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

Co-residence Matching

Naijia Guo[1], Xiaoyu Xia[2], and Junsen Zhang[3,*,†]

**Abstract:**

We develop a co-residence model between young adults and the elderly as an application of the Shapley-Shubik-Becker bilateral matching framework. This model captures competition between adult children and between parents and parents-in-law. Using micro data from China, we estimate our model by using a network simulation method to fill in partially unobservable marriage links. We find that our model explains the child-side and parent-side competitions observed in the data better than two alternative multinomial logit models with only one-side competition. In addition, counterfactual experiments quantify the effects of changes in the one-child policy and housing prices on intergenerational co-residence.

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

**Classification:** D1, J1, J2

*Correspondence e-mail: jszhang@cuhk.edu.hk.

†Present address: School of Economics, Zhejiang University, Hangzhou, China.

We are grateful to the editor, anonymous referees, Pierre-André Chiappori, James Heckman, and Mark Rosenzweig for their valuable comments and suggestions. We also thank many seminar participants at the National University of Singapore, University of Melbourne, University of Hong Kong, and University of Chicago, and the conference participants at the ASSA annual meeting, Asian Meeting of the Econometric Society, International Symposium on Contemporary Labor Economics, and CUHK labour workshop for their helpful comments. Junsen Zhang acknowledges partial financial support from the Major Program of the National Social Science Fund of China (Grant No. 20 & ZD076) and the Key Program of the National Natural Science Foundation of China (Grant No. 72034006). All errors are our own.

# 1 Introduction

In China, 36% of married children live with either their parents or parents-in-law, and more than 40% of parents aged 55 to 75 live with their children.[1] 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 be even more important factors in this decision. Since China's housing privatization reform in the mid-1990s, many adult children have been unable to afford home ownership and have chosen to live with their elderly home-owning parents (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 eldercare and childcare needs arise from China's lack of nursing homes and pre-kindergartens.

Cost saving is a potentially important mechanism for the co-residence decision, which is why family members may compete to co-reside. For example, 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. Table 1 illustrates child-side and parent-side competitions using data from the 2010 China Family Panel Study (CFPS, Institute of Social Science Survey, 2015). The probability of a young adult living with his or her own parent is correlated with that young adult's and his/her parents' characteristics (Column 1), and the characteristics of the young adult's siblings and spouse's parents also affect the co-residence likelihood (Columns 2 and 3). Between-sibling competition is reflected by the fact that 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 reflected by the fact that an adult child is more likely to live with his/her own parents if the spouse's parents are highly educated and if the spouse's parents have more children.

In this study, we develop a bilateral matching model to explain this child-side and parent-side competition. After establishing the theoretic framework, we empirically estimate our model using Chinese data to demonstrate how social norms and economic incentives affect living arrangements. Through counterfactual experiments, we use 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 (TU) between parents and adult children. Our matching model is an extension of Shapley and Shubik (1971) that 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, which are captured by congestion costs—that is, reduced utility resulting from living with others that could be affected by the gender and birth order of adult children. In such a TU model, to

[1] See the calculation by the authors that is illustrated in Figures 1 and 2.

solve for the optimal co-residence assignment, it is sufficient for the planner to know only the gross surplus of co-residence for each parent–adult child pair.

Different from Shapley–Shubik–Becker's matching model, the co-residence matching model sets a restriction on the choice set of elderly and young adults. Elderly people can live with one of their adult children but not with any other young adults, whereas young adults can live with their parents or parents-in-law but not with 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 children's marriage links) in a large market. In this paper, we propose a network simulation procedure to fill the gaps in 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 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 for young couples in the CFPS data and use the estimates to impute the marriage links for married children in the bootstrap family sample. Third, using the 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 the coefficients in a co-residence regression.

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

We compare the matching model with two alternative models: 1) a multinomial logit model with a child couple and both sets of parents, which considers the competition between parents and parents-in-law but ignores the competition between siblings, and 2) a multinomial logit model with a parent and multiple adult children, which considers the competition between siblings but ignores the competition between parents and parents-in-law. We show that the matching model dominates the two alternative multinomial logit models in terms of capturing the empirical patterns of parent-side and child-side competition simultaneously.

We conduct counterfactual experiments to analyse 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 likelihood of co-residence for adult children by 3.5% and for parents by 3.5%. In particular, female, first-born, and highly educated children and highly educated parents are more sensitive to housing price changes. 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 likelihood of living with their own parents, from 20.0% to a range of 21.1%–26.2%. However, the co-residence likelihood of parents would drop significantly from 43.9% to a range of 21.1%–26.2% when they have only one child. In particular, less educated parents would experience an even larger decline in 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, a few studies model the strategic interactions between one parent and multiple children, including Hiedemann and Stern (1999), Engers and Stern (2002), Konrad *et al.* (2002), Stern (2014), Maruyama and Johar (2017), Pezzin *et al.* (2007), and Byrne *et al.* (2009). In contrast, our model characterizes the living arrangement as a matching game, which 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 show that co-residence helps adult children insure themselves against poor labour market opportunities (Rosenzweig and Wolpin, 1993, Card and Lemieux, 1997, and Kaplan, 2012) and rising housing costs (Haurin *et al.*, 1993, Ermisch, 1999, and Rosenzweig and Zhang, 2014). Several studies 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, 2018) and 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 contribute to the high rate of cohabitation in numerous countries, including China, as shown by Giuliano (2007), Chu *et al.* (2011), and Zorlu and Mulder (2011).

The remainder of this paper is structured as follows. In Section 2, we describe a one-to-one, two-sided, TU matching model in the co-residence context. In Section 3, we introduce our data sources. In Section 4, we present the identification and estimation of the co-residence matching model. We detail the estimation results in Section 5. In Section 6, we present two alternative models and compare their model fit with the matching model. We present counterfactual analyses in Section 7. In Section 8, we conclude the paper.

# 2 Model

Our model has two stages, where marriage and co-residence decisions are made sequentially. In the first stage, young men and women form marriage matching, and they are forward looking and take into account the expected utility from co-residence as part of the marriage surplus. In the second stage, the child couple and their parents form co-residence matching based on the family network generated by the marriage matching in the first stage.

Within each stage, we model the marriage matching and co-residence matching as a one-to-one, two-sided, TU matching game, similar to that in Shapley and Shubik (1971) and Becker (1973). This matching model has been widely used in studying the marriage decision, but we are the first to apply it to study the co-residence decision. In the co-residence matching, 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.[2] It is also a two-sided match as both parents and adult children must agree to the living arrangement. By assuming TU, parents who are in favour 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.

We solve the two-step matching problem using backward induction. Below, we first introduce the co-residence matching model, followed by the marriage matching model.

## 2.1 Stage II: Co-residence Matching

### 2.1.1 Model Setup

Our model includes two types of decision-makers: parents and adult children.[3] 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." 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, and we abstract away from the bargaining within the young couple. If diverging preferences exists between the husband and the wife, then the congestion cost of the child couple reflects their joint preference on co-residence after their collective bargaining process.

We model co-residence decisions between parents and adult children as a static matching decision. Co-residence is a package that provides both eldercare and childcare. Co-residence is a one-shot game and the decision is made at the time when parents are retired and all adult children are married. When parents and

[2]Even in China, 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).

[3]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).

adult children make their co-residence decisions, they consider the probability of parents getting sick and the probability of adult children having children of their own (a.k.a. "grandchildren") in the future. When parents are relatively young and healthy, they can provide childcare to their co-residing grandchildren (if any); when parents grow old and become sick, adult children can care for their co-residing parents. Therefore, the model captures the intertemporal substitutions between childcare and eldercare over the life cycle.

Suppose all parents live for $N$ years, where $N$ is the life expectancy. When parents reach age $N_0$, parents and all of their adult children make the co-residence decision together. We assume that no sequential co-residence decision is made by siblings and the timing of when adult children marry does not affect the co-residence decision. We also 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.[4] Parents and adult children, if living together, jointly allocate total household income into private and public consumption every year based on a collective bargaining framework (Chiappori, 1992). We assume that people consume all of their income in each period.

We first specify the utility function of living together for a parent–adult child pair. Following 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. 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 $\alpha$ and $1-\alpha$ capturing the effects of private and public consumption on utility, respectively. We follow the literature to assume that the two parties have the same utility function (Chiappori *et al.*, 2017, Chiappori *et al.*, 2018). Such assumption is a necessary condition for keeping the TU property discussed below.

