Learning your Child's Price: Evidence from Data on Projected Dowry in Rural India

Annemie Maertens, University of Pittsburgh* A. V. Chari, RAND†

10 November 2011

Abstract

We combine novel data and methodology to shed light on the contribution to dowry of a composite characteristic that we refer to as child quality. Our findings can be summarized as follows: (1) Dowry values are not determined by household characteristics alone: child quality is a very significant determinant of dowry; (2) Quality is not a homogenous attribute for boys and girls: we distinguish between "high-level" quality (which matters for boys) and "low-level" quality (which matters for girls); (3) High-level quality does not begin revealing itself until the child enters school, whereas low-level quality starts becoming apparent at an earlier stage; (4) An increase in quality in girls appears to increase their dowry values: we argue that this is consistent with the idea that girls marry up and that quality has a horizontal component; (5) For boys, quality gets partially absorbed into educational attainments, whereas for girls quality continues to matter because it does not get translated into educational investments.

*Corresponding author: maertens@pitt.edu. The data for this working paper were gathered in India during 2007–2008, in collaboration with International Crop Research Institute of the Semi-Arid Tropics. The data collection was funded through a NSF Doctoral Dissertation Research Improvement Grant (Grant No. 0649330), an AAEA McCorkle Fellowship, a Mario Einaudi Center for International Studies International Research Travel Grant, a Cornell University Graduate School Research Travel Grant, an International Student and Scholar Office Grant and funds provided by AEM and Chris Barrett. We thank the research assistants on the field: Sanjit Anilesh, Shraavya Bhagavatula, Pramod Bangar, Sana Butool, V.D. Duche, Shital Duche, Madhav Dhere, Anand Dhumale, Nishtha Ghosh, Navika Harshe, Meenal Inamdar, Shilpa Indrakanti, Sapna Kale, Jessica Lebo, Labhesh Lithikar, Nishita Medha, Ramesh Babu Para, Abhijit Patnaik, Gore Parmeshwar, Amidala Sidappa, Nandavaram Ramakrishna, K. Ramanareddy, P.D. Ranganath, Arjun Waghmode and Yu Qin. We are grateful to Achyuta Adhvaryu, Brian Dillon, Marcel Fafchamps, Chris Ksoll, Russell Toth, Zaki Wahhaj, Shing-Yi Wang, Shuang Zhang and seminar participants at Oxford University, MWIEDC and the College of William and Mary for their helpful comments. Any remaining errors and omissions are our own.

†Part of this paper was written while Chari was a post-doctoral fellow at Cornell University.

1 Introduction

The practice of dowry, generally defined as the transfer of assets and goods from the bride's to the groom's family at the time of marriage, is now a pervasive phenomenon in India even though it was declared illegal in 1961. Anderson (2007B) reports that dowry is being practiced in about 93% percent of Indian marriages, and has even replaced brideprice in many communities. Attention has been drawn in the literature to a parallel phenomenon of dowry inflation that has resulted in dowry payments that are often several multiples of a household's annual income (for a discussion see, among others, Epstein 1973, Rajaram 1983, Caldwell et al. 1983, Billig 1992, Rao 1993A/B, Sautman 2009, Edlund 2000 and 2006, Botticini and Siow 2003 and Anderson 2003 and 2007A/B).

Understanding how dowries paid and received relate to the characteristics of bride and groom and their respective families is fundamental to determining the exact function of dowry[1] (and in turn explaining the twin phenomena of increasing prevalence and dowry inflation). While there has been a profusion of work by anthropologists, geographers and sociologists on the correlates of dowry in India (see, among others, Caldwell et al. 1983, Caplan 1984, Srinivas 1984, Billig 1992), there is remarkably little firm empirical evidence on this matter. This state of affairs is due in large part to the paucity of data on dowry transactions.[2]

In this study, we analyze a unique dataset that we collected in three Indian villages in 2008. We asked households to indicate the maximum amount that they were willing to pay as dowry for each of their daughters, and the minimum amount they were willing to accept as dowry for each of their sons (hereafter, we will refer to these amounts as projected dowry or simply dowry, wherever the gender distinction is not essential). In addition, we also asked each household to indicate their educational aspirations for each child, as well as their projection of what the child would be able to earn if he/she were to complete each of a set of educational levels.

We combine these data with a novel empirical methodology to analyze the contribution to dowry of a composite child characteristic that we refer to as child quality. The essential idea is that the cross-sectional dispersion in projected dowry values for a cohort of children should increase over time as households learn about their children's latent quality. Further, concavity in the learning process implies that the dispersion should be a concave function of age. This is exactly what we find in the data. This observation is then matched to a learning model to obtain an estimate of the dowry price of the underlying quality, as well as the parameters governing the learning process. This method of pricing an unobservable (to the econometrician) characteristic is, to the best of our knowledge, novel and may prove useful in other applications.

This study represents an improvement over the existing literature on dowry in two important respects. The existing literature relies on retrospective data on marriages and dowry payments. Because it is difficult to obtain recall data on characteristics at the time of marriage (especially for marriages that took place many years ago), researchers must usually resort to relating dowry to measures of current quality of bride and groom, which may be quite different. In contrast, our use of prospective data has the advantage that the dowry amounts elicited are projections based on each household's forecast of what its child's quality will be at time of marriage. Secondly, measurable dowry-relevant characteristics are typically few in number and are measured with error. The advantage of our methodology is that it can price aspects of quality that are difficult to observe, and it does so by identifying their effects on an observable quantity, namely dowry

[1] There are three main theories speculating on the role of dowry. In Becker's (1991) model, dowry arises as a compensation to the other partner due to a fixed division of surplus within the household (see also Rao 1993A/B, Deolalikar and Rao 1999, Dalmia and Lawrence 2005, Ambrus et al. 2009). Alternatively, if there are institutional or legal barriers to women's ability to inherit property, dowry may emerge as a culturally-sanctioned method of bequest (Zhang and Chan 1999, Arunachalam and Logan 2008). Or a pre-mortem bequest to daughters might serve as a method of maintaining sons' incentive to exert full effort in maintaining their parents' farm (Botticini and Siow 2003). Bargaining models assume that the transfers made between households during (and after) marriage can play a role (through between and within household signaling and cooperative bargaining) in marriage duration, domnestic violence, bargaining power of the newlywed daughter-in-law and overall satisfaction of the newlyweds (Zhang and Chan 1999, Bloch and Rao 2002, Platteau and Gaspart 2007, Gaspart and Platteau 2010). See Fafchamps and Quisumbing (2008) for a general introduction to marriage in developing countries.

[2] The International Crop Institute for the Semi-Arid Tropics (ICRISAT) collected dowry data among 120 households in six villages in 1983 (see Rao 1993 A/B). To our knowledge there are two recent datasets which have been collected or are in the process of being collected, one by NCAER in collaboration with the World Bank and the University of Maryland collected in 1994 in Karnataka and Utter Pradesh (see Dalmia and Lawrence 2005, Sautman 2010) and one panel survey currently being collected by the Center for Microfinance in collaboration with the Economic Growth Centre in Tamil Nadu. In addition, the national representative Indian Human Development Survey collected in 2005 by NCAER in collaboration with the University of Maryland contains some general information on marriage markets and dowries.

dispersion. As we discuss below, our results provide a substantially more nuanced picture of how dowry relates to individual qualities of bride and groom.

Our findings indicate that dowry values are not determined by household characteristics alone: child quality is a very significant determinant of dowry, as evidenced by the high dowry price of quality. Specifically, a one-standard-deviation increase in child quality changes projected dowry by as much as 60-70%. Overall, we find that child quality can explain at least 45- 65% of the variance in dowry. This may be a surprising result because marital alliances in India are principally agreed upon by the parents of the children, which raises the possibility that household characteristics may be the primary determinants of dowry.[3] The existing literature is divided on this issue: whereas Deolalikar and Rao (1998) and Rao (1993A, 1993B) find that the individual attributes in their data seem to play little role in determining dowry, Dalmiya and Lawrence (2005) report that dowry appears to equalize differences in individual characteristics (we should however note that both studies, being based on retrospective data, have limited information on individual attributes and are moreover not necessarily comparable, since they correspond to different regions and time periods).[4] Anthropological and demographic evidence also points to the importance of the individual attributes of the bridegroom and (according to some) to a lesser extent the attributes of the bride in determining the dowry (see, for example, Caplan 1984, Srinivas 1984, Caldwell et al. 1992).

Interestingly, although our estimates of the dowry price of quality are of comparable magnitude for boys and girls, quality itself correlates with different characteristics for boys and girls. We distinguish between "high-level" and "low-level" quality: the former refers to ability as captured by returns to higher education, while the latter refers to ability as captured by returns to "lower" (high-school and below) education. It turns out that only high-level quality matters for boys on the marriage market, whereas only low-level quality matters for girls.

There is also an important element of differential treatment of boys and girls that emerges from our analysis. For boys, differences in quality appear to get partially absorbed into differences in educational investment so that it becomes difficult to distinguish the dowry value of quality per se from the dowry value of educational attainments. In contrast, girls' quality continues to matter because it does not get translated into educational investments.

The distinction between high- and low-level quality, along with the finding of differential investments by gender, are congruent with the social context of rural India. Because girls are largely restricted to domestic production after marriage, high-level ability is irrelevant to their desirability on the marriage market. This result mirrors that of Behrman et al. (1999), who note that while female literacy correlates with dowry value, possibly because literacy increases a woman's productivity in home teaching, schooling achievement by women beyond levels that enable literacy are not associated with enhanced value in the marriage market (see also Srinivas 1984, Kapadia 1993, Dalmia and Lawrence 2005). In addition, because girls also tend to marry young (most girls get married by the age of 19), there is little scope for even their low-level abilities to induce greater educational investment. Put differently, the benefits of extra education are over-ridden by the costs of delaying marriage. The restrictive effect of norms relating to age of marriage on educational attainments has also been confirmed by Field and Ambrus (2008), who have instrumented for age of marriage using the timing of menarche to show that, in Bangladesh, social pressures on women to marry young cause them to have low levels of schooling (see also Maertens 2011).

The two kinds of quality are also distinguished by the rate at which they reveal themselves. Our analysis of the learning process suggests that high-level quality (which is what matters for boys) does not begin revealing itself until the child enters school, whereas low-level quality (which matters for girls) starts becoming apparent at an earlier stage. While our data are not particularly suitable for testing the implications of differential learning rates, one may nonetheless speculate that if investments in children are driven by parents' perceptions of their quality, there is a potential for greater divergence in child investments (and measurable outcomes such as health, etc) at an early age in the case of girls (assuming that investments reinforce quality differences). In contrast, such differential investments may not manifest until a later age

[3] For example, Rosenzweig and Stark (1989) have found that marriage serves a risk diversification function in rural India: households that have more variable streams of income tend to marry their daughters to grooms in more distant villages.

[4] In a slightly different context, Ambrus et al. (2008), using recall data of about 1,300 marriages from Bangladesh, find that a more educated groom receives a higher dowry (in contrast to the education of a bride, which appears to have no effect on the dowry). Arunanchalam and Logan (2008) find, in a similar context, that both bride and groom's education relate in a positive manner to dowry.

for boys, by which time the child has already had a chance to develop, possibly resulting in a lower extent of inequality among boys. We believe this is an intriguing hypothesis to pursue in future research.

A last, and surprising, finding is that an increase in quality in girls appears to increase, rather than reduce the dowry payments their parents are willing to make. While the finding of a positive price of (presumably) valuable female traits has been reported in other studies (see Edlund 2006, Arunachalam and Logan 2008) and may reflect a desire for parents of a high-quality girl to find her a "more appropriate" match, the nature of our data allows us to invoke the envelope theorem to argue that this may actually reflect an increase in what the girl's family is willing to pay to the "old" groom (i.e. the one they had in mind before the perceived increase in the girl's quality). If quality possesses "horizontal" as well as "vertical" characteristics (to use the terminology of Banerjee et al. 2010), and if girls tend to marry up, then an increase in the girl's quality reduces the extent of horizontal mismatch in her "old" relationship, and thereby increases the amount the household is willing to pay to the "old" groom.

