

Love and Money by Parental Match-Making: Evidence from Chinese Couples*

Fali Huang
School of Economics
Singapore Management University
90 Stamford Road, Singapore 178903
fhuang@smu.edu.sg

Ginger Jin
University of Maryland & NBER
Department of Economics
College Park, MD 20742
jin@econ.umd.edu

Lixin Colin Xu
World Bank
1818 H Street, N.W.
Washington, D.C., 20433
lxu1@worldbank.org

July 19, 2010

Abstract

How match-making methods affect marriage outcomes is an understudied topic. Using a survey of Chinese couples in 1991, this paper examines whether parental involvement in match-making affects a couple's marriage harmony and joint income. We find that, compared with those married through self search, couples relying on parental help in match-making have on average less harmony and less income in rural areas. This finding holds after we use the area-specific tradition of parental match-making as an instrumental variable (IV) for parental involvement. In comparison, urban areas show a similar negative effect of parental match-making on marriage harmony, but the effect on income is slightly positive in OLS and becomes more significant with the IV. These results are consistent with our theoretical prediction about the agency cost of using parents as matchmakers in the marriage market.

JEL: J12, D82, D83.

Keywords: Marriage Outcome, Match-Making Method, Parental Involvement, China, Agency Cost.

*We wish to thank William Evans, Vijayendra Rao, Seth Sanders, Mary Shirley, Jeffrey Smith, Yi Lu, Yuyu Chen and the participants at U. of Maryland workshop and the 2003 PAA conference for constructive comment and suggestions. Huang gratefully acknowledges the financial support of SMU Research Grant 10-C244-SMU-002. The views expressed here do not implicate the World Bank and the countries that it represents.

1 Introduction

Marriage formation is typically modeled in the literature as a matching process similar as labor market matching between workers and firms, where males and females meet each other randomly or assisted by commercial agents.¹ This modeling approach, though working quite well in modern western society, has not addressed a unique feature of marriage matching, i.e., marriage is not a simple matter of two individuals forming a new family, it also connects the existing families of their parents and perhaps of relatives. Such externality has been very important in history and is still so in today's developing countries, where parents play a significant role in the marriage of children.²

In this paper we argue that parental involvement in match-making may have a distorting effect on marriage outcomes. This goes beyond the typical principal-agent problem in matching such as that between a house seller and a real estate agent, because parents that help match-making in the marriage formation process also have a long-term relationship with the couple after matching is done; such future interactions may distort incentives in the matching process and thus affect matching quality. For example, parents who expect to receive financial support from their son after his marriage may care less about how attractive his wife is to him but more about how able she is in contributing to family incomes, and as a result, the best wife candidate in the parents' eye can differ from what is optimal to the son, even though parents are altruistic and care dearly about the son's welfare.

This idea is formally analyzed in a simple model where the son chooses whether to search a potential wife himself or to delegate his parents to conduct the search for him.³ Relying on parental involvement in match-making has two types of effects on his welfare. The negative effect arises from the agency cost due to the above-mentioned conflict of interests between parents and the son. The positive effect comes from parents' expertise, since parents with the help of relatives and friends may get access to a larger pool of marriageable singles than the son, and they may have better capabilities of evaluating the quality of marriage candidates. And in many cases, parents often pay the searching cost themselves if they do the search, as searching involves not only monetary expenditure, but also time and effort as well as social capital that cannot be easily compensated by the son even if he is willing to. In this sense, parental involvement saves cost for the son compared with his own search. So the net effect of parental involvement on the son's welfare is not obvious and may vary across time and areas.

This framework is tested using the data of more than 8000 Chinese couples surveyed in 1991 across 7 provinces. In this dataset, about 58% in the rural area and 19% in the urban area were married by parental involvement, and the rest by self match. To measure marriage outcomes, we focus on two variables – the degree of domestic harmony and the joint income of husband and wife. The former measures the extent to which a couple has conflicts in their marital life, while the latter measures a couple's economic wellbeing.

The comparison across the two groups of match-making shows that couples with parents as matchmakers have less harmonious marital life and lower income in the rural area, while the pattern in the urban area is the same in terms of marital harmony but the opposite in joint income. This result is further confirmed in regressions that feature many control variables and the treatment of the endogeneity problem of parental involvement.

A challenge in estimating the effect of parental involvement on marriage outcomes is the endogenous choice

¹ See the survey by Weiss (1997) for the related literature.

² See, for example, Davidson and Ekelund (1997) and Anderson (2003) for the changing marriage formation in European history and India and Cheung (1972) in pre-modern China.

³ The model applies also to women searching for husbands; the man's case is used only to simplify the exposition.

of match-making means. First, those with lower human capital or higher individual-specific search costs are more likely to delegate the search process to their parents (“adverse selection”). Second, individuals with more competent parents or whose parents have larger stakes in the marriage outputs are more likely to use parental involvement (“positive selection”). Since the factors that determine the choice between parental involvement and self search also affect life quality after marriage, if some of them are unobservable to researchers, the estimated effect of parental intervention on marriage outcomes to be biased, the direction of which can be either way, depending upon whether the adverse selection of individuals or the positive selection of parents is stronger. For example, if only those who have difficulty in finding a partner by themselves rely on parental help, their average joint couple incomes may be lower even though parental involvement has a positive effect by itself.

We use an instrumental variable approach to address this issue. The basic idea is as follows. Suppose parents are more likely to help match-making where the tradition of doing so is stronger. In the theory section, we show that a tradition of parental match-making reflects an advantage of search cost applicable to every parent independent of individual characteristics. Thus, couples in an area with a strong parental match-making tradition are more likely to rely on parental help than others even though they are identical otherwise. We use the percentage of parent-involved matches in the previous cohort in an area to measure the tradition of parental involvement. After controlling for area characteristics, this measure is unlikely to have any direct effect on an individual couple’s marriage outcomes, and thus satisfies the exclusion criterion of a valid instrumental variable for the endogenous choice of match-making method.

Overall, we find that, compared with couples married by self search, the emotional dimension of marriage outcome is lower for those with parental involvement, but their joint incomes may be higher or lower depending on how parents and children differ in preferences and search costs. The different effects on the emotional and monetary dimensions of marriage outcomes are consistent with the notion that love between husband and wife is consumed privately within the couple, thus it generates much less positive externality on others than their incomes, which can be shared or transferred among members in the extended family.

This paper contributes to the marriage literature by generating new insights on how match-making methods may affect marriage outcomes. Unlike the classical focus on the effects of sex ratio (Angrist 2002), divorce law (Chiappori 2002), or educational composition on marriage outcomes, we highlight the institutional details of how the match is accomplished in the first place. In particular, our theory suggests that parental involvement in match-making may distort the optimal spouse choice in that they tend to emphasize more on the potential spouse’s earning abilities than the match quality or attraction between the couple.

In a related paper, Edlund and Lagerlof (2006) use a simple model to show that the shift from parental to individual consent in marriage allows the son instead of his parents to receive the bride price and thus facilitates economic growth. In comparison, our focus is not on who controls the resources in marriage, but on the agency cost of relying parents as matchmakers, and the two match-making methods in our paper are both under the individual consent. The trade-off between love and money is also explored by Fernandez et al. (2005) but from a different perspective of marriage sorting on skills and its relationship with income inequality; they do not discuss different match-making methods at all, though it seems likely that the prevalence of parental involvement should be positively associated with income inequality.

The roles of dowries and bride prices in the traditional marriage formation process have been examined extensively in the literature⁴. Such financial transfers between families of bride and groom, though not explicitly analyzed in this paper, are implicitly embedded in the searching cost of our model. When they are

⁴See, for example, Becker (1973, 1981), Zhang and Chan (1999), Botticini and Siow (2003), Suen et al. (2003) and Anderson (2007).

large and therefore have to be paid by parents due to credit constraints faced by young people, parents are more likely to take charge of the match-making process even when its effects are negative on the marriage outcomes, which seems to be indeed the case in the rural area in our data.

The rest of the paper is organized as follows. Section 2 analyzes a simple theoretical model on the choice of match-making methods and develops key empirical implications. Section 3 describes the data and presents some preliminary evidence consistent with the model. The main empirical results are shown in Section 4, and some conclusive remarks are offered in Section 5.

2 A Theoretical Model: Effects of Search Methods on Marriage Outcomes

To guide the empirical exercise, we construct a simple theoretical model on the choice of marriage searching methods. The model focuses on the marital decision of a young man, who has finished schooling and started working to earn a living. The search process for a potential wife can either be managed directly by the young man's parents or by himself, and the process that yields a higher net expected utility to him will be implemented. This set up is meant to capture the current practice in China, where marriage cannot be forced upon by parents, and males are usually the one who initiates and proposes a deal for marriage. It is useful to note that the same model can be used to study the search process of a young woman, and it can also be readily extended to the case where parents arrange the marriage without consent of children.

Marriage Output. An individual's benefit from marriage can be categorized into two dimensions; one is the economic gain from joint household production, and the other is the emotional support. The total benefit is affected not only by the human capital of husband and wife, but also by their matching quality.

Let $h_m \geq 0$ denote the young man's human capital level, which presumably affects his earning and also his intra-household productivity, and thus may include, for example, his innate ability, character, years of schooling, communication skills, etc. Similarly, let $h_f \geq 0$ denote his potential wife's human capital level. h_m and h_f jointly determine the total household production $f(h_f, h_m)$, which reflects both the couple's joint income earned from markets and home production. We assume $f(0, 0) > 0$, $f_i > 0$, $f_{ij} > 0$, and $f_{ii} \leq 0$ for $i, j \in \{1, 2\}$.