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 TU property of a matching model will be satisfied if each individual's preferences satisfy the ACIU property.[5] 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.[6]

[4] To verify this assumption, we select married adult children who are local, living in the urban area, and having parents with age ranging from 55 to 75 in the 2010 CFPS and merge them with the following-up survey in 2012. We find 2,144 married children who are successfully tracked by the CFPS. Among 369 married children who lived with their parents or parents-in-law in 2010, 30 moved out in the 2012 survey. Among 1,775 married children who lived alone in 2010, 33 moved in with their parents or parents-in-laws by 2012. Therefore, only 2.9% of married children changed their co-residence status over a two-year period.

[5] 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.

[6] 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.

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:

$$u_{it} = \frac{1}{c_{ij}^{p}} q_{it}^{\alpha} Q_t^{1-\alpha} \tag{1}$$

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. Private goods refer to non-sharable consumption goods, such as food, clothing, and personal care products. Public goods include sharable consumption goods, such as furniture and appliances. $c_{ij}^{p}$ is the parent's congestion costs of co-residence, which captures the utility loss of shared residence due to the loss of privacy as well as the non-pecuniary utility gain/loss from culture and social norms. For example, parents may enjoy living with sons or first-born children. Therefore, the congestion cost can be larger or smaller than 1, which depends on parent's and child's characteristics.

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

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

where $q_{jt}$ is the private consumption of the adult child $j$. $c_{ij}^{c}$ is the congestion cost of adult children when they live with their parents, which can be different from the congestion cost of parents. We assume that the adult child shares the same public good, $Q_t$, as the parent.

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

$$p_{it}^{ec} p_{1jt}^{cc} Z^{cc} + q_{it} + q_{jt} + Q_t = Y_{it} + Y_{jt} \equiv Y_{ijt}$$

$Y_{it}$ and $Y_{jt}$ are the expected incomes of parent $i$ and adult child $j$ in period $t$ when they live together, respectively. $Y_{ijt}$ is the total expected income of the co-residing parent–child pair. In this model, incomes are expected incomes not actual incomes, because when parents and adult children make the co-residence decision, they cannot fully predict their future incomes.

We consider the welfare of the grandchildren in terms of savings in childcare costs instead of public goods.[7] $p_{it}^{ec}$ is parent $i$'s probability of being sick in period $t$. $p_{1jt}^{cc}$ and $p_{0jt}^{cc}$ denote the probability of having a young grandchild aged below 6 in period $t$ for child couples living with parents and those living alone, respectively. We assume that when parents are healthy, they can take care of grandchildren and child couples do not need to pay childcare costs when living with parents. When parents are unhealthy, adult children

[7] We use the CFPS data to check whether grandchildren's cognitive and non-cognitive skills are affected by whether they live with their grandparents. As shown in Appendix Table B9, with the use of the number of siblings as an instrumental variable for the co-residence status of the child couple, both OLS and IV regressions do not find evidence that living with parents affects the cognitive or non-cognitive skills of grandchildren.

provide informal care to parents and parents do not need to pay eldercare costs. However, sick parents cannot take care of grandchildren, so adult children need to pay childcare costs.[8]

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

$$U_{ij} = \max_{q_{it}, q_{jt}, Q_t, t \in \{0, N-N_0\}} \sum_{t=0}^{N-N_0} \beta^t \left[ \frac{1}{c_{ij}^c} q_{it}^{\alpha} Q_t^{1-\alpha} + \mu \frac{1}{c_{ij}^p} q_{jt}^{\alpha} Q_t^{1-\alpha} \right] \tag{3}$$

$$\text{s.t.,} \quad q_{it} + q_{jt} + Q_t = Y_{ijt} - Z_{1ijt}$$

$$q_{it}, q_{jt}, Q_t \geq 0$$

where the expenditure of living together is $Z_{1ijt} = p_{it}^{ec} p_{1jt}^{cc} Z^{cc}$. The Pareto weight of adult children is $\mu$. Note that the utility functions of parents and adult children satisfy the TU property. It follows that any efficient allocation maximizes the sum of utilities under the budget (Chiappori *et al.*, 2017, Chiappori *et al.*, 2018). Therefore, $\mu$ must be 1.

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)):

$$U_{ij} = \sum_{t=0}^{N-N_0} \beta^t \max_{q_{it}, q_{jt}, Q_t} \left[ \frac{1}{c_{ij}^p} q_{it}^{\alpha} Q_t^{1-\alpha} + \frac{1}{c_{ij}^c} q_{jt}^{\alpha} Q_t^{1-\alpha} \right] \tag{4}$$

$$\text{s.t.,} \quad q_{it} + q_{jt} + Q_t = Y_{ijt} - Z_{1ijt}$$

$$q_{it}, q_{jt}, Q_t \geq 0$$

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

$$q_{it} = \frac{D_{ij}^q}{1 + D_{ij}^q + D_{ij}^Q}(Y_{ijt} - Z_{1ijt})$$

$$q_{jt} = \frac{1}{1 + D_{ij}^q + D_{ij}^Q}(Y_{ijt} - Z_{1ijt})$$

$$Q_t = \frac{D_{ij}^Q}{1 + D_{ij}^q + D_{ij}^Q}(Y_{ijt} - Z_{1ijt})$$

where $D_{ij}^q = (\frac{c_{ij}^p}{c_{ij}^c})^{\frac{1}{\alpha-1}}$ and $D_{ij}^Q = \frac{1-\alpha}{\alpha}((\frac{c_{ij}^p}{c_{ij}^c})^{\frac{1}{\alpha-1}} + 1)$.

The total utility of living together for a parent–child pair is

$$U_{ij} = \sum_{t=0}^{N-N_0} \beta^t D_{ij}(Y_{ijt} - Z_{1ijt})$$

[8]In China, children can go to primary school when they reach age 6. Primary school belongs to the compulsory education programme. It is free and lasts all day from 8 a.m. to 5 p.m. Therefore, we assume that childcare is costly only for children below age 6.

where $D_{ij} = \left(\frac{1}{c^p_{ij}}\left(\frac{D^q_{ij}}{1+D^q_{ij}+D^Q_{ij}}\right)^\alpha + \frac{1}{c^c_{ij}}\left(\frac{1}{1+D^q_{ij}+D^Q_{ij}}\right)^\alpha\right)\left(\frac{D^Q_{ij}}{1+D^q_{ij}+D^Q_{ij}}\right)^{1-\alpha}$. The total utility of living together depends on the total income of parent $i$ and adult child $j$, but not the transfer between $i$ and $j$.

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 sharable and non-sharable consumption goods ($Q$ and $q$, respectively) 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

$$\begin{aligned} v_{it} = \max_{q_{it}, Q_{it}} \quad & q_{it}^\alpha Q_{it}^{1-\alpha} \\ \text{s.t.,} \quad & q_{it} + Q_{it} = Y_{it} - Z_i^{hc} - p_{it}^{ec} Z^{ec} \end{aligned} \tag{5}$$

Note that $Z_i^{hc}$ is the extra housing cost paid by parents if they live alone, because housing cost does not appear in the budget constraint of parents and children living together. We assume that housing does not enter the utility function as a public good, but only serves as an expenditure. With probability $p_{it}^{ec}$, 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.

Solving the first-order conditions with respect to $q_{it}$ and $Q_{it}$, we obtain

$$\begin{aligned} q_{it} &= \alpha(Y_{it} - Z_i^{hc} - p_{it}^{ec} Z^{ec}) \\ Q_{it} &= (1-\alpha)(Y_{it} - Z_i^{hc} - p_{it}^{ec} Z^{ec}) \end{aligned}$$

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

$$v_{it} = \alpha^\alpha (1-\alpha)^{1-\alpha}(Y_{it} - Z_i^{hc} - p_{it}^{ec} Z^{ec})$$

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

$$\begin{aligned} v_{jt} = \max_{q_{jt}, Q_{jt}} \quad & q_{jt}^\alpha Q_{jt}^{1-\alpha} \\ \text{s.t.,} \quad & q_{jt} + Q_{jt} = Y_{jt} - Z_j^{hc} - p_{0jt}^{cc} Z^{cc} \end{aligned} \tag{6}$$

where $Z_j^{hc}$ is the housing cost paid by adult children if they live alone. $p_{0jt}^{cc}$ is the probability of having a young grandchild (i.e., below age 6) when adult children are not living with parents. If adult children live alone and have young grandchildren, then they incur childcare costs, $Z^{cc}$.

