A matching model of co-residence with a family network: Empirical evidence from China

Naijia Guo, Xiaoyu Xia, and Junsen Zhang*

This version: September 9, 2019

Abstract

We develop a co-residence model between young adults and the elderly as a new application of Shapley–Shubik–Becker bilateral matching framework. This model captures competition between adult children as well as between parents and parents-in-law. It extends the Shapley–Shubik–Becker model by restricting potential matching choices in a family network. Using data from the China Family Panel Study, we estimate our model using a network simulation method to fill in marriage links that are not directly observed in the data. We find that the child-side and parent-side competition predicted by our model match well with the patterns observed in the data. In addition, counterfactual experiments allow us to quantify the effects of changes in housing prices and China's family planning policies on intergenerational co-residence arrangements.

JEL Classification: D1, J1, J2

Keywords: Intergenerational co-residence; matching model; competition for co-residence; family network.

*The Chinese University of Hong Kong. Guo: guonaijia@cuhk.edu.hk; Xia: xiaoyuxia@baf.cuhk.edu.hk; Zhang: jszhang@cuhk.edu.hk. We thank the editor and anonymous referees for providing extremely helpful comments on an earlier version of this paper.

1 Introduction

In China, 36% of married children are living with either their parents or parents-in-law, and more than 40% of parents aged 55 to 75 live with their children, see Figures 1 and 2. Why do such a large portion of married adults in China live with their parents?

While social norms and family values may lead families to live together, economic reasons can also be important factors in this decision. Following China's housing privatization reform in the mid-1990s, many adult children have been unable to afford home ownership and choose to live with their elderly parents who own homes (Rosenzweig and Zhang, 2014). In addition, many parents live with their adult children to receive informal care and save on eldercare costs (Hoerger et al., 1996), while other parents may live with their adult children to help care for their grandchildren (Maurer-Fazio et al., 2011). Such needs of eldercare and childcare arise out of the lack of nursing homes and pre-kindergartens in China.

Because cost savings is a potentially important mechanism for the co-residence decision, competition to co-reside might exist between family members. Adult children may compete with their siblings to live with their parents, and parents and parents-in-law may compete to live with their children.[1] Table 1 illustrates child-side and parent-side competitions using data from the 2010 China Family Panel Study (CFPS). While the probability of a young adult living with his or her own parent is correlated with that young adult's and their parents' characteristics (Column 1), the characteristics of the young adult's siblings and spouse's parents also affect the co-residence likelihood (Columns 2 and 3). Among-sibling competition is reflected in two facts: an adult child is more likely to live with his/her parents if the adult child has fewer siblings, and if his/her siblings are highly educated. Parent-side competition is also reflected in two facts: an adult child more likely to live with his/her own parents if the spouse's parents are highly educated and also if the spouse's parents have more children.

In this study, we develop a bilateral matching model to explain such child-side and parent-side competitions. By using Chinese data, we estimate the model to demonstrate how social norms and economic incentives affect living arrangements. Through counterfactual experiments, we apply the model to examine the effects of housing prices and family planning policy on the co-residence decisions of different types of parents and adult children.

We model the living arrangement as a one-to-one, two-sided matching game with transferable utility between parents and adult children. Our matching model is an extension of Shapley and Shubik (1971) and it incorporates elements from the marriage matching model in Becker (1973) and the collective model of the household in Chiappori (1992). In our model, the co-residence surplus between a parent and a young adult consists of two parts. The first pertains to savings on housing, eldercare, and childcare costs, and the second pertains to preferences, captured by congestion costs — reduced utility caused by living with others.

[1] In some cases, while adult children may not particularly want to live with their parents, the parents may provide financial transfers to their adult children to "buy" co-residence. Therefore, adult children may compete for these transfers.

Different from Shapley–Shubik–Becker's matching model, we set a restriction on the choice set of participants on each side of our co-residence matching model. Elderly people can live with one of their adult children but not any other young adults; young adults can live with their parents or parents-in-law but not any other elderly people. This restriction poses challenges for model solutions, as researchers usually cannot observe the family network of the entire population (including parent-child links and marriage links of children) in a large market. In this paper, we propose a network simulation procedure to fill in the missing links of the partial family network observed in the data.

We estimate our model using the 2010 CFPS, which has basic demographic information on every family member regardless of residence. The CFPS contains complete links between parents and all of their children, which is essential for our model estimation. However, it only provides spousal information for young adults living in the surveyed households. To cope with the incomplete family network in the CFPS, we carry out our estimation in three steps. First, we bootstrap a sample of families for each city from our national family sample, which contains information on parents and all of their married children. Second, we estimate a parametric marriage-matching model with transferable utility among young couples in the CFPS data using a moment matching estimator suggested by Galichon and Salanié (2010). We then use the estimation results to simulate the marriage links for married children in the bootstrap family sample. Lastly, based on simulated family networks of the bootstrap sample, we estimate the co-residence matching model using indirect inference by solving for the optimal co-residence assignments and then matching coefficients in a co-residence regression.

Our model provides several predictions related to intergenerational living arrangements that are later verified by the data: 1) the co-residence likelihood increases as housing, eldercare, and childcare costs increase, revealing the importance of economic factors in living arrangements; 2) when the congestion costs are sufficiently large, the co-residence likelihood declines with the income and education of parents and adult children, suggesting that co-residence is likely an inferior good; 3) the co-residence likelihood with one's own parents for a married adult declines as the number of his/her siblings increases or as the siblings' educational achievements increase due to the potential competition between siblings; and 4) the co-residence likelihood of parents living with their own children increases when their in-laws (parents of their children's spouse) have higher educational achievement or more children due to the potential competition between parents and parents-in-law.

We conduct counterfactual experiments to analyze how changes in housing prices and China's family planning policy affect the co-residence likelihood of different types of adult children and parents. For example, a 20% increase in the housing price increases the co-residence likelihood by 3.2% for adult children, and by 3.2% for parents overall. In particular, subgroups who originally benefit less from co-residence, highly-educated, female or older children and highly educated parents, are more sensitive to housing price changes. They experience larger increases in the co-residence likelihood as the housing price increases. We also predict the co-residence pattern when all parents have only one child. Compared to the baseline level, we show that adult children would experience a small increase in the probability of living with their own parents, from 19.9% to 21.0%. However, the co-residence likelihood of parents would drop significantly from 43.9% to 21.0% when they can only have one child. In particular, less educated parents would experience an even larger decline in the co-residence likelihood.

Our co-residence model is the first bilateral matching model between parents and adult children. Existing studies often model co-residence as a decision between one parent and one adult child in a cooperative or non-cooperative framework. For example, McElroy (1985), Pezzin and Schone (1999), and Hu (2001) use a Nash bargaining framework, and Rosenzweig and Wolpin (1993, 1994), Sakudo (2007), and Kaplan (2012) consider co-residence as a non-cooperative game. In addition, Konrad et al. (2002) and Maruyama and Johar (2013) model the strategic interactions between one parent and multiple children.[2] Our model, however, characterizes the living arrangement as a matching game, as it allows us to incorporate both the competition between adult children and the competition between parents and in-laws.

The mechanisms emphasized in our model are motivated by previous co-residence studies that focus on cost-sharing channels. Many studies have shown that co-residence helps adult children to insure themselves against poor labor market opportunities (Rosenzweig and Wolpin, 1993, Card and Lemieux, 2000, and Kaplan, 2012) and rising housing costs (Haurin et al., 1993, Ermisch, 1999, and Rosenzweig and Zhang, 2014). Several studies also suggest that co-residence is a way for parents to receive informal care from their adult children (e.g., Hoerger et al., 1996, Konrad et al., 2002, Dostie and Léger, 2005, Byrne et al., 2009, and Barczyk and Kredler, 2017) and a way for adult children to receive childcare support from parents (e.g., Compton and Pollak, 2014, Ma and Wen, 2016, and Garcia-Moran and Kuehn, 2017). We also include a flexible preference component in our model as taste and identity have been found to be contributing factors to the high rate of cohabitation in countries including China as shown by Giuliano et al. (2004), Manacorda and Moretti (2006), Xie and Zhu (2009), and Yasuda et al. (2011).

The rest of this paper is structured as follows. In Section 2, we describe a one-to-one, two-sided, transferable utility matching model in the co-residence context. We also discuss the importance of a family network in solving co-residence assignments and how we deal with an incomplete family network by developing a marriage matching model to predict missing marriage links. In Section 3, we introduce our data sources. In Section 4, we present the identification and estimation of the marriage and co-residence matching models. In Sections 5 and 6, we detail the estimation results and the counterfactual analyses, respectively. In Section 7, we conclude the paper.

[2] In addition, Pezzin et al. (2007) provide a theoretical framework for living arrangements and caregiving decisions in a two-stage model in which the living arrangement is determined in the first stage through a non-cooperative game between one parent and multiple children. Hiedemann and Stern (1999) and Engers and Stern (2002) model the interaction between siblings for long-term care decisions.

2 Model

We model the living arrangement as a one-to-one, two-sided, transferable utility matching game, similar to that in Shapley and Shubik (1971) and Becker (1973). In this model, parents are on one side of the market and adult children are on the other: each adult child can either live alone, with his/her own parents or with his/her parents-in-law; each parent can live alone or with at most one of his/her own children. Therefore, co-residence is a one-to-one match.[3] It is also a two-sided match as both parents and adult children must agree to the living arrangement. By assuming transferable utility, parents who are in favor of co-residing with one specific adult child can "buy" co-residence from this child by providing a sufficient monetary transfer; likewise, an adult child can win the co-residence competition against siblings by providing a monetary transfer to his or her parents.

2.1 Setup

There are two types of decision-makers in our model: parents and adult children.[4] All parents and adult children form two finite, disjointed sets of agents, a "parent" set and an "adult child" set. i and j denote a parent and an adult child, respectively. Note that we treat a parent couple or the only living parent as one decision unit, referred to as "parent."[5] All adult children in our model are married and we treat a child couple as a decision unit, referred to as "adult child." We use this simplification because our focus is the interactions between parents and adult children, thus we abstract away from the bargaining within the young couple.

We model co-residence decisions between parents and adult children as a static matching decision. Suppose all parents live for N years, where N is the life expectancy. When parents reach age N_0, adult children get married and decide whether to co-reside with their parents or parents-in-law. We assume that parents and adult children do not change their living arrangements once the decision is made at age N_0. While it is a restrictive assumption, this simplification is consistent with the relatively low frequency of changes in living arrangements that parents and adult children experience in China.[6] We further assume that parents and adult children, if living together, jointly allocate total household income into private and public consumption in every year based on a collective bargaining framework (Chiappori, 1992). We assume people consume all of their income in each period.

We first specify the utility function of living together for a parent-adult child pair. Fol-

[3]It is rare for one young couple to live with both sets of parents (only 0.03% of families in the CFPS) or for two married couples to live with the same parent (1.3% of families in the CFPS).

[4]Because grandchildren do not make decisions, they are not treated as a generation explicitly. Throughout the paper, we use parents as a benchmark, and therefore, when we use the term "grandchildren", we mean the grandchildren of parents (of those adult children).

[5]We also refer to them as the "elderly." We exclude families with divorced or separated parents. Only 0.2% of parents are divorced or live separately in our family sample of the CFPS.

[6]To verify this hypothesis, we merged the two waves of the CFPS data (2010 and 2012) and identified 2,733 married children. In 2010, among 1,160 married children who lived with their parents or parents-in-law, 109 moved out by 2012. Among 1,573 married children who lived alone in 2010, only 91 moved in with their parents or parents-in-laws by 2012.

lowing Chiappori (1992), a parent and an adult child choose their private goods (q_{it} and q_{jt}, respectively) and shared public goods (Q_t) in a way that maximizes a weighted sum of their individual utilities. Similar to Chiappori et al. (2017), we assume that the family's total utility depends on the multiplication of private and public consumption, i.e. $u_i = q_i^{\alpha} Q^{1-\alpha}$ with α and $1-\alpha$ capturing the effects of private and public consumption on utility, respectively.

This utility function is a special case of affine conditional indirect utility (ACIU), see Chiappori and Gugl (2015) and Chiappori and Mazzocco (2017). Chiappori and Gugl (2015) have proved that the transferable utility (TU) property of a matching model will be satisfied if each individual's preferences satisfy the ACIU property.[7] In other words, when the utilities of parents and adult children follow ACIU, the total utility of living together for a parent-adult child pair depends only on their combined income and not on the transfers between the parent and the adult child. Within the ACIU family, we choose the multiplication functional form so that the co-residence decision is affected by the parent's and child's income.[8]

Now, we present the utility functions of parents and adult children. For a parent i, the utility of living with an adult child j in period t, is expressed as follows:

$$u_{it} = \frac{1}{c_i} q_{it}^{\alpha} Q_t^{1-\alpha} \tag{1}$$

where q_{it} is the private consumption of parent i and Q_t is the public good shared between the parent and the adult child. c_i is the parent's congestion costs of co-residence, which captures the utility loss of shared residence due to the loss of privacy (when $c_i > 1$). We also allow for the case $c_i < 1$ in the model, when parents enjoy living together with their adult children.

For an adult child j, the utility of co-residing with one side of parents i in period t, is symmetric to that of their parents:

$$u_{jt} = \frac{1}{c_j} q_{jt}^{\alpha} Q_t^{1-\alpha} \tag{2}$$

where q_{jt} is the private consumption of the adult child j. c_j is the congestion cost of adult children when they live with their parents. We assume that the adult child shares the same public good, Q_t, as the parent.

[7]TU property is satisfied when there exists a cardinal representation of each individual's preferences such that the Pareto frontier is a straight line with slope equal to -1 for all prices and incomes. ACIU has the following form: $v_m(Q, q_m, p) = \alpha(Q, p)q_m + \beta_m(Q, p), m \in \{i, j\}$, where m represents the two parties in the matching game; Q is public consumption; q_m is the private consumption which varies by m; and p is the price of a private good.

[8]We have considered another candidate utility function that satisfies ACIU: $u_m = c_m + q_m$, where c_m is the congestion cost and q_m is the private consumption for $m \in \{i, j\}$. In this case, the co-residence surplus $S = c_i + c_j + Z$ (where Z represents the difference in the total expenditure between living together and living alone) does not depend on the level of income for either the parent or the adult child, which is inconsistent with our empirical findings that the co-residence likelihood decreases with the education of parents and adult children, which serves a proxy of their income, as shown in Table 1.

When the parent i and the adult child j live together, their budget constraint is

$$q_{it} + q_{jt} + Q_t = Y_{it} + Y_{jt} \equiv Y_{ijt}$$

Y_{it} and Y_{jt} are the incomes of parent i and adult child j in period t, respectively. Incomes in this model have a deterministic part $\tilde{Y}_{mt}$ and an idiosyncratic part ζ_{mt}, so

$$Y_{mt} = \tilde{Y}_{mt} + \zeta_{mt} \quad m \in \{i, j\}$$

where $\zeta_{it} \sim N(0, \sigma_p^2)$ and $\zeta_{jt} \sim N(0, \sigma_c^2)$.

When parents and adult children live together, they solve the following maximization problem:

$$\begin{aligned} U_{ij} = & \max_{q_{it}, q_{jt}, Q_t, t \in \{0, N-N_0\}} \sum_{t=0}^{N-N_0} \beta^t \left[\frac{1}{c_i} q_{it}^{\alpha} Q_t^{1-\alpha} + \mu \frac{1}{c_j} q_{jt}^{\alpha} Q_t^{1-\alpha} \right] \\ \text{s.t.,} \quad & q_{it} + q_{jt} + Q_t = Y_{ijt} \end{aligned} \tag{3}$$

where the Pareto weight of adult children is μ.