In terms of general implications for dowry in India, the finding that the willingness-to-accept dowry is strongly influenced by the quality of the son seems compatible with a price model of dowries, and less compatible with the notion that dowry is purely a pre-mortem bequest to the daughter, or, relatedly, an altruistic model (as in Botticini 1999) where the parents of the bride compensate for a "less qualified" groom by giving their daughters more dowry. The importance of individual quality as a determinant of dowry and its correlation with returns to higher education for boys also suggests the possibility that dowry inflation may be related to the rising returns to education in India.

The remainder of this article is structured as follows. The next section introduces the data and discusses some descriptive statistics. Section 3 presents some preliminary analysis of the dowry data. Section 4 outlines the learning model that will be mapped to the data; Section 5 discusses the econometric issues involved. Section 6 reports the results and section 7 concludes.

2 Data description

We collected data in three villages in South and West India in 2008: Dokur in the Telangana region in Andhra Pradesh, and Kalman and Shirapur in the Solapur district in Maharashtra. These three villages were carefully selected in 1975 by the International Crop Research Institute of the Semi-Arid Tropics (ICRISAT) as part of their Village Level Studies (VLS) program as to represent (albeit not statistically) the semi-arid tropics in India. Issues of generalizability are discussed in Walker and Ryan (1990).[5] As a general matter, working with households who have been part of a sample for almost 40 years has the advantage that a bond of trust has been established between researchers and respondent, something which is conceivably of great importance when collecting data on activities such as giving and receiving dowry which are technically illegal. However, because this is rural India and both enforcement as well as knowledge of the law are lacking, the respondents cheerfully and freely discussed the dowry payments that they had made and received, as well as those they planned to make and receive in the future.

To obtain information on financial transfers at the time of wedding, we resurveyed 339 ICRISAT-VLS households. Only one individual was interviewed in each household, the main decision maker with regard to the education of the individuals up to the age 25 years. In most cases this is the father of the children, but in some cases it is the mother, grandfather or uncle. In the remainder of this article we will refer to this decision maker as the "parent" of the child.

In this survey we included questions on household composition, income, wealth, employment, education, and marriage related practices and social norms. The respondents' subcaste or jati are recorded, which we used to divide up the families as belonging to one of three categories: (1) Forward caste, (2) Scheduled caste (SC) and (3) Other Backward Castes (OBC). We have not attempted a detailed analysis of the data based on caste, due to the limited number of observations available.

The survey elicited (for each child) the parent's educational aspirations as well as the parent's expectation of the returns to various levels of education (unconditional on the nature of employment). The educational aspirations with regard to each child (currently enrolled in school or planning to go to school) were established

[5] This ongoing program collects detailed household and plot level agricultural data among a sample of households in six villages in semi-arid India. The sample selected is representative for each village in terms of landholding size. For an overview of the ICRISAT-VLS program see Bantilan et al. (2006), Rao and Charyulu (2007), Singh et al. (1985) and Walker and Ryan (1990).

as answers to the questions: "What is the minimum amount of education you want this particular child to obtain?" and "What is the maximum amount of education you would allow this particular child to complete?"[6]

Regarding the returns to education, we first elicited the minimum and maximum earnings the respondent thought the child would earn after finishing particular schooling milestones, for instance, 12^{th} standard. Then, we made three boxes, evenly distributed between this minimum and maximum and asked the respondent to use 20 stones (each stone representing a 5% probability) to form an earnings density function (see Delavande et al. 2011A/B on various methods to elicit beliefs in developing countries). This question was repeated for the various levels that the child still had ahead of him/her, and could include, 8^{th} standard, 10^{th} standard, 12^{th} standard, diploma, bachelor's, engineering, medical doctor, master's.[7]

Among the individuals who were below the age of 25 (including those who had married and moved out and excluding those who had married in), we collected detailed data on marriage and dowry. Among the individuals who were married already, this included retrospective information on the kind of marriage (love versus arranged marriage), the spousal selection process, income, wealth and education of both families at the time of wedding, the various transfers made at the time of wedding, and the cost of the ceremony and remaining debts with regard to the wedding.

Among the individuals who were not yet married (again below the age of 25), we asked the parent the maximum amount he or she was willing to pay for dowry at the time of the wedding in the case of girls and the minimum amount he or she was willing to accept in the case of boys.[8] In this article we analyze the latter data, i.e., the projected transfers at the time of wedding.

Note that as dowry is frequently given in the form of home-produced goods, jewelry, household appliances, and assets, such as a bicycle or three-wheeler or land, and often in several installments (Sautman 2009), the dowries elicited should be interpreted as the total value of these various goods and assets, including gifts and cash.

While we recognize the complexity of the various transfers that take place at wedding time (the bride's family offers gifts, assets, goods and cash to the groom's family and vice versa; both families transfer gifts, assets, goods and cash to the bride and groom; and both families contribute to the expenses of the wedding ceremony), in this measure of projected dowry we do not take into account the gifts transferred from the groom's family to the bride's family, or distinguish the *stridhan* component of the dowry from the price component of the dowry.[9] We are therefore measuring what Edlund (2006, 2009) refers to as "gross dowry", rather than "net dowry".[10] From the retrospective data, however, it appears that the median value of the gifts etc. transferred from the groom's family to the bride's family is only about 8% of the value of the amount spent by the bride's family. The *stridhan*, on the other hand, is a large component of the total amount spent by the parents of the bride (for the median household, *stridhan* is about 40% of gross dowry).[11] Qualitative interviews preceding the data collection however indicated that the gifts to the bride consist of clothes, gold and household appliances, and the latter are often used by the entire family-in-law. The fact that the stridhan appears to depend on whether or not the groom lives with his parents speaks to this fact (based on regression analysis of retrospective dowry data, results available on request). This phenomenon, where the line between the *stridhan* and the price component of the dowry is blurred, has also been documented by anthropologists (see for instance, Miller 1981, den Uyl 2005). For this reason, we think

[6] In some cases, the minimum and maximum aspirations are identical, but in many cases the decision maker anticipates receiving more information with regard to the child's ability and the future financial situation of the household, and the minimum and maximum represent, respectively, the worst and best case scenario in the decision maker's mind.

[7] Thus, for a child currently enrolled in 11^{th} standard, one was asked to predict the returns to completing 12^{th} standard, diploma, bachelors, engineering, medical doctor and masters, but not for 8^{th} or 10^{th} standard.

[8] During the qualitative and trial round preceding the data collection, it appeared that the respondents were unable to answer the question: "how much do you expect to pay in terms of dowry", and responded in terms of how much they are willing to set aside (and/or borrow) in the case of girls, and for how much they were willing to "let their sons go" in the case of boys. Hence, the questions were rephrased to reflect this.

[9] Historically, dowry, was perceived to consist of two components, the *stridhan,* which serves as a pre-mortem inheritance for women, and hence defined as a transfer from the bride's natal family to the bride at the time of wedding, and the *dakshina*, defined as the transfers given directly to the family-in-law, as a way of compensating the groom's family for the economic support they had to provide their new family member (Srinivas 1996).

[10] Anthropologists and sociologists also traditionally focus on gross dowry (see Srinivas 1984 and Tambiah 1989).

[11] This is in contrast with what Ambrus et al. (2008) find in the Bangladeshi context where the dowry given from parents to the bride (which they call "bequest dowry") is small compared to the dowry given to the groom's household (which they call "gift dowry").

treating the *stridhan* component separately from payments to the groom's family may not be justified.

3 Preliminary analysis

3.1 Summary statistics

Table 1 introduces the sample. With a total number of households of 1,720 and a sample size of 339 households (a total of 1,876 individuals), the sampling rate is almost 20%. The average size of a household is 5.6 members, the average *kharif* income is 51,176 Rupees (about $1,280 at the time of the survey)[12], and the average education level of the respondent is low, 4.8 years.

The sample includes a total of 838 individuals up to the age of 25 years (referred to as the "children" henceforth). Of these, 719 are unmarried: 429 boys and 290 girls. The difference between the number of boys versus girls in this sub-sample is partially due to a difference in average age of marriage: girls typically get married earlier and move out of the natal household with their husband, a practice known as patrilocality. The data confirms this. The (average) "socially acceptable" age of marriage (as reported by the households) is 18.3 years in the case of girls and 22.7 years in the case of boys.

The median projected dowry for both boys and girls is 50,000 Rs (about $1100 at the time of the survey). The average dowry willing to accept for boys is 70,671 Rs and the average dowry willing to pay for girls is 79,130 Rs. Comparing these numbers with the average kharif (or rainy season) income in Table 1, it is evident that these amounts are significant. It should be noted that these projected dowry statistics exclude the children with zero dowry[13] and the children for whom the parent could not answer the dowry question.[14]

3.2 The role of child quality

Figure 1 overlays the kernel density estimates of the projected dowry for boys and girls. Remarkably, the two distributions are practically identical (a Kolmogorov-Smirnov test cannot reject equality between the two distributions, with a p-value of 0.74). An intuitively obvious matching scheme would simply be to match households on the basis of their willingness to pay and accept dowry: that is to say, a household that is willing to pay no more than 5000 rupees for their daughter would be matched up with a household that is willing to accept no less than 5000 rupees for their son. To verify this conjecture, Figure 2 adds in the kernel density of the gross dowries that were actually paid by households of married girls in our data as a comparison (the dowries received by boys' households are not comparable with the projected dowry data because the former do not contain the *stridhan* component). One can see that the distributions match up reasonably well in the lower values of dowry but the retrospective data do not share the long tail on the right-hand side (again, even with the tail, a Kolmogorov-Smirnov test cannot reject equality between the two distributions, with a p-value of 0.25). This is due to sample selection: the retrospective data only includes data on dowries paid and received for children up to the age of 25, but the prospective sample includes children who will go on to complete higher education and get married at a later age for a higher dowry. Comparing, for instance, 18 year old girls who are married with 18 year old girls who are not married, it appears that the unmarried counterparts are more likely to belong to a higher caste, have a higher educated father, and come from wealthier families (these results are available on request).

In short, the dowry amounts elicited appear to correspond to equilibrium payments. In addition, the fact that these amounts also represent maximum and minimum willingnesses to pay and accept implies that in equilibrium most (if not all) participants in the marriage market are forced to their reservation value. While this interpretation is by no means essential to our analysis, it is certainly remarkable.

[12] The *kharif* season is the rainy season, and the main agricultural season in the semi-arid tropics of India.

[13] There are 38 boys and 8 girls for whom the respondent mentioned that they will not accept or give any dowry. As we mentioned earlier, the illegality of dowry came up not even once during the qualitative discussions preceding the data collection, so we are inclined to take these responses at face-value. Several of the respondents who told us that they would not accept dowry mentioned in the same breath that they were a "modern family" who would opt for a "registered marriage".

[14] Among the unmarried children, there are 45 boys and 12 girls for whom the respondent reported "don't know", or "it will depend" when asked about their dowry expectations. These children are excluded from the analysis. In terms of observable characteristics, there are few significant differences between this group of children and the rest of the sample.

Figure 3 graphs (separately for boys and girls) the kernel density of the logarithm of projected dowry[15] for each of three age groups: pre-school (0-6 years), primary and middle-school (7-14 years) and high school and above (15 years and upwards).[16] What is immediately apparent for boys is that the distribution of projected dowry is highly peaked for children who are in their pre-school years, but flattens significantly and spreads out thereafter. This phenomenon is strongly suggestive of a learning process, whereby new information that comes in over time improves households' forecasts of the quality of their children. We also note that this phenomenon does not appear to extend to girls, although we must caution that at this point we have not controlled for confounding differences between households.

We briefly discuss here some alternative explanations for this divergence in dowry expectations with age. The first possibility is that the phenomenon is related to the dispersion of household characteristics. As we will show in Section 6, the dispersion in dowry predicted by differences in household characteristics has no relation to child age. A second possibility is that households are not really learning anything new, but that differences in child-specific investments cumulate over time to create a deterministic divergence in children's quality. It is worth emphasizing at this point that what is being elicited is the household's forecast of the child's *future* (i.e. at time of marriage) dowry value. Presumably, this forecast takes into account future investments planned by the household. Thus, the evaluation of *current* quality may diverge in the cross-section in a deterministic way, but this by itself would not affect the distribution of forecasts of *future* quality.