Another important element in marriage is the match quality, denoted by α , which is idiosyncratic to the specific couple and is thus not readily observed by others; it can be interpreted as love or attraction between two persons, which is often unpredictable based on commonly observed characteristics; this implies that α can be treated as uncorrelated with h_f . Given our assumption that marriage is always implemented with mutual consent by the young couple, it implies that the emotional output of marriage is positive and thus $\alpha > 0$ is assumed.⁵

For a young man with h_m , the overall gain from marrying a wife with h_f and α is $(\beta + \alpha)f(h_f, h_m)$, where $\beta > 0$. One may think of the parameter β as capturing the husband's share of material gain from the marriage, while α captures the degree of emotional benefit. The parents' gain from their son being married to a wife with characteristics (α, h_f) also contains two parts: one is the public good component $\gamma f(h_f, h_m)$ received by them, and the other is the altruistic component $\delta(\beta + \alpha)f(h_f, h_m)$ because they care about the welfare of their son, where $\gamma > 0$ and $\delta \in (0, 1)$. Note that the love α between the husband and wife is by definition consumed privately by the couple themselves, and is thus not affecting the parents' welfare directly. The wife's characteristics that may affect the whole family, such as pleasant personality and beauty,

⁵This assumption is for simplicity only, as the same results can be derived for the case with $\alpha \leq 0$.

are already indicated by the wife's human capital h_f , which as mentioned earlier is broadly defined and not restricted to formal schooling.

Searching Costs. Marital search is costly. If the son conducts the search himself, he has to bear the search cost, which is $\eta_m c(\alpha, h_f, h_m) > 0$, where $\eta_m, c_1, c_2 > 0$ and $c_3, c_{31}, c_{32} < 0$. This means that it is more costly for a man with a given h_m to find and persuade a better quality woman (with higher α or h_f) to become his wife, and the search cost for a given quality of wife is lower if the man's h_m is higher. The parameter η_m indicates the effect of some common elements on the cost of searching by oneself for all individuals in a marriage market and is thus not dependent on idiosyncratic conditions.

If the parents are dealing with the search process, they will bear the search cost, which presumably depends on how intelligent they are in assessing α and how well they are connected with relevant social networks that have access to potential candidates. The parents' degree of competence in this matter is denoted by $h_p \geq 0$, and the parental search cost is $\eta_p s(\alpha, h_f, h_p) > 0$, where $\eta_p, s_1, s_2 > 0$ and $s_3, s_{31}, s_{32} < 0$. Similar as η_m , the parameter η_p also denotes some common factor that affects the cost of searching by all parents. To capture the idea that the match quality α is couple idiosyncratic, we assume that, in order to achieve the same level of α , the parents' search cost cannot be too low compared with the direct search by their son, i.e., $\eta_p s_1 \geq \delta \eta_m c_1$ for any given h_m, h_f , and h_p .

The Optimal Choice of Search Methods. The son decides whether to search for his marriage partner himself or delegate the search to his parents. If he searches himself, his objective function is

$$U^* \equiv \max_{\alpha, h_f} (\beta + \alpha) f(h_f, h_m) - \eta_m c(\alpha, h_f, h_m).$$

The corresponding optimal choices of his potential wife's characteristics resulted from searching by himself are denoted by α^* and h_f^* , which are characterized by the following first order conditions

$$f(h_f^*, h_m) - \eta_m c_1(\alpha^*, h_f^*, h_m) = 0, \tag{1}$$

$$(\beta + \alpha^*) f_1(h_f^*, h_m) - \eta_m c_2(\alpha^*, h_f^*, h_m) = 0. \tag{2}$$

Proposition 1 (1) Effects of Parental Involvement: The emotional output and the overall match quality are lower under parental involvement, i.e., $\alpha^* f(h_f^*, h_m) > \alpha^{**} f(h_f^{**}, h_m)$ and $(\beta + \alpha^*) f(h_f^*, h_m) \geq (\beta + \alpha^{**}) f(h_f^{**}, h_m)$ hold, respectively, but it is possible that the couple's joint household production is higher, i.e., $f(h_f^*, h_m) \leq f(h_f^{**}, h_m)$ may be true.

(2) Adverse Selection of Individuals: There exists a unique threshold value $h_m^\#$ of the son's human capital level such that he will choose to search marriage partner himself if $h_m \geq h_m^\#$, while he will delegate his parents to do the search for him if $h_m < h_m^\#$, where $h_m^\#$ increases with h_p , γ and η_m .

(3) Positive Selection of Parents: There exists a unique threshold value $h_p^\#$ of the parents' competence level such that they will be delegated to do the search if $h_p > h_p^\#$, where $h_p^\#$ increases with h_m but decreases with γ and η_m .

This proposition suggests that the effects of parental involvement in marriage search can be different for the two dimensions of marriage output: It is always negative for the emotional output, which is driven by the fact that the matching quality α is idiosyncratic to the couple and thus not easily observed or shared by others; the effect on the economic output, however, can be either negative or positive. The reason for a positive effect is because the household output can be shared among family members and thus parents have more incentives to care about the potential wife's human capital. On the contrary, parental involvement could have a negative effect on the economic output and is still an optimal choice from the son's perspective if parental matchmaking leads to substantial savings in search cost.

Proposition 1 also suggests that parental involvement in marital search is endogenous; it is more likely to happen when the son's human capital level h_m is lower or the searching cost η_m is higher, and when his parents benefit more from the household public good (when γ is higher) and have lower searching costs (when h_p is higher and η_p is lower). In other words, in a mixed marriage market, there are two sources of self-selection in the choice of marital search methods: one is from the son and the other is from the parents; a young man with a lower human capital and more capable or more motivated parents is more likely to rely on his parents to search for wife.

Figure 1 illustrates the positive relationship between $h_m^\#$ and h_p and how their combination affects the endogenous choice of marital searching methods. In the graph, a young man with human capital h_m' and parents' effectiveness h_p' , for example, will optimally choose to rely on his parents to search for a potential wife because his human capital is lower than the threshold level $h_m^\#$ corresponding to his parents' effectiveness h_p' . This choice can also be understood from the alternative perspective: given his human capital level h_m' , his parents are competent enough (since h_p' is higher than the corresponding threshold level $h_p^\#$) to find a good wife for him so that he does not need to bother searching by himself.

The general equilibrium effect of the marriage market matching, though not directly considered, is implicitly taken into account in the model. Specifically, the search cost may reflect such effect. For example, the change in the common cost parameters η_m and η_p should not affect an individual's optimal choices of wife quality given his search method; the reason is that, since the marriage market (i.e., the pool of eligible wives) is mixed, the same change in search cost for everyone does not affect any individual man's relative ranking of desirability in the eye of eligible wives (in other words, the feasible set of his optimization problem is still the same as before), and thus should not affect an individual's optimal choice of wife.

Empirical Implications. If individuals make different choices of match-making methods randomly, then the OLS regressions should yield unbiased estimates for the negative effects of parental involvement on the emotional output and for ambiguous effects on joint incomes, after controlling for individual characteristics as well as relevant marriage market conditions that may also directly affect marital outcomes. The potential

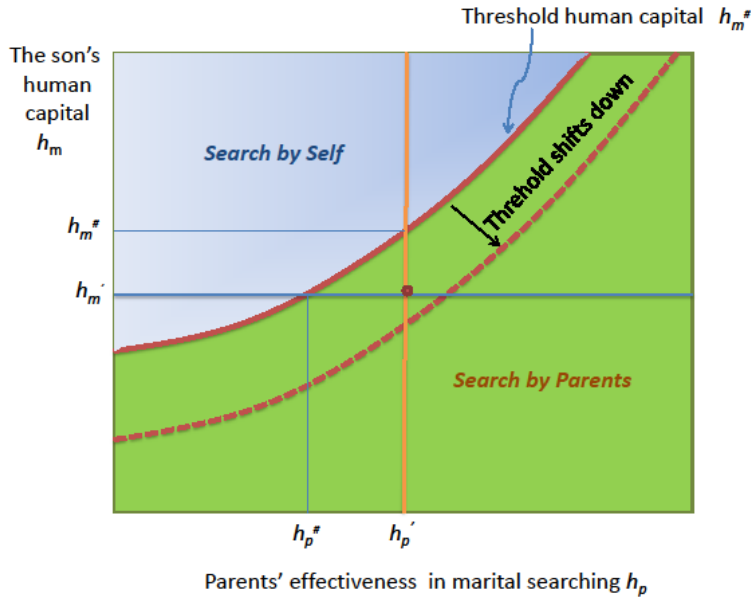


Figure 1: Endogenous Choice of Match-Making Methods

problem, however, is that the choice of search methods may not only be affected by random elements, but also by the individual's and his parents' characteristics as reflected by the adverse and positive selection problems in Proposition 1.

If we can perfectly control parents' characteristics (h_p, γ) , then the average marital quality of husbands with parental involvement must be lower than others even when their wives are of exactly the same quality because the husbands in the former group have lower human capital ($h_m < h_m^\#$); this is the adverse selection effect of sons. In contrast, when the husband's characteristics are fully controlled, those with parental involvement must have had more competent parents ($h_p > h_p^\#$) with respect to searching, which implies that their wife's overall quality, especially their human capital level h_m^{**} , may be higher than others', and hence their marital quality may also be higher; this is the positive selection effect of parents. And so without properly accounting for these two sources of the endogeneity problem, the OLS estimate of the effect of parental involvement in match-making can be either higher or lower than the true effect, depending on which effect is dominant.