Given that the consumption allocation decisions (q_{it}, q_{jt}, and Q_t) in period t do not affect the income or consumption allocation decisions in period $t+1$, the maximization problem can degenerate from a dynamic problem (Equation (3)) to a static problem (Equation (4)):

$$\begin{aligned} U_{ij} = & \sum_{t=0}^{N-N_0} \beta^t \max_{q_{it}, q_{jt}, Q_t} \left[\frac{1}{c_i} q_{it}^{\alpha} Q_t^{1-\alpha} + \mu \frac{1}{c_j} q_{jt}^{\alpha} Q_t^{1-\alpha} \right] \\ \text{s.t.,} \quad & q_{it} + q_{jt} + Q_t = Y_{ijt} \end{aligned} \tag{4}$$

Solving the first-order conditions with respect to q_{it}, q_{jt}, and Q_t, we obtain the following:

$$\begin{aligned} \frac{q_{it}}{q_{jt}} &= \left(\frac{\mu c_i}{c_j} \right)^{\frac{1}{\alpha - 1}} \\ \frac{Q_t}{q_{jt}} &= \frac{1-\alpha}{\alpha} \left(\left(\frac{\mu c_i}{c_j} \right)^{\frac{1}{\alpha - 1}} + 1 \right) \end{aligned}$$

Defining

$$\begin{aligned} q_{it} &= A q_{jt} \\ Q_t &= B q_{jt} \end{aligned}$$

where $A = (\frac{\mu c_i}{c_j})^{\frac{1}{\alpha-1}}$ and $B = \frac{1-\alpha}{\alpha}((\frac{\mu c_i}{c_j})^{\frac{1}{\alpha-1}} + 1)$, we can solve the private consumptions and public consumption of both the adult child and the parent,

$$q_{it} = \frac{A}{1+A+B}Y_{ijt}$$
$$q_{jt} = \frac{1}{1+A+B}Y_{ijt}$$
$$Q_t = \frac{B}{1+A+B}Y_{ijt}$$

The utility of living together for a parent-child pair in period t is

$$U_{ij} = DY_{ijt}$$

where $D = \left(\frac{1}{c_i}\left(\frac{A}{1+A+B}\right)^{\alpha} + \frac{\mu}{c_j}\left(\frac{1}{1+A+B}\right)^{\alpha}\right)\left(\frac{B}{1+A+B}\right)^{1-\alpha}$. The total utility of living together depends only on the total income of parent i and adult child j, but not the transfer between i and j.[9]

Next, we turn to the case in which the parent and the adult child live separately. When living alone, parents and adult children maximize their utilities by choosing their private and public consumptions in each period conditional on their income. We normalize the congestion cost of living alone for parents and adult children to be 1. The utility of living alone for parent i in period t is

$$\begin{aligned} v_{it} = \max_{q_{it}, Q_t} \quad & q_{it}^{\alpha} Q_t^{1-\alpha} \\ \text{s.t.,} \quad q_{it} + Q_t = & \; Y_{it} - Z_i^{hc} - p^{ec}(A_{it}, E_i) Z^{ec} \end{aligned} \tag{5}$$

Notice that Z_i^{hc} is the housing cost paid by parents if they live alone. We assume that housing costs do not enter the utility function as a public good, but only serve as an expenditure. This assumption predicts that a rise in the housing price has a positive effect on the co-residence surplus, while the assumption that housing is a public good predicts that a change in the housing price has no effect on the co-residence surplus, which is inconsistent with our empirical evidence.[10] With probability $p^{ec}(A_{it}, E_i)$, parents get sick and need to pay Z^{ec} as the eldercare cost (such as going to a nursing home or hiring a helper) when living alone. If they live with their adult children, they do not need to pay Z^{ec} because adult children can provide informal care. The probability of needing eldercare is a function of a parent's age and education (A_{it} and E_i). In particular, the parent's age is $A_{it} = N_0 + t$.

[9] Transfers are implicit in the budget constraint. Parents and adult children pool their income together (Y_{ijt}) and decide their own private (q_{it} and q_{jt}) and public consumption (Q_t). Therefore, an agent can transfer money to the other to support his/her private consumption. Moreover, they agree to use part of the total income to buy public goods.

[10] Table 1 shows that the co-residence likelihood increases as city-level housing price rises.

Solving the first order conditions with respect to q_{it} and Q_t, we get

$$q_{it} = \alpha(Y_{it} - Z_i^{hc} - p^{ec}(A_{it}, E_i)Z^{ec})$$
$$Q_t = (1-\alpha)(Y_{it} - Z_i^{hc} - p^{ec}(A_{it}, E_i)Z^{ec})$$

Therefore, the utility of living alone for parent i in period t is

$$v_{it} = \alpha^\alpha(1-\alpha)^{1-\alpha}(Y_{it} - Z_i^{hc} - p^{ec}(A_{it}, E_i)Z^{ec})$$

Similarly, we calculate the utility of living alone in period t for an adult child j as follows:

$$v_{jt} = \max_{q_{jt}, Q_t} \quad q_{jt}^\alpha Q_t^{1-\alpha} \tag{6}$$
$$\text{s.t.,} \quad q_{jt} + Q_t = Y_{jt} - Z_j^{hc} - p^{cc}(A_{jt}, E_j)Z^{cc}$$

where Z_j^{hc} is the housing cost paid by adult children if they live alone. Here, an adult child's age is $A_{jt} = N_0 + t - G_{ij}$, where G_{ij} is the age gap between parent i and adult child j. We assume that adult children can have young children (those children who are too young to enter primary school, i.e. below age 6 in China) with probability $p^{cc}(A_{jt}, E_j)$, which is a function of the adult child's age A_{jt} and education E_j. If adult children live alone and have grandchildren, they incur childcare costs, Z^{cc}. These adult children will not pay for childcare if they live with their parents.

After solving the first order conditions with respect to q_{jt} and Q_t, the utility of living alone for adult child j in period t is

$$v_{jt} = \alpha^\alpha(1-\alpha)^{1-\alpha}(Y_{jt} - Z_j^{hc} - p^{cc}(A_{jt}, E_j)Z^{cc})$$

By combining the utility of living alone for both parties, we derive the total utility of living alone for an adult child-parent pair as Equation (7).

$$V_{ij} = \sum_{t=0}^{N-N_0} \beta^t(v_{it} + v_{jt}) = \sum_{t=0}^{N-N_0} \beta^t \alpha^\alpha(1-\alpha)^{1-\alpha}(Y_{ijt} - Z_{ijt}) \tag{7}$$

where the expenditure of living alone is $Z_{ijt} = Z_{ij}^{hc} + p^{ec}(A_{it}, E_i)Z^{ec} + p^{cc}(A_{jt}, E_j)Z^{cc}$, which includes housing costs, eldercare costs, and childcare costs. The total housing cost is defined as the sum of the housing costs of parents and adult children, i.e., $Z_{ij}^{hc} = Z_i^{hc} + Z_j^{hc}$. In this model, it does not matter whether the parent or the adult child owns the house because only the total housing costs of parents and adult children Z_{ij}^{hc} matter.

Once we have specified all of the utility functions in each scenario, the realized co-residence surplus is defined as the difference between the utility of living together and the utility of living alone for both parties, plus an idiosyncratic preference shock ϵ_{ij}, which follows a normal distribution $N(0, \sigma^2)$.

$$\begin{aligned}\tilde{S}_{ij} &= U_{ij} - V_{ij} \\ &= \sum_{t=0}^{N-N_0} \beta^t \left[DY_{ijt} - \alpha^\alpha (1-\alpha)^{1-\alpha}(Y_{ijt} - Z_{ijt})\right] + \epsilon_{ij}\end{aligned}$$

When parents and adult children make the co-residence decision, they cannot fully predict their future incomes. We define the ex-ante co-residence surplus (referred to as "co-residence surplus" from now on) as the expected surplus when the co-residence decision is made and the expectation is over the future income shocks.

$$\begin{aligned}S_{ij} &= E[\tilde{S}_{ij}] + \epsilon_{ij} \\ &= \sum_{t=0}^{N-N_0} \beta^t [D\tilde{Y}_{ijt} - \alpha^\alpha (1-\alpha)^{1-\alpha}(\tilde{Y}_{ijt} - Z_{ijt})] + \epsilon_{ij} \qquad (8)\end{aligned}$$

Based on Equation (8), a family's co-residence decision depends on the congestion costs that capture the reduced utility of living together (reflected in D) and expense savings in housing costs, eldercare costs, and childcare costs (reflected in Z_{ijt}).

2.2 Solution

The co-residence surplus, S_{ij}, predicts the living arrangement in a family. If $S_{ij} \leq 0$, the parent–child pair will not live together. Given that people from different families would not live together, we assume that the co-residence surplus is negative infinity among those who are not related, i.e. $S_{ij} = -\infty$, when i, j are not from the same family.

With the above assumption, we convert our model back to a general TU model as that in Shapley and Shubik (1971) and Becker (1973). We can solve the model by considering the following linear programming (LP) problem:

$$\begin{aligned}\text{maximize} \sum_{i,j} S_{ij} \cdot x_{ij} \qquad \qquad & \text{(LP)} \\ \text{subject to } (a) \quad \sum_i x_{ij} \leq 1 & \\ (b) \quad \sum_j x_{ij} \leq 1 &\end{aligned}$$

Definition 1. A **feasible assignment** for parents, P, adult children, C, and surplus, S, is a matrix $x = (x_{ij})$ (of zeros and ones) that satisfies (a) and (b).

We can say that $x_{ij} = 1$ if i and j form a co-residence and $x_{ij} = 0$ otherwise. If $\sum_j x_{ij} = 0$, then i lives alone. If $\sum_i x_{ij} = 0$, then j also lives alone.

Definition 2. A feasible assignment, x, is **optimal** for (P, C, S) if it solves the preceding LP problem.

With the assumption that the co-residence surplus shared between a pair of elderly people and young adults who are not in a parent–child relationship is negative infinity, the model becomes a general transferable utility matching model, and we can apply the results from Shapley and Shubik's (1971) assignment model to our context. Hence, all of the properties of a general transferable utility matching model hold for our co-residence matching framework. The co-residence assignment problem always has a solution, as the number of assignments is finite. Furthermore, the optimal assignment is unique as long as the surplus, S_{ij}, has no discrete mass.[11]

2.3 Predictions

The model also provides a useful framework for understanding how the co-residence surplus is affected by parent's and adult children's income and education, and by the parent–child age gap. In this section, we discuss the model predictions through the effects of housing, eldercare, and childcare costs on the co-residence surplus.

Recall that the co-residence surplus for a parent, i, and an adult child, j, is expressed as follows:

$$S_{ij} = \sum_{t=0}^{N-N_0} \beta^t [D\tilde{Y}_{ijt} - \alpha^\alpha (1-\alpha)^{1-\alpha} (\tilde{Y}_{ijt} - Z_{ijt})] + \epsilon_{ij}$$

where $D = \left(\frac{1}{c_i} \left(\frac{A}{1+A+B} \right)^\alpha + \frac{\mu}{c_j} \left(\frac{1}{1+A+B} \right)^\alpha \right) \left(\frac{B}{1+A+B} \right)^{1-\alpha}$, $A = (\frac{\mu c_i}{c_j})^{\frac{1}{\alpha-1}}$, $B = \frac{1-\alpha}{\alpha}((\frac{\mu c_i}{c_j})^{\frac{1}{\alpha-1}} + 1)$, and $Z_{ijt} = Z_{ij}^{hc} + p^{ec}(A_{it}, E_i) Z^{ec} + p^{cc}(A_{jt}, E_j) Z^{cc}$. According to the above equation, the co-residence surplus increases if there is an increase in housing costs, Z_{ij}^{hc}, eldercare costs, Z^{ec}, or childcare costs, Z^{cc}.

To examine the effect of the expected total income of a parent–child pair, $\tilde{Y}_{ijt}$, on the co-residence surplus, S_{ij}, we take the derivative of S_{ij} with respect to $\tilde{Y}_{ijt}$.

$$\frac{\partial S_{ij}}{\partial \tilde{Y}_{ijt}} = \beta^t (D - \alpha^\alpha (1-\alpha)^{1-\alpha})$$

If $(D - \alpha^\alpha(1-\alpha)^{1-\alpha}) < 0$, we obtain $\frac{\partial S_{ij}}{\partial \tilde{Y}_{ij}} < 0$, meaning that the co-residence surplus decreases with expected total income. Since D decreases with the congestion costs c_i and c_j, we will predict that the co-residence likelihood declines with income when the congestion costs exceed a certain cutoff.

We next discuss how the education of parents and adult children affects the co-residence

[11] Here are some other properties: if x is an optimal assignment, then it is compatible with any stable payoff (u, v). A stable payoff has two properties: 1) individual rationality, which reflects that a player always has the option of living alone; 2) no blocking pairs, which suggests that two pairs of elderly people and young adults do not want to break up their present co-residency and form a new co-residency by switching partners.

surplus. For parents, if education is positively associated with higher income and lower probability of needing eldercare, an increase in parental education leads to an increase in $\tilde{Y}_{it}$ and a decline in Z_{ijt}, and therefore, a decline in the co-residence surplus. For adult children, if an increase in education associates with higher income, then it negatively correlates with the co-residence surplus.[12] In sum, this model predicts that an increase in the parent's or child's education leads to a reduction in the co-residence likelihood.

The parent–child age gap also influences the co-residence surplus. We assume that the probability of having a young grandchild positively correlates with the parent–child age gap G_{ij}, given that families with younger married children are more likely to have young grandchildren. As our model assumes that parents make the co-residence decision at the same age and have the same life expectancy, we attribute more the effect of age-gap through the channel of childcare instead of eldercare. In addition, if younger adult children are more likely to have lower income, they will also value more the cost savings from co-residing with parents. Therefore, this model predicts that co-residence surplus increases with the parent–child age gap.[13]

In sum, the model generates the following predictions which we can test in the data:

Proposition 1. Co-residence likelihood increases with housing costs, eldercare costs, and childcare costs.

Proposition 2. When $(D - \alpha^{\alpha}(1-\alpha)^{1-\alpha}) < 0$, co-residence likelihood declines as the income and educational achievement of parents increases, it also declines as adult children have higher income or educational achievement. Moreover, the co-residence likelihood increases as the parent–child age gap increases.

2.4 Empirical Specification

In this section, we list all the empirical specifications used when we estimate the co-residence model. For simplicity, we set the parent's age when making the co-residence decision, N_0, to be 55. Life expectancy N, is set to be 75. As we cannot separately identify c_i, c_j, μ and α due to unobservable private and public consumption as well as the transfer between parents and adult children, we impose the following assumptions: 1) parents and adult children have the same congestion costs ($c_i = c_j = c_{ij}$); 2) the Pareto weight of adult children is one ($\mu = 1$), so parents and adult children have the same weight; and 3) the effects of private and public consumption on the utility are the same ($\alpha = 0.5$). The co-residence surplus for a parent-child pair then becomes

$$S_{ij} = \sum_{t=N_0}^{N} \left(\left(\frac{0.71}{c_{ij}} - 0.5 \right) Y_{ijt} + 0.5 Z_{ijt} \right) + \epsilon_{ij}$$

[12] The education of adult children may also affect the probability of having young grandchildren, which in turn affects the need of childcare, but we find no such correlation in our CFPS data, as shown in Table 4.

[13] If we relax the assumption that all parents make the co-residence decision at the same age, it is still likely that the larger parent–child age gap associates with higher co-residence surplus because older parents are in more need of eldercare.

In this surplus function, we specify the housing cost as a function of the housing price at the city level P_{ij},

$$Z_{ij}^{hc} = \delta_0 + \delta_1 P_{ij} \tag{9}$$

We allow the congestion cost to be different based on whether a young couple lives with the husband's parents or with the wife's parents to incorporate the "culture" component.[14]

$$c_{ij} = \kappa_0 + \kappa_1 H_j \tag{10}$$

where H_j is an indicator of whether lives with the husband's parents.