One may also wonder about the role of uncertainty regarding dowry prices: perhaps households only begin to find out about dowry values once their children start nearing marriageable age. However, this implies that (absent any differences in child quality) the dispersion in price forecasts would fall over time and thereby lower dowry dispersion, rather than raise it. Similarly, if households are adjusting their forecasts for inflation in the price of quality (or other dowry-relevant characteristics), this would also tend to reduce rather than increase dowry dispersion for older children.[17] In sum, it is difficult to explain the phenomenon of increasing dowry dispersion in terms of factors other than innovations in child quality.

While child quality is by and large unobservable (to us), we nonetheless have a rough measure of quality in the form of the expected returns to different levels of education that the households report for each of their children. We aggregate these expected returns into two categories: (1) The average return to higher education (defined as post-high-school education), and (2) The average return to "lower" education (defined as high-school and below). Columns 1 and 3 in Table 2 report (separately for boys and girls) the results from a regression of projected dowry on these two measures of quality, controlling for various household characteristics.[18]

As one would expect, household income and wealth are important correlates of willingness to pay and accept dowry. Importantly, child quality (as measured by returns to education) matters for boys as well as girls, but the differences are interesting: the returns to higher (but not lower) education are important for boys, while the opposite is true for girls. To add some nuance to this picture, we repeat the estimation but this time controlling for the educational aspirations that parents have for their children (recall that these are measured as the minimum and maximum number of years of education desired for each child). The results are reported in Columns 2 and 4 of Table 2. We now find that in the case of boys, the coefficient on child quality shrinks and loses its significance whereas child quality continues to matter strongly in the case of girls (in fact the coefficients on the quality variables are practically unchanged for girls).

In short, the evidence suggests that (1) In terms of dowry value, "high-level" ability matters more than "low-level" ability for boys and that the opposite is true for girls, and (2) Differences in ability for boys appear to get (partially) transformed into differences in educational investment, whereas differences in ability for girls do not.

Notice that the number of observations in the regressions in Table 2 is significantly smaller than the full set of unmarried children. There are three reasons for this: Firstly, the returns data are incomplete because some parents could not guess the returns to all of the different levels of education. Secondly, if a child had

[15] Because the dowry distribution appears to be log-normal, we have taken the logarithm of dowry. Normality will be an important element in the analysis in Section 4.

[16] Recall that this is a cross-sectional comparison since we only have one year of data.

[17] Our sense from the interviews, however, is that households are not factoring inflation into their forecasts.

[18] We have also run these regressions with household fixed effects, but there appears to be insufficient within-household variation in the returns variables to precisely identify the effects of these variables on dowry. These results are available on request.

already completed one of the levels of education in our survey, the parents were not asked about the returns to that level of education. To create a comparable measure of returns across individuals, we have restricted the regression sample to children whose parents were able to guess the full set of returns. However, the results turn out to be robust to including all the children and creating measures of returns using the limited information provided by their parents: these results are reported in Appendix Table A1. Lastly, the returns and aspirations data were not obtained for children of school-going age who were not currently enrolled in school and did not plan to return to school in the future (presumably, the attrition in these variables is driven partly by low returns to education and partly by household characteristics/circumstances that limit educational aspirations). Because we are using returns and aspirations as regressors, we believe selection bias is not an issue in these regressions (however, because aspirations and projected dowry are simultaneously chosen by the household, we do not attempt to interpret the coefficients on the former).

This speaks to another issue: One could also think of extracting a child-specific quality component from the "returns to education" variables. Unfortunately, any positive effect of age on the cross-sectional dispersion of this measure of quality is swamped by the countervailing sample selection effect due to children dropping out of school. This attrition begins early, as revealed in Figure 4 which graphs the fraction of children in school at each age. Figure 5 shows that the (logarithm of the) cross-sectional dispersion of the education returns variables actually seems to decrease with age, as the children at the lower end of the distribution gradually drop out. For this reason, we limit our use of these variables to situations where the sample selection does not induce any obvious biases.

One last point remains to be discussed. It is perhaps not surprising that higher quality increases the minimum dowry that a boy's household is willing to accept, but it may seem unusual that higher quality would also increase the maximum dowry that a girl's household is willing to pay. However, this finding is consistent with the idea that an increase in the quality of a girl would induce her parents to look for a better match for her, and that this better match would perhaps entail paying a higher dowry. Once again an interesting nuance to this interpretation is obtained if we consider the fact that elicited dowry amounts represent maximum and minimum willingnesses-to-pay and -accept. We explain this briefly with a simple setup. This will also allow us to introduce some notation that we will carry forward into later sections.

3.3 Interpreting the dowry price of quality

We outline the argument in terms of the maximum dowry that a girl's household is willing to pay. The setup is analogous for the minimum dowry that a boy's household is willing to accept. Consider a girl's household that anticipates that at time of marriage $\mathbf{x}^F$ will be the vector of household characteristics and $\mathbf{z}^F$ will be the set of individual (i.e. the girl's) characteristics. To keep matters simple, we will restrict attention to a single individual characteristic z^F that we will refer to as "quality".

Suppose, for the moment, that only individual qualities matter. We can therefore ignore households in what follows. Denote by $V(z_M; z_F)$ the relationship value (in monetary terms) to the girl whose quality is given by z_F arising from the match with the groom whose quality is denoted by z_M. Denote by $R(z_F)$ the autarkic (reservation, i.e., "remaining single") value of the girl (we will assume that it increases in z_F). The maximum transfer that she is willing to make to the groom is given by:

$$G^{\max}(z_M; z_F) = V(z_M; z_F) - R(z_F) \tag{1}$$

For concreteness, consider the following functional form for V (we will revert to a general specification for the dowry function in the following sections):

$$V(z_M; z_F) = a + b(z_M^2 + z_F^2) - c(z_M - z_F)^2, \quad b, c > 0 \tag{2}$$

We are assuming that quality is partly "vertical" in the sense that $b > 0$ and and partly "horizontal" in the sense that $c > 0$. (This terminology is borrowed from Banerjee et al's (2011) discussion of the vertical and horizontal dimensions of caste in Indian marriage). We think of this as follows: from the perspective of the girl, higher groom quality is desirable to the extent that it increases the value of household production, but a mismatch between the qualities of bride and groom is not good for the relationship.

Clearly, the girl will be willing to pay the most to the groom who produces the greatest match value, V_F. Maximizing $G^{\max}(z_M; z_F)$ with respect to z_M, we can show that the optimal groom quality is given by:

$$z_M^* = \frac{c}{c-b} z_F \tag{3}$$

Assuming that $c > b$, it follows that $z_M^* > z_F$, i.e. all girls would prefer to "marry up" (for a discussion on hypergamy versus endogamy, see, for instance, Caplan 1984, Srinivas 1984, Anderson 2003).[19] Notice that if $c < b$ then girls would instead be marrying down. We consider the latter to be an unlikely scenario in our setting. We will therefore assume that $c > b$.[20]

Substituting optimal groom quality into (1), we obtain the (unconditional) maximum willingness to pay as a function of z_F alone (i.e. the value function of the optimization problem), which we denote by $G^*(z_F)$.

$$G^*(z_F) = a + b(z_M^{*2} + z_F^2) - c(z_M^* - z_F)^2 - R(z_F) \tag{4}$$

Now consider a marginal increment in z_F. By the envelope theorem, we can measure the associated change in dowry WTP holding fixed the groom quality at z_M^*:

$$\frac{dG^*}{dz^F} = \{2bz_F - R'(z_F)\} + 2c(z_M^* - z_F) \tag{5}$$

The increment in z^F increases the girl's production value and hence her value in the relationship with (the groom indexed by) z^{M*}. The increment in quality presumably also raises her autarkic value. Even if the latter increase cancels out the former (i.e. the term in curly brackets is zero), her WTP with respect to z_M^* still tends to increase because the term $2c(z^{M*} - z^F)$ is positive. The reason is that the girl was marrying up: the increment in her quality therefore reduces the extent of horizontal mismatch vis-a-vis z^{M*}. Conversely, because boys are marrying down, an increment in a boy's quality will increase the horizontal mismatch between him and his potential partner, which will make him demand greater compensation. Thus, the dowry price of quality may be positive for both girls as well as boys.

The framework outlined above is Beckerian in the sense that relationship values are determined independently of dowry, the latter playing the role of equilibrating transfers (see Becker 1981 for a model of marriage market transfers). Alternative, one could conceive of a framework in which dowry itself determines relationship value. For example, the amount of dowry that a bride brings with her may determine her bargaining power, and hence her own welfare in the new household (Zhang and Chan 1999). If the marginal effect of dowry on bargaining power (and post-marriage bridal welfare) increases with girl's quality, then the envelope theorem may again imply that the dowry price of girl's quality is positive.

4 Learning and the evolution of dowry variance

We propose to use the evolution in dowry variance with age to estimate the price of child quality, the latter being defined as the change in willingness-to-pay (or accept in the case of boys) due to an increment in child quality. We adopt the following convention: because child quality is unobservable, we will define it so that its price (as defined above) is always positive. Thus, for example, suppose an increase in child quality of a girl, z^F, lowers the maximum willingness-to-pay. We would then re-define z^F as the negative of child quality so that its price becomes positive.

We will assume that the logarithm of the $G^*()$ function, denoted by $g(\cdot)$, takes the following form:

$$g(\mathbf{x}_t^i, z_t^i) = f(\mathbf{x}_t^i) + pz_t^i \tag{6}$$

where the subscript t indexes the age of the child and i indexes the child (we have dropped gender superscripts to reduce notational clutter). The vector $\mathbf{x}$ indexes household characteristics while p is the price of quality. We have indexed z by child age to emphasize that quality is learnt over time. z_t^i therefore refers to the expectation of child quality, based on information available at time t.

[19] The relationship value function V must take a different form for boys in order to prevent them from marrying up as well. A simple expedient (and one that may have some psychological validity) is to assume that boys have a greater aversion to marrying up than to marrying down. In the case of girls, we have assumed a symmetric form for the horizontal match component for the sake of algebraic covenience, but we could also have assumed that girls are more averse to marrying down than to marrying up.

[20] If $c = 0$, then all girls will have the same ranking of grooms. Specifically, the ranking of grooms will correspond to their (i.e. the grooms') ranking of quality. This would also imply that the elicited maximum WTP for all girls is with reference to the first-ranked groom. Similarly, the elicited minimum WTA for all boys is with reference to the first-ranked girl. We think this is unlikely given that the distribution of WTP and WTA also map up with *actual* dowries paid and received.

Our goal is to develop an estimable model that relates learning about quality to changes in dowry dispersion. The essential theoretical complication is that quality is not completely pre-determined: while it is plausible that some part of quality is "given", there is no doubt that purposeful investments such as schooling, time spent with the child, etc can influence overall quality. These investments, in turn, are (a) sensitive to new information about latent child quality and (b) constrained by household resources. This makes it extremely difficult (if not impossible) to derive any clear predictions about how the distribution of overall quality evolves over time. To proceed, we therefore focus on a component of quality that is not affected by household resources or constraints. To be clear, this component of quality may still be correlated with household resources to the extent that innovations in quality may cause adjustments in household assets (for example), but this is not a problem in terms of theoretical modeling (although it does constitute an econometric problem, as we discuss in the next section).

To flesh this out a little more formally, suppose that there is a (perceived) production function for overall quality:

$$Z_t^i = \Phi(\mathbf{x}_t^i, Q_t^i) \tag{7}$$

where the production function Φ depends on an exogenously given child-quality term denoted by Q_t^i and household resources $\mathbf{x}_t^i$. Exogenous child-quality is fixed, but is gradually revealed over time, with Q_t^i representing its expected value at time t. We are assuming that this production function is multiplicatively separable as follows:

$$Z_t^i = K(Q_t^i)H(\mathbf{x}_t^i) \tag{8}$$

In this formulation, K subsumes that part of household investment that does not depend on $\mathbf{x}$. Using lower-case letters to denote logarithms, we can write:

$$z_t^i = q_t^i + h(\mathbf{x}_t^i) \tag{9}$$

where $q_t^i = log(K(Q_t^i))$. We will henceforth refer to q_t^i as quality, and model its evolution.

