace{\pi(t, Ed)p_S(t, Ed, q = 1) [EV(t + 1, Ed + 1, inschool) + E(\varepsilon_t^H | d^* = S, t, Ed, inschool, q = 1)]}_{\text{high offer, school}} \\
&+ \underbrace{(1 - \pi(t, Ed))p_H(t, Ed, q = 0) [EV(t + 1, Ed, outschool) + E(\varepsilon_t^H | d^* = H, t, Ed, inschool, q = 0)]}_{\text{low offer, home}} \\
&+ \underbrace{\pi(t, Ed)p_H(t, Ed, q = 0) [EV(t + 1, Ed, outschool) + E(\varepsilon_t^H | d^* = H, t, Ed, inschool, q = 1)]}_{\text{low offer, home}} \quad (D.12)
\end{aligned}$$

where $p_H(t, Ed, q = 0) = 1 - p_M(t, Ed, q = 0) - p_S(t, Ed, q = 0)$. We can rearrange to find the remaining $\pi(t, Ed)$ in states when daughters are still in school.

D.2 Case 2: $\pi(\bar{\omega}_t, N) > 0$, $K = 2$

We next show how the inclusion of instruments which affect preferences but not beliefs can identify beliefs in models where marriage offers may not be received every period.

Let z be the instrument vector which affects the utility associated with different paths taken through the model: $u(t, Ed, q, z') \neq u(t, Ed, q, z'')$, $\forall z' \neq z''$. We require that these instruments are included in both the ex-post and ex-ante experiments. In this case, the ex-post experiment can identify $u(t, Ed, q, z), \forall z \in Z$ while in the ex-ante experiment we will observe: $p_d(t, Ed, q, z), \forall z \in Z$. Following the same process as in Step 1 above, we can thus combine these objects to identify expected value functions conditional on each level of the instrument, $EV(t, Ed, ., z)$, and the expected value of the preference shocks conditional on optimal actions, $E(\varepsilon_t^d | d^*, t, Ed, z, q, z)$.

Take the model with two groom types but where an offer is not guaranteed, i.e. where $\pi_t^0(\cdot) + \pi_t^1(\cdot) \leq 1$. Dropping dependence on schooling for ease of notation, we have:

$$\begin{aligned}
EV(t, z) = & \underbrace{\pi_t^0 p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)]}_{\text{low offer, accept}} \\
& + \underbrace{\pi_t^1 p_M(t, q = 1, z) [u(t, q = 1, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 1, z)]}_{\text{high offer, accept}} \\
& + \underbrace{\pi_t^0 (1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)]}_{\text{low offer, reject}} \\
& + \underbrace{\pi_t^1 (1 - p_M(t, q = 1, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 1, z)]}_{\text{high offer, reject}} \\
& + \underbrace{(1 - \pi_t^0 - \pi_t^1) [EV(t + 1, z)]}_{\text{no offer}}
\end{aligned} \tag{D.13}$$

Stacking over two distinct $z' \neq z''$, we get:

$$\begin{bmatrix} EV(t, z') - EV(t + 1, z') \\ EV(t, z'') - EV(t + 1, z'') \end{bmatrix} = \begin{bmatrix} a(z') & b(z') \\ a(z'') & b(z'') \end{bmatrix} \begin{bmatrix} \pi_t^0 \\ \pi_t^1 \end{bmatrix}$$

where:

$$\begin{aligned}
a(z) & \equiv p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)] \\
& + (1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)] \\
& - EV(t + 1, z) \\
b(z) & \equiv p_M(t, q = 1, z) [u(t, q = 1, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 1, z)] \\
& + (1 - p_M(t, q = 1, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 1, z)] \\
& - EV(t + 1, z)
\end{aligned}$$

Thus, we can invert and recover π_t^q whenever the instrument shifts utility profiles.

D.3 Case 3: $K > 2$

Very similarly to in case 2, when there are more than two groom types the inclusion of preference-shifting instruments is needed to identify the model. As in case 2, we can identify utility functions and expected value functions

conditional on each value of the instrument z . Taking the example as three groom types as an illustration, then, the recursive relationship between expected value functions becomes:

$$\begin{aligned}
EV(t, z) = & \underbrace{(1 - \pi_t^1 - \pi_t^2)p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)]}_{\text{low offer, accept}} \\
& + \underbrace{\pi_t^1 p_M(t, q = 1, z) [u(t, q = 1, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 1, z)]}_{\text{medium offer, accept}} \\
& + \underbrace{\pi_t^2 p_M(t, q = 2, z) [u(t, q = 2, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 2, z)]}_{\text{high offer, accept}} \\
& + \underbrace{(1 - \pi_t^1 - \pi_t^2)(1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)]}_{\text{low offer, reject}} \\
& + \underbrace{\pi_t^1 (1 - p_M(t, q = 1, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 1, z)]}_{\text{medium offer, reject}} \\
& + \underbrace{\pi_t^2 (1 - p_M(t, q = 2, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 2, z)]}_{\text{high offer, reject}}
\end{aligned} \tag{D.14}$$