After the first-order conditions with respect to $q_{jt}$ and $Q_{jt}$ are solved, the utility of living alone for adult child $j$ in period $t$ is

$$v_{jt} = \alpha^\alpha (1-\alpha)^{1-\alpha}(Y_{jt} - Z_j^{hc} - p_{0jt}^{cc} Z^{cc})$$

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).

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

where the total income of the adult child–parent pair when living alone is $Y_{ijt} = Y_{it} + Y_{jt}$. The expenditure of living alone is $Z_{0ijt} = Z_{ij}^{hc} + p_{it}^{ec} Z^{ec} + p_{0jt}^{cc} Z^{cc}$, 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 expected co-residence surplus is defined as the difference between the utility of living together and the utility of living alone for both parties.

$$\begin{aligned} \tilde{S}_{ij} &= U_{ij} - V_{ij} \\ &= \sum_{t=0}^{N-N_0} \beta^t \left[ D_{ij}(Y_{ijt} - Z_{1ijt}) - \alpha^\alpha (1-\alpha)^{1-\alpha} (Y_{ijt} - Z_{0ijt}) \right] \end{aligned} \tag{8}$$

The utility of living together and the utility of living alone are different in three dimensions: 1) co-residence can save living costs ($Z_{ijt}$), 2) living together is subject to a congestion cost $c_{ij}$, and 3) co-residence has economy of scale, represented by the difference between the coefficients in front of $Y_{ijt}$ ($D_{ij}$ vs $\alpha^\alpha(1-\alpha)^{1-\alpha}$).

When parents and adult children make the co-residence decision, the realized co-residence surplus is the sum of the expected co-residence surplus and an idiosyncratic preference shock $\epsilon_{ij}$, which follows a standard type-I extreme value distribution. The idiosyncratic preference shock captures the non-pecuniary co-residence utility, which may be driven by non-economic characteristics of adult children and their parents.

$$\begin{aligned} S_{ij} &= \tilde{S}_{ij} + \epsilon_{ij} \\ &= \sum_{t=0}^{N-N_0} \beta^t \left[ D_{ij}(Y_{ijt} - Z_{1ijt}) - \alpha^\alpha (1-\alpha)^{1-\alpha} (Y_{ijt} - Z_{0ijt}) \right] + \epsilon_{ij} \end{aligned} \tag{9}$$

Based on Equation (9), a family's co-residence decision depends on the congestion costs that capture the reduced utility of living together (reflected in $D_{ij}$), expense savings in housing costs, eldercare costs, and childcare costs (reflected in $Z_{0ijt}$ and $Z_{1ijt}$), and a preference shock $\epsilon_{ij}$.

#### 2.1.2 Discussion of Alternative Models

We model co-residence arrangement as a TU model. An alternative is to consider an imperfectly transferable utility (ITU) model, which allows for a more flexible form of transfer between parents and adult children, as one unit of monetary transfer provided by parents may not equal one unit of monetary transfer received by adult children. However, estimating an ITU model requires accurate information on the transfers between parents and adult children or observing multiple markets (Galichon *et al.*, 2019). In most survey data (including the one used in our paper), it is difficult to define and observe complete transfers between parents and adult children, especially when they live together. Moreover, we do not have representative and large enough samples for each market (defined as a city in our analysis). Estimating a TU model only requires information on the matching outcomes. In addition, solving the optimal co-residence arrangement in a TU model is equivalent to solving a social planner's problem to maximize the co-residence surplus of a society, which simplifies the model's solution.

Another potential concern of the TU model is that it predicts that transfers exist only between adult children and parents who live together, but not for those living alone. Empirically, the average amounts of different types of net transfers from non-coresiding adult children to parents (transfer from adult children to parents subtracted by that from parents to adult children) are very small, as shown in Appendix Table B8. The annual net monetary transfer from parents to non-coresiding adult children is only 56 RMB (0.2% of adult children's annual income and 0.8% of parent's annual income) and the annual net total transfer from adult children to parents is only 508 RMB (2.0% of adult children's annual income and 7.4% of parent's annual income). Moreover, we find that the means of net transfers from adult children to parents do not vary by children's co-residence status. The differences in all six types of transfers between non-coresiding and coresiding children listed in Appendix Table B8 are not statistically significant. This finding suggests that the transfers are likely to be driven by altruistic motives, as both co-residing and non-coresiding children should be altruistic toward parents. Therefore, we assume that the competition in the matching model is independent of an individual's altruistic motive, and thus, our model prediction does not interact with the altruistic transfers between parents and adult children. In other words, the matching model captures the benefit of co-residence on top of the transfers generated from the altruistic motive. Similarly, adult children could provide eldercare to parents and parents could provide childcare to grandchildren even when they do not live together. Our co-residence matching model captures the additional benefit of providing eldercare and childcare when living together on top of the service received when living alone.

### 2.1.3 Solution to the Co-residence Matching Model

The co-residence matching model is similar to a standard TU model except that there are some restrictions on who can live together with whom. For example, the elderly people only live with one of their children, not any young adult, and young couple only live with either side of parents, not any other elderly people. 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:

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

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.*

All the properties of a general TU matching model hold for our co-residence matching framework. In particular, solving the optimal co-residence arrangement is equivalent to solving a social planner's problem to maximize the co-residence surplus of a society. Moreover, the co-residence assignment problem always has a solution, as the number of assignments is finite. The optimal assignment is unique as long as the surplus, $S_{ij}$, has no discrete mass.[9]

The above discussion shows the existence and uniqueness of the model's solution conditional on the information on the complete family network. In the next section, we show some comparative statistics of the co-residence matching model.

### 2.1.4 Predictions of the Co-residence Matching Model

The model provides a useful framework for understanding how the co-residence surplus is affected by parent's and adult children's demographic characteristics and city-level characteristics. In this section, we discuss the model predictions through the effects of housing, eldercare, and childcare costs, as well as parent's and adult children's income on the co-residence surplus.

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

$$\tilde{S}_{ij} = \sum_{t=0}^{N-N_0} \beta^t [D_{ij}(Y_{ijt} - Z_{1ijt}) - \alpha^\alpha (1-\alpha)^{1-\alpha} (Y_{ijt} - Z_{0ijt})]$$

where $D_{ij} = \left( \frac{1}{c^p_{ij}} \left( \frac{D^q_{ij}}{1+D^q_{ij}+D^Q_{ij}} \right)^\alpha + \frac{1}{c^c_{ij}} \left( \frac{1}{1+D^q_{ij}+D^Q_{ij}} \right)^\alpha \right) \left( \frac{D^Q_{ij}}{1+D^q_{ij}+D^Q_{ij}} \right)^{1-\alpha}$, $D^q_{ij} = (\frac{c^p_{ij}}{c^c_{ij}})^{\frac{1}{\alpha-1}}$, $D^Q_{ij} = \frac{1-\alpha}{\alpha}((\frac{c^p_{ij}}{c^c_{ij}})^{\frac{1}{\alpha-1}} + 1)$, $Z_{1ijt} = p^{ec}_{it} p^{cc}_{1jt} Z^{cc}$, and $Z_{0ijt} = Z^{hc}_{ij} + p^{ec}_{it} Z^{ec} + p^{cc}_{0jt} Z^{cc}$. According to the above equation, when $0 < \alpha < 1$,

[9] 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.

the co-residence surplus increases if there is an increase in housing costs, $Z_{ij}^{hc}$ or eldercare costs, $Z^{ec}$.[10] The effect of childcare costs ($Z^{cc}$) on co-residence surplus depends on the relative size of $D_{ij}p_{it}^{ec}p_{1jt}^{cc}$ and $\alpha^{\alpha}(1-\alpha)^{1-\alpha}p_{0jt}^{cc}$. Empirically, we find that $p_{it}^{ec}p_{1jt}^{cc}$ is small, i.e., parents are healthy when adult children have young grandchildren. Therefore, the childcare expenditure is larger when living alone compared with that when living together, and the co-residence surplus also increases in childcare costs.

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

$$\frac{\partial \tilde{S}_{ij}}{\partial Y_{ijt}} = \beta^t (D_{ij} - \alpha^{\alpha}(1-\alpha)^{1-\alpha})$$

If $(D_{ij} - \alpha^{\alpha}(1-\alpha)^{1-\alpha}) < 0$, then we obtain $\frac{\partial \tilde{S}_{ij}}{\partial Y_{ijt}} < 0$, meaning that the expected co-residence surplus decreases with total income. $D_{ij}$ decreases with the congestion costs $c_{ij}^{p}$ and $c_{ij}^{c}$; thus, we will predict that the co-residence likelihood declines with income when the congestion costs exceed a certain cutoff. This comparative statics analysis suggests that, as the congestion costs go up, co-residence changes from a normal good to an inferior good, which explains why the co-residence can increase with income in the case of low congestion costs (e.g., Italy) while it decreases with income when congestion costs are high (e.g., China).[11]

In sum, the model generates the following predictions:

Prediction 1. When $0 < \alpha < 1$, co-residence likelihood increases with housing costs and eldercare costs.