To model the evolution of cross-sectional variance of q_t^i (i.e. $Var(q_t^i|t)$), we consider a simple learning model. We assume that before a child is born, his or her parents have a prior belief about his or her quality. Denote by θ^i the true (unobserved) quality of child i. In particular, we assume that the prior is normally distributed with mean and variance θ and σ_θ^2, and that this prior corresponds exactly to the actual distribution of child quality in the population. Thereafter, the parents receive independent signals in each period about the child's productivity which they then use to update their priors. These signals are drawn from a normal distribution with mean θ^i and variance σ_ϵ^2:

$$s_t^i = \theta^i + \epsilon_t^i, \qquad \epsilon_t^i \sim N(0, \sigma_\epsilon^2) \tag{10}$$

In addition, we assume that the signal errors ϵ_t^i are uncorrelated over time and across individuals. Using the notation introduced above, denote by q_t^i the expected quality of child i at age t. Denote now by ρ_t the precision of this belief after the t-th signal has been received - this is the inverse of the variance of the belief distribution. Similarly we denote by ρ_ϵ the precision of the signal (i.e. the inverse of the variance of the signal, σ_ϵ^2). Bayesian updating implies:

$$q_{t+1}^i = (1 - \alpha_{t+1})q_t^i + \alpha_{t+1}s_{t+1}^i \tag{11}$$

where s_{t+1}^i denotes the signal of child i's quality at age $t+1$ and $\alpha_{t+1} = \frac{\rho_\epsilon}{\rho_t + \rho_\epsilon}$. A standard result (see for example Chamley 2003) is that with normal priors and signals, the precision of the belief increases deterministically in a linear way, i.e. $\rho_{t+1} = \rho_t + \rho_\epsilon$.

Using the recursivity of (11), we can show that:

$$\begin{aligned} q_t^i &= \frac{\rho_0}{\rho_t} q_0^i + \frac{\rho_\epsilon}{\rho_t} \sum_{\tau=1}^{t} s_\tau^i && (12) \\ &= \frac{\rho_0}{\rho_t} q_0^i + \frac{\rho_\epsilon}{\rho_t} (t\theta^i + \sum_{\tau=1}^{t} \epsilon_\tau^i) && (13) \end{aligned}$$

where q_0^i denotes the expected quality immediately after birth. Recall that q_0^i equals θ because when the child is born, its expected quality is θ. We can now write:

$$q_t^i = \frac{\rho_0}{\rho_t}\theta + (1 - \frac{\rho_0}{\rho_t})\theta^i + (1 - \frac{\rho_0}{\rho_t})\frac{1}{t}\sum_{\tau=1}^{t}\epsilon_\tau^i \tag{14}$$

As t increases, the precision of the belief, ρ_t increases unboundedly. Because the signal errors are mean-zero, it follows that the expected quality, q_t^i, converges to true quality θ^i in the probability limit. Thus the parents' belief about their child's quality collapses to true quality in the limit.

In contrast, the cross-sectional dispersion in expected qualities increases over time. To see this, we start by writing the cross-sectional variance of q_t^i:

$$Var(q_t^i|t) = (\frac{\rho_\epsilon}{\rho_t})^2[t^2\sigma_\theta^2 + t\sigma_\epsilon^2] \tag{15}$$

where we have used (10) and the assumption that parents have common priors before the child is born. With a little bit of algebra, this can be rewritten as:

$$Var(q_t^i|t) = \frac{\sigma_\theta^2 ct}{1 + ct} \tag{16}$$

where $c = \frac{\rho_\epsilon}{\rho_0}$. A more revealing form of this expression is given by:

$$Var(q_t^i|t) = \sigma_\theta^2 - \sigma_t^2 \tag{17}$$

where σ_t^2 is the variance of the belief distribution at t. Recalling that $\sigma_0^2 = \sigma_\theta^2$ and that σ_t^2 falls with t, the expression above reveals that the cross-sectional dispersion in quality increases from 0 and asymptotes to σ_θ^2, the rate of convergence being governed by the parameter c. It can be verified from (16) that $Var(q_t^i|t)$ increases at a diminishing rate, i.e. $Var(q_t^i|t)$ is concave.

5 Identification and econometric methodology

When quality is priced in dowry terms, the variance in dowry due to quality (assuming for the moment that projected dowry is solely a function of quality) is given by:

$$Var(pq_t^i|t) = \frac{p^2\sigma_\theta^2 ct}{1 + ct} \tag{18}$$

The quantity $p\sigma_\theta$ represents the standardized price of quality: it is the percentage increase in projected dowry (recall that we are now working with the logarithm of dowry) due to a one-standard deviation increase in quality.

Figure 6 plots the function given in (18) for some selected values of $p\sigma_\theta$ and c. The figure confirms that $p\sigma_\theta$ and c are separately identifiable: $p^2\sigma_\theta^2$ is the value that $Var(pq_t^i|t)$ asymptotes to, while c and $p^2\sigma_\theta^2$ together control the curvature (notice that we cannot identify the sign of the price).

However, $Var(pq_t^i|t)$ is not directly observed because q_t^i is unobservable. We must therefore distinguish q_t^i from other time-varying determinants of dowry by isolating it as a residual. The difficulty with this approach is that q_t^i is possibly correlated with household variables because innovations in expected quality change the lifetime wealth of the family and may result in the adjustment of household assets, etc (recall, however, that by definition, q does not change in response to changes in household resources). As long as the correlation between q and household variables is not a function of age, however, it is possible to consistently estimate c and estimate a lower bound on the price of quality, as we now show.

We adopt a linear regression framework in what follows. We are dispensing with the time and individual subscripts in this section because they are not essential for the following discussion. Using (6) and (9), we consider the projected dowry function:

$$g(\mathbf{x}, q) = \{f(\mathbf{x}) + ph(\mathbf{x})\} + pq \tag{19}$$

Suppose that we can approximate $\{f(\mathbf{x}) + ph(\mathbf{x})\}$ by a linear combination of the variables in $\mathbf{x}$. Denote by e the residual from a least-squares regression of projected dowry on household characteristics $\mathbf{x}$. By the properties of linear projections, we can write:

$$e \approx pq - \tilde{x} \tag{20}$$

where $\tilde{x}$ is the linear projection of pq on $\mathbf{x}$ (implying that $cov(e, \mathbf{x}) = 0$). The conditional variance of e is given by:

$$\begin{aligned} Var(e|t) &= Var(pq|t) + Var(\tilde{x}|t) - 2cov(pq, \tilde{x}|t) && (21) \\ &= Var(pq|t) - Var(\tilde{x}|t) - 2cov(e, \tilde{x}|t) && (22) \end{aligned}$$

Suppose now that $cov(e, \tilde{x}|t) = 0$. Notice that this is not guaranteed, although the unconditional covariance $cov(e, \tilde{x}) = 0$. However, a sufficient condition for this to hold is that the variables in $\mathbf{x}$ are interacted with the full set of time dummies. We can then write (21) as:

$$Var(e|t) = Var(pq|t)[1 - R_t^2] \tag{23}$$

where R_t^2 denotes the R-squared from a regression of pq on $\mathbf{x}$. If $\mathbf{x}$ were a scalar, R_t^2 would simply be the correlation between pq and $\mathbf{x}$ at time t. Then we can write:

$$Var(e|t) = \frac{(1 - R_t^2)p^2\sigma_\theta^2 ct}{1 + ct} \tag{24}$$

Suppose further that R_t^2 does not change with t. That is, the adjustment in $\mathbf{x}$ following a change in q does not affect the proportion of variation in pq that can be explained by $\mathbf{x}$ (in the scalar case, this means that the correlation between pq and x does not change with t). In that case, an inspection of (24) suggests that we can estimate $(1 - R^2)p^2\sigma_\theta^2$. We would necessarily be underestimating $p\sigma_\theta$ unless changes in q do not induce any adjustments in $\mathbf{x}$ (i.e. q and $\mathbf{x}$ are uncorrelated). The latter condition cannot be directly tested, although as we discuss in Section 6 there is reason to think that it may be approximately true in our data.

Finally, having obtained the residual e, we estimate price by casting (24) in estimable form as follows:

$$Var(e|t) = \frac{\tilde{p}^2 ct}{1 + ct} + \xi_t \tag{25}$$

where $\tilde{p}^2 = (1 - R^2)p^2\sigma_\theta^2$ and ξ_t represents sampling error. We estimate the coefficients $\tilde{p}$ and c by non-linear least squares. To account for the fact that the variance in each age-group is constructed from a different number of observations, we also implement a weighted version of this regression. In practice, there turns out to be virtually no difference between the estimates from the weighted and unweighted regressions.

6 Results

We begin by isolating the quality component of dowry for boys and girls separately, using the regression framework outlined in Section 5. To recapitulate the discussion in the preceding sections, we will obtain the residuals from a regression of projected dowry on a set of household characteristics. The regression equation is:

$$g_t^i = \boldsymbol{\beta}_t \mathbf{x}_t^i + u_t^i \tag{26}$$

where g_t^i is the logarithm of the willingness-to-pay (accept, in the case of boys) and $\mathbf{x}$ is a vector of observable household characteristics. As discussed in Section 5, we are interacting household-level variables with age in order to force the age-wise covariance between the residual and the $\mathbf{x}$ variables to be zero. Because this constitutes a demanding specification, we narrow down the set of household variables those that appeared to be significant correlates of projected dowry in Table 2 (we will also present estimates from a less demanding specification which includes all the household variables in levels, i.e. not interacted with age). Specifically, for boys we retain the following variables: (1) Logarithm of household income, (2) Logarithm of total value of household assets, (3) Education level of decision maker, (4) Number of adults in the family, (5) Number of unmarried boys in the family, and (6) Overall (i.e. not within-gender) child order. For girls, we replace the variable "number of unmarried boys in the family" with "number of unmarried girls in the family". The age-wise variance in the residual will then be mapped to the learning model to estimate the price if quality, $\tilde{p}$, and the learning parameter c.

At this point we limit the regression sample to include only children up to the age of 18, in order to limit the effect of attrition from unmarried status on the age-wise dowry variance. Attrition actually begins a few years earlier for women, but because the final estimation of parameters will be based on age-wise variance we retain women until the age of 18 in order to keep the estimation sample from being too small. While this is our base specification, we will also present the results for different age-cutoffs, to examine the robustness of the estimates.

Figure 7 graphs the age-wise variance of actual, predicted and residual dowry from the base specification. Because the variance in each age-category is constructed from a different number of observations, the graphs are weighted. An immediate observation is that for boys as well as girls, actual and residual dowry dispersion appear to increase in a concave fashion, strongly supporting the idea of a concave learning effect.

Also notable is that the part of dowry dispersion that is predicted by household variables does not seem to be strongly related to age. This is interesting because one might intuitively expect that if there is a correlation between child quality and household variables, then increasing dispersion in child quality would also show up as increasing dispersion in household variables, and possibly an increasing dispersion in predicted dowry. We conjecture that the weak correlation between household variables and child quality is due to the fact that many households have multiple children. As many as 72% of unmarried boys and 60% of unmarried girls are in households with more than one unmarried boy and more than one unmarried girl, respectively. Nearly 94% of unmarried children are in households with at least one other unmarried child (of either gender). If children in the same household have independent innovations in quality, then this might conceivably dampen the adjustment of household variables to any particular child's innovation.

We now turn to the analysis of residual dowry, i.e. the quality component. Figure 8 graphs the kernel densities of residual dowry for each of the three age categories; 0-6 years, 7-14 years and 15 years and upwards. In the figures, the spreading-out of the dowry distribution with age, that was noted in Section 3, is even more starkly apparent, and this time for girls as well.

The figure reinforces an important fact that we had also noted in Section 3. The peakedness of the quality (in dowry terms) distribution changes dramatically for boys going from pre-school to school, suggesting that sending boys to school may reveal a substantial component of their quality. In contrast, the peakedness of the quality distribution for girls does not exhibit as sharp a change when girls reach school age. Formally, a Kolmogorov-Smirnov test for equality of distributions rejects the equality of pre-school and primary-middle school quality distributions in the case of boys (with a p-value of 0.04), but fails to reject equality in the case of girls (with a p-value of 0.42). Figure 9 replaces the 0-6 category with the finer categories 0-3 years and 3-6 years (due to small sample sizes, we are not able to precisely estimate densities at finer levels of aggregation). There appears to be no evidence of learning for boys in the pre-school years but learning happens rapidly once the boys enter school-age; however, parents appear to be learning about girls' quality at a steady rate throughout. A similar observation emerges from a visual inspection of the residual dowry variance graphs in Figure 7: in the pre-school years, the variance does not show any perceptible increase for boys whereas it does appear to be increasing for girls. As we will see shortly, the estimated learning rates are correspondingly quite different for boys and girls.