We then further assume both parent's and young couple's income are functions of education $(E_{i,j})$ and age $(A_{it,jt})$, as well as city fixed effects (ψ_k^p and ψ_k^c).

$$Y_{it} = \alpha_0^p + \alpha_1^p E_i + \alpha_2^p A_{it} + \alpha_3^p A_{it}^2 + \psi_k^p + \zeta_{it} \tag{11}$$

$$Y_{jt} = \alpha_0^c + \alpha_1^c E_j + \alpha_2^c A_{jt} + \alpha_3^c A_{jt}^2 + \psi_k^c + \zeta_{jt} \tag{12}$$

In addition, we assume that the expected probability of parents needing eldercare is a function of parents' average age and education, and similarly, the expected probability of having a young grandchild (aged below 6) is a function of young couple's average age and education, as well as city fixed effects (ϕ_k). Note that we focus on whether there is a young grandchild rather than the number of young grandchildren in family because grandchildren are subject to the one-child policy and each family can have at most one grandchild.

$$p^{ec}(A_{it}, E_i) = \begin{cases} 0 & \gamma_0^p + \gamma_1^p A_{it} + \gamma_2^p A_{it}^2 + \gamma_3^p E_i < 0 \\ 1 & \gamma_0^p + \gamma_1^p A_{it} + \gamma_2^p A_{it}^2 + \gamma_3^p E_i > 1 \\ \gamma_0^p + \gamma_1^p A_{it} + \gamma_2^p A_{it}^2 + \gamma_3^p E_i & \text{otherwise} \end{cases} \tag{13}$$

$$p^{cc}(A_{jt}, E_j) = \begin{cases} 0 & \gamma_0^c + \gamma_1^c A_{jt} + \gamma_2^c A_{jt}^2 + \gamma_3^c E_j + \phi_k < 0 \\ 1 & \gamma_0^c + \gamma_1^c A_{jt} + \gamma_2^c A_{jt}^2 + \gamma_3^c E_j + \phi_k > 1 \\ \gamma_0^c + \gamma_1^c A_{jt} + \gamma_2^c A_{jt}^2 + \gamma_3^c E_j + \phi_k & \text{otherwise} \end{cases} \tag{14}$$

[14] The traditional norm in China is that parents live with sons. In our CFPS data, 29% of young couples live with a husband's parents while only 6% of young couples live with a wife's parents. Thus, we suspect that congestion costs differ according to gender.

2.5 Family Network and Marriage Matching Model

In addition to the above empirical specifications, we need information on the family network to solve the co-residence matching problem. This is because the assignment game in our model places a restriction on the potential matching choices, which differs from the original Shapley–Shubik–Becker framework. Solving our co-residence matching model requires knowledge of the family network for the entire society being studied — a family network identifies whether two individuals are in a parent–child relationship or marital relationship. Often the data available does not contain a complete family network, so the application of our model requires us to provide a practical estimation method based on an incomplete family network.

Figure 3 illustrates an example of a family network in which individuals are represented by nodes and their social interactions (parent–child or marriage links) are represented by edges. Edges connect parent nodes to their adult children. Edges also connect adult child nodes to their parents and spouses. In this example, the society contains six parents, $i \in \{A, B, ...\}$. Adult children are indicated by ij, $j \in \{1, 2, ...\}$. For example, parent A has two adult children, $A1$ and $A2$.

The links in the middle of the diagram represent marriages of adult children. For example, child $A2$ is married to child $E1$, and, hence, parent A and parent E are connected through their children's marriage. Similarly, parent A relates to parent C through the marriages of $(A2, E1)$ and $(E2, C1)$. The family network shows that the co-residence decision of any individual in one of the six families could be affected by the co-residence decisions of all the other members in these six families.[15] The family network also tells how certain co-residence arrangements cannot be achieved, e.g. parent A would not want to live with child $B1$ because they are not in a parent–child or in-law–child relationship.

To estimate the family network, we find a comprehensive data set that contains all of the parent–child links among the individuals who are surveyed. However, that data only contains partial information on the marital links. Figure 4 illustrates the information available from our data, which are highlighted and include all of the links between parents and adult children and some of the links between husbands and wives. We observe the spouses of adult children living in the surveyed households but not those of adult children living elsewhere.[16] In this example, we observe couples $(A1, D1)$ and $(C1, E2)$, but we do not observe couples $(A2, E1)$, $(B1, F1)$, or $(C2, F2)$.

We therefore develop a procedure to recover the missing marriage links in the family network. To do this, we: 1) estimate a marriage matching model using data on the observed married couples; and 2) use the estimates from the marriage matching model to predict the

[15]Information on the family network of the society is necessary when we solve the co-residence decision, but for each family, only the pertinent network matters. In this example, these six families form a closed network, and the pertinent network of the six families only includes members from the six families, not the entire society.

[16]The absence/presence of marriage links is due to the way the survey was conducted. All children are surveyed, regardless of residence. However, information on the spouses of adult children is only available for adult children living in the surveyed household.

unobserved marriage links for married children. In below, we develop the theoretic marriage matching model (theory part for Step 1), then discuss the estimation of the marriage matching model in Section 4.2 (empirical part for Step 1) and the imputation of marriage links in Section 4.3 (Step 2).

We develop a transferable utility matching model for the marriage matching process. Let p (resp., q) denote the marginal distribution of types of males (resp., females) such that $p(x)$ (resp., $q(y)$) is the probability of being type x for a male (resp., y). Let $\Pi(p, q)$ denote the set of feasible matches between p and q. Then for any $\pi \in \Pi(p, q)$, we use $\pi(x, y)$ to denote the probability that a randomly chosen couple has male type x and female type y. Since we only focus on married men and women, $\sum_y \pi(x, y) = p(x)$ and $\sum_x \pi(x, y) = q(y)$ always hold.

The marriage surplus function, $\tilde{\Phi}(\tilde{x}, \tilde{y})$, is expressed as:

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = \Phi(x, y) + \chi(\tilde{x}, y) + \xi(\tilde{y}, x)$$

The marriage surplus function is composed of two parts, observable marriage surplus $\Phi(x, y)$, and marriage surplus unobservable to econometricians (but observable to agents) $\chi(\tilde{x}, y) + \xi(\tilde{y}, x)$. Here we denote $\tilde{x}$ and $\tilde{y}$ as the full types of men and women, and x and y as the observable types of men and women. We assume that the χ and ξ follow central Gumbel (type-I extreme value) distribution with scale factor σ_1 and σ_2, respectively. We impose a separability assumption (also used by Choo and Siow (2006), Chiappori et al. (2008), and Galichon and Salanié (2010)) on our marriage surplus function, which excludes interactions between the unobservable types of partners.

The observable part of marriage surplus between a man m and a woman w is defined as follows

$$\Phi(x, y) = f(x, y) + h(x, y) \tag{15}$$

The matching surplus for the pairing of a man with observable type x and a woman with observable type y, comes from two parts. The first part is the direct marriage utility $f(x, y)$, which captures the child couple's joint utility when they live without parents. The second part is the co-residence utility $h(x, y)$, which is non-zero if the couple co-resides with either side of parents.

In our marriage model, x and y include education, age, parental education, parental age, and number of siblings of the husband and wife, respectively. In the marriage matching literature, the demographic characteristics most frequently used to determine marriage surplus are age and education (e.g. Choo and Siow, 2006 and Chiappori et al., 2012). In addition, we include parental characteristics as they affect the direct marriage utility of the couple.[17] Moreover, a couple's number of siblings may affect the couple's direct marriage utility. Psychology

[17] See the discussion of parental involvement in the chapter "Household Formation and Marriage Markets in Rural Areas" in the Handbook of Development Economics vol. 4, 2007.

literature suggests that individuals from families of similar sizes often have similar personalities, which can relate to marriage compatibility (see Kelly and Conley, 1987 and Blake et al., 1991).[18]

Characteristics of the adult child couple and their parents also affect co-residence utility, $h(x, y)$. We make a few assumptions here. Some parents may want to influence their children's marriage by providing certain transfer so that they can benefit from co-residence after children get married; we assume away such strategic interactions between parents and adult children when adult children make their marriage decision. In addition, we assume that adult children cannot observe the entire family network, and therefore, adult children make expectations about the co-residence utility based on their own characteristics and their spouse's characteristics as well as the characteristics of their parents and their spouse' parents. These two assumptions simplify the model and avoids the non-existence of stable matching. Based on our co-residence model, the co-residence utility depends on the adult child's education, the parent's education, the parent-child age gap, and the adult child's number of siblings, which are captured by $h(x, y)$.

Note that we do not need to identify f and h separately because it is sufficient to identify Φ to estimate the marriage surplus and impute the marriage links. Therefore, we parametrize our marriage matching model by approximating the observable part of the marriage surplus, $\Phi(x, y)$, by a linear expansion over some known basis functions ϕ^k, with unknown weights λ:

$$\Phi(x, y) = \sum_k \lambda_k \phi^k(x, y) \tag{16}$$

In our case, we choose a quadratic approximation, following Fox (2018).[19] Therefore, the basic functions include the interactions between a man's characteristics and a woman's characteristics.

$$\begin{aligned}\Phi(x, y) =& \lambda_1 E_m E_w + \lambda_2 E_m A_w + \lambda_3 E_m PE_w + \lambda_4 E_m PA_w + \lambda_5 E_m N_w \\ &+ \lambda_6 A_m E_w + \lambda_7 A_m A_w + \lambda_8 A_m PE_w + \lambda_9 A_m PA_w + \lambda_{10} A_m N_w \\ &+ \lambda_{11} PE_m E_w + \lambda_{12} PE_m A_w + \lambda_{13} PE_m PE_w + \lambda_{14} PE_m PA_w + \lambda_{15} PE_m N_w \\ &+ \lambda_{16} PA_m E_w + \lambda_{17} PA_m A_w + \lambda_{18} PA_m PE_w + \lambda_{19} PA_m PA_w + \lambda_{20} PA_m N_w \\ &+ \lambda_{21} N_m E_w + \lambda_{22} N_m A_w + \lambda_{23} N_m PE_w + \lambda_{24} N_m PA_w + \lambda_{25} N_m N_w \end{aligned} \tag{17}$$

where E denotes own education, A own age, PE parent's education, PA parent's age, and N number of siblings.

In this quadratic approximation, the first order terms (such as E_m and E_w) and the quadratic

[18] Notice that we cannot use other time-variant variables, such as income or self-reported health status, to estimate marriage matching as this information is not observed at the time of marriage.

[19] In the early version of Fox (2018), the study uses the quadratic functional form in the empirical application, but this part is not included in the published version.

terms (such as E_m^2 and E_w^2) do not affect marriage sorting but only the probability of being single, while the interaction terms affect both.[20] Because we use this model to predict the marriage sorting behavior of married individuals, we only keep the interaction terms. This functional form can account for both vertical preference (i.e. E_m, E_w, people prefer partners who have higher educations) and horizontal preference (i.e., $(A_m - A_w)^2$, people prefer partners that are similar in age).[21] This general functional form includes a flexible set of the interaction terms: 1) that between an adult child's characteristics and his/her spouse' characteristics, 2) that between a parent's characteristics and parent-in-law's characteristics, and 3) that between an adult child's characteristics and his/her parent-in-law's characteristics.

In sum, we incorporate the characteristics of not only young men and women but also their parents into our model. These characteristics may affect a couple's marriage surplus through their co-residence decision. Our specification allows us to capture the positive assortative matching in education, age, number of siblings, parental education, and parental age. After we estimate the $\lambda' s$, we can use λ to predict the missing marriage links in our co-residence model.

3 Data

3.1 Source and Sample Construction

Our main analysis combines a micro-level dataset, the CFPS, and city economic indicators from the "City Statistical Yearbooks," published by the National Bureau of Statistics of China. The CFPS surveys a nationally representative sample of Chinese households, covering 25 provinces and municipal cities in China. It uses a multi-stage probability sampling method to randomly survey households within each county or community. All members over age 9 in a sampled household are interviewed. Family members are defined as financially dependent immediate relatives, or non-immediate blood/marital/adoptive relatives who have lived in the household for more than three consecutive months and are financially related to the sampled household.

The 2010 CFPS includes approximately 15,000 households and 33,600 individuals. For each household, the CFPS provides demographic information for every family member regardless of residence location. It also provides a complete family relationship map. The basic demographic information of non-resident members was collected through relatives who lived in the surveyed households. The relationship map helps us identify the parent-child and marriage relationships in a family network for the individuals in the sample. In particular, we observe all of the parent–child links of those surveyed. We also observe spouses of adult children living in

[20]Consider two men x_i, x_j and two women y_k, y_l, whether $\Phi(x_i, y_k) + \Phi(x_j, y_l)$ is bigger than $\Phi(x_i, y_l) + \Phi(x_j, y_k)$ does not depend on the first order terms and the quadratic terms. Therefore, the first order terms and the quadratic terms do not affect the marriage sorting behavior.

[21]The quadratic terms in the horizontal preference is omitted in our surplus function because they do not affect the marriage sorting behavior of married children. Given that we only observe matching patterns, we cannot distinguish between vertical and horizontal preferences. Both types of preferences predict positive assortative matching.

the surveyed households but not those of adult children living elsewhere.

As the CFPS is a household-based survey but our model is family-based, we re-organize the data to construct the samples used in the estimation. First, we construct a couple sample, which is used to estimate our marriage matching model. Next, we expand the couple sample to a family sample, which is used to estimate our co-residence matching model.

The couple sample includes all observed married couples that meet the following criteria: 1) the husband and wife are both urban local residents; 2) the ages of parents of the husband and wife are between 55 and 75; and 3) all of the siblings of the husband and wife are married. An observation in the couple sample is referred to as "focal couple". The restriction on parental age makes sure that our focal couples are from the adult child generation, not from the parent generation or the grandchild generation. With this restriction, the grandchildren in a family are young; most parents of focal couples are retired and have time to provide childcare to grandchildren by age 55.[22] We then further restrict couples to be urban and local. A couple is identified as urban and local if they live in the urban area with parents and siblings in the same city. Lastly, we only keep adult children who are married and whose siblings are all married, because the co-residence incentive of single adult children may be quite different from that of married adult children. A total 659 couples satisfy our sample restrictions. In this couple sample, we have complete information about the demographics of the husbands and wives, their siblings, and parents.

We then construct a family sample by pooling families where the husbands and wives are from. A family is defined as a parent couple and all of their children. In other words, we expand the couples to include their parents and siblings so the number of families should be the number of couples times two. We find two families with two adult child couples living in the same family so we only keep the oldest child couples in the two families to avoid double counting when doing the expansion. We eventually have 1,314 families that form the set of observations in the family sample. Last, we combine all the couples and their married siblings to form an adult child sample composed of 3,709 married individuals. Although individual questionnaires are only available for adult children residing in the households interviewed by CFPS representatives, we can still use family-reported data in the CFPS to track down basic demographics such as sex, age, education, and marital status for those who do not complete individual questionnaires.[23] Notice that the way we construct the family sample and the adult child sample oversamples adult children with more siblings. This is because if an individual is one of the N children in a family, the probability that he/she appeared in the family sample/adult child sample is equal to the probability that at least one of the children in his/her family is sampled, which is proportional to the number of children in the family (N). As a result, when

[22]Retirement in China is mandatory for most people. Male workers mostly retire by 60 and female workers mostly retire by 50-55, depending on their profession.

[23]The CFPS survey team attempts to find adult children if they do not reside in the surveyed household but live in the same city as the surveyed household. Therefore, some of the adult children who are not living in the surveyed household also answer an adult questionnaire.

we use the family sample or the adult child sample, we use $1/N$ as a weight to adjust the oversampling problem.