Stacking over two distinct $z' \neq z''$, we get:

$$\begin{bmatrix} EV(t, z') - c(z') \\ EV(t, z'') - c(z'') \end{bmatrix} = \begin{bmatrix} d(z') & e(z') \\ d(z'') & e(z'') \end{bmatrix} \begin{bmatrix} \pi_t^1 \\ \pi_t^2 \end{bmatrix}$$

where:

$$\begin{aligned}
c(z) &\equiv p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)] \\
&\quad + (1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)] \\
d(z) &\equiv -p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)] \\
&\quad + \pi_t^1 p_M(t, q = 1, z) [u(t, q = 1, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 1, z)] \\
&\quad - (1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)] \\
&\quad + (1 - p_M(t, q = 1, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 1, z)] \\
e(z) &\equiv -p_M(t, q = 0, z) [u(t, q = 0, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 0, z)] \\
&\quad + \pi_t^1 p_M(t, q = 2, z) [u(t, q = 2, z) + E(\varepsilon_t^M | d_t^* = M, t, q = 2, z)] \\
&\quad - (1 - p_M(t, q = 0, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 0, z)] \\
&\quad + (1 - p_M(t, q = 2, z)) [EV(t + 1, z) + E(\varepsilon_t^H | d_t^* = H, t, q = 2, z)]
\end{aligned}$$

Again, so long as z is relevant for utility, we can invert to recover the vector of π s. The proof extends analogously for $K > 3$, requiring an increased support of the instruments.

E Heterogeneity by Observed Characteristics

We look at heterogeneity in the reduced form patterns of responses to the choice experiments to guide our structural specification for two reasons. First, our structural results are identified from patterns that are evident in the reduced form; Testing a null hypothesis of no heterogeneity in the reduced form is therefore equivalent to testing a null of no heterogeneity in the full structural model. Second, machine-learning methods exist that are effective at finding the most relevant predictors of heterogeneity in reduced form models when heterogeneity potentially exists along many dimensions. We thus make use of these methods to guide our search for important dimensions of heterogeneity, which we can then incorporate into the structural estimation.

To motivate what dimensions of observable respondent heterogeneity to include in our structural results, we interact a set of eighteen demographic characteristics with vignette characteristics given in Table 1. This is a high-dimensional problem as these demographic variables must be interacted with all of the characteristics described

in the vignettes. To prevent over-fitting, we employ a Lasso model to assess the most important predictors of heterogeneity. To this end, we first start by specifying linear probability models of respondent choice in the ex-post and the ex-ante experiments. We use a linear probability model and introduce interactions between the scenario characteristics and the vector of observed characteristics of respondent i , Z_i :

$$Y_{ir} = \delta H_{ir} + \alpha H_{ir} Z_i + \zeta_{ir} \quad (\text{E.1})$$

In the ex-post experiment, $Y_{ir} = 1$ if a respondent chooses the first option and $H_{ir} = H_{ir1} - H_{ir2}$ with H_{irj} giving the characteristics of option $j = \{1, 2\}$ for respondent i in round r . In the ex-ante experiment, $Y_{ir} = 1$ if a respondent chooses marriage and H_{ir} gives the characteristics described in the vignette.

Lasso minimises the mean-squared error of the prediction subject to a penalty on the absolute size of the estimated coefficients. We impose the inclusion of the scenario characteristics themselves, H_{ir} and only reduce model complexity in the interactions with observed demographic heterogeneity. Thus, we only penalise coefficients associated with the interaction terms $H_{ir} Z_i$. Following the notation of (Ahrens, Hansen, and Schaffer 2019), the estimator is:

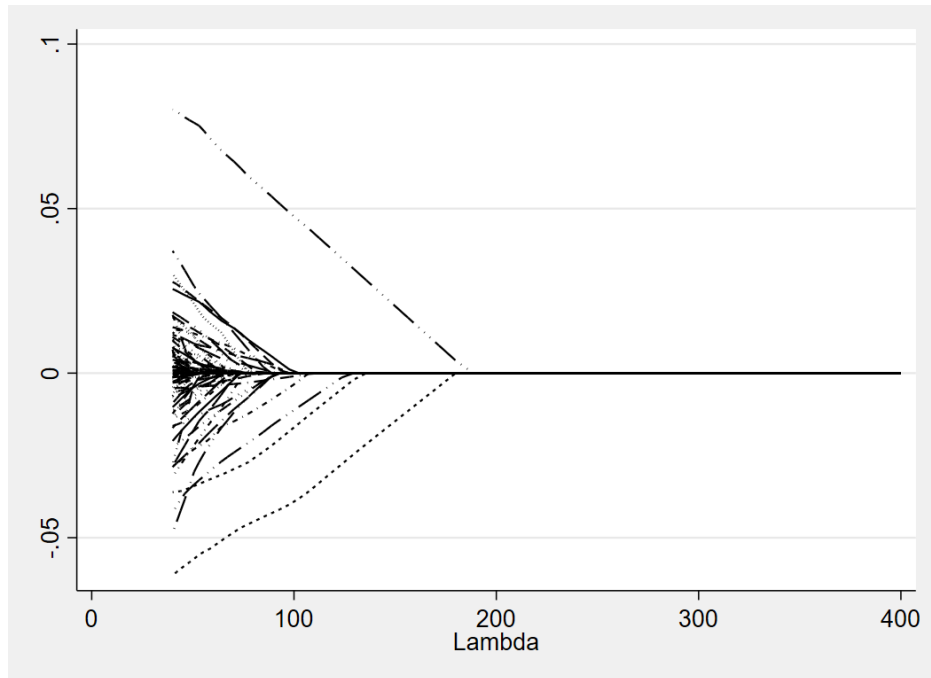
$$\left[\hat{\delta}_{Lasso}(\lambda), \hat{\alpha}_{Lasso}(\lambda) \right] = \arg \min_{\delta, \alpha} \frac{1}{NR} \sum (Y_{ir} - \delta H_{ir} - \alpha H_{ir} * Z_i)^2 + \frac{\lambda}{NR} \sum_k \psi_k |\alpha_k| \quad (\text{E.2})$$