This result bears an intuitive correspondence with the descriptive analysis in Section 3, where we commented on the fact that "high-level" ability seems to matter for boys but not for girls, whereas the converse is true of "low-level" ability. If "high-level" ability is only revealed by performance in school whereas "low-level" ability begins to become apparent much earlier, this might explain our findings above.

Figures 7-9 indicate that the principal benefit of controlling for household variables is that by doing so we remove a significant amount of variation in dowry values that is unrelated to quality and masks the effects of the latter. At the same time, Figure 7 also suggests that the correlation between these variables and quality may be quite weak. If so, we could also carry out our estimation of the structural parameters without adjusting for the effects of household variables (i.e. using the raw dowry data) and hope to obtain similar (albeit more noisy) estimates. We will test this conjecture shortly.

We now present the results from our estimation of the learning model. To compare estimates for boys and girls, we pool the samples and jointly estimate separate parameters for the two groups. Table 3 reports the estimated coefficients from a variety of specifications, and also tests whether the differences in estimated coefficients for boys and girls are statistically significant.

We describe the results briefly before discussing their interpretation. The estimates in Column 1 indicate that both price as well as the learning parameter c are comparable for boys and girls, although the latter is not very precisely estimated. Column 2 repeats the estimation, but this time restricting the price to be the same for boys and girls. As we noted above, however, it appears that learning about quality begins at a later age for boys. We therefore redo the estimations, setting age 6 (the first year of primary school) as the starting year for boys. Column 3 reports the results. Prices are again comparable between the groups, but this time there appears to be a large difference in learning rates: the rate of learning is greater for boys than for girls. In Column 4 we again restrict the price to be equal across the two groups (because this restriction appears to be generally valid, we impose it throughout in the remaining estimations). In this specification, the difference in learning rates is statistically significant. Columns 5 and 6 repeat the specification in Column 4 but this time restricting the sample to children up to the ages of 17 and 16, respectively, in order to assess the sensitivity of the estimates to the age-cutoff. Column 7 estimates the model coefficients but does so using the residuals from a regression of dowry on all time-varying household variables in levels, i.e. not interacted with age. Finally, Column 8 uses the specification in Column 4 but this time performs the estimation on the raw dowry data (i.e. without first removing the effects of the household variables).

In all specifications, the price of quality is very strongly significant, both in economic as well as in statistical terms. In particular, a one-standard deviation increase in quality would appear to increase dowry by about 60%-70% for both boys and girls.[21] These are very significant amounts and speak to the importance of child attributes in determining dowry values. To get a better sense of these magnitudes, recall that, assuming that quality is perfectly observed, the square of the standardized price can also be interpreted as the variance in dowry that is due to variation in quality (not accounting for any unobservable correlation between quality and the other determinants of dowry). Given that the variance in (the logarithm of) projected dowry at age 18 (when quality has been largely revealed and the dowry variance levels off) is 0.80 for boys and 0.58 for girls, quality variation can account for about 46% of dowry variance for boys and about 64% of the dowry variance for girls (and to the extent that our estimates of price are downward biased, these figures represent lower bounds).

The significant difference in estimated learning rates for boys and girls may be understood as follows. The equality of price of quality for boys and girls implies that the dowry variance due to quality must converge to the same value for the two groups. Given that quality only starts to be revealed at a later age for boys, it follows that once learning begins, the rate of learning must be greater for boys than for girls.

Turning to our robustness checks, the estimates in Columns 5 and 6 confirm that the results are quite robust to the age-cutoff, indicating that attrition from married status is not significantly biasing our estimates. Finally, performing the estimation using (a) the residuals from the regression of dowry on all household variables in levels (Column 7), and(b) raw dowry data (Column 8), we find that the estimated price is comparable to our earlier estimates, while the learning parameters are poorly estimated.

7 Concluding discussion

Using a novel dataset that contains elicited child-specific expectations of dowry payments, as well as expected educational returns and aspirations among households in three Indian villages, we have analyze the contribution to dowry of a set of characteristics that we refer to as "child quality". We hypothesize that certain aspects of this child quality are only revealed gradually over time, as the child grows older. If this is the case, the cross-sectional dispersion in projected dowry values for a cohort of children bears a concave relation to their age, and this relationship can be used to fit a learning model to the data to estimate the price of quality as well as the parameters governing the learning process.

We have found that child quality is an important determinant of dowry, and that parents learn about this quality as children grow older. However, we show that "quality" correlates with different characteristics for boys and for girls. In particular, returns to higher (but not lower) education matter for boys on the

[21] That the price of quality is comparable for boys and girls seems intuitive to us: if boys and girls are being matched on the basis of their qualities, then it would make sense that an increase in quality would have the same equilibrium price (in absolute terms) for both sexes.

marriage market, while the returns to lower (but not higher) education matter for girls. These results are consistent with the qualitative interviews we conducted in the villages at the start of this study. According to the respondents' own account a girl's physical appearance, "homely nature" and more general ability with regard to household work and management are characteristics which are highly valued on the marriage market. Note that our results do not imply that girl's education is not valued on the marriage market. Over 90% of the children under the age of 15 years in our study are enrolled in school, with little difference between the sexes. As Behrman et al (1999) note using the 1983 ICRISAT data, a literate wife is associated with a lower dowry, but any education beyond that does not affect dowry. Literate women, in their data, are more effective home teachers and command a premium in the marriage market. In the same way, it is plausible that today, girls who received 10 years of education are considered more effective home makers, while beyond that, girls bump into marriage and labor market related social norms. In the case of boys, respondents mentioned education, and more generally, the ability to earn a living, as the most important quality. Several respondents indicated that the potential groom's educational attainment matters for dowry even if he does not have a job.

In terms of quantitative studies set in India, our results can be compared to Dalmia and Lawrence (2005) who, using retrospective dowry data from almost 1,900 households collected in 1995 in Karnataka and Utter Pradesh, find that both groom's schooling and bride's schooling increase the net dowry the bride family pays. Deolalikar and Rao (1998) and Rao (1993 A/B), on the other hand, using a dataset of 120 households collected retrospectively in 1983 in the same set of villages as we consider in this study, find that none of the individual traits of the bride and groom such as schooling or height seems to matter. However, as we already noted, Behrman et al (1999) have reported the dowry value of female literacy using the same data. While the difference in their results is driven by choice of specification, the differences with our own results are more likely to relate to the fact that there are 25 years between the two surveys, and the returns to education have changed drastically during this period.

Finally, the differences in the rates of learning for boys and girls offers an interesting avenue for future research. The fact that quality is revealed earlier for girls may disadvantage them relative to boys, particularly if investments in children are indeed driven by parents' perceptions of quality and if such investments tend to reinforce existing differences among children.

References

[1] Ambrus, A., E. Field.and M. Torero. 2008. "Muslim Family Law, Prenuptial Agreements and the Emergence of Dowry in Bangladesh." Working Paper Harvard University and IFPRI.

[2] Anderson, S. 2003. "Why Dowry Payments Declined with Modernization in Europe but Are Rising in India,"*Journal of Political Economy*, 111: 2, pp. 269-310.

[3] Anderson, S. 2007A. "The Economics of Dowry and Brideprice," *Journal of Economic Perspectives*, 21:4, pp. 151-174.

[4] Anderson, S. 2007B. "Why the Marriage Squeeze cannot cause Dowry Inflation," *Journal of Economic Theory*, 137:1, pp.140-152

[5] Arunachalam, R. and T.D. Logan. 2008. "On the Heterogeneity of Dowry Motives," NBER Working Paper No 12630.

[6] Banerjee, A., E. Duflo, M. Ghatak and J. Lafortune. 2010. "Marry for What? Caste and Mate Selection in Modern India." Working Paper MIT, LSE and the University of Maryland.

[7] Bantilan, M.C.S., P. Anand Babu, G.V. Anupama, H. Deepthi, and R. Padmaja. 2006. "Dryland Agriculture: Dynamic Challenges and Priorities." Research Bulletin No. 20, GT-IMPI, ICRISAT.

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

[9] 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, pp. 682–714.

[10] Billig, M.S. 1992. "The Marriage Squeeze and the Rise of Groomprice in India's Kerala State," *Journal of Comparative Family Studies* 23:2, pp.197-216.

[11] Botticini, Maristella. 1999. "A Loveless Economy? Intergenerational Altruism and the Marriage Market in a Tuscan Town, 1415-1436, "*Journal of Economic History,* 59:1, pp. 104-121.

[12] Botticini, M. and A. Siow. 2003. "Why Dowries?," *American Economic Review*, 93:4, pp. 1385-1398.

[13] Caldwell, J.C., P.H. Reddy and P. Caldwell. 1983. "Causes of Marriage Change in South India," *Population Studies* 37:3, pp. 343-361

[14] Caplan, L. 1984. "Bridegroom Price in Urban India: Class, Caste and 'Dowry Evil' Among Christians in Madras," *Man*, 19:2, pp. 216-233.

[15] Chamley, C. 2003. *Rational Herds.* Cambridge University Press, Cambridge, UK.

[16] Dalmia, S.and P.G. Lawrence. 2005. "The Institution of Dowry in India: Why it Continues to Prevail," *The Journal of Developing Areas,* 38:2*, pp. 71-93.*

[17] Dasgupta, I. and D. Mukherjee. 2003. "'Arranged' Marriage, Dowry and Female Literacy in a Transitional Society," CREDIT Research Paper.

[18] Delavande, A., X. Giné, and D. McKenzie. 2011A. "Measuring Subjective Expectations in Developing Countries: A Critical Review and New Evidence," *Journal of Develment Economics*, 94:2, pp. 151-163 .

[19] Delavande, A., X. Giné, and D. McKenzie. 2011B. "Eliciting Probabilistic Expectations with Visual Aids in Developing Countries: How Sensitive are Answers to Variations in Elicitation Design?," *Journal of Applied Econometrics*, 26:3, pp. 479-497.

[20] Den Uyl, M.. 2005 "Dowry in India: Respected Tradition and Modern Monstrosity," in *The Gender Question in Globalization: Changing Perspectives and Practices,* T. Davids and F. Van Driel (eds.), Ashgate Publishing, pp. 143-158

[21] Deolalikar, A. and V. Rao. 1998. "The Demand for Dowries and Bride Characteristics in Marriage: Empirical Estimates for Rural South-Central India," in *Gender, Population and Development*, M. Krishnaraj, R. Sudarshan, and A. Shariff (eds), Oxford University Press.

[22] Edlund, L.. 2000. "The Marriage Squeeze Interpretation of Dowry Inflation: A Comment," *Journal of Political Economy*, 108:6, pp. 1327-1333.

[23] Edlund, L.. 2006. "The Price of Marriage: Net Versus Gross Flows and the South Asian Dowry Debate", *Journal of the European Economic Association Papers and Proceedings*, 4:2-3, pp. 542-551.

[24] Epstein, T. S. 1973. *South India: Yesterday, Today and Tomorrow.* Macmillan, London.

[25] Field, E.and A. Ambrus. 2008. "Early Marriage, Age of Menarche, and Female Schooling Attainment in Bangladesh." *Journal of Political Economy*, 116:5, pp. 881-929.

[26] Fafchamps, M. and A.R. Quisumbing. 2008. "Household Formation and Marriage Markets in Rural Areas," in P.Schultz and J.A. Strauss (eds.), *Handbook of Development Economics*, vol. 4, number 5, pp. 3188-3247.

[27] Gaspart, F.and J. Platteau. 2010. "Strategic Behavior and Marriage Payments: Theory and Evidence from Senegal," *Economic Development and Cultural Change,* 59:1, pp. 149-185.

[28] Kapadia, K. 1993. “Marrying Money: Changing Preference and Practice in Tamil Marriage,” *Contributions to Indian Sociology,* 27:1, pp. 25-51.