One particular restriction imposed on our sample is location, because we want to take location choice out of the picture in the model. Although rural-urban migration is an important phenomenon in the Chinese labor market, our model cannot feasibly incorporate this additional location decision.[24] Adding a location choice would significantly complicate our model because it would require a model that jointly captures marriage decisions, co-residency choices, and location decisions.[25] We use 2010 micro-level Census data to show how our sample's characteristics differ from the population characteristics. The average years of education of urban local residents aged between 18 and 55 is 11.9, while that of the rest of China is 9.4. The average number of offspring among females in our sample is 1.2, while that of the rest of China is 1.6. However, the urban-local restriction does not bias our results because a majority of local urban residents marry other local residents. In our data, only 16% of local young adults married non-local residents. We therefore assume that the urban marriage market is a closed market for local residents and focus on the co-residence decision of local urban residents.

We use a supplementary dataset called the China Health and Retirement Longitudinal Study (CHARLS) to predict parent's health status as the CFPS lacks of accurate information on parent's health status. The CHARLS has a nationally representative sample of Chinese residents aged 45 years and above. We use the baseline national wave in 2011, which includes approximately 10,000 households and 17,500 individuals in 150 counties or districts to match the timing of the CFPS data. The CHARLS provides rich information on the health status and the demographics of the elderly. By restricting to local urban residents aged between 55 and 75, we use CHARLS data to predict elderly parents' health statuses. Summary statistics of CHARLS data are shown in Appendix Table D8.

3.2 Descriptive Statistics

Table 2 reports the summary statistics of the key variables. We first present statistics at the individual level. We report the weighted mean of the couple sample and the family sample using the inverse of the number of adult children as the sample weight. In our weighted adult child sample, 20.4% of the 3,709 urban married adults live in the same housing unit as their own parents. The average age of these urban adults is 37.5 and they have an average of 10.2 years of education. The average parent–child age gap is 27.3 years. In our weighted family sample, the parents of adult children are, on average, 63.7 years old, have 6.2 years of education, and have 2.2 adult children. As most of the individuals in our adult child sample were born before China's one-child policy was implemented, we observe that most parents have more than one

[24]Urban-to-urban migration is relatively rare in China. According to the 2010 Census micro-sample, 13.8% of individuals with non-agricultural residence-permits (hukou) migrated to other cities.

[25]There are two papers that jointly model marriage and location choices. Dupuy (2018) uses a matching model to capture marriage and location choices jointly. Gemici (2011) uses a dynamic model with intra-household bargaining to jointly model marriage and location choices, abstract from competition.

adult child. In our couple sample, the 659 married couples have 0.3 young children under six years old on average and they are more likely to live with the husband's parents than the wife's parents.[26] Specifically, 29% co-reside with the husband's parents and only 6% co-reside with the wife's parents. Lastly, we show summary statistics related to housing prices at the city level collected from the statistical yearbook. The average price per square meter is 3,988 Yuan (approximately $600 USD). We scale the housing price by multiplying the average housing price per square meter to the average area per capita in that city and then dividing the product by GDP per capita in the specified city.

We further check the correlation between an adult child's probability of living with his/her own parents against child's, parent's, and city-level characteristics. Column (1) of Table 1 presents the results using the CFPS data. Overall, men are more likely to live with their parents than women. Children are less likely to live with highly educated parents, as such parents have more income and are more likely to be healthy. When adult children have higher education attainment themselves, their likelihood of co-residing decreases due to the negative income effect. We also observe that a larger age gap between parents and adult children increases the chance of co-residing. This is because as the age gap between an adult child and his or her parents increases, the family is more likely to have young grandchildren. Lastly, high housing costs are associated with a high probability of co-residing with one's own parents.

These reduced-form results help us to verify the key components of our model. For example, we observe an adult child's likelihood of co-residing with his or her own parents decreases as his or her parents' predicted income rises, though Manacorda and Moretti (2006) find that in Italy, a parent's likelihood of co-residence increases with his or her income. This difference suggests that it is necessary to have a co-residence model with flexible relationship between parents' income and their likelihood of co-residence with their children. In China's case, the income effect is negative, indicating that co-residence is an inferior good.

4 Identification and Estimation

4.1 Identification of the Co-residence Matching Model

Our co-residence model includes sets of both endogenous and exogenous parameters. In this section, we explore the identification of endogenous parameters, which include 1) congestion costs, 2) housing costs, eldercare costs, and childcare costs, and 3) the distribution of idiosyncratic shocks. We defer the discussion of the estimation of the exogenous parameters, including parameters in the income equations of parents and adult children, parent's health status, and the fertility of adult child couples, to Section 4.3.

To prove the identification of our model, we show how a subset of data can identify the

[26]In the CFPS data, when parents are aged 55-57, the fraction of adult child couples with children older than six is 30%. Given that parents need to work before the retirement stage, they cannot take care of their grandchildren if the grandchildren need childcare before they reach retirement age. Therefore, by the time parents make the co-residence decision, around 30% of young couples no longer need childcare because their children are already too old (older than six).

above three sets of endogenous parameters. As the parameters in the co-residence model do not depend on whether adult children have siblings or not, we use a group of couples in which neither the husband nor the wife has any siblings to show identification. The advantage of focusing on this subsample is that there is no competition among siblings, and therefore, our matching model degenerates into a multinomial choice model. Within this special couple sample, young couples choose whether to live alone, with their own parents, or with their spouse's parents.[27] Below, we introduce how to achieve the identification via three steps in this multinomial choice model.

We start by proving that the congestion cost parameters in our model are identifiable. According to Equation (18), the marginal effect of the total expected income of ($\tilde{Y}_{ijt}$) on the co-residence surplus can be mapped one-to-one into the congestion cost (c_{ij}).

$$\frac{\partial S_{ij}}{\partial \tilde{Y}_{ijt}} = \beta^t \left(\frac{0.71}{c_{ij}} - 0.5 \right) \tag{18}$$

where β, the annual discount rate is set to be 0.95. Equation (10) indicates that congestion costs (c_{ij}) include two specific parameters: an average effect (κ_0) and a gender difference parameter indicating whether a young couple prefers to live with the husband's parents or wife's parents (κ_1). A stronger income effect on the co-residence likelihood suggests a larger κ_0. A larger gap between the co-residence rate of living with husband's parents and that of living with wife's parents indicates a larger κ_1.

Next, we describe how we identify housing cost, eldercare cost, and childcare cost parameters. First, the total housing cost (Z^{hc}_{ij}) contains two parameters: national average housing cost (δ_0) and city-specific housing price (δ_1), as shown in Equation (9). The average co-residence likelihood among all families pins down the national average housing cost, δ_0, as higher average housing costs are associated with larger co-residence surpluses. Then as shown in Equation (19):

$$\frac{\partial S_{ij}}{\partial P_{ij}} = \sum_{t=0}^{N-N_0} 0.5\beta^t \delta_1, \tag{19}$$

the marginal effect of city housing price on the co-residence likelihood identifies the city-specific housing price parameter.[28]

[27]In our sample, 44 out of our total 659 couples fall into this special case. The summary statistics of these 44 couples are shown in Appendix Table D9. The small sample is not a problem for identification, as identification is purely a theoretical issue. We use a larger sample (all 659 married couples) in our later estimation of the co-residence model.

[28]In the estimation, a single city is considered a closed market for both marriage and co-residence decisions. Our estimation allows each city to have a different income, fertility rate (through city fixed effects when we estimate income and fertility outside the model), and average housing price. Thus, we allow for the possibility that cities with higher housing prices also have higher incomes and/or lower fertility rates. We can get an unbiased estimate of the housing price parameter (δ_1) if housing price differences across cities are driven by some unobserved factors (e.g., land supply), which do not directly affect co-residence likelihood.

Second, the eldercare cost, Z^{ec}, corresponds to the marginal effect of parent's education on the co-residence likelihood as shown in Equation (20).

$$\frac{\partial S_{ij}}{\partial E_i} = \sum_{t=0}^{N-N_0} \beta^t \left(\left(\frac{0.71}{c_{ij}} - 0.5 \right) \alpha_1^p + 0.5 \gamma_1^p Z^{ec} \right) \tag{20}$$

This equation shows that parental education affects the co-residence surplus through the income (captured by α_1^p) and health (captured by γ_3^p) of parents. Since c_{ij} has been identified in the first step and α_1^p and γ_3^p are exogenous parameters, the eldercare cost is identified from Equation (20).

Third, the childcare cost, Z^{cc}, is identified through the marginal effect of parent–child age gap on the co-residence likelihood. Our model assumes that parents make the co-residence decision at the same age so that the parent–child age gap only affects the child's age. Parents can always enjoy eldercare from age 55 to 75, which is not affected by the parent–child age gap. By taking the partial derivative of the co-residence surplus with respect to the parent-child age gap (see Equation (21)), we can see that the parent-child age gap affects the co-residence surplus through the income and childcare cost of adult children. Once c_{ij} has been identified and exogenous parameters, $\alpha_2^c, \alpha_3^c, \gamma_2^c$, are separately estimated, the childcare cost parameter (Z^{cc}) is one-to-one mapped into the marginal effect of the parent–child age gap on the co-residence likelihood.

$$\frac{\partial S_{ij}}{\partial G_{ij}} = -\sum_{t=0}^{N-N_0} \beta^t \left(\left(\frac{0.71}{c_{ij}} - 0.5 \right) (\alpha_2^c + 2\alpha_3^c G_{ij}) + 0.5 \gamma_2^c Z^{cc} \right) \tag{21}$$

Finally, we use the fraction of living arrangement that cannot be explained by the model's parameters to estimate the distribution of the idiosyncratic shock on the co-residence surplus ($\epsilon_{ij} \sim N(0, \sigma^2)$).

4.2 Marriage Matching Estimation

As discussed earlier, solving our co-residence matching model requires us to fill in the missing marriage links in the family networks that are unobserved in the data. To complete all the marriage links, we estimate a marriage matching model with transferable utility. However, we face several challenges in this estimation. First, we have a relatively small sample – there are 659 couples in our couple sample. As we have a marriage model with multiple characteristics and the data have limited sample size, we cannot achieve the ideal precision if we estimate the marriage model non-parametrically following the seminal work of Choo and Siow (2006).[29] Second, we cannot follow Fox (2010, 2018) to estimate the parametric model using the max-

[29] In our marriage matching model, a man/woman has multiple characteristics, including education, age, number of siblings, parental education, and parental age. We define a combination of education, age, number of siblings, parental education, and parental age as the man/woman's type. If we want to adopt the method of Choo and Siow (2006), we need to calculate the frequency of matches for each type of man and woman, but the number of men who belong to a specific type is too small in our sample.

imum score estimator either, because this approach does not estimate the distribution on the error term. However, knowing the dispersion of the shocks is necessary when we simulate the marriage matching to estimate the co-residence model.[30]

To overcome these challenges, we follow Galichon and Salanié (2010) by using a moment matching estimator and set up a parametric marriage surplus function as shown in Equation (17). Galichon and Salanié (2010) provide a parametric procedure to estimate a marriage matching model with multiple characteristics, which allows us to overcome the small sample problem and estimate the scale of the unobserved shocks (relative to the observable surplus).[31]

Galichon and Salanié (2010) provide a moment estimator to estimate the λ's. In particular, the following covariations serve as our targeted moments. Consider any matching π. Define the covariations $C(\pi) = \big(C^1(\pi), C^2(\pi), \dots\big)$ such that:

$$C^k(\pi) = \sum_{x,y} \pi(x,y)\phi^k(x,y) \ .$$

We normalize $\sigma_1 + \sigma_2 = 1$.[32] We estimate λ's by minimizing the distance between the observed covariations and predicted covariations. Galichon and Salanié (2010) provide a fast and stable algorithm to predict the optimal matching and match the covariations. The details are shown in Appendix B. Standard errors are calculated using the bootstrap method.

Given that this approach only applies to characteristics with discrete values, we make the following adjustments to convert our variables into discrete values: 1) the education of adult children ranges from one to four, corresponding to less than junior high school or drop-outs, junior high graduates, high school graduates, and college graduates; 2) the age of adult children ranges from one to four, corresponding to below 34 years of age, between 34 and 38, between 39 and 43, and 43 and above; 3) the number of siblings ranges from one to two, corresponding to zero or one sibling, and more than one sibling; 4) the education of parents ranges from one to three, corresponding to less than primary school, primary school graduates, and junior high graduates; and 5) the age of parents ranges from one to two, corresponding to age below 65, and age above 65. Each type of men x (women y) is a combination of these five characteristics. In total, we have 192 types. Since we only have 659 couples in our couple sample, we cannot have too many types. However, the approach of Galichon and Salanié (2010) is quite robust for small samples.

One potential concern for our marriage link estimation is that the procedure is subject to selection bias, as we only observe a fragment of the marriage links in our data. Selection

[30] Although Fox's method is more general in the sense that it does not impose any functional assumptions on the error term, it cannot estimate the distribution of the shocks. Fox's method normalizes the coefficient of the first attribute (λ_1) to be one.

[31] We would like to thank the anonymous referee for pointing out this reference. The approach proposed by Galichon and Salanié (2010) does not need to normalize the coefficient of the first attribute (λ_1) to be one, and therefore allows us to estimate the scale of the unobserved shocks relative to the observable surplus.

[32] As pointed out in Galichon and Salanié (2010), we cannot separately identify σ_1 and σ_2.

bias may exist if the observed marriage is not a representative sample of all marriages in the population. We now discuss why the observed marriages in the couple sample is not a selective sample for the purpose of imputing missing marriage links and estimating the co-residence model.

Observed marriages in the original couple sample is a random sample of all marriages in the population. Married children often live in separate households, and each household has the same probability of being surveyed in the CFPS because the CFPS provides a random and representative sample of households.[33] As a result, each child couple is sampled with equal probability. In addition, the probability of each couple being sampled does not depend on whether the couple lives with either side of the parents. We use Figure 4 to illustrate an example, where Parent A has two adult children, A1 and A2. Suppose A1 lives with Parent A and A2 lives alone. $(A, A1, D1)$ and $(A2, E1)$ form two households and they will be sampled with equal probability.[34] Therefore couples $(A1, D1)$ and $(A2, E1)$ will be sampled with equal probability, which does not depend on the co-residence status.

Next, we impose restrictions on the original couple sample in terms of hukou status, parental age, and siblings' marital status, as mentioned in the Data section. Therefore, our restricted couple sample is no longer a representative sample of the national population, but the selection bias is of second-order importance given the nature of the restrictions.[35] More importantly, the restricted couple sample is consistent with the family sample and the bootstrap sample. All married adults in the restricted couple sample and the family/bootstrap sample are selected via the same set of criteria. As introduced earlier, the married adult children in the family/bootstrap sample are just husbands and wives in the restricted couple sample as well as their siblings.[36] The siblings of focal couples also satisfy the same restrictions as focal couples in terms of hukou status, parental age, and siblings' marital status. As a result, the restricted couple sample can provide consistent estimates for the marriage matching model that allows us to impute the missing marriage links in the bootstrap sample.

In summary, although our restricted couple sample is a selective sample of the Chinese population, there is no selection bias in terms of the co-residence model estimation. This is because observed marriages (of focal couple) are drawn from the same distribution of unobserved marriages (of focal couples' siblings who were not surveyed). Therefore, we can use the restricted couple sample to estimate the marriage matching model, and then use the estimates to impute the marriage links in the bootstrap sample.

---

[33] In our data, only 2 out of 1,316 families have two married children living in the same household.

[34] In this example, we observe $(A, A1, D1)$ but not $(A2, E1)$. Although we obtain the basic demographic information of A2, but we do not observe E1.

[35] The restricted couple sample is introduced in the Data section. We believe that these are mild restrictions because we want to analyze the co-residence decision of urban local married adult children and their parents.

[36] The bootstrap sample is sampled from the family sample, so the two samples have the same distribution of the characteristics of parents and adult children.