The main challenge in the empirical strategy is how to tease out the endogenous selection effects from the true effect of parental involvement on marital quality. Our approach is to use an instrumental variable that affects the choice of search method but not the marital outcomes directly. Consider two identical marriage markets A and B that are mutually exclusive. Due to some exogenous shocks, the threshold level of the son's human capital $h_m^\#$ as a function of parents' characteristics h_p shifts down in market A but not in B. This can be achieved in the model, for example, by a lower η_m , which affects the search costs of all individuals in a marriage market. As one can see in Figure 1, this downward shift in market A will induce a group of young men, who are between the new and old threshold curves, to change their search method from parental involvement to self-search; as a result, identical individuals make different choices: those in market A have

parental involvement while those in market B adopt self-search. Comparing their difference in marital quality will thus lead to the true effect of parental involvement on marital quality. This is the main prediction of the model, which will be tested using data.

3 Data Description and Preliminary Evidence

We use the Study of the Status of Contemporary Chinese Women (SSCCW), a data set collected jointly by the Population Institute of Chinese Academy of Social Science and the Population Council of United Nations in 1991 (Institute of Population Studies, 1993). SSCCW collects information on personal traits, marriage characteristics, fertility, work, intra-family arrangements, and gender norms. Husband and wives were interviewed separately for the same questions.

The survey used stratified random sampling to select households from seven areas: one municipality (Shanghai) and 6 provinces (Guandong, Sichuan, Jilin, Shandong, Shanxi, and Ninxia). They scatter across China in the southeast, south, southwest, northeast, east, middle and north, respectively, and are often far away from each other given the size of China. As migration across different provinces was not common in China by 1991, each province can be regarded as a separate marriage market. Another important dimension that cuts across areas is the urban-rural distinction in that the rigid Hukou system effectively blocked people from migrating between cities and countryside. Furthermore, the economic structure and life styles are dramatically different between urban and rural areas, which will in many important ways affect both match-making methods and marriage outcomes.

3.1 Measures of Matching-Making Methods and Marital Outcomes

The question on match-making methods asked how an individual met his or her spouse initially. There are four original categories in the data, where 35.2% of the sample were introduced by parents or relatives, 36.6% by friends, 27.3% by themselves, and 0.8% by other means. We treat the first two categories as searching by parents for two reasons; one is technical, as the further distinction between introduction by parents and by relatives is not available in the data, and the other is conceptual, because relatives are an integrated part of the parents' social networks to facilitate the search process. A perhaps more debatable decision is that we do not differentiate couples initially introduced by friends from those who met by themselves. The main reason is that these two groups are similar in terms of our theoretical model: In both cases, it is the young people themselves, not their parents, that conducted the search process and bore the search cost; whether explicitly introduced by friends or not at the initial stage is thus not essential to our purpose. And indeed, empirically these two groups are very similar.⁷ And finally, since our main focus is on the effect of parental involvement versus self search in the match-making process, we drop the few couples introduced through other means. We then construct a dummy variable Parental Involvement that equals 1 if a couple were introduced by parents or relatives and 0 if by friends or themselves.

Marriage outcomes are represented by money and love as in the model. The economic output is measured by the joint couple income; the emotional aspect of the marriage output, however, is difficult to accurately quantify. In the following, we discuss in detail an indicator of harmony within a couple and argue it is a plausible proxy for the love dimension of marriage.

⁷One may conjecture that the human capital of couples introduced by friends are lower than that by self-match; if this is the case, then treating them together should introduce downward biases to the scale of estimated effects of parental involvement and thus make it less likely to find any significant results.

The survey question most closely related to the emotional aspect of marriage asked: "How do you usually reconcile with your spouse when you have conflicts?" We define a harmony index as follows: it is equal to 2 if the couple reported no conflicts, 1 if conflicts are usually solved by mutual compromise, and 0 if conflicts are solved by either unilateral compromise or third-party mediation by family members, relatives or friends. Third-party involvement in conflict solution is a rare event in the data (only 3% reported so) so we do not distinguish it from unilateral compromise. From the best to the worst, the composition in the whole sample is 26%, 49%, and 25%, respectively.

with parental involvement ranges from 29% in Guangdong to 64% in Ningxia province. The corresponding average harmony levels in these two places are 1.04 and 0.60, respectively, while their average log incomes are 9.45 and 7.97, respectively. Across cohort, younger couples (those below 40 years old) on average are less likely to rely on parental involvement, though the trend is not monotonic; in contrast, there is a clear pattern that younger cohorts are less harmonious and earn lower incomes, which may not be surprising given their age differences.

The largest gap, however, lies in the rural-urban comparison, where 58% couples in the rural area had parental involvement in match-making, while only 19% in the urban area did so. The income gap is also the largest, where the log income per couple is 7.90 in rural and 9.92 in urban areas, which reflects the underlying gaps in economic structures and human capital levels as well as other fundamentals. Interestingly, they do not differ much in terms of marital harmony level, which is 0.99 in rural and 1.02 in urban areas. These differences are statistically significant.

The lower panel in Table 1 contrasts the two groups with different match-making methods. Across all areas, couples with parental involvement are on average less harmonious (0.97 versus 1.03) and having lower incomes (8.26 versus 9.19) than those that underwent searching themselves, and these differences are highly significant. The same pattern applies to the rural area, where the between-group contrast is 0.96 versus 1.02 in harmony, and 7.80 versus 8.03 in log income. In the urban areas, couples with parental involvement also have lower harmony levels (0.98 versus 1.03), but slightly higher incomes (9.95 versus 9.91) than others.

This first look at the data suggests some systematic association between parental involvement and marriage outcomes. These primitive results turn out to be robust to controlling various variables and the treatment of the endogeneity issue.

3.3 Endogeneity of Parental Involvement

Table 2 shows that the choice of match-making methods differs systematically across individuals. In particular, individuals with lower human capital and whose parents gain more from household production are more likely to rely on parental involvement, which is consistent with Proposition 1.

The top panel in Table 2 shows that, compared with those who have conducted spouse search themselves, individuals with parental involvement in match-making have on average about 2.5 years less in schooling, were married about 2 years earlier, have parents with fewer years of schooling, and are much more likely to live with parents after marriage. These differences are all statistically significant. Note that the effects of parental education on the match-making choice may come from opposite directions. For example, parents with better education tend to rely less on married children for financial support, which should reduce the likelihood of parental involvement in match-making; but the opposite effect can happen if parental education increases parents' competence in helping children find a good spouse. And so the overall effect of parent education on the choice of match-making methods can be ambiguous. Nonetheless, the evidence is quite clear that parental involvement in match-making is not random.

Since it is impossible to measure and control all the relevant individual and parental characteristics that may affect the choice of match-making, the estimated effect of parental involvement is likely to be biased in the OLS regressions due to the problem of omitted variables. To address this issue, we use the tradition of parental involvement in a marriage market as the instrumental variable for an individual's choice of parental involvement. Specifically, the tradition is measured by the prevalence of parental involvement in the earlier cohort (those who are three to six year older and of same gender) in the same area, where an area is a specific province-urban combination. For instance, for a husband j of 30 years old in the rural area of province i ,

the proportion of all husbands in the cohort of 33-36 years old in the rural area of the same province that adopted parental involvement will be used as the tradition measure for j . Since the group size drops a lot after age 55, the measure of tradition for individuals older than 55 by gender-urban-province becomes less justifiable, and we no longer use them in our regression analysis; as this group is quite small, about 100 couples, our results are not likely to be affected by dropping them.⁹ The summary statistics for this final sample are in Table A1 in the appendix.

As shown in the lower panel of Table 2, an individual's choice of parental involvement in match-making is indeed significantly and positively affected by the tradition of parental involvement in the relevant marriage market, and the estimated effect of such tradition is larger for women. The coefficients of other variables are also quite intuitive: an individual's schooling level has a negative and significant effect on parental involvement for both men and women, while the effect of good health status is negative for both but significant only for men. Parents' education is not significant, which does not necessarily mean no effects, because, as discussed earlier, their multiple influences can operate in opposite directions. And no significant difference is found across age or cohort. Males in urban areas are less likely to use parental involvement, while the urban-rural difference is much smaller and not significant for females.

The tradition of parental involvement as defined above increases the probability of parental involvement possibly due to social learning or the inclination to follow the social custom (Cheung 1972). Both mechanisms suggest that in a society with a stronger tradition of parental matchmaking, parents have a greater advantage in search cost and this advantage applies to every parent regardless of the parent's or child's individual characteristics. The exclusion criterion for the tradition to be a valid instrumental variable for parental involvement is also likely to hold. Except through parental involvement indirectly, it should not have a direct effect on an individual couple's marriage outcomes, especially after controlling individual characteristics, area characteristics and cohort fixed effects. It is useful to note that this area-specific tradition is not affected by individual characteristics because Chinese people had little freedom in changing their residential location at least up to 1991 due to the strict enforcement of the Hukou system in particular and the planning economy in general. This is also evident in our data, where 91% have never changed residence since age 12, while only 7% moved once and 1.5% twice, and such change of location is mostly due to parents' or spouse' change of job.