[29] Mani, A.. 2011. "Mine, Yours or Ours: The Efficiency of Household Investment Decisions – an Experimental Approach," Working Paper University of Warwick.

[30] Miller, B. 1981. *The Endangered Sex: Neglect of Female Children in Rural North India*. Cornell University Press, Ithaca, NY.

[31] Maertens, Annemie. 2011. "High returns? Low Aspirations. Social Norms and Education in Village India". Working Paper Cornell University.

[32] Platteau, J. and F. Gaspart. 2007. "The Perverse Effects of High Brideprices," *World Development,* 35:7, pp. 1221–1236.

[33] Rajaram, I. 1983. "Economics of Bride Price and Dowry," *Economic and Political Weekly,* Feb 19, pp. 275-278.

[34] Rao, K.P.C, 2008. "Documentation of the Second Generation Village Level Studies 2001-2004". GT-IMPI, ICRISAT.

[35] Rao, K.P.C. and Kumara D. Charyulu. 2007. "Changes in Agriculture and Village Economies." Research Bulletin no 21, GT-IMPI, ICRISAT.

[36] Rao, V. 1993A. "The Rising Price of Husbands: A Hedonic Analysis of Dowry Increases in Rural India," *Journal of Political Economy*, 101:4, pp. 666–677.

[37] Rao, V. 1993B. "Dowry ‘Inflation’ in Rural India: A Statistical Investigation," *Population Studies*, 47:2, pp. 283-293.

[38] Rosenzweig, M.R., and O. Stark, 1989. “Consumption Smoothing, Migration, and Marriage: Evidence from Rural India,” *Journal of Political Economy*, 97:4, pp. 905-926.

[39] Sautman, A.. 2009. "Partner Search and Demographics: The Marriage Squeeze in India". Job market paper, NYU.

[40] Singh, R.P., Hans P. Binswanger, and N.S. Jodha. 1985. "Manual of Instructions for Economic Investigators in ICRISAT’s Village Level Studies (Revised)." GT-IMPI, ICRISAT.

[41] Srinivas, M.N. 1984. *Some Reflections on Dowry.* Oxford University Press.

[42] Srinivas, M.N., 1996. *Village, Caste, Gender and Method: Essays in Indian Social Anthropology.* Oxford University Press.

[43] Tambiah, S.J. 1989. "Bridewealth and Dowry Revisited – the Position of Women in Sub-Saharan Africa and North India," *Current Anthropology,* 30:4, pp. 413-435.

[44] Walker, T.S. and J.G. Ryan. 1990. *Village and Household Economies in India's Semi-Arid Tropics.* John Hopkins University Press, Baltimore and London.

[45] Zhang, Junsen and William Chan, 1999. "Dowry and Wife's Welfare: A Theoretical and Empirical Analysis,"*Journal of Political Economy,* 107: 4, pp. 786-808.

Table 1: Selected descriptive statistics

Number of households in village	1,720
Number of households in sample	339
Number of married children[1]	119
Number of unmarried children[1]	719
Average age of married children	21
Average age of unmarried children	11
Average number of household members	5.55
Average *Kharif* income (Rs)[2]	51,176
Average education level of respondent (in years)[3]	4.77
Average dowry willing-to-accept (boys) (Rs)[4]	70,671
Average dowry willing-to-pay (girls) (Rs)[4]	79,130
Average dowry willing-to-accept (boys) (Rs)[4]	50,000
Average dowry willing-to-pay (girls) (Rs)[4]	50,000

Notes: [1]A child is defined as an individual up to the age of 25. Note that the adults also include the daughters-in-law under 25 who did not received their education in the household; [2]The Kharif season is the rainy season; [3] respondent is the main decision-maker with regard to the education of the children under 25 year in the household; [4] Excludes children with zero dowry and the children for whom the parent could not answer the dowry question.The information of one child in Dokur is missing from the sample because the enumerators by accident skipped his/her questionnaire.

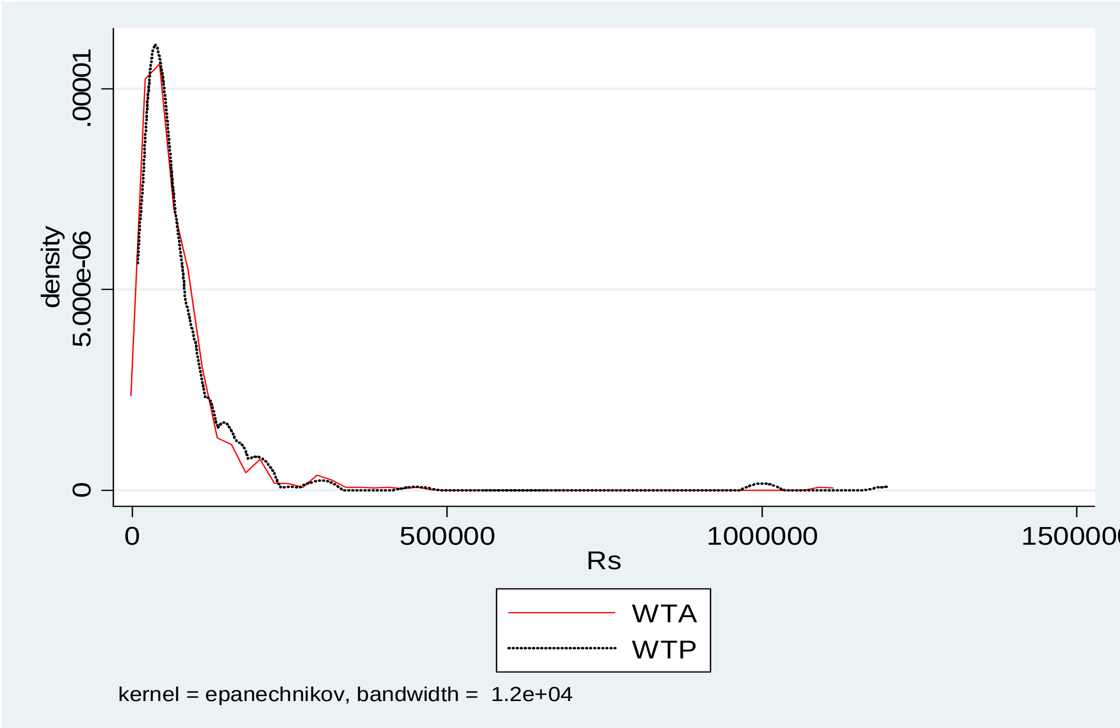

Figure 1: The kernel density of projected dowry for boys and girls

Note: Excludes children with zero dowry and the children for whom the parent could not answer the dowry question

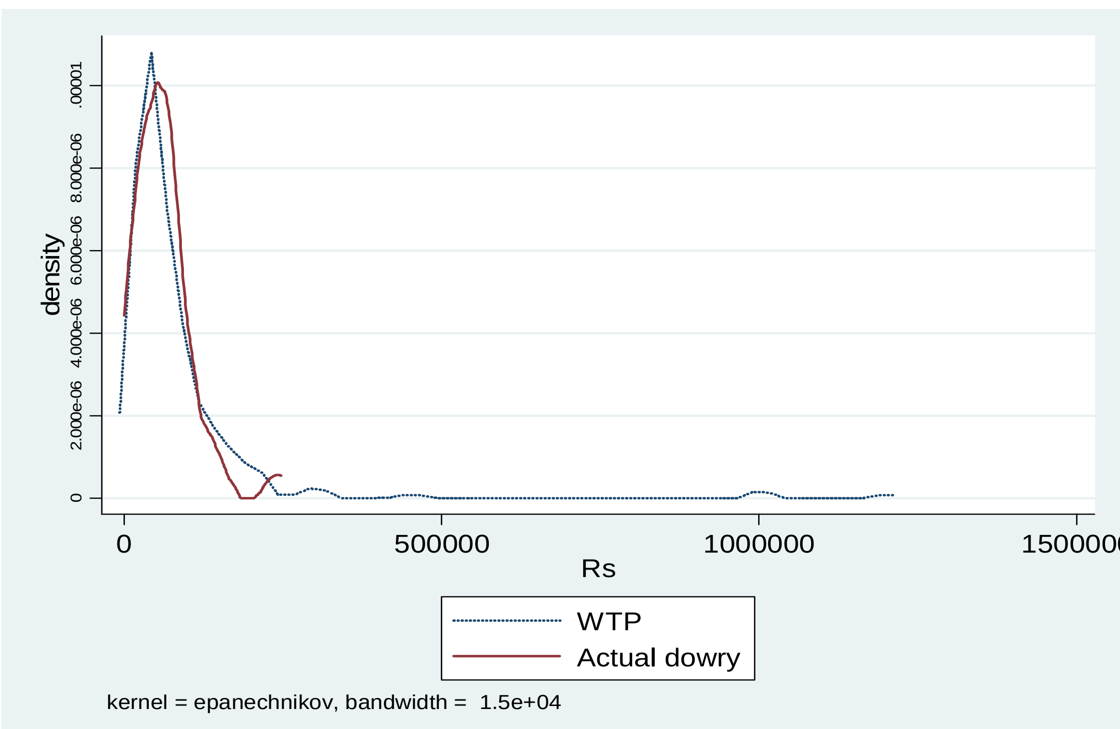

Figure 2: The kernel densities of projected dowries for girls as well as the kernel density of actual dowries from the retrospective data

Note: Excludes children with zero dowry and the children for whom the parent could not answer the dowry question

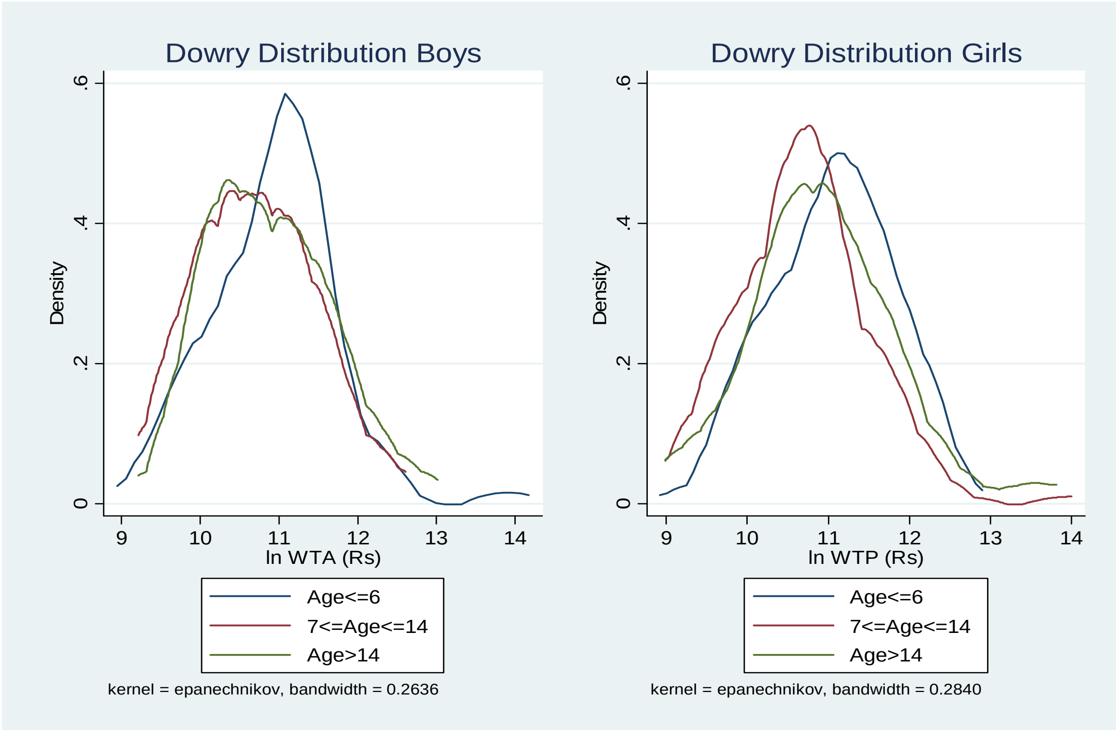

Figure 3: The kernel density of the logarithm of projected dowry for each of the three indicated age categories. The figure on the left panel refers to the results for boys and the panel on the right contains the results for girls