4.3 Co-residence Matching Estimation

Once we prove the identification of our co-residence model and the feasibility of our marriage matching estimation, we introduce the co-residence model estimation in the following four steps.

1. Sample 1,314 families for each of the 104 cities from the CFPS national family sample to form the bootstrap sample, which only contains the links between parents and adult children, not the marriage links.

2. Based on the estimates of the marriage matching model, complete the family networks in each city by simulating the marriage links for all married children in the bootstrap sample.

3. Estimate the parameters that are exogenous to our model, including parameters in the income equations of parents and adult children, parent's health status, and the fertility of adult child couples.

4. Simulate optimal co-residence assignments in each city using the bootstrap sample with simulated marriage links.

The first step is the sampling step, where we bootstrap a family sample for each of the 104 cities from the CFPS national family sample (1,314 families). As each city has only 2–128 families, we use sampling to expand our data set.[37] For each city, we randomly draw, with replacement, 1,314 families from the national family sample. We refer to this as a "bootstrap sample." When we sample the families, we use the inverse of the number of adult children in the family as a weight because the adult child sample oversamples adult children with more siblings. We only bootstrap the links between parents and adult children, not the marriage links. We choose to sample 1,314 families for each city because we would like the bootstrap sample for each city to have the same sample size as the national sample, to maintain the statistical power of the original sample. As a result, the bootstrap sample for each city has the same distribution of family characteristics in terms of age, education, number of siblings, parental age, and parental education for adult children. Notice that each city has different realized family characteristics due to the randomness in the sampling process.

Our second step is to simulate marriage matching within each city to complete the family networks in our sample. We assign a spouse for all young men and women who report being married in the city-level bootstrap sample. The assignment is based on the demographics of young men and women, the marriage matching estimation results obtained in Section 4.2 (λ's), and simulated marriage matching shocks (ξ) for each pair of potential spouses. We ensure that

[37] Using only the existing city-level sample in our estimation poses two potential problems. First, the city-level sample may not be representative. Second, there may be too few men and women that are eligible to date and become married within each city.

no two individuals are assigned the same spouse by simulating marriage matching shocks from the type I extreme value distribution for each potential husband-wife pair; the shocks guarantee that no two pairs have the same marriage matching surplus.[38] Hence, there is a unique solution to the optimal marriage matching, which maximizes the social marriage surplus. We omit city-level heterogeneity in marriage matching because the marriage model is estimated using the national couple sample.[39] Each city can still have different realized couple characteristics due to the randomness in the simulation process.

In the third step, we estimate the parameters that are exogenous to our model. We estimate adult children's income, following Equation (12), using a sample of urban, local residents with parents aged 55 to 75 from the CFPS data. We estimate parental income, following Equation (11), using a sample of urban local residents aged 55 to 75 from the CFPS data. We estimate parents' health, following (13), using a sample of urban, local residents aged 55 to 75 from the CHARLS data. We estimate the number of grandchildren (number of children for each child couple), following Equation (14), using our couple sample. We also estimate the city fixed effects of incomes of adult children and parents using our CFPS data, and estimate the city fixed effects of the number of grandchildren using the 2010 Census data.[40] Notice that we predict parents' income, adult children's income and the number of grandchildren for each city's bootstrap sample using individual demographics with city fixed effects. This method allows each city to have different average levels of income and number of grandchildren. We also predict parents' health status using the parent demographics in the CHARLS data. We cannot add the city fixed effects in the prediction of parents' health status because the sampling cities are different between the CHARLS and the CFPS.

In the last step, we simulate the optimal co-residence assignments in each city based on the family demographics obtained from step 1, the marriage link constructed in step 2, and the predicted income, health, and fertility outcomes from step 3, as well as the city-level housing price observed in the data and simulated co-residence surplus shocks (ϵ_{ij}).

We repeat the above steps 100 times to obtain 100 simulated samples for each city. In other words, we use Monte Carlo simulation to integrate the shocks in marriage matching and co-residence matching. We then use indirect inference to estimate the co-residence model. The target of the indirect inference comes from the regression result based on the CFPS adult child sample. More specifically, we regress the indicator of an adult child living with his/her own parents on his/her demographic characteristics, parents' characteristics, and the local city's characteristics. We then use our co-residence model to predict the co-residence decisions of

[38] Based on our model, the marriage surplus is subject to two shocks ($\chi(\tilde{x}, y)$ and $\xi(\tilde{y}, x)$) and we assume that the two shocks follow a central Gumbel distribution with scale factors σ_1 and σ_2. We normalize $\sigma = \sigma_1 + \sigma_2 = 1$ and we cannot separately identify σ_1 and σ_2. Therefore, we choose $\sigma_1 = \sigma_2 = 0.5$.

[39] As we only have between 1 and 74 couples in each city, we cannot estimate the marriage model separately for each city.

[40] Given that we only have around 659 couples in our couple sample, we cannot precisely estimate the city fixed effect of the number of grandchildren using the CFPS data. Therefore, we estimate the city fixed effect using the 2010 Census data and merge it with the CFPS data.

individuals in the bootstrap sample and run the same reduced-form regression. Indirect inference sets parameters so that the co-residence regression coefficients implied by the model are as close as possible to the coefficients we obtain directly from our data. A weighted squared deviation between the sample statistics and their simulated analogs is minimized with respect to the model's parameters. The weights are the inverse values of the standard errors of the regression coefficients. We use the bootstrap method to calculate the standard errors of our parameter estimates.

In sum, our co-residence estimation imposes two assumptions. First, it assumes that the distribution of family characteristics is the same across cities. Second, it assumes that the distribution of married couples is the same across cities. However, we allow each city to vary when it comes to housing price, parental income, adult children's income, and the number of grandchildren. The variations are shown in the last panel of Table 2. As a result, the co-residence patterns differ across cities.

Note that the approaches of Choo and Siow (2006) and Fox (2018) are not applicable to the co-residence matching model. Due to the small sample, we cannot recover the co-residence surplus non-parametrically as in Choo and Siow (2006). Fox (2018) requires information on every co-residing pair. However, part of the family network is missing in our data. In particular, we do not observe whether adult children live with parents-in-law if they do not live in the surveyed household. Therefore, we must simulate the family network. In our case, we use indirect inference rather than the simulated method of moments (SMM) because indirect inference allows us to target the co-residence regression coefficients rather than the co-residence likelihood conditional on a single characteristic.[41] For example, we find that the co-residence rate has a non-monotone relationship with adult children's education (see Appendix Table D7) because the number of siblings is negatively correlated with an adult child's education and highly educated children face less competition from their siblings. Once we control for the number of siblings in the co-residence regression, the coefficient on an adult child's education becomes negative (see Table 1), which reflects the marginal effect of an adult child's education. By targeting regression coefficients, we can analyze the marginal effect of one characteristic holding other characteristics constant.

5 Estimation Results

We summarize the estimation results of marriage matching and co-residence matching in this section. To fill in the missing marriage links in the data, we estimate the marriage matching model with transferable utility based on our couple sample in the CFPS. Table 3 shows the estimation results using the moment matching estimator with five dimensions of couples' characteristics: education, age, number of siblings, parental education, and parental age. We find that the parameters on the diagonal are all significant, suggesting positive assortative matching

[41]The co-residence likelihood conditional on the joint characteristics of adult children, parents, and the city is imprecise because of the small sample size.

in all five dimensions.

Additional evidence in Appendix Tables C1 to C5 evaluates the fit of our marriage matching model. In Panel A of each table, we calculate the empirical probability of marriage conditional on husband's and wife's characteristics using the couple sample. In Panel B of each table, we use our model's predicted marriages to calculate the same conditional probability based on the couple sample as that in Panel A. The closeness in the marriage probability between Panel A and Panel B indicates that our model predicts the marriage patterns well. In Panel C of each table, we predict the marriage matching of all adult children in the bootstrap sample. By comparing Panels A, B and C, we can see that bootstrapping procedure preserves the marriage patterns observed in Panel A and B.

We then present the co-residence model estimation results in two steps. In the first step, we estimate parameters that are exogenous to our model. These include the incomes of parents and adult children, parental health status, and the fertility status of the adult child couple. In the second step, we estimate the endogenous parameters, which include 1) congestion costs, 2) housing costs, eldercare costs, and childcare costs, and 3) the distribution of idiosyncratic shocks.

We use a Mincerian-type regression to predict the incomes of parents and adult children, following Equations (11) and (12).[42] More specifically, we regress the log income of parents (summing up the income of a father and mother) on their average age and education plus city fixed effects, see Columns (1) and (2) in Table 4. We run a similar income regression for young married couples (using their joint income), and the results of which are available in Columns (3) and (4) of Table 4. The estimation results suggest that parental income increases with education and decreases with age. In addition, an adult child's income increases as his/her education increases and has a hump-shape life-cycle profile. We use the estimates in Column (1) to predict a parent's income and the estimates in Column (4) to predict an adult child's income in the next step's estimation.

We use a linear probability model to predict a parent's health status based on the parent's age and education, following Equation (13). We define a parent as unhealthy if his/her self-reported health status is poor.[43] Columns (5) and (6) of Table 4 show that the proportion of parents who are in poor health increases with age and decreases as their education levels increase. For our co-residence model estimation, we use Column (5) to predict the likelihood that parents are in poor health and therefore need eldercare.[44] To predict whether a family has a grandchild under six years old, we use a linear probability model following Equation (14), and the regressors include the average age and education level of the couple and city fixed effects.

[42]The incomes of parents and adult children include both labor income and asset income.

[43]The five categories of the health status are "excellent", "very good", "good", "fair" and "poor". The self-reported "poor" health status accounts for 24% of the selected sample from the CHARLS data.

[44]We have investigated whether the probability of being in poor health is correlated with the co-residence status using the CHARLS data but find no significant correlation. We believe that a parent's health status largely depends on exogenous shocks and predetermined factors rather than co-residence status.

Columns (7) and (8) of Table 4 show that a family unit's probability of having a young child declines with couple's age but does not vary by their educational achievement. We use the estimates in Column (7) to predict the likelihood of having grandchildren in our co-residence model estimation.

In the second step, we estimate endogenous parameters in our model. Table 5 shows the estimation results of the co-residence model using indirect inference. The first two rows show the estimates of the congestion costs, which reflect adult children's and parents' preferences for living together. The larger the coefficient, the higher the congestion cost and the less preferable co-residency is. As the male specific congestion cost (κ_1) is negative, parents prefer living with their sons over living with their daughters. We then show our estimates for housing costs, eldercare costs, and childcare costs in the next three rows of Table 5. All cost estimates are annual costs, e.g. the average national housing costs (which represents the total housing expenditure of living alone for an adult child-parent pair) are 830 Yuan (approximately 119 USD). Our estimate of δ_1 indicates that if the housing price of a city is 100% higher than the national average housing price, the housing cost for a family in that city would be 160 Yuan (19%) above the national average. In addition, we find that annual eldercare costs are 1,945 Yuan, and childcare costs are 702 Yuan. These cost estimates suggest that savings on housing costs, eldercare costs, and childcare costs are important incentives for families to live together. We last report on the standard deviation of the co-residence shock, which is 4,884. This accounts for around 15% of the average co-residence surplus.

To evaluate the predictive power of our model, we compare two sets of moments in Table 6. The odd columns of Table 6 show the regression results using our CFPS data (called "data"), and the even columns of Table 6 show the regression results using our bootstrap sample (called "model"). The first set of moments (in Columns (1) and (2)) are the coefficients of the co-residence regression, which are targeted in our model estimation. Column (1) presents the weighted regression coefficients using our adult child sample, showing how an adult child's probability of living with his/her parents depends on his/her own, parents' and local city's characteristics.[45] Our estimated coefficients from the bootstrap sample are shown in Column (2), which are close to the empirical regression coefficients in Column (1). According to the regression results, our model predicts that adult children are more likely to co-reside with their own parents if they are male, less educated, have fewer siblings, have a larger parent–child age gap, or have parents that are less educated. In addition, we find that the co-residence likelihood increases as city-level housing prices increase.

The second set of moments are those not targeted in our estimation, which are presented in Columns (3) to (6) in Table 6. Columns (3) and (4) check the child-side competition by estimating the effect of average education among siblings on the co-residence likelihood.[46] The regression results suggest that the co-residence likelihood increases with the average ed-

[45] We use the inverse of the number of adult children in the family as weight for the regression.

[46] Column (3) use the adult child sample with the inverse of the number of adult children in the family as weight.

ucation of siblings. This is because siblings with higher education are less likely to live with parents due to the income effect. This pattern is both observed in the data and predicted by our model. Columns (5) and (6) check the parent-side competition by estimating the effects of a spouse's parental education and number of siblings on the co-residence likelihood.[47] Both columns restrict the sample to couples living with either side of parents so that we can focus on the parent-side competition. We find that the likelihood of co-residence with one's own parents increases as their spouse's parental education increases. Likewise, it also increases as the number of siblings of a spouse increases. This is true in both the data and the model. When an adult child's parents-in-law have higher levels of education, they are less willing to live with their children due to the income effect; hence, the likelihood of living with an adult child's own parents increases. When one's spouse has more siblings, the spouse's parents are more likely to live with the spouse's siblings; therefore, the likelihood of living with one's own parents increases.

As supplementary results, we show the model fit in Appendix Table D6 and Appendix Table D7. Table D6 presents how parents' co-residence likelihood varies based on their own education levels. Consistent with our earlier discussion, highly educated parents have a lower co-residence likelihood than less educated parents due to income and health effects. Appendix Table D7 shows how an adult child's co-residence likelihood varies by his or her own characteristics. Given that parents have a lower congestion cost for living with sons than living with daughters, married couples are most likely to live with the husband's parents. The co-residence likelihood has a hump-shape relationship with adult children's education levels, without controlling for other factors. The non-monotonicity arises possibly because highly educated adult children also have fewer siblings. When we control for the number of siblings in our regression, the effect of an adult child's education becomes negative. Our model predicts the marginal negative effect of education on the co-residence of adult children (as shown in Column 2 of Table 6) and the unconditional hump-shape relationship between co-residence and an adult child's education (as shown in the second panel of Appendix Table D7). In addition, we also find a positive relationship between co-residence and the parent–child age gap, possibly because the demand for childcare becomes stronger as the parent-child age gap increases.

6 Counterfactual Analysis

We conduct two counterfactual experiments based on our model estimates. Our first experiment focuses on housing costs as the surging price of housing in China has been at the center of policy debates over the past decade. To understand how fluctuations in the housing market affect family living arrangements, we adjust city-level housing prices and predict the co-residence rates for different groups of parents and adult children. Our second experiment predicts the co-residence pattern when parents only have one adult child. The Chinese government implemented the one-child policy from 1979 to 2016 and most of the adult children in our data were

[47]Column (5) uses the couple sample.

born before the implementation of the one-child policy. The second experiment allows us to predict what co-residence will be like for families in a society where all adult children have no sibling. With the help of counterfactual analysis, we can predict the co-residence pattern among future cohorts.

We first explore the effect of changes in the housing prices on co-residence. The housing prices range from a decrease of 50% to an increase of 50%. Based on the estimates in Table 5, a 10% increase in the housing price translates to a 0.70% increase in housing cost in the surplus function.[48] The upper panel of Table 7 summarizes the key results in this counterfactual experiment. At the baseline level of housing price, the average probability of co-residence for adult children is 19.9%.[49] When the housing price increases by 20%, the co-residence likelihood increases by 3.2% or 0.6 percentage point (ppt) for adult children. We find some groups are more sensitive to a change in housing price than others. For example, female children experience a larger increase in the probability of living with their own parents than male children (4.2% vs. 3.2%) when housing prices increase. Highly educated adult children (those with a high school education or above) experience a larger increase in the co-residence likelihood compared to less educated adult children (3.8% vs. 2.9%). In addition, adult children with a smaller parent–child age gap are more likely to respond to a 20% housing price increase (3.4% vs. 3.1%). From the parent's perspective, a 20% increase in the housing price leads to a 3.2% (1.4 ppt) increase in the co-residence likelihood and the marginal increase is larger for highly educated parents (those with junior high educations or above) than for less educated parents (3.7% vs. 3.1%). Appendix Figure 5 shows the counterfactual results for different levels of housing prices.