4 Estimation Results

4.1 The Benchmark Results

The estimated effects of parental involvement on marriage outcomes across all areas are shown in Table 3. The upper left and right panels are based on husband's and wife's information, respectively, where the regressions do not control for the spouse's characteristics because they are endogenous to the match-making choices. The common control variables across these regressions include an individual's age, years of schooling, health status, political party membership (whether the individual is a member of communist party, communist youth league, or democratic party), religion (Muslim, Christian or catholic, Buddhist), ethnic (Han, Hui, Korean, Manchurian), type of the first job (state-owned sectors, individual farms or firms, collective firms, joint ventures of foreign firms), years of schooling of both parents, and location characteristics including whether the average education of parents is above the sample average, whether the average couple income

⁹We have indeed checked this by comparing the means of the sample and the OLS estimates with and without this oldest cohort, and there is not much difference.

is higher than the sample average, and whether it is an urban or rural area.

The overall results across these two panels are very similar: The estimated coefficients of parental involvement are negative and statistically significant for both harmony and income and across both OLS and IV specifications. And their scales in the IV estimation are also quite large. For example, based on the IV results for husbands in the left panel, the estimated effect of parental involvement on harmony is -0.92 of one standard deviation (SD), while the corresponding effect on income is -0.76 SD of log income. The corresponding effects for wife in the right panel are -0.44 SD of harmony and -0.79 SD of log income. In the first stage regressions, the instrumental variable is highly significant, and the F-statistics are quite high (49 to 144), which minimizes the concern of weak instruments.

The absolute scales of IV estimates are much larger than the OLS estimates, which seems to suggest that the adverse selection problem of sons is minor relative to the positive selection problem of parents in the choice of parental involvement; in other words, the OLS estimates will underestimate the negative effects of parental involvement when those who eventually choose parental involvement are likely to have more competent parents. Indeed, the positive selection of parents may be more salient given that very little information about parents is controlled due to availability of data; as a matter of fact, only their education levels are controlled.

The estimated coefficients of other variables are also quite consistent across panels. For example, urban couples on average have lower harmony levels but much higher incomes. The estimated effects of years of schooling on marital harmony form a U-shape, where the bottom is reached at about 12 years for husbands and 8 years for wives based on the IV results in the left and right panels, respectively; in contrast, their effects on income form an inverted U-shape, where the top is reached at about 12 years for husbands and 8.5 years for wives. Another noticeable result is that couples in relatively richer provinces are more harmonious. All of these effects are statistically significant. The effects of good health are positive on both harmony and income, which are highly significant for wives, but for husbands they are significant only for incomes.

One may have issues with the choice of control variables. Some may worry about the possible endogeneity of certain variables such as one's years of schooling,¹⁰ health status, political and religious membership as well as current job type, while others may think even more variables should be controlled. These concerns are addressed in the bottom part of Table 3, where the left panel shows estimation results excluding all of the potentially endogenous variables (years of schooling, health status, political party membership, religion, and job), while the right panel shows results adding extra controls including detailed information of an individual's spouse, differences between husband and wife, and spouse selection criteria at marital searching stage. For such robustness check purpose, we only show results based on the husband's information, since those on the wife are very similar. In both panels, the estimated coefficients of parental involvement are still negative and highly significant as before, where the IV estimates have even larger magnitudes, and the F-statistics in the first stage are also high enough. So our main results are robust to specific choices of control variables.¹¹

¹⁰The possibility that a person's education level is negatively correlated with one's unobserved attractiveness in the marriage market is suggested by Boulier and Rosenzweig (1984) for Philippines.

¹¹Another potential problem is that the data do not have working hours and labor force participation status, which may particularly bias the joint income regressions of urban couples. However, in the context of China, this is not likely to be a serious matter because in China almost all women engaged in full time work at least up to the survey time 1991. For relevant evidence on the almost full labor market participation of women see, for example, Huang et al. (2009).

4.2 The Rural-Urban Differences

Table 4 presents estimation results highlighting differences between rural and urban areas. The control variables are the same as in the main result of Table 3. The upper panel adds a cross term “parental involvement*urban” to the original specification in Table 3, allowing the effects of parental match-making to be different across rural and urban areas. The lower panel shows separate regression results for rural and urban couples, allowing all coefficients to be different. Since results are again similar when using different sets of control variables for both specifications, they are not shown in the table.

The estimated effects of parental involvement in the upper panel are very similar to the benchmark results in Table 3: They are negative and significant for both marital harmony and joint couple incomes, for both husbands and wives, and across OLS and IV specifications. For marital harmony, the coefficients of the cross term are negative and often insignificant, suggesting little difference across rural and urban areas. In contrast, in the income regressions the coefficients of the cross term are not only positive, but also having magnitudes large enough to turn the overall effect of parental match-making to positive for incomes of urban couples; this is true in both OLS and IV specifications, where the F-statistics in the first stage regressions range from 19 to 73. Based on the IV results in the left upper panel, the estimated effects of parental involvement for rural husbands are -0.88 SD of harmony and -1.13 SD of log income. The corresponding effects for urban husbands are -1.08 SD of harmony and 1.34 SD of log income.

Similar results are obtained in the lower panel when we separate the rural and urban couples into two samples. That is, the effects of parental match-making are negative for both harmony and incomes in the rural area, while in the urban area it is negative for harmony but positive for couple incomes. The instrumental variable is still significant in the first stage regressions, though the F-statistics (around 14) are relatively low for the urban sample, indicating the possibility of the tradition variable being a weak instrument in the urban area.¹²

Although the magnitudes of IV estimates are always larger than those of OLS (as in Table 3), the OLS and IV difference is more striking for urban income. We suspect this is because the percent of parental involvement is much lower in urban and there is more income heterogeneity in urban areas. Both could make the identification of the urban-specific effect less robust. That being said, all the specifications show different signs of estimates for rural and urban income, which supports the theoretical prediction that parental involvement could have positive or negative impacts on a couple's joint income, though the impact on marriage harmony is always negative.

One explanation for the rural-urban difference in joint income is that urban parents are more competent in identifying potential spouses with higher earning capabilities for their children, partly because urban parents are better educated. Another possibility is that it may be easier for seasoned seniors to predict a young person's earnings in urban areas, at least up to the survey time, which are determined more by the type of firms they belong to than by their individual qualifications, given that most urban firms were run by the state and job security was very high (under the nick name “iron bowl”). Furthermore, the incomes in the rural area are not easy to measure, especially when the couple lives together with other family members in the same household; this is evident from the much lower explanation power of income regressions in the rural area; for example, the R-squared in the OLS regression of incomes in the lower panel is 0.152 in the rural area, while the corresponding number is 0.447 in the urban area. The fact that the magnitude of the effect for the income equation is much larger in the IV regressions than OLS seems to suggest that the adverse selection problem is more severe in the income regression of urban areas; and this is consistent with

¹²Results for wives are similar and thus not shown.

the observation that only 19% urban couples rely on parental involvement, while 58% rural couples do so (see Table 1).

In summary, the effect of parental involvement is negative on marital harmony in both rural and urban areas, while its effect on joint couple incomes is negative in the rural area but positive in the urban area. More competent urban parents, higher heterogeneity of earning potential, and more severe adverse selection problem for urban individuals may account for the different effects of parental involvement on joint couple incomes across rural and urban areas. Given that parental involvement is more prevalent but incomes are much lower in the rural area, when combined together in the same regression as we have done in Table 3, it is not surprising that the average effect of parental involvement on income is negative and significant; that is, the overall result across all areas is not only driven by within-area heterogeneity but also by large rural-urban gaps.

4.3 Robustness Check

In addition to using different control variables, we have also conducted various tests to check the robustness of our main results in Table 3 and 4. Specifically, we've tried constructing alternative measures of parental involvement tradition and using different measures of the marriage harmony index. Our main results are robust to these changes.

4.3.1 An Alternative Measure of Tradition as IV

Since there is no definite way to measure the tradition of parental involvement, the specific measure used in our estimation may seem a little arbitrary. To check whether our main results are affected by the choice of IV, we construct an alternative measure of the tradition as the new IV, which is possibly more arbitrary and much less nuanced than the one used before.

Given the age distribution of the sample, we divide individuals into 8 age cohorts in each marriage market by province-urban units: the youngest cohort is of ages 18-25, the second youngest cohort is 26-30 years old, followed by cohorts aged 31-35, 36-40, 41-45, 46-50, 51-55, and finally 56 and above. Then we use the proportion of parental involvement among individuals with the same gender of cohort $i + 2$ as the new measure of tradition for everyone in the current cohort i . For instance, the proportion of husbands adopting parental involvement in the cohort of 41-45 years old will be used as the measure of tradition for all husbands of 31-35 years old. Since there is no corresponding measure of tradition for the two oldest cohorts in our sample, we have to drop them (those older than 50) in our regression analysis. And so the identification comes from heterogeneity across 14 marriage markets and 6 cohorts.

The estimation results using this new IV are shown in Table 5, where the control variables are the same as in our main results. Comparing the top panel of the overall results across areas with that in Table 3, there is little difference between the two, though not surprisingly, this new IV has lower explanation power than the old one. The middle panel adds the cross term "parental involvement*urban" to the pooled sample, while the lower panel contrasts rural and urban areas for husbands; the regression results are qualitatively the same and quantitatively similar as their counterparts in Table 4, where the negative effect on marital harmony for urban husbands is much more significant than before. It is also useful to note that the F-statistics are quite high in the urban regressions, which to some degree may alleviate the worry of weak instrument for the urban area. In other words, our main results can be obtained by using a much cruder measure of tradition as IV.