Note: Excludes children with zero dowry and the children for whom the parent could not answer the dowry question

Table 2: Correlates of ln(WTP) and ln(WTA)

	Girls		Boys	
	1	2	3	4
Expected returns to higher education	-0.15	-0.16	0.19	0.10
(in thousands of Rs)	(0.12)	(0.14)	(0.09)**	(0.10)
Expected returns to lower education	1.25	1.20	-0.18	-0.11
(in thousands of Rs)	(0.31)***	(0.33)***	(0.44)	(0.41)
Logarithm of household income (Rs)	0.19	0.20	0.03	-0.04
	(0.07)***	(0.08)**	(0.06)	(0.06)
Logarithm of household wealth (Rs)	0.10	0.09	0.19	0.16
	(0.05)**	(0.05)*	(0.04)***	(0.03)***
Education level of decision-maker (in years)	0.02	0.01	0.03	0.02
	(0.01)*	(0.01)	(0.01)***	(0.01)***
Number of unmarried boys	-0.06	-0.04	-0.16	-0.18
	(0.07)	(0.07)	(0.07)**	(0.06)***
Number of unmarried girls	-0.10	-0.08	0.06	0.04
	(0.06)*	(0.06)	(0.04)	(0.04)
Number of adults	0.09	0.09	0.10	0.12
	(0.10)	(0.10)	(0.08)	(0.08)
Within-gender child order	-0.09	-0.10	-0.13	-0.14
	(0.06)	(0.06)	(0.05)***	(0.05)***
Overall child-order	0.06	0.05	0.07	0.10
	(0.04)	(0.04)	(0.04)**	(0.04)***
Minimum education desired		-0.01		0.09
(in years)		(0.04)		(0.03)***
Maximum education desired		0.05		0.03
(in years)		(0.03)**		(0.03)
Observations	142	138	155	149
R-squared	0.71	0.73	0.75	0.79

Notes: *** p<0.01, ** p<0.05, * p<0.1; Standard errors have been clustered at household level. The dependent variable in columns 1 and 2 is the log of the maximum dowry that the girl's family is willing to pay. The dependent variable in columns 3 and 4 is the log of the minimum dowry that the boy's family is willing to accept. All regressions include caste and village dummies (coefficients not reported). Adults include the married boys and girls in the household. Household wealth includes durables, land, animals, stock, equipment, savings and net-lendings. Expected returns to higher education computed as the average return to higher education. Expected return to lower education computed as the average return to highschool and below.

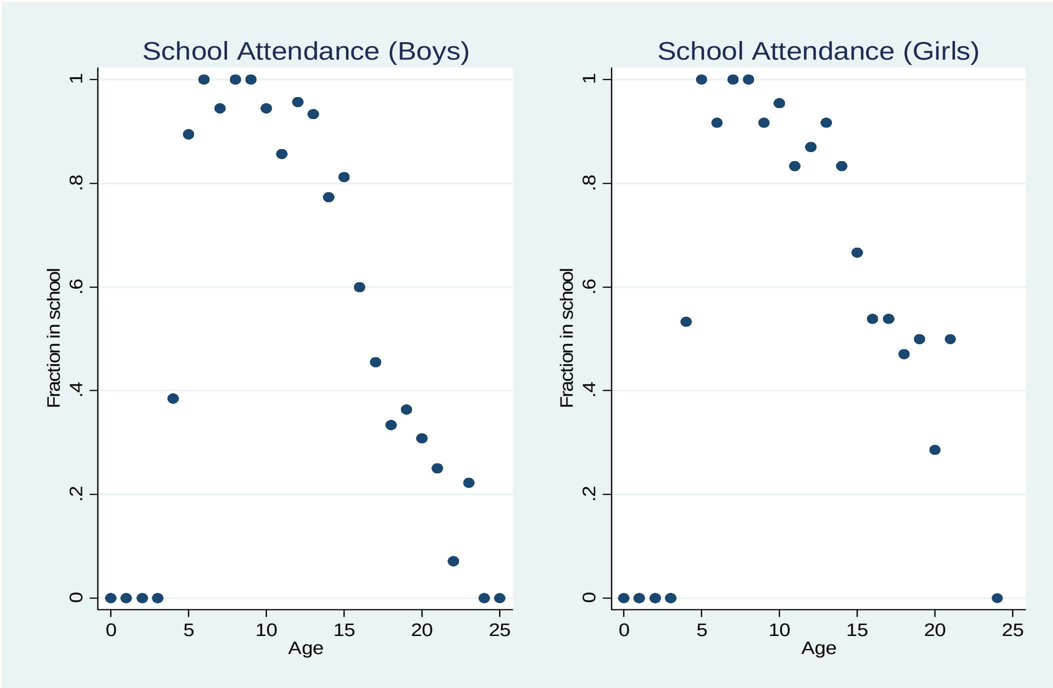

Figure 4: The fraction of unmarried children currently attending school at each age

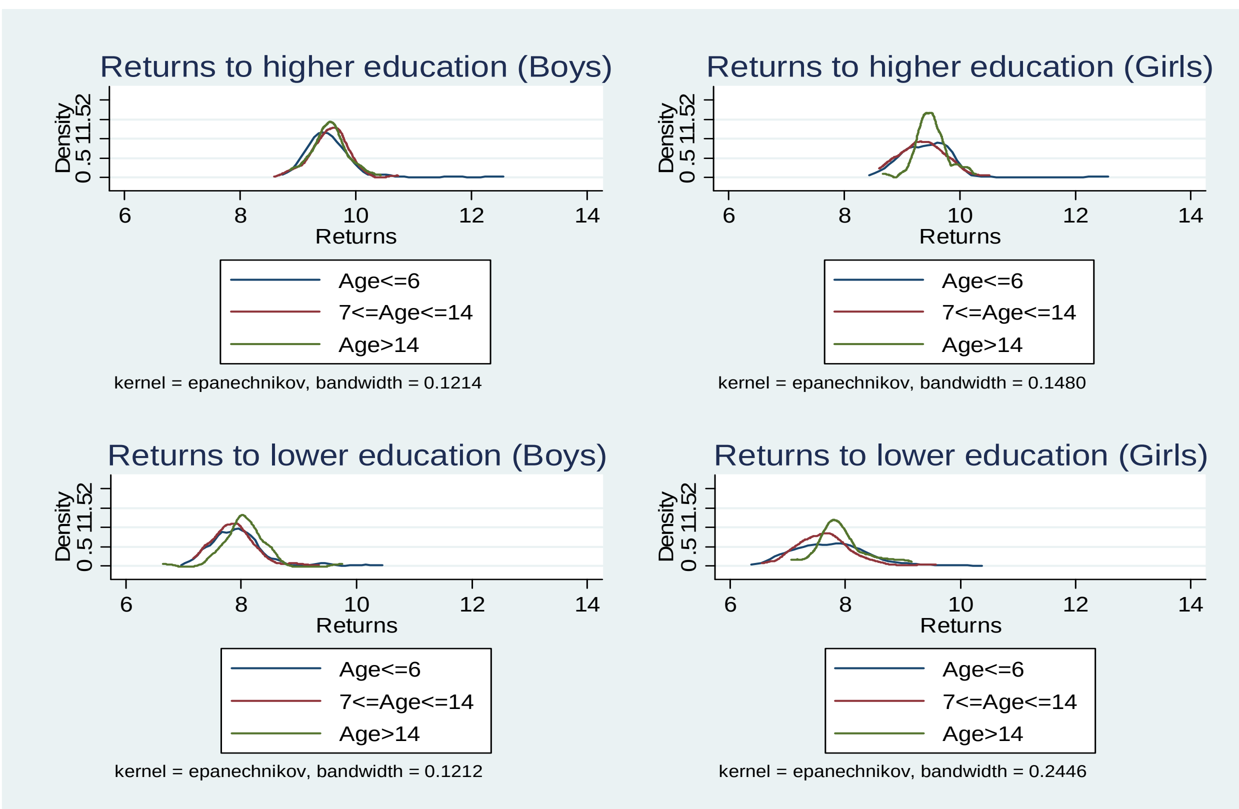

Figure 5: The distribution of expected returns to higher and lower education change for each of the indicated age categories (for boys and girls separately). The upper two figures refer to returns to higher education and the lower two figures refer to returns to lower education

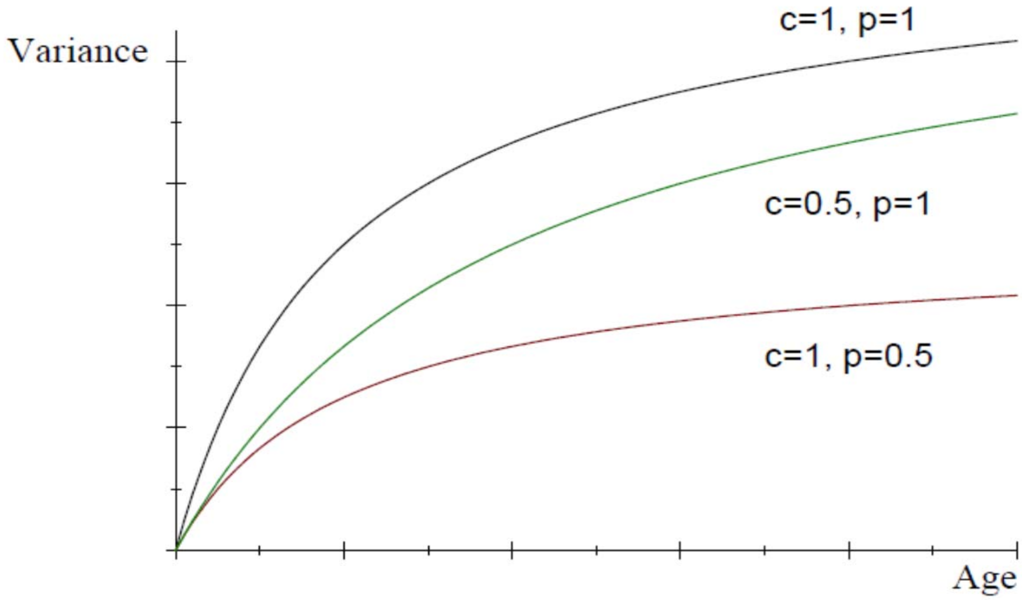

Figure 6: Dowry variance as a function of age for various parameter values of the price of quality and the learning parameter, c.

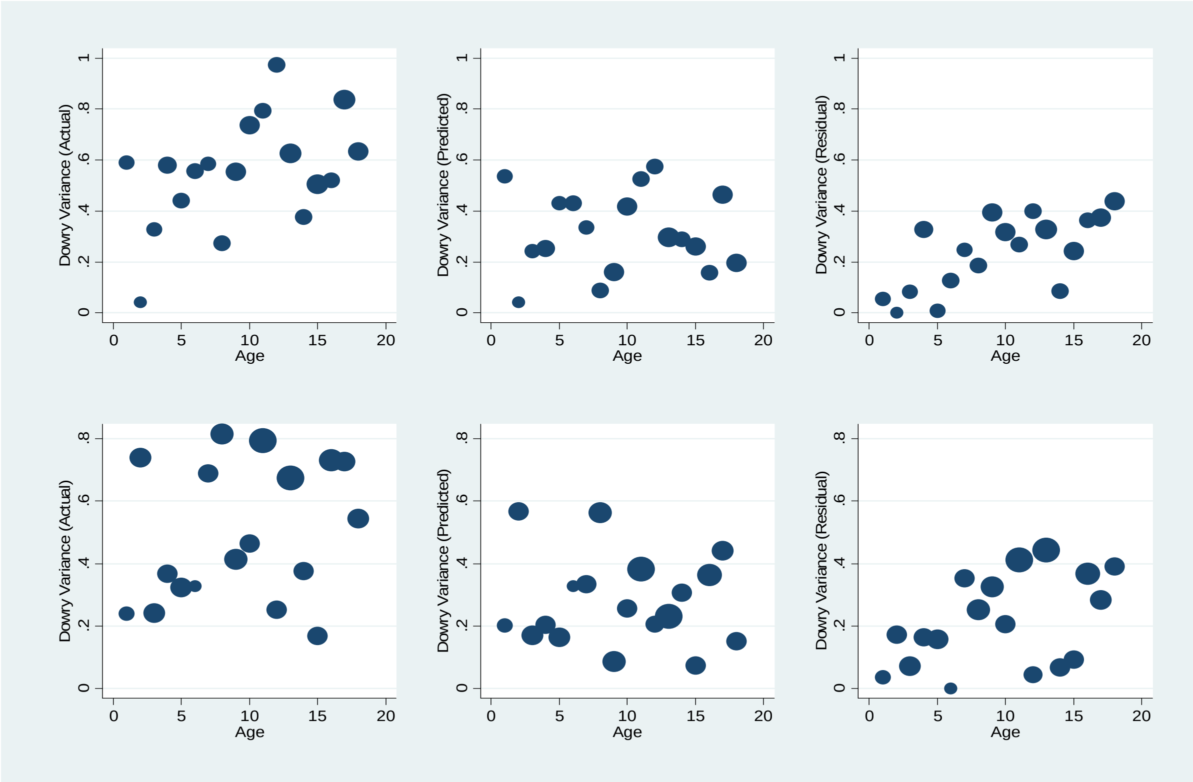

Figure 7: The age-wise variance of actual, predicted and residual dowry, where the predicted and residual values are obtained from a regression of projected dowries on household variables. The top panel contains the results of the exercise for boys and the bottom panel contains the results of the exercise for girls