The second counterfactual experiment predicts the co-residence likelihood in the near future, when all families only have one adult child, which will likely be the case for the next twenty years. In this case, parents face much stronger competition from in-laws because a young couple bears the burden of providing eldercare to both sides of parents. In this counterfactual experiment, we keep the oldest married child from each family in our current sample and re-simulate the marriage and co-residence decisions for these adult children. The bottom panel of Table 7 presents the results of this counterfactual experiment. We predict an increase in the co-residence likelihood of adult children, from 19.9% (with siblings) to 21.0% (without siblings). Among different subgroups, there is a 10.5% (3.2 ppt) increase in the probability of male children living with their own parents, and a 19.1% (1.4 ppt) increase in the co-residence rate of female children living with their own parents. Similar to the first counterfactual experiment, we find that highly educated children and children with a smaller parent–child age gap would experience a larger increase in the co-residence likelihood. As for parents, we would observe a huge decline in the co-residence likelihood, from 43.9% to 21.0% when each family

[48] We assume that the housing price does not affect marriage matching in a city.

[49] The average co-residence likelihood for adult children is calculated with the adult child sample using the inverse of the number of adult children in the family as weight.

has one child instead of 2.2. In addition, we find that less educated parents experience a larger decline in the co-residence likelihood compared to highly educated parents (53.8% vs. 48.1%). Considering that only half of the cities in our CFPS sample had nursing homes as of 2010, our results reveal an upcoming demand for formal elder care in China.

7 Conclusions

We develop a bilateral matching model with transferable utility to analyze the intergenerational co-residence decisions of adult children and their parents. The model allows us to capture competition not only between multiple children but also between parents and in-laws. This study is the first application of the Shapley–Shubik–Becker model to a co-residence context. The model captures several reasons for co-residence: social norms that are reflected in preferences and savings on housing costs, eldercare costs, and childcare costs. However, estimating the model is challenging, as family networks restrict co-residence matching choices but we cannot observe the complete network with our data. Therefore, we propose a network simulation method to estimate a marriage matching model. This method allows us to fill in missing marriage links not directly observed in the data. We estimate our co-residence model using the CFPS data and conduct counterfactual experiments to analyze the effects of both housing prices and the fertility policy in China on the co-residence likelihood of different types of adult children and parents.

Our model provides a few implications that can be tested directly in the data. Using the CFPS data, we show how the probability of co-residence for adult children depends on gender, education, parental education, local housing prices, and the age gap between them and their parents. In addition, we provide evidence of the competition both between adult children and between parents and in-laws. We find that on top of an adult child's and a parent's characteristics, the characteristics of siblings and parents-in-law of adult children can affect the co-residence probability of a parent-child pair. Specifically, we show that the number of siblings, the average education of the siblings, the spouse's parents' education, and the spouse's number of siblings affect the co-residence likelihood of an adult child, and the effects are consistent with the model's predictions. These results provide some justification for the choice of a matching model. Our counterfactual analysis shows that an increase in the housing price increases the co-residence likelihood, but the effects differ across subgroups. Greater sensitivity to housing price changes is observed in subgroups that benefit less from co-residence, including adult children that are female, highly educated, or older and parents that are highly educated. We further demonstrate that in the near future, when all families in China will have only one adult child, the co-residence likelihood of adult children will increase while the co-residence likelihood of parents will substantially decline. In particular, less educated parents will experience a larger decline in the co-residence likelihood.

Our study can be extended in two directions. First, to maintain a tractable analysis, the current model and estimation consider only local urban residents and exclude migrants. To include migrants, we must model another decision: where an adult child chooses to live. In the simplest case, adult children have four options: working in a big city and leaving their parents at home, working in a big city and bringing their parents with them, working in their hometown and living with their parents, or working in their hometown and not living with their parents. There are a few complications that would occur if we were to add a location choice. First, different locations provide different employment opportunities. In addition, parents without hukou registration in a big city pay much higher eldercare costs when they are treated at a hospital in that city. Lastly, allowing adult children to move opens up the marriage market and extends it from a local (currently at the city level) to a national market. This is an interesting and important direction in which to extend our work, but it will require more data and a more complicated model.

Another limitation of our work is that we take marriage matching as given when we consider co-residence matching. People may make marriage and co-residence decisions simultaneously. We handle this problem by adding some variables in the marriage matching model that may not be directly related to the utility of marriage but that reflect the utility of co-residence, such as the number of siblings, parent's age, and parent's education. However, if we want to explicitly model the joint decisions of marriage matching and co-residence matching, the model becomes incredibly complicated. First, this scenario no longer constitutes bilateral matching but rather matching between three parties: sons, daughters, and their parents. This theoretical question has not yet been studied in the literature. Second, in a non-bilateral matching setup, we need to redefine stable matching, which may not exist. Lastly, solving such a model would be computationally infeasible.[50] As this analysis is beyond the scope of the current study, we leave it for future research.

[50] Appendix A provides more details on these arguments.

References

Barczyk, Daniel and Matthias Kredler (2017). Evaluating long-term-care policy options, taking the family seriously. *The Review of Economic Studies 85*(2), 766–809.

Becker, Gary S (1973). A theory of marriage: Part I. *Journal of Political Economy 81*(4), 813–846.

Blake, Judith, Barbra Richardson, and Jennifer Bhattacharya (1991). Number of siblings and sociability. *Journal of Marriage and the Family*, 271–283.

Byrne, David, Michelle S Goeree, Bridget Hiedemann, and Steven Stern (2009). Formal home health care, informal care, and family decision making. *International Economic Review 50*(4), 1205–1242.

Card, David and Thomas Lemieux (2000). *Adapting to Circumstances: The Evolution of Work, School, and Living Arrangements Among North American Youth*. University of Chicago Press.

Chiappori, Pierre-André (1992). Collective labor supply and welfare. *Journal of Political Economy 100*(3), 437–467.

Chiappori, Pierre-André and Elisabeth Gugl (2015). Transferable utility and demand functions: a complete characterization. *Manuscript*.

Chiappori, Pierre-Andre and Maurizio Mazzocco (2017). Static and intertemporal household decisions. *Journal of Economic Literature 55*(3), 985–1045.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque (2012). Fatter attraction: anthropometric and socioeconomic matching on the marriage market. *Journal of Political Economy 120*(4), 659–695.

Chiappori, Pierre-Andre, Bernard Salanié, and Yoram Weiss (2008). Assortative matching on the marriage market: a structural investigation. Technical report, mimeo Columbia University.

Chiappori, Pierre-André, Bernard Salanié, and Yoram Weiss (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review 107*(8), 2109–67.

Choo, Eugene and Aloysius Siow (2006). Who marries whom and why. *Journal of Political Economy 114*(1), 175–201.

Compton, Janice and Robert A Pollak (2014). Family proximity, childcare, and women's labor force attachment. *Journal of Urban Economics 79*, 72–90.

Dostie, Benoit and Pierre Thomas Léger (2005). The living arrangement dynamics of sick, elderly individuals. *Journal of Human Resources 40*(4), 989–1014.

Dupuy, Arnaud (2018). Migration in China: to work or to wed? *Manuscript*.

Engers, Maxim and Steven Stern (2002). Long-term care and family bargaining. *International Economic Review 43*(1), 73–114.

Ermisch, John (1999). Prices, parents, and young people's household formation. *Journal of Urban Economics 45*(1), 47–71.

Fox, Jeremy T (2010). Identification in matching games. *Quantitative Economics 1*(2), 203–254.

Fox, Jeremy T (2018). Estimating matching games with transfers. *Quantitative Economics 9*(1), 1–38.

Galichon, Alfred and Bernard Salanié (2010). Matching with trade-offs: Revealed preferences over competing characteristics.

Garcia-Moran, Eva and Zoe Kuehn (2017). With strings attached: Grandparent-provided child care and female labor market outcomes. *Review of Economic Dynamics 23*, 80–98.

Gemici, Ahu (2011). Family migration and labor market outcomes. *Manuscript, New York University*.

Giuliano, Paola et al. (2004). On the determinants of living arrangements in Western Europe: Does cultural origin matter? Technical report, mimeo, UC Berkeley.

Haurin, Donald R, Patric H Hendershott, and Dongwook Kim (1993). The impact of real rents and wages on household formation. *Review of Economics and Statistics 75*(2), 284–293.

Hiedemann, Bridget and Steven Stern (1999). Strategic play among family members when making long-term care decisions. *Journal of Economic Behavior & Organization 40*(1), 29–57.

Hoerger, Thomas J, Gabriel A Picone, and Frank A Sloan (1996). Public subsidies, private provision of care and living arrangements of the elderly. *Review of Economics and Statistics 78*(3), 428–440.

Hu, Wei-Yin (2001). Welfare and family stability: Do benefits affect when children leave the nest? *Journal of Human Resources 36*(2), 274–303.

Kaplan, Greg (2012). Moving back home: Insurance against labor market risk. *Journal of Political Economy 120*(3), 446–512.

Kelly, E Lowell and James J Conley (1987). Personality and compatibility: A prospective analysis of marital stability and marital satisfaction. *Journal of personality and social psychology 52*(1), 27.

Konrad, Kai A, Harald Künemund, Kjell Erik Lommerud, and Julio R Robledo (2002). Geography of the family. *American Economic Review 92*(4), 981–998.

Ma, Sen and Fangqi Wen (2016). Who coresides with parents? An analysis based on sibling comparative advantage. *Demography 53*(3), 623–647.

Manacorda, Marco and Enrico Moretti (2006). Why do most Italian youths live with their parents? Intergenerational transfers and household structure. *Journal of the European Economic Association 4*(4), 800–829.

Maruyama, Shiko and Meliyanni Johar (2013). Do siblings free-ride in 'being there' for parents? *UNSW Australian School of Business Research Paper* (2013-06).

Maurer-Fazio, Margaret, Rachel Connelly, Lan Chen, and Lixin Tang (2011). Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000. *Journal of Human Resources 46*(2), 261–294.

McElroy, Marjorie B (1985). The joint determination of household membership and market work: The case of young men. *Journal of Labor Economics 3*(3), 293–316.

Pezzin, Liliana E, Robert A Pollak, and Barbara S Schone (2007). Efficiency in family bargaining: Living arrangements and caregiving decisions of adult children and disabled elderly parents. *CESifo Economic Studies 53*(1), 69–96.

Pezzin, Liliana E and Barbara Steinberg Schone (1999). Intergenerational household formation, female labor supply and informal caregiving: A bargaining approach. *Journal of Human Resources 34*(3), 475–503.

Rosenzweig, Mark and Junsen Zhang (2014). Co-residence, life-cycle savings and intergenerational support in urban China. Technical report, National Bureau of Economic Research.

Rosenzweig, Mark R and Kenneth I Wolpin (1993). Intergenerational support and the life-cycle incomes of young men and their parents: Human capital investments, coresidence, and intergenerational financial transfers. *Journal of Labor Economics 11*(1), 84–112.

Rosenzweig, Mark R and Kenneth I Wolpin (1994). Parental and public transfers to young women and their children. *American Economic Review 84*(5), 1195–1212.

Sakudo, Mari (2007). Strategic interactions between parents and daughters: Co-residence, marriage and intergenerational transfers in Japan. *Ph.D. thesis, University of Pennsylvania*.

Shapley, Lloyd S and Martin Shubik (1971). The assignment game I: The core. *International Journal of Game Theory 1*(1), 111–130.

Xie, Yu and Haiyan Zhu (2009). Do sons or daughters give more money to parents in urban China? *Journal of Marriage and Family 71*(1), 174–186.

Yasuda, Tokio, Noriko Iwai, Chin-Chun Yi, and Guihua Xie (2011). Intergenerational coresidence in China, Japan, South Korea and Taiwan: Comparative analyses based on the east asian social survey 2006. *Journal of Comparative Family Studies 42*(5), 703–722.

Table 1: A child's co-residence probability with his/her own parents

	Baseline	Child-side competition	Parent-side competition
	(1)	(2)	(3)
Male	0.2445***	0.2591***	0.6100***
	(0.0154)	(0.0133)	(0.0384)
Years of education	-0.0042*	-0.0110***	-0.0016
	(0.0025)	(0.0031)	(0.0058)
Parent's education	-0.0044*	-0.0045*	-0.0184***
	(0.0026)	(0.0023)	(0.0064)
Parent–child age gap	0.0071***	0.0074***	-0.0024
	(0.0018)	(0.0016)	(0.0040)
Number of siblings	-0.0501***	-0.0398***	-0.0793***
	(0.0064)	(0.0055)	(0.0155)
Logged housing price	0.0485**	0.0413**	0.0122
	(0.0193)	(0.0165)	(0.0403)
Average education of siblings		0.0067*	
		(0.0038)	
Spouse' parental education			0.0201***
			(0.0059)
Spouse' number of siblings			0.0876***
			(0.0133)
Constant	-0.4624**	-0.4283**	0.1301
	(0.2106)	(0.1823)	(0.4562)
Observations	3526	3328	437
R-squared	0.1258	0.1368	0.5201

We use a linear probability model and the dependent variable is whether the adult child co-resides with his or her own parents. The first column uses our adult child sample, which is restricted to urban local residents whose parents are aged 55 to 75 and whose siblings (including themselves) are all married. The second column also uses the adult child sample and is further restricted to those with siblings. The last column uses the couple sample in which the characteristics of a spouse's parents are observed and is further restricted to couples who live with one set of parents. The first two columns use the inverse of the number of adult children in the family as a weight. The last column does not need to be weighted. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Data source: 2010 CFPS.

Table 2: Summary statistics

Variable	Raw Mean	Weighted Mean	S.D.	Min	Max
Adult Child Level, N=3,709					
Co-residence with own parents	17.5%	20.4%	0.4	0	1
Male	50.0%	51.5%	0.5	0	1
Age	38.6	37.5	5.9	18	55
Age gap with parents	27.5	27.3	4.4	16	47
Years of education	9.7	10.2	3.6	0	22
Parent Level, N=1,314					
Co-residence with any child	47.6%	43.7%	0.5	0	1
Parents age	64.8	63.7	5.6	55	75
Parents education	5.8	6.2	3.5	1	16
Number of adult children	2.8	2.2	1.3	0	9
Adult Child Couple Level, N=659					
Co-residing with husband's parents	29.0%	29.0%	0.5	0	1
Co-residing with wife's parents	6.0%	6.0%	0.2	0	1
Number of children under six years old	0.3	0.3	0.5	0	3
City Level, N=104					
Housing price (Yuan per square-meter)	3,988	3,988	2,528	1,454	17,315
Housing size per person (in square-meter)	31.0	31.0	5.9	19.7	58.0
Adjusted housing price (via GDP per capita)	4.1	4.1	2.1	1.3	12.9
Child income city fixed effect	10.0	10.0	0.4	8.9	11.1
Parent income city fixed effect	8.8	8.8	0.9	5.8	10.5
Number of grandchildren fixed effect	0.2	0.2	0.05	0.1	0.3

Table 3: Marriage matching model estimation results

	Years of education	Age	Number of siblings	Parental education	Parental age
Years of education	0.713	-0.171	-0.091	0.312	0.006
	(0.121)	(0.115)	(0.063)	(0.124)	(0.074)
Age	0.073	1.174	0.115	0.037	0.315
	(0.097)	(0.344)	(0.048)	(0.063)	(0.062)
Number of siblings	-0.048	0.176	0.124	-0.061	0.016
	(0.064)	(0.060)	(0.064)	(0.055)	(0.031)
Parental education	0.182	-0.103	-0.052	0.291	0.087
	(0.118)	(0.066)	(0.058)	(0.081)	(0.100)
Parental age	0.009	0.370	0.021	0.084	0.106
	(0.065)	(0.114)	(0.033)	(0.062)	(0.030)

Standard errors are shown in parentheses. Columns show the characteristics of the husband and rows show the characteristics of the wife.