4.3.2 An Alternative Measure of Harmony Index

With regard to the harmony index we used to measure the emotional output of marriage, one may argue the ranking between mutual compromise and other ways of conflict resolution is irrelevant as long as the conflict is resolved. To address this concern, we construct an alternative measure of harmony index, a dummy variable that equals one if an individual reports no conflict in marriage life. The Probit results are presented in Table 6, where the standard errors in the IV specification are obtained by bootstrapping. The top panel shows results across all areas, where the coefficients of parental involvement are negative and significant for both husbands and wives, the same as our main results. The next two panels contrast rural with urban areas, which exhibit the same pattern as before. These results suggest that our findings are robust to the exact measure of the harmony index.

In our main results, we treated the harmony index as a cardinal variable in order to use the standard IV method, even though it is by definition an ordinal variable. To check the sensitivity of our results in this regard, we redo the regression results for harmony by ordered probit; as shown in the bottom of Table 6, they are indeed similar to the corresponding OLS results on harmony in Table 3 and Table 4.

4.4 Discussions and Further Evidence

Our main empirical results fit well with the theoretical predictions of the model. The key prediction about the agency cost of using parents as matchmakers is the negative effect of parental involvement on the emotional output of marriage, and it is indeed found across the board in our empirical analysis, which is robust to various specifications and regardless of whether the rural and urban areas are pooled together or separated. The predicted effect of parental involvement on the economic output of marriage, in contrast, can be either negative or positive, depending on how competent parents are in searching. Both possibilities are realized in the empirical results, where the effect on joint couple incomes is negative in the rural area but positive in the urban area.

A mechanism that may give rise to these results, as suggested by our theoretical model, is that match-making by parental involvement tends to overemphasize the earning capability of a potential spouse, which directly affects household public goods that can be shared among extended family members, than what is deemed optimal if one searches spouse on one's own. If this is true, then the spouse selection criteria should also differ systematically by match-making methods. There is indeed supporting evidence on this matter, which is shown in Table 7.

The seven variables listed at the top of Table 7 are spouse selection criteria considered at the match-making stage. They are dummy variables equal to 1 if the specific trait was deemed one of the two most important characteristics in selecting the marriage partner. Most people (75% of the sample) regard the character of a spouse as one of the top two characteristics, the temperament of a spouse comes next (41%) in popularity, good look is ranked the third (23%), and family background is the fourth (14%), which is then followed by the education level, occupation, and political party membership in the order of popularity. The next two rows compare spouse selection criteria between the two groups with different match-making methods. The results show that individuals with parental involvement are more likely to treat family backgrounds and good look as the top two selection criteria, while less likely to treat character and temperament as well as others as the top two. These differences are statistically significant.

Further confirmation to such differences is provided in the lower panel of Table 7, where the probit regression results on these selection criteria have controlled many variables including individual, parental, and provincial characteristics. The estimated effects of parental involvement are positive on family backgrounds

while negative on character and temperament, and they are highly significant; in contrast, the coefficients of parental involvement on other selection criteria are much smaller and statistically insignificant. The coefficients of other control variables in the probit regressions are also intuitive, suggesting that the analysis does capture something relevant to decisions on spouse selection. For example, individuals in urban areas emphasize much less on a potential spouse's family backgrounds and good looks but more on the other dimensions. Better educated individuals concern more about education but less about family backgrounds. Compared with women, men emphasize more on good looks and temperament as well as character but less on the other traits.

5 Conclusion

This paper examines a new aspect of marriage market - the match-making means - and investigates its effects on marital harmony and joint couple income. The method of match-making matters because it often leads to a different choice of spouse and therefore a different life after marriage.

Specifically, we show that parents' involvement in the match-making process may distort the optimal spouse choice in that they tend to emphasize more on the potential spouse's earning abilities than match quality or attraction between the couple. The rationale is that the joint income of married children can be shared among extended family members more easily than mutual attraction or happiness felt by the couple themselves. Put differently, parents are often more willing to substitute love for money than the individuals themselves. We find supporting evidence for this prediction using a unique sample of Chinese couples in the early 1990s: the estimated effect of parental involvement is indeed negative for marital harmony in both urban and rural areas; in comparison, its effect on income is negative for rural couples but positive for urban couples. These results are already evident by comparing the means of the two groups using different match-making methods, and are further confirmed in various regressions. We also use an instrumental variable approach to address the self-selection issues of both individuals and parents, and obtain qualitatively similar results.

In future research, it would be interesting to investigate whether parental involvement also affects other aspects of marriage life. For example, it seems possible that, given the strong preference of grandchildren and especially grandsons in China, couples married with parents as matchmakers may be more likely to have more children and in particular more sons; and they should also marry and start to have children at younger ages. Another interesting topic would be to explore the potentially positive relationship between the prevalence of parental involvement and income inequality across area and over time.

Reference

- Anderson, Siwan [2003] "Why Dowry Payments Declined with Modernization in Europe but Are Rising in India," *Journal of Political Economy* 111(2), 269-310.
- Anderson, Siwan [2007] "The Economics of Dowry and Brideprice," *Journal of Economic Perspectives* 21(4), 151-174.
- Angrist, Joshua [2002] "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation," *Quarterly Journal of Economics* 117(3), 997-1038.
- Baker, Hugh [1979] *Chinese Family and Kinship*, New York: Columbia University Press.
- Barron, John and Mellow, Wesley [1979] "Search Effort in the Labor Market" *Journal of Human Resources* 14, 389-404.

- Becker, Gary [1973] "A Theory of Marriage: Part I," *Journal of Political Economy* 81(4), 813-46.
- Becker, Gary [1974] "A Theory of Marriage: Part II," *Journal of Political Economy* 82(2), S11-S26.
- Becker, Gary [1981] *A Treatise on the Family*, Cambridge, Mass: Harvard University Press.
- Blood, Robert [1967] *Love Match and Arranged Marriage*, New York: Free Press.
- Botticini, Maristella and Aloysius Siow [2003] "Why dowries?" *American Economic Review* 93(4), 1385-1398
- Boulier, Bryan and Mark Rosenzweig [1984] "Schooling, search, and spouse selection: Testing economic theories of marriage and household behavior," *Journal of Political Economy* 92(4), 712-732
- Chiappori, Pierre-Andre, Bernard Fortin, and Guy Lacroix [2002] "Marriage Market, Divorce Legislation and Household Labor Supply," *Journal of Political Economy* 110(1), 37-72.
- Cheung, Steven [1972] "The Enforcement of Property Rights in Children, and the Marriage Contract," *Economic Journal* 82, 641-657.
- Chao, Paul [1983] *Chinese Kinship*. Kegan Paul International.
- Cohen, Myron [1992] "Family Management and Family Division in Contemporary Rural China," *China Quarterly* 130: 357-377.
- Audrey, Davidson and Robert Ekelund [1997] "The Medieval Church and Rents from Marriage Market Regulations," *Journal of Economic Behavior and Organization* 32, 215-245.
- Edlund, Lena, and Nippe Lagerlof [2002] "Individual versus Parental Consent in Marriage: Implications for Intra-Household Resource Allocation and Growth," *The American Economic Review* 96(2), 304-307.
- Fernandez, Raquel, Guner Nezih, and Knowles John [2005] "Love and Money: A Theoretical and Empirical Analysis of Household Sorting and Inequality," *Quarterly Journal of Economics* 120(1), 273-344.
- Foster, Andrew, and Mark Rosenzweig [2000] "Missing Women, the Marriage Market and Economic Growth," working paper, University of Pennsylvania, August.
- Gilford, Roslie, and Vern Bengtson [1979] "Measuring Marital Satisfaction in Three Generations: Positive and Negative Dimensions," *Journal of Marriage and the Family* 41(2), 387-398.
- Grossbard-Shechtman, and Shoshana Amyra [1993] *On the Economics of Marriage: A Theory of Marriage, Labor and Divorce*, Boulder and Oxford: Westview Press.
- Huang, Chong, Li Hongbin, Liu Pak Wai and Zhang Junsen [2009] "Why Does Spousal Education Matter for Earnings? Assortative Mating and Cross-Productivity," *Journal of Labor Economics* 27(4)
- Institute of Population Studies, CASS. *Sampling Survey Data of Women's Status in Contemporary China*. Beijing: International Academic Publishers.
- Johnson, David, Lynn White, John Edwards and Alan Booth [1986] "Dimensions of Marital Quality: Toward Methodological and Conceptual re..nement," *Journal of Family Issues* 71(3), 31-50.
- Korenman, Sanders and David Neumark [1991] "Does Marriage Really Make Men More Productive?" *Journal of Human Resources* 26(2), 282-307.
- Suen, Wing, William Chan and Junsen Zhang [2003] "Marital Transfer and Intrahousehold Allocation: A Nash-Bargaining Analysis" *Journal of Economic Behavior Organization* 52, 133-146.
- Weiss, Yoram [1997] "The Formation and Dissolution of Families: Why Marry? Who Marries Whom? And What Happens Upon Divorce?" Chapter 3, in M.R. Rosenzweig and O. Stark eds., *Handbook of Population and Family Economics*, Elsevier Science.
- Weiss, Yoram, and Robert Willis [1997] "Match Quality, New Information and Marital Dissolution," *Journal of Labor Economics* 15(1), S293-329.
- Xu, Lixin Colin, Christine Zhen-Wei Qiang, and Limin Wang [2003] "The Timing of Marriage in China" *Annals of Economics and Finance* 4(2): 343-357.