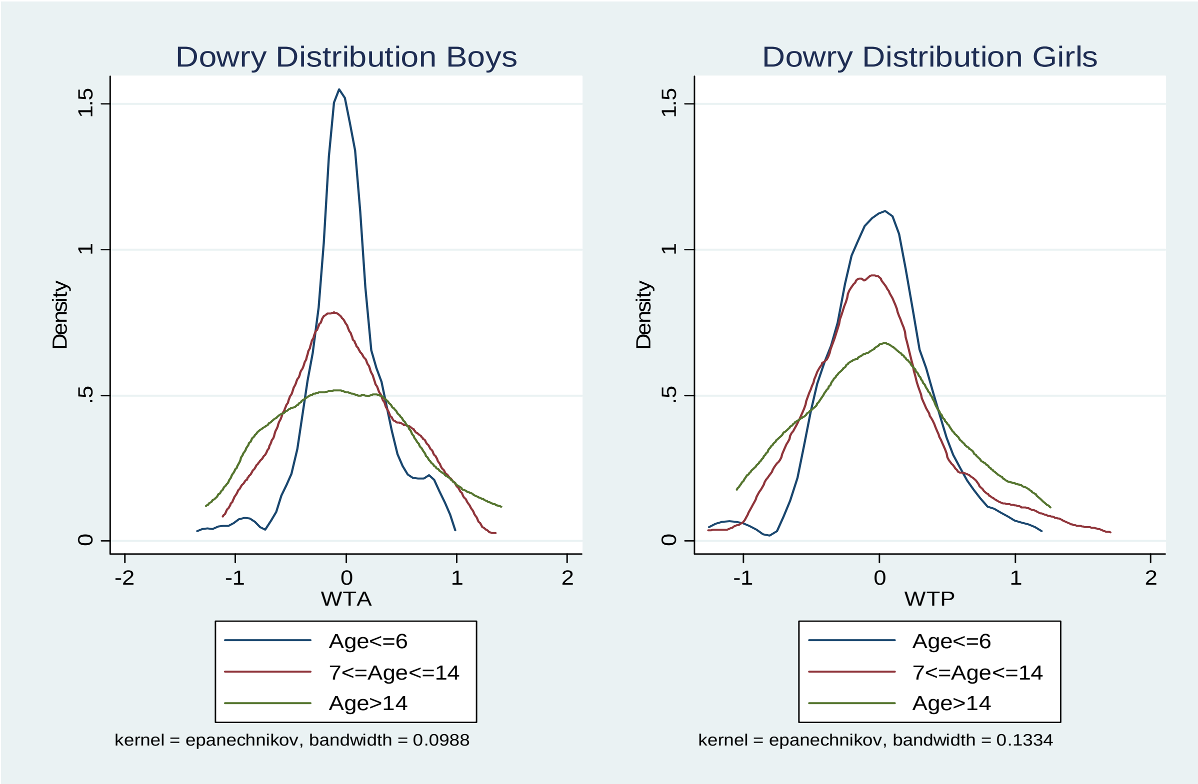

Figure 8: The kernel density of residual dowry for each of the three indicated age categories. The residual dowry was obtained from a regression of the logarithm of projected dowry on a set of household variables. The figure on the left panel refers to the results for boys and the panel on the right contains the results for girls

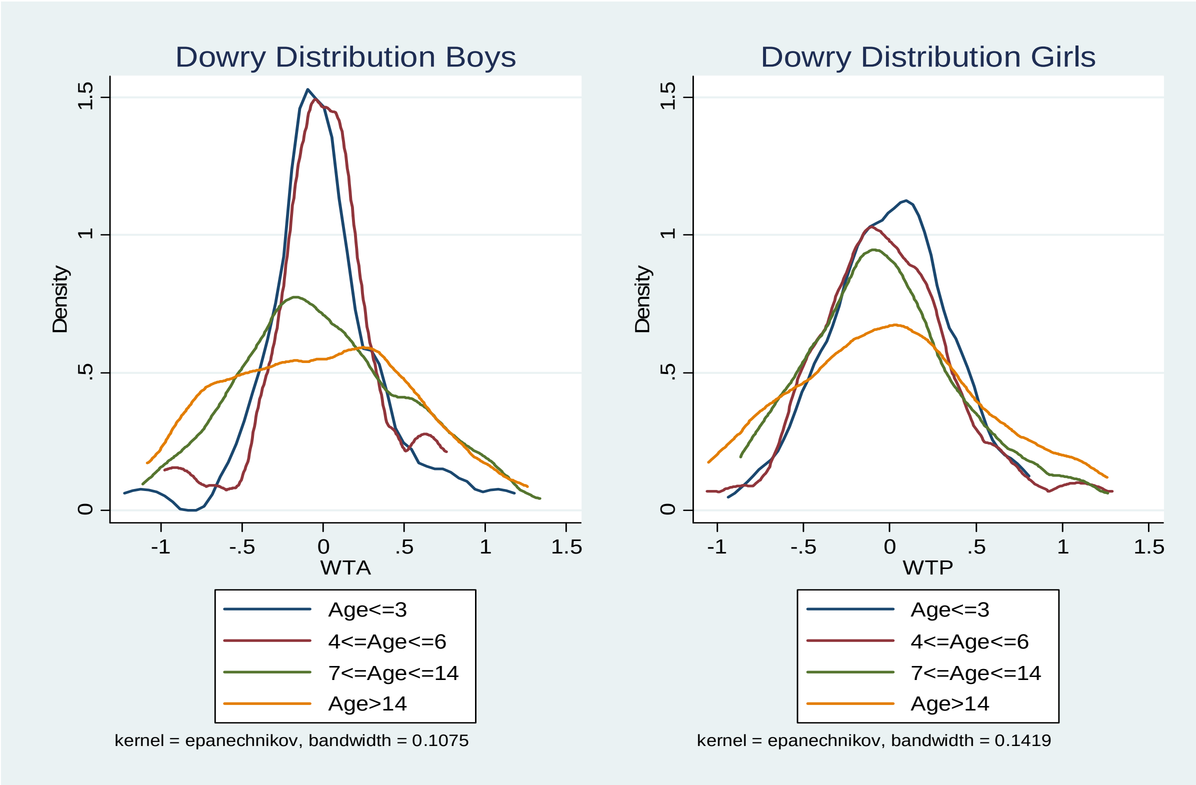

Figure 9: The kernel density of residual dowry for each of the four indicated age categories. The residual dowry was obtained from a regression of the logarithm of projected dowry on a set of household variables. The figure on the left panel refers to the results for boys and the panel on the right contains the results for girls

Table 3: Estimates from the structural model

	Base Specifications				Robustness			
	1	2	3	4	5	6	7	8
Price (girls)	0.67	0.79	0.67	0.61	0.61	0.60	0.71	0.80
	(0.17)***	(0.14)***	(0.17)***	(0.04)***	(0.04)***	(0.04)***	(0.04)***	(0.04)***
Price (boys)	0.83	0.79	0.60	0.61	0.61	0.60	0.71	0.80
	(0.18)***	(0.14)***	(0.04)***	(0.04)***	(0.04)***	(0.04)***	(0.04)***	(0.04)***
c (girls)	0.08	0.05	0.08	0.13	0.13	0.12	0.24	0.58
	(0.08)	(0.03)*	(0.08)	(0.05)**	(0.05)**	(0.05)**	(0.12)*	(0.41)
c (boys)	0.06	0.07	0.77	0.71	0.73	0.84	1.36	4.53
	(0.05)	(0.05)	(0.43)*	(0.37)*	(0.40)*	(0.47)*	(0.94)	(5.50)
Price(boys)-Price(girls)	0.16		-0.07					
	(0.26)		(0.18)					
c(boys)-c(girls)	-0.02	0.02	0.69	0.58	0.60	0.71	1.12	3.95
	(0.09)	(0.02)	(0.43)	(0.34)*	(0.37)	(0.45)	(0.88)	(5.31)
R-sq	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.89
Obs	38	38	33	33	32	31	33	33

Notes: Robust standard errors are reported in parentheses. The estimated parameters were derived from a non-linear least-squares regression. In Columns 1 and 2, the dependent variable is the age-specific variance in the residual dowry, the latter being obtained from a regression of dowry on a set of household variables interacted with age. In Columns 3 and 4, we set age 6 to be the starting year for boys. In Column 5, we restrict the sample to children up to the age of 17. In Column 6 we restrict the sample to children under the age of 16. In Column 7 the dependent variable is the variance in residual dowry, the latter obtained from a regression of dowry on the full set of household variables in levels. In Column 8, the dependent variable is the variance in the raw dowry.

Appendix Table A1: Correlates of ln(WTP) and ln(WTA)

	Girls		Boys	
	1	2	3	4
Expected returns to higher education	-0.02	-0.03	0.17*	0.10
(in thousands of Rs)	(0.12)	(0.12)	(0.09)	(0.09)
Expected returns to lower education	0.64**	0.55*	-0.09	-0.11
(in thousands of Rs)	(0.34)	(0.34)	(0.38)	(0.39)
Logarithm of household income (Rs)	0.24*	0.23*	0.08	0.02
	(0.08)	(0.09)	(0.07)	(0.07)
Logarithm of household wealth (Rs)	0.11	0.09	0.17***	0.13***
	(0.06)	(0.06)	(0.04)	(0.04)
Education level of decision-maker (in years)	0.03***	0.02**	0.04***	0.03***
	(0.01)	(0.01)	(0.01)	(0.01)
Number of unmarried boys	0.04	0.04	-0.10*	-0.09*
	(0.07)	(0.07)	(0.06)	(0.05)
Number of unmarried girls	-0.07	-0.05	0.04	0.02
	(0.07)	0.07	(0.05)	(0.04)
Number of adults	0.00	0.01	-0.01	0.02
	(0.04)	0.04	(0.04)	(0.03)
Within-gender child order	0.10	0.09	0.08	0.07
	(0.09)	0.09	(0.06)	(0.06)
Overall child-order	-0.10**	-0.11*	-0.09*	-0.08*
	(0.06)	0.06	(0.05)	(0.05)
Minimum education desired		0.04		0.07**
(in years)		0.04		(0.03)
Maximum education desired		0.02		0.04
(in years)		0.03		(0.02)
Observations	190	185	202	195
R-squared	0.67	0.69	0.77	0.80

Notes: *** p<0.01, ** p<0.05, * p<0.1; Standard errors have been clustered at household level. The dependent variable in columns 1 and 2 is the log of the maximum dowry that the girl's family is willing to pay. The dependent variable in columns 3 and 4 is the log of the minimum dowry that the boy's family is willing to accept. All regressions include caste and village dummies (coefficients not reported). Adults include the married boys and girls in the household. Household wealth includes durables, land, animals, stock, equipment, savings and net-lendings. Expected returns to higher education computed as the average return to higher education. Expected return to lower education computed as the average return to highschool and below. These average returns haven been computed for all children for whom the parents were able to guess returns for at least one level of schooling attainment.