Table 4: Out-of-model estimation

	Parent's income		Adult child's income		Parents in poor health		With grandchildren	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	-0.059***	-0.060	-0.007**	0.080***	0.003*	0.019	-0.040***	-0.158***
	(0.013)	(0.221)	(0.003)	(0.030)	(0.002)	(0.046)	(0.002)	(0.019)
Age square		0.000		-0.001***		-0.000		0.002***
		(0.002)		(0.000)		(0.000)		(0.000)
Years of education	0.051**	0.051**	0.057***	0.057***	-0.007***	-0.007***	-0.003	-0.002
	(0.021)	(0.021)	(0.006)	(0.006)	(0.003)	(0.003)	(0.004)	(0.004)
Constant	3.296***	3.311	-0.314**	-1.946***	0.086	-0.422	1.548***	3.751***
	(0.876)	(6.973)	(0.151)	(0.588)	(0.123)	(1.478)	(0.094)	(0.360)
City FE	Yes	Yes	Yes	Yes	No	No	Yes	Yes
Observations	545	545	1,026	1,026	1,438	1,438	1,064	1,064
R-squared	0.097	0.097	0.097	0.105	0.008	0.008	0.270	0.297

We use ordinary least square to estimate columns (1) to (4) and linear probability models to estimate columns (5) to (8).

Columns (1) to (4), (7), and (8) use the CFPS data and Columns (5) and (6) use the CHARLS data.

For columns (1), (2), (5), and (6), the sample is restricted to urban local residents aged 55 to 75. For columns (3), (4), (7), and (8), the sample is restricted to urban local residents whose parents are 55 to 75.

Robust standard errors in parentheses

***p<0.01, * *p<0.05, * P<0.1

Table 5: Co-residence matching model estimation results

Parameter	Explanation	Estimate	Standard Error
κ_0	constant	1.453	0.502
κ_1	co-residing with husband's parents	-0.029	0.007
δ_0	average housing costs	830	151
δ_1	effect of log housing price on housing costs	160	41
S_{ec}	eldercare costs	1945	314
S_{cc}	childcare costs	702	312
σ	standard deviation of congestion cost shock	4884	1221

Table 6: Model fit: A child's co-residence probability with his/her own parents

	Baseline		Child-side competition		Parent-side competition	
	Data	Model	Data	Model	Data	Model
	(1)	(2)	(3)	(4)	(5)	(6)
Male	0.2445***	0.244***	0.2591***	0.2332***	0.6100***	0.6047***
	(0.0154)	(0.0016)	(0.0133)	(0.0215)	(0.0384)	(0.0027)
Years of education	-0.0042*	-0.0040***	-0.0110***	-0.0033***	-0.0016	-0.0008*
	(0.0025)	(0.0003)	(0.0031)	(0.0003)	(0.0058)	(0.0005)
Parent's education	-0.0044*	-0.0044***	-0.0045*	-0.0036***	-0.0184***	-0.0062***
	(0.0026)	(0.0003)	(0.0023)	(0.0003)	(0.0064)	(0.0005)
Parent–child age gap	0.0071***	0.0071***	0.0074***	0.0072***	-0.0024	-0.0054***
	(0.0018)	(0.0002)	(0.0016)	(0.0002)	(0.0040)	(0.0003)
Number of siblings	-0.0501***	-0.0248***	-0.0398***	-0.0228***	-0.0793***	-0.0169***
	(0.0064)	(0.0006)	(0.0055)	(0.0008)	(0.0155)	(0.0011)
Logged housing price	0.0485**	0.0458***	0.0413**	0.0407***	0.0122	0.0003
	(0.0193)	(0.0018)	(0.0165)	(0.0019)	(0.0403)	(0.0032)
Average education of siblings			0.0067*	0.0070***		
			(0.0038)	(0.0005)		
Spouse' parental education					0.0201***	0.0153***
					(0.0059)	(0.0011)
Spouse' number of siblings					0.0876***	0.0620***
					(0.0133)	(0.0046)
Constant	-0.4624**	-0.4893***	-0.4283**	-0.4472***	0.1301	0.0574
	(0.2106)	(0.0202)	(0.1823)	(0.0215)	(0.4562)	(0.0350)
Observations	3526	23,663,900	3328	20,015,400	437	8,669,800
R-squared	0.1258	0.1087	0.1368	0.1012	0.5201	0.3766
Root MSE	0.3766	0.3776				

Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Data columns use the CFPS data to run OLS regressions, the same as Table 1. Model columns use simulated co-residence decisions in the bootstrap sample to run the same regressions.

Table 7: Counterfactual results on co-residence probability (%)

	Child	Child's gender		Child's education		Parent–child age gap		Parent	Parental education	
	total	male	female	<12	≥ 12	<27	≥ 27	total	<9	≥ 9
0. Baseline	19.9	30.3	7.1	20.5	19.1	18.0	21.6	43.9	47.7	34.0
1. Changes in housing price										
-20%	19.3	29.3	6.8	19.9	18.3	17.3	20.9	42.4	46.1	32.6
20%	20.6	31.3	7.4	21.1	19.8	18.6	22.2	45.3	49.1	35.3
2. Each family only has one child										
	21.0	33.5	8.4	21.3	20.5	20.2	23.2	21.0	22.0	17.7

Child's co-residence probability refers to the probability of living with the child's own parents (not the spouse's).
Parent's co-residence probability refers to the probability of living with any of their adult children.

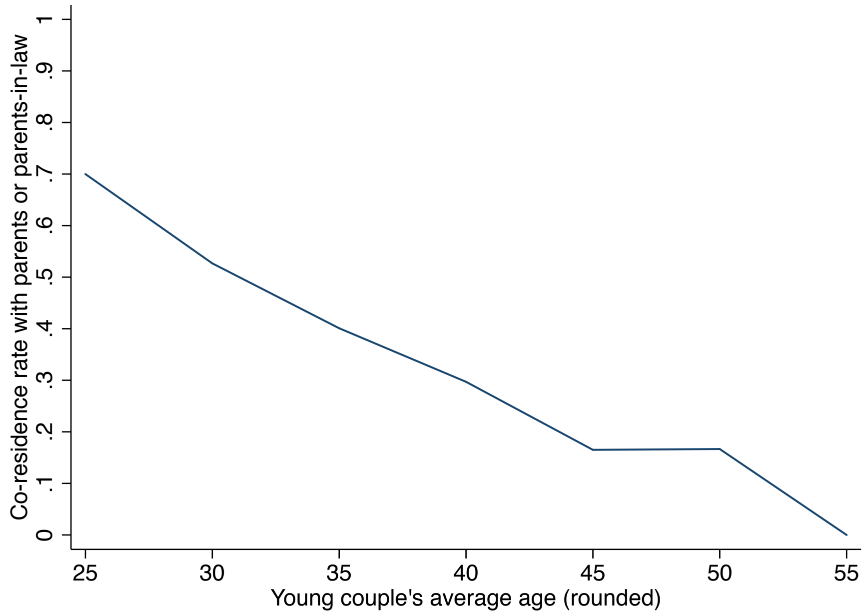

Figure 1: The probability of co-residence with parents or parents-in-law for married couples by age. Data source: China Family Panel Study 2010

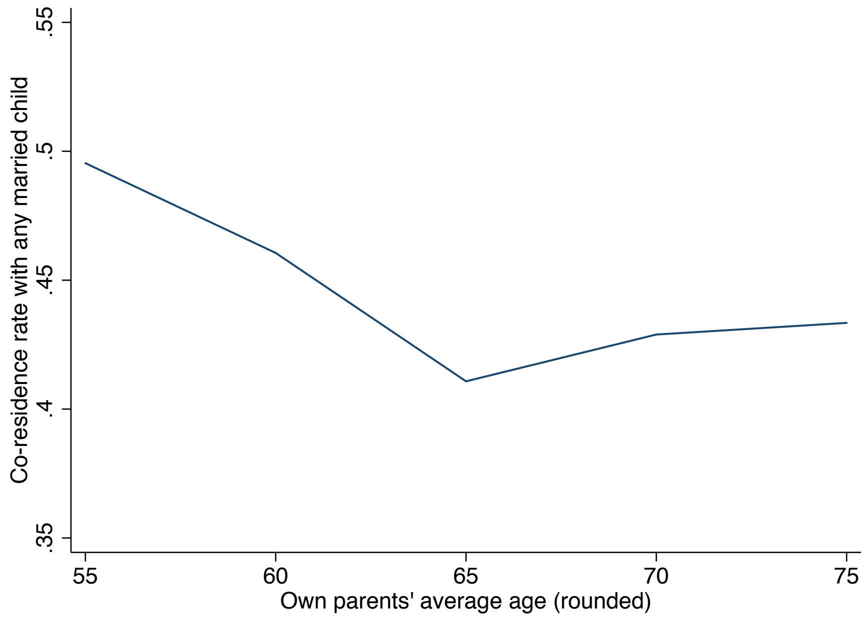

Figure 2: The probability of co-residence with adult children for parents by age. Data source: China Family Panel Study 2010

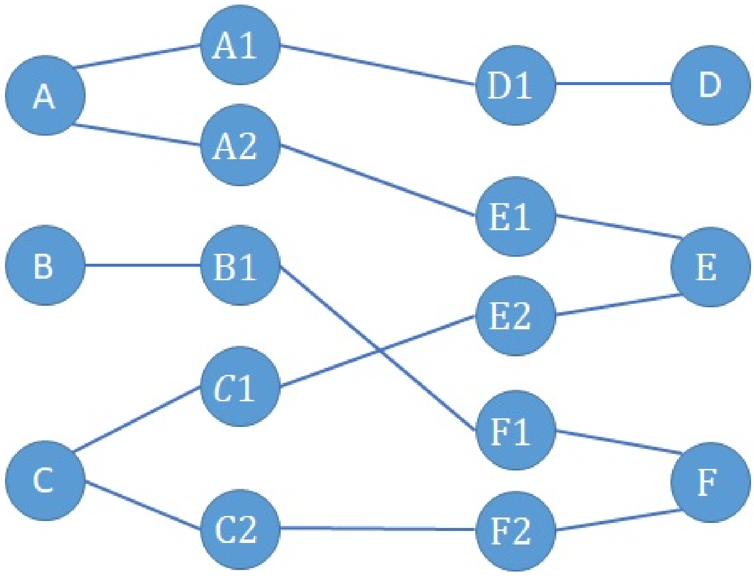

Figure 3: Family Network

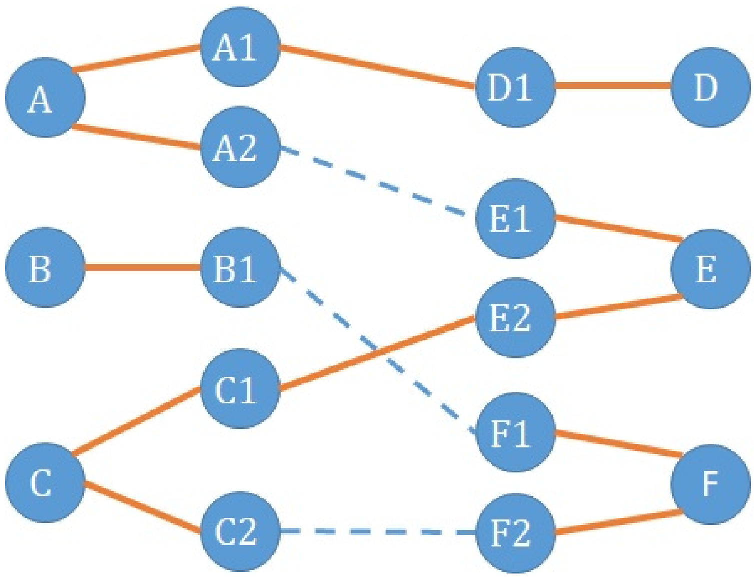

Figure 4: Observed Family Network

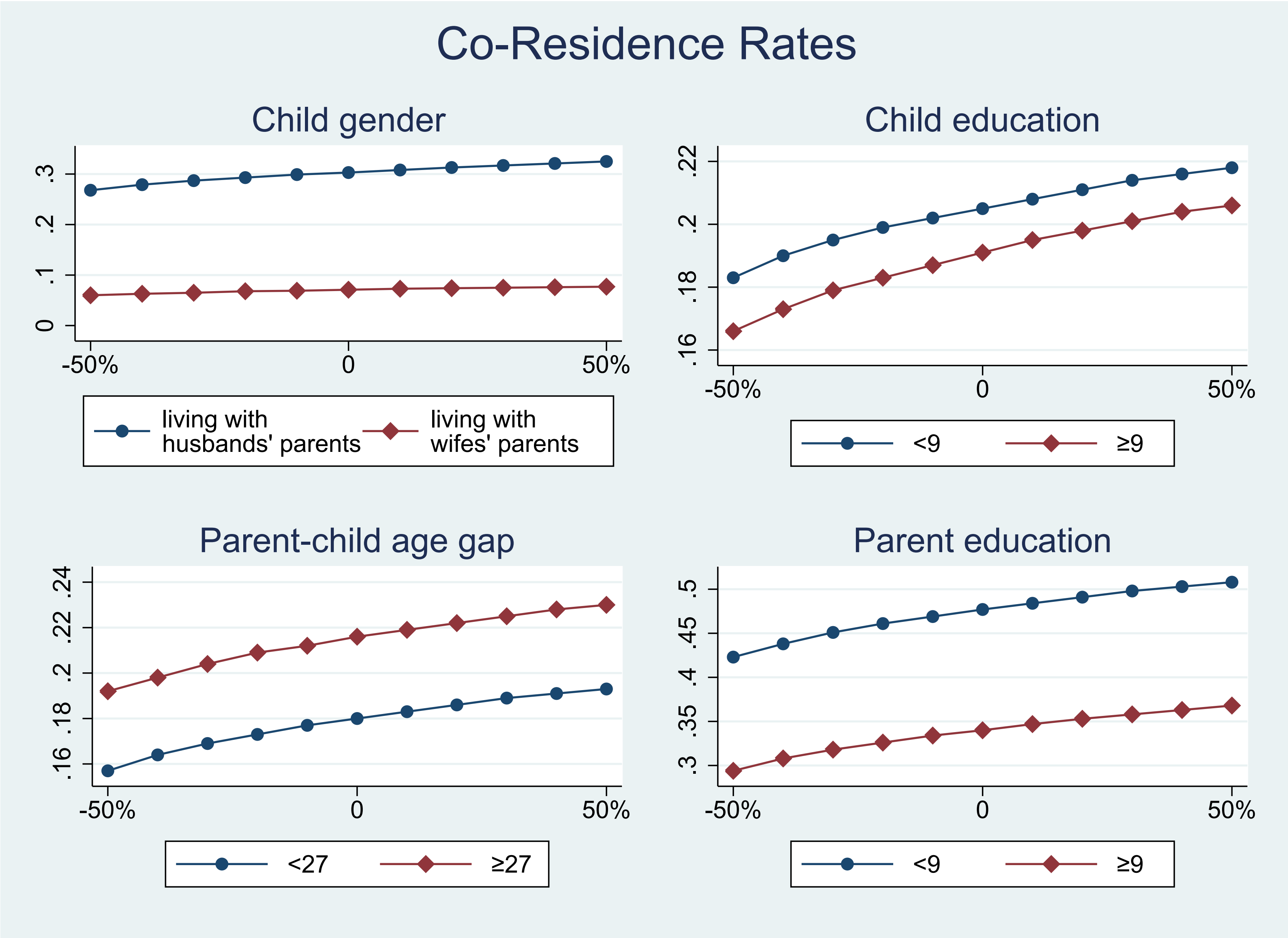

Figure 5: Effects of Changes in Housing Price

APPENDIX

Appendix A Jointly Model Co-residence and Marriage Decisions

If we consider a model where marriage matching and co-residence decision are made simultaneously, the problem becomes extremely difficult to solve for two reasons. Theoretically, the stable matching may not exist. Computationally, the problem cannot be solved because we cannot find the solution in a reasonable amount of time.