where NR are the number of observations and ψ_k standardise the magnitude of the coefficients.⁵⁰ λ is a tuning parameter that controls the degree of penalisation. We estimate coefficients under a range of values for λ and consider the “optimal” lambda as dictated by the Extended Bayesian Information Criterion (EBIC) and Rigorous Lasso methods to determine which dimensions of heterogeneity should be included.

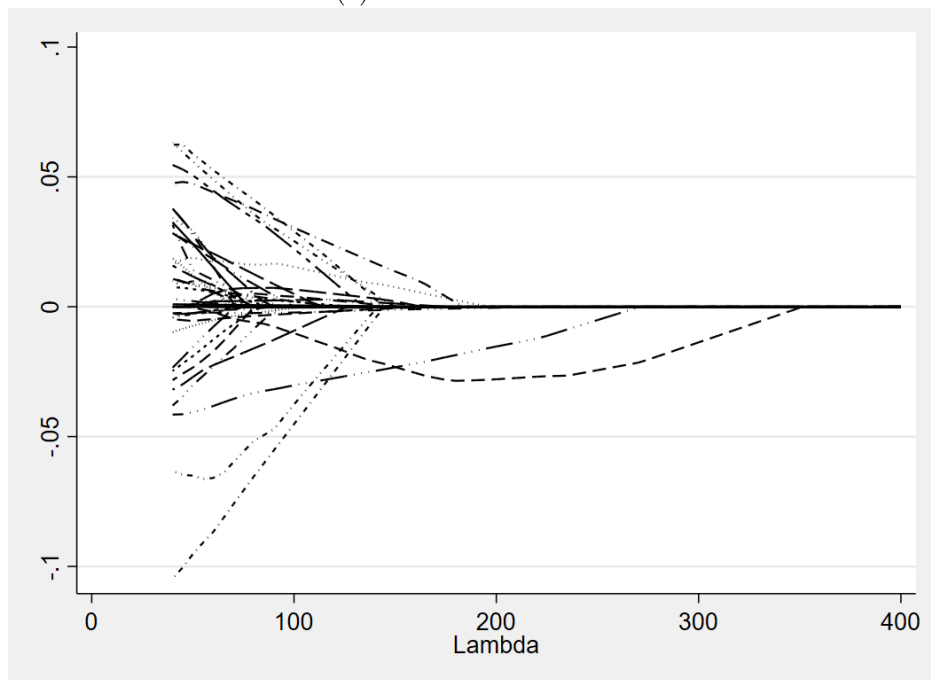
Figure A.4 plots $\hat{\alpha}_{Lasso}(\lambda)$ with the the penalty parameter λ . As we increase the severity of the penalization, more coefficients are estimated to be zero. In both the ex-post and ex-ante experiments, the EBIC and Rigorous Lasso approaches agree that no interaction terms should be included and thus that there is no heterogeneity in response patterns by observed characteristics. In the ex-post experiment, both the EBIC and Rigorous Lasso approaches agree that the optimal λ is greater than 189. Specifically, the EBIC is maximised for any $\lambda > 189$ while the rigorous method suggests the optimal λ is 315. In the ex-ante experiment, the EBIC is maximised for models

⁵⁰In particular, $\psi_k = [\sum_{ir} (\{H_{ir} * Z_i\}_k)^2]^{1/2}$ where $\{H_{ir} * Z_i\}_k$ is the k th component of $H_{ir} * Z_i$

Figure A.4: $\hat{\alpha}_{lasso}(\lambda)$ with Variation in λ



(a) Ex-Post Coefficients



(b) Ex-Ante Coefficients

where $\lambda > 351$, i.e. models where observed characteristics play no role.

F Online Appendices

Online Appendix A contains full details of the option generation process and the distribution of characteristics in our discrete choice experiments. Online Appendix B gives the details and transcripts of our qualitative focus group discussions that were used to inform the design of our vignettes and experiments.