Xiaohe Xu [1996] "Measuring the Concept of Marital Quality as Social Indicators in Urban China," *Social Indicators Research* 37(2), 189-206.

Xu, Xiaohe and Martin Whyte [1990] "Love Matches and Arranged Marriages: A Chinese Replication," *Journal of Marriage and the Family* 52(3), 709-722.

Zeng, Yi and Deqing Wu [2000] "Regional Analysis of Divorce in China since 1980," *Demography* 37(2), 215-219.

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

Zimmer, Zachary and Julia Kwong [2003] "Family Size and Support of Older Adults in Urban and Rural China: Current Effects and Future Implications," *Demography* 40(1), 23-44.

APPENDIX

Proof. The optimal values of α^* and h_f^* are jointly determined by (1) and (2). Note that α^* can be solved from condition (1) as a function of h_f^* , which can then be plugged in (2) to solve h_f^* . The optimal values of α^{**} and h_f^{**} are jointly determined by (3) and (4), based on which we get $\partial\alpha^{**}/\partial h_p > 0$ and $\partial h_f^{**}/\partial h_p > 0$ by Cramer's rule:

$$\frac{\partial\alpha^{**}}{\partial h_p} = \frac{\begin{matrix} (-) \\ \eta_p s_{13} \end{matrix} \quad \begin{matrix} (+) \\ \delta f_1 - \eta_p s_{12} \end{matrix}}{\begin{matrix} (-) \\ \eta_p s_{23} \end{matrix} \quad \begin{matrix} (-) \\ [\gamma + \delta(\beta + \alpha^{**})]f_{11} - \eta_p s_{22} \end{matrix}} \quad /|H| > 0,$$

$$\frac{\partial h_f^{**}}{\partial h_p} = \frac{\begin{matrix} (-) \\ -\eta_p s_{11} \end{matrix} \quad \begin{matrix} (-) \\ \eta_p s_{13} \end{matrix}}{\begin{matrix} (+) \\ \delta f_1 - \eta_p s_{12} \end{matrix} \quad \begin{matrix} (-) \\ \eta_p s_{23} \end{matrix}} \quad /|H| > 0,$$

where $|H|$ is the determinant of Hessian matrix

$$|H| = \begin{matrix} (-) & (+) \\ -\eta_p s_{11} & \delta f_1 - \eta_p s_{12} \\ \delta f_1 - \eta_p s_{12} & [\gamma + \delta(\beta + \alpha^{**})]f_{11} - \eta_p s_{22} \end{matrix} \quad ,$$

and $|H| > 0$ is assumed for the existence of optimal solutions. We can also get $\partial\alpha^{**}/\partial\gamma > 0$ and $\partial h_f^{**}/\partial\gamma$ in a similar way. We assume $\delta f_1 - \eta_p s_{12} \geq 0$, which essentially means that $\partial^2\mathcal{E}/\partial\alpha\partial h_f \geq 0$; and then based on (3) we have

$$\frac{\partial\alpha^{**}}{\partial h_f} = \frac{\delta f_1(h_f^*, h_m) - \eta_p s_{12}(\alpha^*, h_f^*, h_m)}{s_{11}} > 0.$$

Comparing conditions (1) and (3), we can see that $\alpha^*(h_f) > \alpha^{**}(h_f)$ must hold, conditional on the same level of h_f ; the reason is that the ..rst terms in both conditions are independent of α while the second terms are strictly increasing in it, which combined with the assumption $\eta_p s_1 \geq \delta\eta_m c_1$ will lead to $\alpha^*(h_f) > \alpha^{**}(h_f)$. Given the same α , we may have $h_f^* > h_f^{**}$ if $\gamma < (1 - \delta)(\beta + \alpha)$ and/or $\eta_m c_2 \leq \eta_p s_2$, in other words, if parents do not bene..t too much from the daughter-in-law's human capital h_f or if their marginal searching cost with respect to h_f is not lower than the son's. The opposite result $h_f^* \leq h_f^{**}$ is otherwise possible.

Since the searching cost is always positive,

$$(\beta + \alpha^*)f(h_f^*, h_m) > (\beta + \alpha^{**})f(h_f^{**}, h_m) \quad (6)$$

must hold, which then implies $\alpha^* > \alpha^{**}$ and $\alpha^*f(h_f^*, h_m) > \alpha^{**}f(h_f^{**}, h_m)$; the reason is as follows. (i) If $h_f^* > h_f^{**}$, then $\alpha^*(h_f^*) > \alpha^{**}(h_f^*) > \alpha^{**}(h_f^{**})$ holds, and it implies $\alpha^* > \alpha^{**}$, where the ..rst inequality follows $\alpha^*(h_f) > \alpha^{**}(h_f)$ and the second follows $\partial\alpha^{**}(h_f)/\partial h_f > 0$. And $\alpha^*f(h_f^*, h_m) > \alpha^{**}f(h_f^{**}, h_m)$ follows directly from $h_f^* > h_f^{**}$ and $\alpha^* > \alpha^{**}$. (ii) If $h_f^* \leq h_f^{**}$, then we have

$$(\beta + \alpha^*)f(h_f^{**}, h_m) \geq (\beta + \alpha^*)f(h_f^*, h_m) > (\beta + \alpha^{**})f(h_f^{**}, h_m),$$

where the ..rst and third terms imply $\alpha^* > \alpha^{**}$; the ..rst inequality holds because $h_f^* \leq h_f^{**}$, while the second inequality is based on (6). And following similar arguments we can derive $\alpha^*f(h_f^*, h_m) > \alpha^{**}f(h_f^{**}, h_m)$ by comparing the ..rst and third terms in

$$\beta f(h_f^{**}, h_m) + \alpha^*f(h_f^*, h_m) \geq (\beta + \alpha^*)f(h_f^*, h_m) > (\beta + \alpha^{**})f(h_f^{**}, h_m).$$

Based on (5), the utility difference between self and parents' searching is

$$\pi \equiv U^* - U^{**} = (\beta + \alpha^*)f(h_f^*, h_m) - \eta_m c(\alpha^*, h_f^*, h_m) - (\beta + \alpha^{**})f(h_f^{**}, h_m),$$

which is strictly decreasing in h_p because

$$\partial\pi/\partial h_p = -f(h_f^{**}, h_m)\partial\alpha^{**}/\partial h_p - (\beta + \alpha^{**})f_1(h_f^{**}, h_m)\partial h_f^{**}/\partial h_p < 0.$$

We get $\partial\pi/\partial h_m > 0$ for the following reason. Note that

$$\partial\pi/\partial h_m = \frac{\partial U^*}{\partial h_m} - \partial(\beta + \alpha^{**})f(h_f^{**}, h_m)/\partial h_m,$$

where

$$\frac{\partial U^*}{\partial h_m} = \frac{\partial(\beta + \alpha^*)f(h_f^*, h_m)}{\partial h_m} - \frac{\partial\eta_m c(\alpha^*, h_f^*, h_m)}{\partial h_m} \quad (7)$$

$$= (\beta + \alpha^*)f_2(h_f^*, h_m) - \eta_m c_3(\alpha^*, h_f^*, h_m) > 0 \quad (8)$$

by the Envelop Theorem. Since

$$\frac{\partial^2 U^*}{\partial h_m \partial \alpha} = \alpha^* f_2(h_f^*, h_m) - \eta_m c_{31}(\alpha^*, h_f^*, h_m) > 0,$$

we have

$$\frac{\partial U^*}{\partial h_m} \Big|_{(\alpha^*, h_f^*)} > \frac{\partial U^*}{\partial h_m} \Big|_{(\alpha^{**}, h_f^*)} > \partial(\beta + \alpha^{**})f(h_f^*, h_m)/\partial h_m, \quad (9)$$

where the first inequality is because $\alpha^* > \alpha^{**}$, and the second inequality is because the second term of $\frac{\partial U^*}{\partial h_m}$ in (8) is positive. (i) If $h_f^* > h_f^{**}$, then $\partial^2(\beta + \alpha)f(h_f, h_m)/\partial h_m \partial h_f = f_{12}(h_f, h_m) > 0$ implies

$$\partial(\beta + \alpha^{**})f(h_f^*, h_m)/\partial h_m > \partial(\beta + \alpha^{**})f(h_f^{**}, h_m)/\partial h_m,$$

which combined with the inequality in (9) implies

$$\frac{\partial U^*}{\partial h_m} \Big|_{(\alpha^*, h_f^*)} > \partial(\beta + \alpha^{**})f(h_f^{**}, h_m)/\partial h_m,$$

and this leads to

$$\partial\pi/\partial h_m = \frac{\partial U^*}{\partial h_m} - \partial(\beta + \alpha^{**})f(h_f^{**}, h_m)/\partial h_m > 0.$$

(ii) When $h_f^* \leq h_f^{**}$ is the case, the result can be derived in a similar way due to

$$\frac{\partial(\beta + \alpha^{**})f(h_f^{**}, h_m)}{\partial h_m} < \frac{\partial(\beta + \alpha^*)f(h_f^{**}, h_m)}{\partial h_m} - \frac{\partial\eta_m c(\alpha^*, h_f^*, h_m)}{\partial h_m} < \frac{\partial U^*}{\partial h_m} \Big|_{(\alpha^*, h_f^*)},$$

where the first inequality holds because of $\alpha^{**} < \alpha^*$ and $-\partial\eta_m c(\alpha^*, h_f^*, h_m)/\partial h_m > 0$, while the second inequality holds because (α^*, h_f^*) is the optimal choice to maximize U^* than (α^*, h_f^{**}) ; comparing the first and the third terms we get $\partial\pi/\partial h_m > 0$.