Let us first define stable matching in a setup with simultaneous marriage and co-residence matching. The matching problem involves three parties: men, women, and parents. A natural extension of the stability conditions would be the following.

- There is no benefit for any adult children to deviate from their current status to become single and live alone.
- There is no benefit for any adult child–parent pair to deviate from their current status to become single (for the adult child) and live together.
- There is no benefit for any parents to deviate from their current status to live alone.
- There is no benefit for any man–woman pair to deviate from their current status to marry and not live with either side's parents.
- There is no benefit for any man–woman pair and man's parent to deviate from their current status to form a marriage (between the man and the woman) and co-reside (together with the man's parent).
- There is no benefit for any man–woman pair and woman's parent to deviate from their current status to form a marriage (between the man and the woman) and co-reside (together with the woman's parent).

The matching is stable as long as none of the above six conditions is violated.

After defining stable matching, we next show a counter example where stable matching may not exist. In this example, parent 1 (P_1) has two adult children, a son M_1 and a daughter W_1. Besides this family, there is one man M_2 and one woman W_2 in the market, and they could never live with their own parents, who have better outside options. Assume that P_1, M_2, and W_2 each have two units of income, while M_1 and W_1 have 0 unit of income. For simplicity, we assume that the total utility of N agents linked together through marriage or co-residence is Y^2, where Y is the total income of N agents linked together. With this quadratic functional form assumption, marriage is always weakly better than being single, and co-residing with parents is always weakly better than living alone (for both adult children and parents). We focus on the marriage/co-residence relationship among P_1, W_2, and M_2, because they are the only three persons with positive income. W_1 and M_1 have no bargaining power in this game because they have zero income, and we can assume that they always receive zero regardless of the marriage and co-residence arrangements. Given that P_1, W_2, and M_2 cannot all live together, we discuss the following three interesting cases in detail.

1. No two persons among P_1, W_2, and M_2 live together. In this case, the maximum utility W_2 can receive is 4 by marrying M_1.[51] Similarly, the maximum utility M_2 can receive is 4. W_2 and M_2 will have the incentive to deviate from the current arrangement and get married because the marriage will generate a total surplus of 16 and can provide each of them more than 4.

2. W_2 marries M_2. The total utility of the $\{W_2, M_2\}$ marriage is 16. One member of the couple will get utility no more than 8. Without loss of generality, we assume that W_2 obtains less than 8. Then $\{P_1, M_1\}$ can approach W_2 by offering a marriage between $\{M_1, W_2\}$ and a co-residence arrangement between $\{P_1, M_1, W_2\}$. The couple–parent pair $\{P_1, M_1, W_2\}$ generates a total utility of 16. $\{P_1, M_1\}$ can offer something slightly higher than 8 to make W_2 happy while improving their own utility to slightly less than 8. Note that the maximum utility $\{P_1, M_1\}$ can receive when W_2 marries M_2 is 4 through living together.

3. M_1 marries W_2 and they live with parent P_1. As discussed above, the total utility of this couple–parent pair is 16. Either W_2 or P_1 will receive not more than half of the total utility 8.

   (a) Suppose that W_2 receives not more than 8. M_2 can approach W_2 by offering a marriage between $\{W_2, M_2\}$ with total utility 16. In this case, both W_2 and M_2 are better off. M_2 can offer W_2 something slightly more than 8, while himself obtain slightly less than 8, which is better than what can receive in the current situation (receive 4 by staying single or marrying W_1).

   (b) Suppose that P_1 receives not more than 8. M_2 can offer to marry W_1 and live with P_1. The total utility of the couple–parent pair $\{P_1, W_1, M_2\}$ is 16. In this case, both P_1 and M_2 are better off. M_2 can offer P_1 something slightly higher than 8, while himself obtain slightly less than 8, which is better than what can receive in the current situation (receive 4 by staying single or marrying W_1).

The case that W_1 marries M_2 and lives with P_1 is symmetric to Case 3. Therefore, under all possible matches, there are always better arrangements that violate stable matching conditions. Hence, this counter example proves that stable matching may not exist.

In addition, even if stable matching exists, finding the optimal assignment in a reasonable amount of time may be impossible. This matching setup involves three parties, therefore becoming a 3D-matching problem, which is well-defined in computer science theory. Computer

[51] Staying single and marrying M_1 give W_2 the same utility of 4.

scientists view this problem as an NP-hard problem, which indicates that no evident short-cut is available to find a solution apart from the brute-force approach. In the brute-force approach, we need to search for all possible marriage matches. Given each marriage matching assignment, we obtain the network to solve the co-residence matching. Then, we can calculate co-residence surplus and add it back to the marriage surplus. Finally, we select the one marriage matching that maximizes the sum of marriage surplus and co-residence surplus. Even if we only have 100 men and 100 women, we need to calculate $100!$ possible marriage matches, which is computationally impossible.

Appendix B Marriage Matching Model Estimation

In this section, we provide more explanations on how we estimate our marriage matching model, following Galichon and Salanié (2010).

Recall that the marriage surplus function is expressed as:

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = \Phi(x, y) + \chi(\tilde{x}, y) + \xi(\tilde{y}, x)$$

Here we assume that the χ and ξ follow central Gumbel (type-I extreme value) distribution with scale factor $\sigma_1 = 0.5$ and $\sigma_2 = 0.5$, respectively. We further approximate the observable surplus function with a linear expansion over some known basis functions ϕ^k, with unknown weights λ:

$$\Phi(x, y) = \sum_k \lambda_k \phi^k(x, y)$$

Then, we estimate λ such that π_λ has covariations matching the observed covariations, i.e.:

$$C(\pi_\lambda) = C(\hat{\pi}) \ . \tag{22}$$

where π_λ is the optimal matching predicted by the model with parameter λ, and $\hat{\pi}$ the observed matching. The covariations $C(\pi) = \big(C^1(\pi), C^2(\pi), \dots\big)$ are defined as:

$$C^k(\pi) = \sum_{x,y} \pi(x, y)\phi^k(x, y) \ .$$

We implement the moment estimator in two steps. In the first step, we predict the optimal marriage matching π_λ conditional on the parameter λ and observed marginal distribution p and q. Galichon and Salanié (2010) prove that the optimal matching maximizes the a linear combination of the observable surplus and of the mutual information:[52]

$$\max_{\pi \in \Pi(p,q)} \left(\sum_{x,y} \pi(x, y)\Phi(x, y) - \sigma \sum_{x,y} \pi(x, y) \log \frac{\pi(x, y)}{p(x)q(y)} \right) \tag{23}$$

Given λ, p, and q, we can find the optimal matching π through an iterative algorithm, the Iterative Projection Fitting Procedure (IPFP), suggested by Galichon and Salanié (2010). They show that the IPFP is fast, stable, and simple.

In the second step, we match the moments in Equation (22) by maximizing:

$$\lambda \cdot C(\hat{\pi}) - W(\lambda) \ . \tag{24}$$

[52] The mutual information captures the case when the unobserved heterogeneity dominates and the matching looks like completely random matching.

where $W(\lambda)$ is the maximum linear combination of surplus and mutual information in Equation (23).

There is a typo in the original paper by Galichon and Salanié (2010) stating Equation (24) as a minimization problem instead of maximization. Therefore, we present the reasoning behind the maximization problem. Note that $W(\lambda)$ as a function of λ can be written as the maximum of linear functions as follows:

$$
\begin{aligned}
W(\lambda) &= \max_{\pi \in \Pi(\hat{p},\hat{q})} \left(\sum_k \lambda_k \sum_{x,y} \pi(x,y) \Phi^k(x,y) - \sum_{x,y} \pi(x,y) \log \frac{\pi(x,y)}{\hat{p}(x)\hat{q}(y)} \right) \\
&= \max_{\pi \in \Pi(\hat{p},\hat{q})} \left(\lambda \cdot C(\pi) - \sum_{x,y} \pi(x,y) \log \frac{\pi(x,y)}{\hat{p}(x)\hat{q}(y)} \right) \qquad (25)
\end{aligned}
$$

This has two implications. First, $W(\lambda)$ is a convex function of λ. Therefore, Equation (24) is a concave function of λ and its maximizer satisfies the first-order condition:

$$C(\hat{\pi}) = \nabla W(\lambda) \, .$$

Second, recall that π_λ denote the matching that maximizes Equation (25). By the envelope theorem, we have:

$$\nabla W(\lambda) = C(\pi_\lambda) \, .$$

Together we get $C(\pi_\lambda) = C(\hat{\pi})$, so that the maximizer of Equation (24) also solves Equation (22).

Appendix C Marriage Model Fit

Table C1: Marriage probability by a couple's education level (%)

A. Data (couple sample)

	Wife's education			
Husband's education	Below middle school	Middle school	High school	College
Below middle school	9.41	3.79	1.06	0.00
Middle school	10.47	18.97	6.98	2.43
High school	1.67	11.08	8.50	3.79
College	0.61	2.73	4.86	13.66

B. Model (couple sample)

	Wife's education			
Husband's education	Below middle school	Middle school	High school	College
Below middle school	8.23	5.09	0.83	0.12
Middle school	11.35	18.68	6.58	2.23
High school	2.26	9.40	7.54	5.83
College	0.28	3.39	6.46	11.72

C. Model (bootstrap sample)

	Wife's education			
Husband's education	Below middle school	Middle school	High school	College
Below middle school	5.83	7.15	1.26	0.17
Middle school	7.26	22.87	8.73	2.94
High school	1.14	8.69	7.60	5.97
College	0.15	3.09	5.81	11.33

Panel A shows the actual marriage matching in the CFPS couple sample. Panel B shows the predicted marriage matching in our model using the CFPS couple sample. Panel C shows the predicted marriage matching in our model using the bootstrap national family sample described in Section 4.3

Table C2: Marriage probability by a couple's age (%)

A. Data (couple sample)

	Wife's age			
Husband's age	≤33	34–38	39–43	≥44
≤33	18.97	2.43	0.00	0.00
34–38	8.50	20.18	2.43	0.00
39–43	1.52	9.86	17.45	1.21
≥44	0.15	0.30	6.98	10.02

B. Model (couple sample)

	Wife's age			
Husband's age	≤33	34–38	39–43	≥44
≤33	15.77	5.23	0.40	0.01
34–38	11.21	14.89	4.71	0.30
39–43	2.17	10.99	13.53	3.36
≥44	0.08	1.69	8.16	7.52

C. Model (bootstrap sample)

	Wife's age			
Husband's age	≤33	34–38	39–43	≥44
≤33	16.50	7.91	0.92	0.02
34–38	7.98	15.00	6.95	0.78
39–43	0.94	7.06	13.18	5.83
≥44	0.02	0.79	5.90	10.22

Panel A shows the actual marriage matching in the CFPS couple sample. Panel B shows the predicted marriage matching in our model using the CFPS couple sample. Panel C shows the predicted marriage matching in our model using the bootstrap national family sample described in Section 4.3

Table C3: Marriage probability by a couple's number of siblings (%)

A. Data (couple sample)

	Wife's no. of siblings	
Husband's no. of siblings	0–1	$\geq$2
0–1	26.25	20.79
$\geq$ 2	15.78	37.18

B. Model (couple sample)

	Wife's no. of siblings	
Husband's no. of siblings	0–1	$\geq$2
0–1	23.77	23.28
$\geq$ 2	18.32	34.64

C. Model (bootstrap sample)

	Wife's no. of siblings	
Husband's no. of siblings	0–1	$\geq$2
0–1	25.52	21.01
$\geq$ 2	21.07	32.40

Panel A shows the actual marriage matching in the CFPS couple sample. Panel B shows the predicted marriage matching in our model using the CFPS couple sample. Panel C shows the predicted marriage matching in our model using the bootstrap national family sample described in Section 4.3

Table C4: Marriage probability by a couple's parental education level (%)

A. Data (couple sample)			
	Wife's parents' education		
Husband's parents' education	Below primary school	Primary school	Middle school
Below primary school	26.71	11.84	5.77
Primary school	10.47	13.20	8.35
Middle school	4.40	8.65	10.62
B. Model (couple sample)			
	wife's parents' education		
Husband's parents' education	Below primary school	Primary school	Middle school
Below primary school	25.25	13.47	5.59
Primary school	11.25	11.89	8.88
Middle school	5.04	8.34	10.29
C. Model (bootstrap sample)			
	Wife's parents' education		
Husband's parents' education	Below primary school	Primary school	Middle school
Below primary school	25.03	12.71	5.65
Primary school	12.75	11.51	8.24
Middle school	5.58	8.29	10.25

Panel A shows the actual marriage matching in the CFPS couple sample. Panel B shows the predicted marriage matching in our model using the CFPS couple sample. Panel C shows the predicted marriage matching in our model using the bootstrap national family sample described in Section 4.3

Table C5: Marriage probability by a couple's parental age (%)

A. Data (couple sample)

Husband's parents' age	Wife's parents' age ≤65	>65
≤65	38.39	11.68
>65	20.49	29.44

B. Model (couple sample)

Husband's parents' age	Wife's parents' age ≤65	>65
≤65	37.90	12.18
>65	21.06	28.86

C. Model (bootstrap sample)

Husband's parents' age	Wife's parents' age ≤65	>65
≤65	38.57	15.01
>65	15.11	31.31

Panel A shows the actual marriage matching in the CFPS couple sample. Panel B shows the predicted marriage matching in our model using the CFPS couple sample. Panel C shows the predicted marriage matching in our model using the bootstrap national family sample described in Section 4.3

Appendix D Co-residence Model Fit

Table D6: Share of parents co-residing with their children

Education	Illiterate	Primary school	Middle school	High school	Total
Data	0.53	0.39	0.41	0.26	0.44
Model	0.52	0.42	0.36	0.30	0.44

Data: CFPS family sample, using the inverse of the number of adult children in the family as weight. Model: bootstrap sample.

Table D7: Share of an adult child co-residing with his/her parents

Gender	Male	Female	Total		
Data	0.32	0.08	0.20		
Model	0.30	0.07	0.20		
Education	Illiterate	Primary school	Middle school	High school	College
Data	0.15	0.17	0.22	0.21	0.17
Model	0.18	0.20	0.21	0.19	0.18
Parent–child age gap	<25	25-29	30-34	≥35	
Data	0.20	0.19	0.22	0.28	
Model	0.17	0.20	0.22	0.26	

Data: CFPS adult child sample using the inverse of the number of adult children in the family as weight. Model: bootstrap national family sample.

Table D8: Summary statistics for the elderly in the CHARLS

	Obs.	Mean	Std. Dev.	Min	Max
In poor health status	1,444	23.5%	0.42	0	1
Age	1,444	63.1	5.9	55	75
Years of education	1,438	7.1	4.4	0	19
Need helper in daily life	1,444	9%	0.29	0	1

Table D9: Summary statistics of the couples without any siblings in the CFPS

	Obs.	Mean	Std. Dev.	Min	Max
Child's years of education	44	11.8	3.3	4	19
Parent-child age gap	44	27.7	4.1	18	39
Parents' years of education	44	6.7	3.3	1	12
Housing price	44	7803	5521	2507	17316