Prediction 2. When $(D_{ij} - \alpha^{\alpha}(1-\alpha)^{1-\alpha}) < 0$, co-residence likelihood declines as the income of parents or adult children increases.

Although these predictions are directly supported by the modeling assumption, we only observe housing costs, but not eldercare costs, childcare costs, or income (when individuals not living in the surveyed household) in the data. We next present some testable results regarding how the co-residence surplus is affected by parent's and adult children's demographic characteristics, which serve as proxies for eldercare costs, childcare costs, or income and can be directly tested in the data.

We start with the education of parents. In our model, when $(D_{ij} - \alpha^{\alpha}(1-\alpha)^{1-\alpha}) < 0$, co-residence likelihood declines as the income of parents or adult children increases. Empirically, we find that parents' education is associated with higher income and lower probability of needing eldercare. An increase in parental education leads to an increase in $Y_{ijt}$ and a decline in $Z_{0ijt}$, and hence, a decline in the co-residence surplus. Therefore, the first result derived from our model's predictions is that an increase in parents' education leads to a reduction in the co-residence likelihood if higher-educated parents have higher income and are less likely to need eldercare compared to less-educated parents.

[10] This prediction depends on the assumption that housing enters the utility function as an expenditure rather than a public good. If we assume that housing is a public good with price $p$, $D_{ij}^{Q}$ becomes $\frac{1-\alpha}{p\alpha}((\frac{c_{ij}^{p}}{c_{ij}^{c}})^{\frac{1}{\alpha-1}}+1)$. Therefore, the model predicts that the co-residence surplus first increases with $p$, and then declines with it.

[11] Manacorda and Moretti (2006) find that in Italy, parents' likelihood of co-residence increases with their income, while we find the opposite using China Family Panel Study.

For adult children, we find that their education is positively associated with higher income, and therefore, negatively correlated with the co-residence surplus.[12] Therefore, the second result of our model is that an increase in adult children's education leads to a reduction in the co-residence likelihood if their income increases with education.

The parent–child age gap can also influence the co-residence surplus. A larger parent–child age gap suggests that adult children are younger. Empirically, we find that younger adult children are more likely to have a young grandchild and they also have lower income compared to older adult children. As discussed above, because the chance of paying childcare costs when adult children living with parents is small, the co-residence surplus increases in childcare costs. Therefore, the co-residence surplus increases with the parent–child age gap due to the need for childcare and the income effect.[13] The last result of our model is that an increase in parent–child age gap leads to a rise in the co-residence likelihood if younger adult children are more likely to have a young grandchild and have lower income than older adult children.

## 2.2 Stage I: Marriage Matching

Now we move to the first stage of the problem, where young men and young women form marriage matching. We develop a TU model of marriage matching where the marriage surplus is a function of the characteristics of young men and women, as well as the characteristics of their parents.

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

$$\tilde{\Phi}(\tilde{x}, \tilde{y}) = \Phi(x, y) + \chi(\tilde{x}, y) + \xi(\tilde{y}, x) \tag{10}$$

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 $\chi$ and $\xi$ follow type-I extreme value distribution with scale factor $\sigma_1$ and $\sigma_2$, respectively. Following Choo and Siow (2006), Siow (2015), and Galichon and Salanié (2021), we impose a separability assumption 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,

$$\Phi(x, y) = f(x, y) + h(x, y) \tag{11}$$

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

[12] The education of adult children may also affect the probability of having young grandchildren, which in turn affects the need for childcare, but we find no such correlation in China, as shown in Table 5.

[13] As our model assumes that parents make the co-residence decision at the same age and have the same life expectancy, we attribute the effect of parent–child age gap more through the channel of childcare instead of eldercare. If we relax the assumption that all parents make the co-residence decision at the same age, then it is also likely that the larger parent–child age gap suggests that parents are older, and the co-residence surplus increases because of the need for eldercare.

joint utility when they live without parents. The second part is the expected co-residence utility $h(x, y)$, which is non-zero if the couple co-resides with either side of parents.

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). 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. While the co-residence utility could depend on the birth order of an adult and the characteristics of his/her siblings, we find that these characteristics only have a slight impact on the marriage surplus empirically. Therefore, we exclude them from the marriage surplus function in the approximation. As a result, the total marriage surplus $\Phi(x, y)$ is a function of education, age, parental education, parental age, and number of siblings of the husband and wife, some of which may not be directly related to the utility of marriage ($f(x, y)$) but reflect the expected utility of co-residence ($h(x, y)$).[14]

Note that we do not need to identify $f$ and $h$ separately because it is sufficient to identify $\Phi$ to predict the marriage matching. 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 $\phi^k$, with unknown weights $\lambda$

$$\Phi(x, y) = \sum_k \lambda_k \phi^k(x, y) \tag{12}$$

In our case, we choose a quadratic approximation, following Fox (2018)[15]

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

where $E$ denotes own education, $A$ denotes own age, $PE$ denotes parent's education, $PA$ denotes parent's age, and $N$ is the number of siblings. In this quadratic approximation, the first-order terms (such as $E_m$ and $E_w$) and the quadratic terms (such as $E_m^2$ and $E_w^2$) do not affect marriage sorting but only the probability of being single, so we keep only 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 who are similar

[14] Note 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.

[15] 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.

in age).[16] This general functional form includes a flexible set of the interaction terms: 1) that between an adult child's characteristics and his/her spousal 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. Our specification allows us to capture the positive assortative matching in education, age, number of siblings, parental education, and parental age.

Our marriage matching model incorporates the co-residence consideration at the time of marriage with a few assumptions. 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 avoid the non-existence of stable matching.

In summary, we incorporate not only the characteristics of young men and women, but also those of their parents into our marriage model. These characteristics may affect a couple's marriage surplus through their co-residence decision. As our main focus is to model the intergenerational co-residence via a novel application of Shapley–Shubik–Becker's seminal work, we take a reduce-form approach to model the marriage surplus function, which contains the expected utility from co-residence.

The marriage matching connects two families — the husband's family and the wife's family. The co-residence surplus and the family network jointly predict the living arrangement in a family. In the next section, we discuss the importance of the family network and the challenge of not observing the complete network.

## 2.3 Family Network

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 available data do 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 of parent $i$ are indicated by $ij$, $j \in \{1, 2, ...\}$. For example, parent $A$ has two adult children, $A1$ and $A2$.

[16]Given that we observe only the matching patterns, we cannot distinguish between vertical and horizontal preferences. Both types of preferences predict positive assortative matching.

The links in the middle of the diagram represent marriages of adult children. In this 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. 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, our 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.[17] 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 impute the missing marriage links in the family network. First, we estimate a marriage matching model using data on the observed married couples. We then use the estimates from the marriage matching model to predict the unobserved marriage links for married children. We relay the discussion of the estimation of the marriage matching model to Section 4.3 and the imputation of marriage links to Section 4.4.

# 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

[17]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.

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 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 reorganize 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 ages of parents of the husband and wife are between 55 and 75; 2) the husband and wife are both urban local residents; and 3) all of the siblings of the husband and wife are married.[18] 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.[19] We then further restrict couples to be urban and local. A couple is identified as urban and local if they live in an 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 very different from that of married adult children. A total of 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. Lastly, 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. Note 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

[18]We include couples with widowed parents and exclude couples with divorced or separated parents. 34% of young couples in our family sample have at least one side of parents who are widows or widowers. Only 0.2% of parents are divorced or live separately in our family sample.

[19]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.

sampled, which is proportional to the number of children in the family ($N$). As a result, when 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 labour market, our model cannot feasibly incorporate this additional location decision. 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.[20] 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 because the CFPS lacks 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 B10.