So the threshold $h_p^\#$ is uniquely determined by

$$\pi = (\beta + \alpha^*)f(h_f^*, h_m) - \eta_m c(\alpha^*, h_f^*, h_m) - (\beta + \alpha^{**}(h_p^\#))f(h_f^{**}(h_p^\#), h_m) = 0.$$

Based on this identity, we get

$$\begin{aligned} \frac{\partial h_p^\#}{\partial h_m} &= -\frac{\partial\pi/\partial h_m}{\partial\pi/\partial h_p} > 0, \\ \frac{\partial h_p^\#}{\partial \gamma} &= -\frac{\partial\pi/\partial \gamma}{\partial\pi/\partial h_p} = -\frac{f(h_f^{**}, h_m)\partial\alpha^{**}/\partial \gamma + (\beta + \alpha^{**})f_1(h_f^{**}, h_m)\partial h_f^{**}/\partial \gamma}{-\partial\pi/\partial h_p} < 0, \\ \frac{\partial h_p^\#}{\partial \eta_m} &= -\frac{\partial\pi/\partial \eta_m}{\partial\pi/\partial h_p} = \frac{-c(\alpha^*, h_f^*, h_m)}{-\partial\pi/\partial h_p} < 0, \end{aligned}$$

The comparative statics for the threshold level $h_m^\#$ can be derived in a similar manner. ■

Table 1: Summary Statistics

Mean (Standard Deviation)

	Number of Observations	Parental Involvement	Harmony Index	Log Income for Couple
The Whole Sample	17330	.40 (.49)	1.00 (.72)	8.81 (1.23)
By Province:				
Guangdong	2822	.29 (.46)	1.04 (.63)	9.45 (1.32)
Shanghai	2966	.30 (.46)	1.13 (.75)	8.48 (.41)
Sichuan	2334	.34 (.47)	.89 (.71)	8.99 (1.24)
Shandong	2574	.39 (.49)	1.18 (.72)	8.99 (1.20)
Shaanxi	2872	.47 (.50)	1.04 (.72)	8.76 (1.38)
Jilin	2192	.50 (.50)	.85 (.72)	8.72 (1.21)
Ningxia	1570	.64 (.48)	.60 (.72)	7.97 (1.21)
By Cohort:				
<30 years old	4227	.41 (.49)	.96 (.72)	8.52 (1.20)
30-40 years old	7172	.38 (.49)	.98 (.71)	8.86 (1.18)
40-50 years	4492	.44 (.49)	1.04 (.71)	8.93 (1.24)
Above 50 years old	1439	.41 (.49)	1.10 (.73)	9.09 (1.40)
By Urban:				
Rural	9502	.58 (.49)	.99 (.71)	7.90 (.68)
Urban	7828	.19 (.39)	1.02 (.73)	9.92 (.76)
Difference		.393*** (.007)	-.039*** (.011)	-.933*** (.018)

Marriage Outcomes by Matchmaking Method

All Areas:			
Parental Involvement		.97 (.009)	8.26 (.013)
Self Search		1.03 (.007)	9.19 (.012)
Difference		-.059*** (.011)	-.227*** (.014)
Rural:			
Parental Involvement		.96 (.71)	7.80 (.67)
Self Search		1.02 (.70)	8.03 (.66)
Difference		-.052*** (.015)	-.227*** (.014)
Urban:			
Parental Involvement		.98 (.73)	9.95 (.71)
Self Search		1.03 (.72)	9.91 (.77)
Difference		-.051** (.021)	.037* (.021)

Note: The unit of observation is individual respondent. The sample excludes couples that one of them remarried or that the matching mode was missing or "others." *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Endogenous Parental Involvement

Individual and Parental Attributes by Matchmaking Method					
Mean (Standard Deviation)					
	Years of Schooling	Age at Marriage	Mother's Schooling	Father's Schooling	Live with Parents after Marriage
Parental Involvement	6.48 (3.90)	22.93 (3.66)	1.40 (2.60)	3.23 (3.49)	.65 (.48)
Self Search	8.93 (3.59)	24.64 (3.53)	2.73 (3.48)	5.00 (3.89)	.46 (.50)
Difference	-2.454*** (.059)	-1.708*** (.056)	-1.341*** (.046)	-1.769*** (.057)	.187*** (.008)

Parental Involvement: Probit		
	Husband	Wife
Tradition of Parental Involvement	1.339*** (.201)	1.934*** (.170)
Urban	-.510*** (.102)	-.047 (.084)
Younger than 35 years old	-.008 (.061)	.048 (.056)
Age	-.004 (.021)	-.001 (.017)
Age Squared	.000 (.000)	.000 (.000)
Years of Schooling	-.034*** (.016)	-.045*** (.012)
Schooling Squared	.001 (.001)	.001 (.001)
Good Health	-.132*** (.039)	-.019 (.033)

Table 3: Benchmark Results

	Husband				Wife			
	Marital Harmony		Couple Income		Marital Harmony		Couple Income	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Parental Involvement	-.063*** (.020)	-.659*** (.217)	-.071*** (.019)	-.925*** (.236)	-.032* (.018)	-.434*** (.124)	-.077*** (.017)	-.973*** (.146)
Urban	-.063* (.036)	-.254*** (.078)	1.795*** (.039)	1.527*** (.083)	-.040 (.030)	-.061 (.044)	1.867*** (.030)	1.638*** (.049)
Age	-.015 (.011)	-.014 (.012)	.021** (.010)	.024** (.012)	-.005 (.010)	-.004 (.010)	.042*** (.009)	.046*** (.010)
Age Squared	.000* (.000)	.000* (.000)	-.000* (.000)	-.000 (.000)	.000 (.000)	.000 (.000)	-.000*** (.000)	-.000*** (.000)
Years of Schooling	-.015* (.009)	-.025** (.010)	.065*** (.009)	.049*** (.010)	-.016*** (.006)	-.027** (.007)	.058*** (.006)	.034*** (.008)
Schooling Squared	.001** (.000)	.001*** (.000)	-.002*** (.000)	-.002*** (.000)	.001*** (.000)	.002*** (.000)	-.003*** (.000)	-.002*** (.000)
Good Health	.039** (.020)	.019 (.022)	.169*** (.019)	.140*** (.023)	.090*** (.017)	.090*** (.017)	.153*** (.016)	.150*** (.018)
Mother Schooling	.002 (.004)	.004 (.004)	-.004 (.003)	-.000 (.004)	-.006* (.003)	-.006* (.003)	-.001 (.003)	-.001 (.003)
Father Schooling	-.001 (.003)	-.001 (.004)	.007** (.003)	.005 (.003)	-.000 (.003)	.000 (.003)	.007*** (.003)	.008*** (.003)
Province w/ Higher Parental Education	-.050** (.024)	-.082*** (.028)	.018 (.022)	-.030 (.028)	-.030 (.023)	-.037 (.024)	.012 (.021)	-.003 (.025)
Rich Province	.062*** (.019)	.050*** (.021)	.488*** (.018)	.477*** (.021)	.105*** (.018)	.094*** (.019)	.484*** (.018)	.461*** (.021)
Observations	6887	6882	7183	7177	7742	7741	8158	8157
Adjusted R2	.021	-	.721	.636	.031	-	.725	.628
First Stage Regression of Parental Involvement								
Tradition of Parental Involvement		.522*** (.070)		.474*** (.068)		.736*** (.061)		.694*** (.059)
F-statistic in the First Stage		56.34		49.39		144.96		137.22
Husband								
	Fewer Control Variables				More Control Variables			
Parental Involvement	-.049*** (.018)	-.913*** (.218)	-.086*** (.017)	-1.433*** (.250)	-.055*** (.020)	-.799*** (.307)	-.047** (.018)	-1.177*** (.358)
Urban	.004 (.020)	-.268*** (.072)	1.902*** (.018)	1.476*** (.080)	-.106*** (.038)	-.272*** (.079)	1.752*** (.041)	1.498*** (.090)
Observations	8051	8046	8462	8456	6887	6882	7183	7177
Adjusted R2	.015	-	.698	.464	.040	-	.738	.592
F-statistic in the First Stage		61.91		64.43		30.36		26.34

Note: The other controls include cohort dummies, political party membership variables (whether the individual is a communist party member, communist youth league, or a democratic party member), religion (Muslim, Christian or catholic, Buddhist), ethnic (Han, Hui, Korean, Manchurian), and the types of first job. The extra controls in the last panel include detailed information on an individual's spouse and selection criteria. *** p<0.01, ** p<0.05, * p<0.1.