## 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 adult child. In our weighted adult child sample, 45.5% of the 3,709 urban married adults are first-born

[20] 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. We find two example papers that jointly model marriage and location choices. Dupuy (2018) uses a matching model to capture marriage and location choices jointly. Gemici (2007) uses a dynamic model with intra-household bargaining to jointly model marriage and location choices, abstract from competition.

children of their parents. In our couple sample, 30% of the 659 married couples have a young child under six years old and they are more likely to live with the husband's parents than the wife's parents.[21] 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 RMB (approximately $600 USD). We scale the housing price through multiplying the average housing price per square meter by 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. These reduced-form results help us verify the key components of our model. Overall, men and first-born children are more likely to live with their parents, suggesting that social norms may play a role in affecting the co-residence arrangement. Adult children's co-residence likelihood decreases with children's and parent's education. This implies that the income effect is negative, and co-residence is an inferior good in China. These findings are consistent with the first two results in Section 2.1.4. We also observe that a larger age gap between parents and adult children increases the chance of co-residing, which supports the last result. Lastly, high housing costs are associated with a larger probability of co-residing with one's own parents, which is consistent with Prediction 1.

# 4 Identification and Estimation

## 4.1 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 55, an age at which most parents are retired. Life expectancy $N$ is set to be 75.

We allow the congestion costs of parents and adult children ($c_{ij}^p$ and $c_{ij}^c$) to differ by the gender and birth order of the co-residing adult child to incorporate the "culture" component.

$$c_{ij}^k = \kappa_0^k + \kappa_1^k H_j + \kappa_2^k O_j \quad \text{for } k \in \{p, c\} \tag{14}$$

where $H_j$ is an indicator of whether the parents live with their son (i.e., whether the child couple lives with the husband's parents), and $O_j$ is an indicator of whether the parents live with their oldest child.

We assume that both parent's and young couple's expected income ($Y_{it}/Y_{jt}$) are functions of education ($E_i/E_j$) and age ($A_{it}/A_{jt}$), as well as city fixed effects ($\psi_k^p/\psi_k^c$). Since most parents in our sample (aged 55

[21] 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). Therefore, 30% of adult children do not need childcare, but they can still provide eldercare to parents and save housing costs when they live with their parents.

to 75) have retired, their income and labour supply are not affected by the living arrangement. We check whether the co-residence status affects a young couple's income because living with parents may increase or decrease adult children's labour supply through the eldercare or childcare channel (Maurer-Fazio *et al.*, 2011, Compton and Pollak, 2014) and we do not find any statistical evidence. Section 5.2 provides more discussions on this issue.

$$Y_{it} = \alpha_0^p + \alpha_1^p E_i + \alpha_2^p A_{it} + \alpha_3^p A_{it}^2 + \psi_k^p \tag{15}$$

$$Y_{jt} = \alpha_0^c + \alpha_1^c E_j + \alpha_2^c A_{jt} + \alpha_3^c A_{jt}^2 + \psi_k^c \tag{16}$$

We further specify the total housing costs of a parent–child pair $(i, j)$ living in city $k$ as a function of the adjusted housing price (by GDP per capita) at the city level $P_k$,

$$Z_{ij}^{hc} = \delta_0 + \delta_1 P_k \tag{17}$$

In addition, we assume that the expected probability of parents needing eldercare ($p_{it}^{ec}$) is a function of parents' average age and education. Section 5.2 examines whether co-residence affects the probability of needing eldercare and finds no evidence. The expected probability of having a young grandchild aged below 6 ($p_{x_{ij}jt}^{cc}$) is a function of a young couple's average age, education, and city fixed effects ($\phi_k$). In addition, we allow the co-residence status ($x_{ij}$) to affect the fertility decision.

$$p_{it}^{ec} = \max\{0, \min\{\gamma_0^p + \gamma_1^p E_i + \gamma_2^p A_{it} + \gamma_3^p A_{it}^2, 1\}\} \tag{18}$$

$$p_{x_{ij}jt}^{cc} = \max\{0, \min\{\gamma_0^c + \gamma_1^c E_j + \gamma_2^c A_{jt} + \gamma_3^c A_{jt}^2 + \gamma_4^c x_{ij} + \phi_k, 1\}\} \tag{19}$$

where $x_{ij}$ is an indicator of whether adult child $i$ lives with parent $j$.

## 4.2 Identification of the Co-residence Matching Model

We identify the parameters in two steps. The first step is to identify the parameters in the income equations of parents and adult children (Equations (15), (16)), parent's health status (Equation (18)), and fertility of adult child couples (Equation (19)). Most of the variables in these equations are exogenous, such as age and education. The only potential endogenous variable is the co-residence status, and we identify the effect of co-residence using an instrumental variable model. In particular, we choose the number of siblings as the instrument for whether the adult child lives with his/her parents in adult children's fertility regression. The intuition is that when adult children have more siblings, they are less likely to live with their parents. Moreover, the number of siblings does not directly affect adult children's fertility decision except through the co-residence decision. Therefore, we can estimate these parameters outside of the co-residence matching model, which we refer to as "exogenous parameters."

In the second step, we identify the rest of the parameters, which include 1) congestion costs, 2) housing costs, eldercare costs, and childcare costs, and 3) the distribution of idiosyncratic shocks, which we refer to as "endogenous parameters." Now we discuss the identification of the endogenous parameters.

Recall that the expected co-residence surplus is

$$\begin{aligned}\tilde{S}_{ij} = \sum_{t=0}^{N-N_0} \beta^t [(D_{ij} - \alpha^\alpha (1-\alpha)^{1-\alpha}) Y_{ijt} \\ - D_{ij} p_{it}^{ec} p_{1jt}^{cc} Z^{cc} + \alpha^\alpha (1-\alpha)^{1-\alpha} (\delta_0 + \delta_1 P_k + p_{it}^{ec} Z^{ec} + p_{0jt}^{cc} Z^{cc})]\end{aligned}$$

where $D_{ij} = \left( \frac{1}{c_{ij}^p} \left( \frac{D_{ij}^q}{1+D_{ij}^q+D_{ij}^Q} \right)^\alpha + \frac{1}{c_{ij}^c} \left( \frac{1}{1+D_{ij}^q+D_{ij}^Q} \right)^\alpha \right) \left( \frac{D_{ij}^Q}{1+D_{ij}^q+D_{ij}^Q} \right)^{1-\alpha}$, $D_{ij}^q = (\frac{c_{ij}^p}{c_{ij}^c})^{\frac{1}{\alpha-1}}$, and $D_{ij}^Q = \frac{1-\alpha}{\alpha}((\frac{c_{ij}^p}{c_{ij}^c})^{\frac{1}{\alpha-1}} + 1)$.

Although the model has the flexibility to allow parents and adult children to have different congestion costs, we can only identify $D_{ij} - \alpha^\alpha(1-\alpha)^{1-\alpha}$ and $\alpha^\alpha(1-\alpha)^{1-\alpha}\delta_0$, but not $c_{ij}^p, c_{ij}^c, \alpha, \delta_0$ separately. Therefore, we impose the following normalizations: 1) parents and adult children have the same congestion costs ($c_{ij}^p = c_{ij}^c = c_{ij}$, i.e., $\kappa_i^p = \kappa_i^c$ for $i = \{0, 1, 2\}$); and 2) the effects of private and public consumption on the utility are the same ($\alpha = 0.5$).[22] Moreover, we normalize the relative scale of the preference shock on the co-residence utility ($\epsilon_{ij}$)—by assuming that it follows a standard type I extreme value distribution so its standard deviation it is one. Because we do not observe the surplus function directly (but only the matching outcomes), we will not be able to separately identify the scale of the utility and the scale of the preference shock. However, all the counterfactual results are invariant to these normalizations.

The expected co-residence surplus for a parent–child pair then becomes

$$\tilde{S}_{ij} = \sum_{t=0}^{N-N_0} \beta^t [(\frac{0.71}{c_{ij}} - 0.5) Y_{ijt} - \frac{0.71}{c_{ij}} p_{it}^{ec} p_{1jt}^{cc} Z^{cc} + 0.5(\delta_0 + \delta_1 P_k + p_{it}^{ec} Z^{ec} + p_{0jt}^{cc} Z^{cc})]$$

To demonstrate the identification of our model, we show how a subset of data can identify the 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.[23] The advantage of focusing on this subsample is that no competition exists between siblings, and therefore, it becomes a two-sided matching problem with a simple network structure. In this specific couple sample, there are three options: 1) the young couple lives with the husband's parents

[22]Note $D_{ij} = \frac{1}{c_{ij}^p} g(\frac{c_{ij}^c}{c_{ij}^p})$, where $g(x) = \left( \left( \frac{x^{\frac{1}{\alpha-1}}}{f(x)} \right)^\alpha + x \left( \frac{1}{f(x)} \right)^\alpha \right) \left( \frac{\frac{1-\alpha}{\alpha}(x^{\frac{1}{\alpha-1}}+1)}{f(x)} \right)^{1-\alpha}$ and $f(x) = 1 + x^{\frac{1}{\alpha-1}} + \frac{1-\alpha}{\alpha}(x^{\frac{1}{\alpha-1}} + 1)$. For any $\alpha$, we can always pick $c_{ij}^p$ and $\delta_0$ to keep $D_{ij} - \alpha^\alpha(1-\alpha)^{1-\alpha}$ and $\alpha^\alpha(1-\alpha)^{1-\alpha}\delta_0$ constant. Therefore, we can normalize $\alpha = 0.5$. In addition, by fixing $\frac{c_{ij}^c}{c_{ij}^p}$ and choosing $c_{ij}^p$, we can have any value of $D_{ij}$. Therefore, we can normalize $c_{ij}^c = c_{ij}^p$.

[23]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 B11. 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.

and the wife's parents live alone ($i = 1$), 2) the young couple lives with the wife's parents and the husband's parents live alone ($i = 2$), and 3) the young couple, the husband's parents, and the wife's parents all live alone ($i = 0$). The optimal living arrangement is the one that maximizes the co-residence surplus of the three parties (the young couple, husband's parents, and wife's parents). Here, our matching model degenerates into a multinomial choice model. Since the idiosyncratic shock to the co-residence surplus is additive and follows the type-I extreme value distribution, the probability of a couple $j$ choosing option $i$ is

$$P(i) = \frac{\exp(\tilde{S}_{ij})}{\sum_{k=0}^{2} \exp(\tilde{S}_{kj})}.$$

Therefore, there is a one-to-one mapping from the observed co-residence patterns to the surplus function (Choo and Siow, 2006). The identification of the endogenous parameters can be shown via the partial derivative of co-residence surplus with respect to these parameters given the invert mapping from the co-residence likelihood to the surplus function. For example, if an increase in the parameter leads to a higher co-residence surplus between a parent–child pair, then it implies that the parent-child pair is more likely to be observed to live together. Below, we introduce how to achieve the identification in this simplified matching model.

We start by demonstrating that the congestion cost parameters in our model are identifiable. According to Equation (20), the marginal effect of the total (expected) income of ($Y_{ijt}$) on the expected co-residence surplus can be mapped one to one into the congestion cost ($c_{ij}$).

$$\frac{\partial \tilde{S}_{ij}}{\partial Y_{ijt}} = \beta^t \left( \frac{0.71}{c_{ij}} - 0.5 \right) \tag{20}$$

where $\beta$, the annual discount rate, is set to 0.95. Equation (14) indicates that congestion costs ($c_{ij}$) include three parameters: an average effect ($\kappa_0$), a gender difference parameter indicating whether parents prefer to live with sons than daughters ($\kappa_1$), and a birth-order effect parameter indicating whether parents prefer to live with the oldest child ($\kappa_2$). A stronger income effect on the co-residence surplus and co-residence likelihood suggests a larger $\kappa_0$. A larger gap between the co-residence rate of living with sons (first-born) than daughters (non-first-born) indicates a larger $\kappa_1$ ($\kappa_2$).

Next, we describe how we identify housing cost, eldercare cost, and childcare cost parameters. First, the total housing cost ($Z_{ij}^{hc}$) contains two parameters: national average housing cost ($\delta_0$) and city-specific housing price ($\delta_1$), as shown in Equation (17). The average co-residence likelihood among all families pins down the national average housing cost, $\delta_0$, as higher average housing costs are associated with larger co-residence surpluses. Then, as shown in Equation (21),

$$\frac{\partial \tilde{S}_{ij}}{\partial P_{ij}} = \sum_{t=0}^{N-N_0} 0.5 \beta^t \delta_1, \tag{21}$$

the marginal effect of city housing price on the co-residence surplus identifies the city-specific housing price parameter.

Second, the eldercare cost, $Z^{ec}$, corresponds to the marginal effect of parent's education on the co-residence surplus as shown in Equation (22).

$$\frac{\partial \tilde{S}_{ij}}{\partial E_i} = \sum_{t=0}^{N-N_0} \beta^t \left( \left( \frac{0.71}{c_{ij}} - 0.5 \right) \alpha_1^p + 0.5 \gamma_1^p Z^{ec} \right) \tag{22}$$

This equation shows that parental education affects the expected co-residence surplus through the income (captured by $\alpha_1^p$) and health (captured by $\gamma_1^p$) of parents. Since $c_{ij}$ has been identified in the first step and $\alpha_1^p$ and $\gamma_1^p$ are exogenous parameters, the eldercare cost is identified from Equation (22).

Third, the childcare cost, $Z^{cc}$, is identified through the marginal effect of parent–child age gap on the co-residence surplus. 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. By taking the partial derivative of the expected co-residence surplus with respect to the parent–child age gap (see Equation (23)), we can see that the parent–child age gap affects the expected co-residence surplus through the income and childcare cost of adult children. Once $c_{ij}$ has been identified and exogenous parameters, $\alpha_2^c, \alpha_3^c, \gamma_2^c$, $\gamma_3^c$, and $p_{it}^{ec}$, 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 surplus (as well as the co-residence likelihood).

$$\frac{\partial \tilde{S}_{ij}}{\partial G_j} = -\sum_{t=0}^{N-N_0} \beta^t \left( \left( \frac{0.71}{c_{ij}} - 0.5 \right) (\alpha_2^c + 2\alpha_3^c G_j) + (0.5 - \frac{0.71}{c_{ij}} p_{it}^{ec})(\gamma_2^c + \gamma_3^c G_j) Z^{cc} \right) \tag{23}$$

## 4.3 Empirical Forecasting of Marriage Links

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 TU model of marriage matching. However, we face several challenges in this estimation. First, we have a relatively small sample; there are 659 couples in our couple sample. We have a marriage model with multiple characteristics and the data have limited sample size, which is why we cannot achieve the ideal precision if we estimate the marriage model non-parametrically following the seminal work of Choo and Siow (2006).[24] Second, we cannot follow Fox (2010, 2018) to estimate the parametric model using the maximum 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.

[24] 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), then 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.

To overcome these challenges, we follow Galichon and Salanié (2021) by using a moment matching estimator. Galichon and Salanié (2021) 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).

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 lower 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 or women 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 (2021) is robust for small samples. To avoid overfitting, instead of estimating the marriage surplus for each of the 192 types, we estimate a parametric marriage surplus function, as shown in Equation (13), which includes 25 parameters.

Following the procedure discussed above, we use our couple sample to estimate the marriage surplus function. We normalize the scale factors of the two shocks ($\sigma_1$ and $\sigma_2$) to 0.5 so that $\sigma_1 + \sigma_2 = 1$. Standard errors are calculated using the bootstrap method. The details of the estimation procedure are discussed in Appendix A.1. In Appendix A.2, we show that our estimation procedure is not subject to selection bias. We use the estimates of the marriage model to impute the missing marriage links, as will be discussed in the next section.

## 4.4 Co-residence Matching Estimation

Once we demonstrate the identification of our co-residence model and the feasibility of our marriage matching estimation, we introduce the estimation procedure of the co-residence model.

First, we estimate the parameters that are exogenous to our model. We estimate parental income, following Equation (15), using parents in our family sample. We estimate adult children's income, following Equation (16), using our adult child sample. We estimate whether the child couple has a young grandchild aged below six (referred to as "fertility decision" below), following Equation (19), using our couple sample. We estimate parents' health, following Equation (18), using a sample of urban, local residents aged 55 to 75 from the CHARLS data. 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 fertility rate using the 2010 Census data.[25]

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

Next, we estimate the endogenous parameters in four steps:

1. A total of 1,314 families for each of the 104 cities from the CFPS national family sample are chosen 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, the family networks in each city are completed by simulating the marriage links for all married children in the bootstrap sample.

3. The income of parents and adult children, parents' health status, and the fertility of adult child couples are predicted by using the exogenous parameters.

4. Optimal co-residence assignments in each city are simulated by 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 (with 1,314 families). As each city has only 2–128 families, we use sampling to expand our data set.[26] For each city, we randomly draw, with replacement, 1,314 families from the national family sample (instead of the city family sample).[27] 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. 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.[28]

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 married male and female adult children in the city-level bootstrap sample. The assignment is based on the demographics of young men and women, the estimates of $\lambda$'s in the marriage matching model, and simulated marriage matching shocks ($\chi$ and $\xi$) for each pair of potential spouses. We ensure that 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. 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.[29]

In the third step, 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

[26] 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.
[27] 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.
[28] Each city has different realized family characteristics due to the randomness in the sampling process.
[29] As we only have between 1 and 74 couples in each city, we cannot estimate the marriage model separately for each city. Each city can still have different realized couple characteristics due to the randomness in the simulation process.

each city to have different average levels of income for parents and adult children, as well as different average number of grandchildren. We also predict parents' health status using the parent demographics, but 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 ($\epsilon_{ij}$).