Table 4: Rural vs. Urban

	Husband				Wife			
	Marital Harmony		Couple Income		Marital Harmony		Couple Income	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Parental Involvement	-.070*** (.025)	-.636*** (.214)	-.132*** (.024)	-1.385*** (.286)	-.027 (.022)	-.375*** (.121)	-.141*** (.021)	-1.337*** (.180)
Parental-Involvement*Urban	-.019 (.041)	-.138 (.385)	.162*** (.037)	3.038*** (.421)	-.014 (.038)	-.833* (.450)	.189*** (.035)	5.133*** (.783)
Urban	-.068* (.038)	-.226** (.100)	1.752*** (.041)	.932*** (.126)	.044 (.032)	.124 (.100)	1.811*** (.032)	.481*** (.164)
Observations	6887	6882	7183	7177	7742	7741	8158	8157
Adjusted R2	.021	-	.722	.487	.031	.	.726	.029
F-statistics in the First Stage								
Tradition of Parental Involvement		28.30		24.69		73.84		69.32
Tradition of Parental Involvement*Urban		23.92		26.04		19.33		20.80
Husband								
	Rural				Urban			
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Parental Involvement	-.067*** (.025)	-.628** (.251)	-.111*** (.022)	-3.262*** (.601)	-.053 (.033)	-.757 (.512)	.033 (.027)	2.846*** (.867)
Observations	3,189	3,189	3,398	3,398	3,689	3,693	3,785	3,779
Adjusted R2	.036	-	.152	-	.016	.	.447	.
First Stage Regression of Parental Involvement								
Tradition of Parental Involvement		.578*** (.095)		.521*** (.092)		.389*** (.104)		.386*** (.103)
F-statistic in the First Stage		36.86		31.98		13.94		14.16

Note: The other control variables are the same as in Table 3. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Different IV

	All Areas							
	Husband				Wife			
	Marital Harmony		Couple Income		Marital Harmony		Couple Income	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Parental Involvement	-.063*** (.021)	-.860*** (.303)	-.067*** (.019)	-.594** (.290)	-.033* (.019)	-.547*** (.174)	-.083*** (.017)	-.964*** (.192)
Urban	-.064* (.037)	-.315*** (.104)	1.802*** (.040)	1.634*** (.097)	-.043 (.031)	-.087 (.055)	1.857*** (.030)	1.625*** (.059)
Observations	6381	6369	6657	6643	7426	7420	7826	7818
Adjusted R2	.021	-	.718	.682	.031	.	.724	.625
First Stage Regression of Parental Involvement								
Tradition of Parental Involvement		.397*** (.069)		.335*** (.067)		.527*** (.058)		.486*** (.056)
F-statistic in the First Stage		33.05		24.98		83.49		75.83
All Areas with Parental-Involvement*Urban								
Parental Involvement	-.068*** (.026)	-.892** (.364)	-.127*** (.025)	-1.545*** (.500)	-.029 (.023)	-.445*** (.164)	-.143*** (.021)	-1.517*** (.258)
Parental-Involvement*Urban	.014 (.043)	.074 (.385)	.168*** (.038)	2.222*** (.433)	-.013 (.039)	-.917 (.4607)	.181*** (.035)	5.180*** (.899)
Urban	-.067* (.039)	-.336** (.159)	1.759*** (.041)	1.017*** (.216)	.046 (.033)	.121 (.123)	1.805*** (.032)	.449*** (.186)
Observations	6381	6369	6657	6643	7426	7420	7826	7818
Adjusted R2	.021	-	.719	.545	.031	.	.725	.031
F-statistics in the First Stage								
Tradition of Parental Involvement		18.16		15.08		44.29		39.31
Tradition of Parental Involvement*Urban		27.73		28.61		12.48		13.78
Husband								
	Rural				Urban			
Parental Involvement	-.067*** (.026)	-1.285** (.636)	-.104*** (.022)	-7.055** (2.892)	-.054 (.035)	-.769** (.382)	.037 (.028)	1.113*** (.362)
Observations	3,025	3,025	3,219	3,219	3,356	3,344	3,438	3,424
Adjusted R2	.038	-	.153	-	.015	.	.436	.173
First Stage Regression of Parental Involvement								
Tradition of Parental Involvement		.310*** (.106)		.247*** (.101)		.486*** (.096)		.457*** (.1095)
F-statistic in the First Stage		8.49		5.91		25.73		23.36

Note: The other control variables are the same as in Table 3 and individuals are not older than 50 years old. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Alternative Measures of Marriage Harmony

Dependent Variable = Have Conflicts or not				
	Husband		Wife	
	Probit	IV	Probit	IV
Parental Involvement	-.089** (.038)	-1.251*** (.248)	-.072** (.036)	-.888*** (.165)
Urban	-.080 (.070)	-.436*** (.098)	.190*** (.059)	-.029 (.062)
Observations	7183	7177	8158	8157
Pseudo R2	.026	-	.037	.
Husband				
	Rural		Urban	
	Probit	IV	Probit	IV
Parental Involvement	-.111** (.049)	-1.093*** (.236)	-.033 (.062)	-.683 (.844)
Observations	3398	3398	3785	3778
Pseudo R2	.056	-	.016	-
Wife				
	Rural		Urban	
	Probit	IV	Probit	IV
Parental Involvement	-.033 (.062)	-.439* (.261)	-.086 (.061)	-2.342*** (.476)
Observations	3785	4237	3906	3905
Pseudo R2	.016	-	.027	.
Ordered Probit: No conflict, mutual compromise, unilateral compromise/third-party intervention				
	All Areas		All Areas	
	Husband	Wife	Husband	Wife
Parental Involvement	-.099*** (.032)	-.051* (.029)	-.110*** (.039)	-.043 (.035)
Parental Involvement* Urban			.028 (.065)	-.022 (.061)
Observations	6887	7742	6887	7742
Pseudo R2	.012	.017	.012	.017

Note: The other control variables are the same as in Table 3. The standard errors in probit IV are obtained by bootstrapping. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Parental Involvement and Spouse Selection Criteria

Spouse Selection Criteria by Matchmaking Method							
	Mean (Standard Deviation)						
	Character	Temperament	Family Background	Good Look	Education	Occupation	Political Membership
All	.75 (.43)	.41 (.49)	.14 (.34)	.23 (.42)	.11 (.31)	.09 (.28)	.05 (.21)
Parental Involvement	.71 (.46)	.38 (.48)	.19 (.39)	.27 (.44)	.07 (.26)	.07 (.25)	.04 (.18)
Self Search	.79 (.41)	.43 (.50)	.10 (.30)	.20 (.40)	.13 (.33)	.10 (.30)	.06 (.23)
Difference	-.082*** (.007)	-.006*** (.008)	.084*** (.006)	069*** (.007)	-.054*** (.005)	-.028*** (.004)	-.022*** (.003)
Regression Results: Probit							
	Character	Temperament	Family Background	Good Look	Education	Occupation	Political Membership
Parental Involvement	-.133*** (.023)	-.133*** (.023)	.133*** (.028)	.021 (.024)	.040 (.033)	-.020 (.033)	-.065 (.041)
Urban	.203*** (.031)	.047* (.022)	-.552*** (.039)	-.552*** (.032)	.274*** (.039)	.410*** (.042)	.255*** (.053)
Years of Schooling	.002 (.004)	-.006* (.003)	-.019*** (.003)	.002 (.004)	.083*** (.005)	.012** (.005)	.016** (.006)
Male	.052** (.023)	.158*** (.021)	-.413*** (.028)	.325*** (.023)	-.324*** (.030)	-.256*** (.030)	-.275*** (.056)
Observations	17119	17119	17119	17117	17119	17119	17119
Pseudo R2	.020	.011	.096	.053	.123	.036	.106

Note: The unit of observation is individual respondent. The other control variables include cohort dummies, mother and father's years of schooling, political party membership variables (whether the individual is a communist party member, communist youth league, or a democratic party member), religion (Muslim, Christian or catholic, Buddhist), ethnic (Han, Huei, Korean, Manchurian), whether the province is rich (with above-average income) and has higher-than-average parental education levels. *** p<0.01, ** p<0.05, * p<0.1.

Table A1: Summary Statistics

Variables	Mean (Std Deviation)	Number of Observations
<u>Marriage Outcomes</u>		
Marital Harmony	1.00 (.72)	16247
Log (joint couple income)	8.81 (1.23)	17119
<u>Searching Method</u>		
Parental involvement	.40 (.49)	17119
<u>Individual Characteristics</u>		
Age	36.25 (8.09)	17119
Years of schooling	7.96 (3.89)	17119
Good health	.67 (.47)	17119
Younger cohort (age<=35)	.48 (.50)	17119
<u>Political Membership</u>		
communist party member	.16 (.37)	17119
communist youth league	.07 (.25)	17119
democratic party member	.004 (.06)	17119
<u>Religion</u>		
Muslim	.06 (.24)	17119
Christian or catholic	.01 (.09)	17119
Buddhist	.04 (.19)	17119
<u>Ethnic</u>		
Huei	.06 (.24)	17119
Korean	.003 (.05)	17119
Manchurian	.01 (.10)	17119
Other minority	.003 (.06)	17119
<u>Type of First Job</u>		
First job in state-owned sector	.28 (.45)	15341
First job in collectively-owned firms	.25 (.43)	15341
First job in individually-owned firms	.06 (.24)	15341
First job in joint venture or foreign firms	.01 (.12)	15341
First job is of other types	.03 (.17)	15341
First job in state-owned sector with no change of jobs	.36 (.48)	15341
<u>Parents Characteristics</u>		
Mother Schooling		
Father Schooling		
<u>Location Characteristics</u>		
Urban	.45 (.50)	17119
With higher-than-average parental education	.48 (.50)	17119
Rich Province with higher-than-average incomes	.61 (.49)	17119
<u>Couple Information</u>		
Log (Time after marriage)	2.37 (.74)	17119
Living together with parents after marriage	.54 (.50)	17119
Same ethnic	.98 (.15)	17119
Same religion	.95 (.22)	17119
Same political membership	.17 (.38)	17119
Difference in ages	2.39 (2.23)	17119
Difference in years of schooling	1.05 (.74)	17119