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, following column (1) of Table 1. As discussed in Section 4.2, these moments are sufficient to identify all the parameters in our model. We then use our co-residence model to predict the co-residence decisions of 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 a bootstrap method to calculate the standard errors of our estimated parameters. In each bootstrap, we re-sample married couples from our couple sample and re-estimate the marriage matching model. We then expand the sampled married couples to include their siblings and parents to construct a family sample, which is used to estimate endogenous parameters in the co-residence matching model. Therefore, our standard errors take into account the uncertainty in the estimation of the marriage model and the co-residence model.

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 their 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.[30] For example, we find that the co-residence rate has a non-monotonic relationship with adult children's education (see Appendix Table B7) 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 analyse the marginal effect of one characteristic holding other characteristics constant.

# 5 Estimation Results

## 5.1 Marriage Matching Model 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 in all five dimensions.

Appendix Tables B1 to B5 evaluate the fit of our marriage matching model. In Panel A of each table, we calculate the empirical probability of marriage conditional on the 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 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 Panels A and B.

## 5.2 Co-residence Model Results: Exogenous Parameters

We then present the co-residence model estimation results in two steps. In the first step, we show the estimated parameters that are exogenous to our model, including the incomes of parents and adult children, parental health status, and the fertility status of the adult child couple. In the second step, we present the estimates of the endogenous parameters, which include congestion costs, housing costs, eldercare costs, and childcare costs. This section shows the estimation results of the first step.

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

We use a Mincerian-type regression to predict the incomes of parents and adult children following Equations (15) and (16). The incomes of parents and adult children include labour and asset income. 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). The results of which are presented 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-shaped 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.

One potential concern for the Mincer regression on child couple's income is that co-residence could affect children's labour supply decision, and therefore, their income. In Appendix Table B12, we examine this possibility by including an indicator of whether the child couple lives with either side of their parents in the regression. In the OLS regression, we find that living with either side of parents reduces the income and labour force participation of male adult children, but increases the income and labour force participation of female adult children. However, the OLS regression result could be biased due to the endogeneity problem or reverse causality problem. We then use the number of siblings as an instrument for co-residence, because having more siblings reduces the likelihood that the adult child lives with his/her parents, but does not directly affects his/her labour market outcomes. In the IV specification, we obtain strong first stage and the effects of co-residence on income and labour supply are no longer significant for male and female adult children in the second stage.

We use a linear probability model to predict a parent's health status based on the parent's age and education, following Equation (18). We define a parent as unhealthy if his/her self-reported health status is poor.[31] Columns (5) and (6) of Table 4 show that the proportion of parents who are in poor health increases with age and decreases with education levels. We also investigate whether co-residence affects parental health. To solve the potential endogeneity problem and reverse causality problem of the co-residence status, we use two instruments, namely, the share of male children and the community co-residence share, and both of them increase the likelihood of living with children but do not directly affect parents' health. Appendix Table B13 shows the OLS and IV regression results, and none of them show that living with children would affect elderly parents' health status. Both instruments have a strong first stage and yield similar second-stage results, which validates our hypothesis that a parent's health status largely depends on exogenous shocks and predetermined factors rather than co-residence status. We also check whether heterogeneous effects of co-residence on parental health exists from parent's education or age and find no significant effect.[32] Therefore, we only use parent's age and education to predict the likelihood that parents are in poor health and therefore need eldercare, as in Column (5) of Table 4.

[31] 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.

[32] Johar and Maruyama (2014) use Indonesian data and find that co-residence reduces elderly parents' health in general, but the effect is positive for parents who had lived independently.

To predict whether a family has a grandchild under six years old, we use a linear probability model following Equation (19), and the regressors include the average age, education level, co-residence status of the couple, and city fixed effects. Again, we use the number of siblings as an instrument for the co-residence status. As shown in Table 5, in both OLS and IV specifications, co-residing with parents increases the chance that the child couple has a young grandchild. Moreover, the chance of having a young grandchild declines with the couple's age but does not vary by their educational achievement. We use the estimates in the IV model (Column (2) of Table 5) to predict the likelihood of having grandchildren for children living with parents ($p^{cc}_{1jt}$) and living alone ($p^{cc}_{0jt}$) in our co-residence model estimation.

### 5.3 Co-residence Model Results: Endogenous Parameters

In the second step, we estimate endogenous parameters in our model. Table 6 shows the estimation results of the co-residence model using indirect inference. The first four rows show the estimates of the congestion costs for first-born male, non-first-born male, first-born female, and non-first-born female children, which reflect adult children's and parents' preferences for living together. Recall that the co-residence surplus in each period is $(\frac{0.71}{c_{ij}} - 0.5)(Y_{ijt} - Z_{1ijt}) + 0.5Z_{0ijt}$. We define $C_{ij} = \frac{0.71}{c_{ij}} - 0.5$ and present $C_{ij}$ instead of $c_{ij}$ because the former directly measures the utility loss/gain from co-residence. The smaller the $C_{ij}$, the higher the congestion cost $c_{ij}$ and the less preferable co-residency is. We find that the $C_{ij}$ is negative for all four types of children, suggesting that co-residence surplus declines with the total expected income of parents and adult children. Moreover, the $C_{ij}$ of female children is much smaller than that of male children, and $C_{ij}$ is also smaller for non-first-born children than for first-born children. To quantify the scale of the congestion costs, we calculate the congestion costs measured by housing price for male and female children. For first-born male children from a median-income family with a total annual income of 30,000 RMB, the congestion costs lead to a loss in the co-residence surplus that is equivalent to that from a 6.9% decline in the housing price.[33] In contrast, the congestion costs of first-born female (non-first-born male) children from a median-income family are equivalent to the loss in the co-residence surplus from a 197.6% (32.7%) decline in the housing price.[34] The estimates of congestion costs suggest that co-residence is affected not only by economic factors, but also by cultural factors that are related social norms, i.e., parents prefer living with sons and the oldest child.

We then show our estimates for housing costs, eldercare costs, and childcare costs in the next four rows of Table 6. Given that the error term is normalized to follow a standard type I extreme value distribution, these costs are also scaled accordingly. Nevertheless, we can compare the relative size of housing costs, eldercare costs, and childcare costs. For parents, the savings in eldercare costs from a 1 percentage point

[33] A decrease in the housing price would make it cheaper for parents or adult children to buy or rent a house when they live alone, which results in a decline in the co-residence surplus.

[34] Consider a median-income family where the total annual income of parents and adult children is 30,000 RMB. For first-born male children, the formula to translate congestion costs into housing price is $0.078 \times 10^{-6} \times 30000/0.034 = 6.9\%$. For first-born female children, the formula is $2.239 \times 10^{-6} \times 30000/0.034 = 197.6\%$. For non-first-born male children, the formula is $0.371 \times 10^{-6} \times 30000/0.034 = 32.7\%$.

(ppt) increase in the probability of getting sick is equivalent to the savings in housing costs from a 16.1% increase (0.548/0.034) in the housing price. For adult children, the savings in childcare costs from a 1 ppt increase in the probability of having a young grandchild is equivalent to the savings in housing costs from a 3.9% increase (0.133/0.034) in the housing price. Figure 5 presents how a continuous change in housing price affects the co-residence rate, with more details to be discussed in the counterfactual analysis in Section 7. On average, a 10% increase in housing price leads to a 1.8% increase in the co-residence rate for adult children and parents. This finding suggests that a 1 ppt increase in the probability of getting sick leads to a 2.9% increase in the co-residence rate and a 1 ppt increase in the probability of having a young grandchild leads to a 0.7% increase in the co-residence rate.[35] In sum, we find that savings on housing costs, eldercare costs, and childcare costs are important incentives for families to live together.

To evaluate the predictive power of our model, we compare two sets of moments in Table 7. The odd columns of Table 7 show the regression results using our CFPS data (called "data"), and the even columns of Table 7 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 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. 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, first-born, less educated, have fewer siblings, have a larger parent–child age gap, have parents who are less educated, or live in a city with higher housing prices.

The second set of moments are those not targeted in our estimation, which are presented in Columns (3) to (6) in Table 7. Columns (3) and (4) check the child-side competition by estimating the effect of average education among siblings on the co-residence likelihood. The regression results suggest that the co-residence likelihood increases with the average education of siblings, 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. 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 correlation is found 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

[35]The formula to calculate the change in the co-residence rate for a 1 ppt increase in the probability of getting sick is $1.8/10 \times 16.1$. The formula to calculate the change in the co-residence rate for a 1 ppt increase in the probability of having a young grandchild is $1.8/10 \times 3.9$.

live with the spouse's siblings; therefore, the likelihood of living with one's own parents increases. The fact that we are able to fit these untargeted moments gives us confidence that the matching model is the correct model that captures two-sided competitions.

As supplementary results, we show the model fit in Appendix Tables B6 and B7. The former presents how parents' co-residence likelihood varies based on their own education levels, and the latter shows how an adult child's co-residence likelihood varies by his or her own characteristics; our model can match all these moments reasonably well.

# 6 Competing Models

In this section, we compare our co-residence matching model with two alternative models. We simplify our matching model by removing child-side competition and then by removing parent-side competition. We estimate these models and compare their model fit with our matching model.

The first model (A1) considers the co-residence arrangement between a child couple and both sets of parents. This approach degenerates the matching problem into a multinomial logit model with three living arrangements: 1) the couple lives with the husband's parents, 2) the couple lives with the wife's parents, and 3) the couple lives without parents. In this model, we capture the parent-side competition by restricting a child couple to living with at most one set of parents. However, we ignore competition from siblings in this model. When an adult child wants to live with his/her parents, his/her siblings may also want to live with these parents. As most of the parents in our sample have multiple children, ignoring child-side competition could result in parents living with multiple children. The husband's parents are denoted as Parent 1 and the wife's parents are denoted as Parent 2. The probability that the couple $j$ lives with parent $i$ is

$$Pr(\text{Child } j \text{ lives with Parent } i) = \frac{\exp(\tilde{S}_{ij})}{1 + \sum_{k=1}^{2} \exp(\tilde{S}_{kj})} \tag{A1}$$

where $\tilde{S}_{ij}$ is the same as that in Equation (8).

The second model considers the co-residence arrangement between one parent couple and multiple adult children. This approach again degenerates the matching problem into a multinomial logit model where parents choose an adult child to live with or live alone. In this case, we capture the child-side competition by restricting one parent to live with at most one child. However, we rule out parent-side competitions in this model. When parents want to live with a child couple, parents-in-law may also want to live with the same child couple. Ignoring parent-side competition could result in a child couple living with both sets of parents. Suppose that parent $i$ has $N$ adult children; then, the probability that the parent lives with adult child $j$ is

$$Pr(\text{Parent } i \text{ lives with Child } j) = \frac{\exp(\tilde{S}_{ij})}{1 + \sum_{k=1}^{N} \exp(\tilde{S}_{ik})} \tag{A2}$$

We estimate the multinomial logit model with both sets of parents (A1) using our couple sample, which contains information on the child couple, the husband's parents, and the wife's parents. We estimate the multinomial logit model with multiple children (A2) using our family sample, which contains information on parents and all of their children.

The last two columns of Table 6 show the estimation results of the two multinomial logit models. If we ignore child-side competition or parent-side competition, then the estimated $C_{ij}$ ($\frac{0.71}{c_{ij}} - 0.5$) for first-born male children becomes positive, suggesting that for first-born male children, the value of co-residence increases with the total income of parents and adult children. These estimates are substantially different from the matching model, where we find negative $C_{ij}$ for all four types of children. In addition, both models overestimate eldercare costs. The matching model estimates eldercare costs to be 0.548, suggesting that a 1 ppt increase in the probability of becoming sick is equivalent to a 16% increase (0.548/0.034) in the housing price. In the multinomial logit model with both sets of parents (Model A1), the estimate of eldercare costs implies that a 1 ppt increase in the probability of becoming sick is equivalent to a 120% increase (2.272/0.019) in the housing price. In the multinomial logit model with multiple children (Model A2), the eldercare cost estimate suggests a 23% increase (0.819/0.036) in the housing price. The two alternative models also underestimate housing costs and overestimate childcare costs compared with the baseline model.

Table 8 shows the model fit of the two multinomial logit models by replicating what we did in Table 7. The regressions of the multinomial logit models use the same simulated data as in the matching model but a different decision framework (without child-side or parent-side competition) to predict the co-residence assignment. In columns (1)-(3), we run baseline regressions, which only include the characteristics of parents and adult children. In columns (4)-(6), we add siblings' attributes to the baseline regression and test whether the two models can capture child-side competition. In columns (7)-(9), we add the attributes of the spouse's parents and test whether the two models can capture parent-side competition.

The model fit of the multinomial logit model with both sets of parents (Model A1) is shown in columns (2), (5), and (8) in Table 8. Overall, the performance of Model A1 is poor. Column (2) shows that the model predicts a positive relationship between children's education and co-residence likelihood, whereas the data indicate a negative relationship. In addition, the model predicts a stronger negative relationship between parents' education and co-residence likelihood than the data indicate. Moreover, the model is not able to predict that the number of siblings has a negative effect on the probability of living with one's own parents. Column (5) shows that the model also fails to predict that the education level of siblings increases the likelihood of co-residence. The failure to match the coefficients of the number of siblings and average education of siblings is a result of the lack of siblings' decision in Model A1. However, column (8) shows that the model can predict that the probability of co-residence increases with the spouse's parental education and the spouse's number of siblings because it directly models the decisions of the husband's and wife's parents. In sum, the multinomial logit model with both sets of parents captures parent-side competition but is unable to predict the patterns of child-side competition.

Columns (3), (6), and (9) in Table 8 show the model fit of the multinomial logit model with multiple children (Model A2). Column (3) shows that the model fits most of the coefficients in the baseline regression. When we include the average education of siblings in the regression, as shown in column (6), the model also fits its coefficient because it directly models the decision of siblings. However, when we add the characteristics of the spouse's parents and the spouse's siblings, the model does not fit, as shown in column (9). The coefficients of the spouse's parental education and the spouse's number of siblings are much smaller than the data indicate because the model does not include the decision of the spouse's parents. These results suggest that the simplified multinomial logit model with multiple children can capture child-side competition but not parent-side competition.

Lastly, we check whether the multinomial logit models generate unrealistic living arrangements. In the multinomial logit model with both sets of parents (A1), we find that 7.02% of parents live with more than one married child. However, only 1.30% of families in the CFPS have two married couples living with the same set of parents. In the multinomial logit model with multiple children (A2), we find that 2.88% of adult children live with more than one set of parents, whereas only 0.03% of families in the CFPS have one young couple living with both sets of parents. Using a one-to-one matching model allows us to restrict one set of parents to live with at most one child and one child couple to live with at most one set of parents.

This exercise shows that the matching model with competition between siblings and competition between parents and parents-in-law dominates the two alternative multinomial logit models that ignore one side of the competition.

# 7 Counterfactual Analysis

We conduct two counterfactual experiments based on our model estimates. Our first experiment focuses on housing costs because the surging price of housing in China has been at the centre 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 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%. In this counterfactual analysis, we assume that the housing price does not affect marriage matching in a city. The upper panel of Table 9 reports the co-residence rate by subgroups when we decrease or increase housing price by 20%, and the first panel of Table 10 reports the corresponding percentage changes. At the baseline level of housing price, the average probability of co-residence for adult children is 20.0%. When the housing price increases by 20%, the co-residence likelihood increases by 3.5% or 0.6 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 (4.9% vs. 3.1%) when housing prices increase. First-born children are more likely to respond to a housing price increase than non-first-born children (3.7% vs. 3.2%). In addition, 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 (3.7% vs. 3.3%). From the parent's perspective, a 20% increase in the housing price leads to a 3.5% (1.6 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 (4.3% vs. 3.2%). Figure 5 shows the counterfactual results for different levels of housing prices.

In Table 10, we compare the counterfactual results on changes in the housing price between the matching model and the two multinomial models. In each panel, we present the model's predicted co-residence rate under the current housing price (referred to as the "baseline") in the first row. In the next two rows, we present the percentage changes in the co-residence rate for an increase and a decrease in the housing price of 20% relative to the baseline. We find that, in most cases, both multinomial logit models underpredict adult children's reaction in the co-residence rate compared to the matching model. Such bias is particularly severe for female, first-born, and highly educated children. For example, the matching model predicts that a 20% increase in the housing price leads to a 4.9% increase in the co-residence rate for female children, but the multinomial logit model with both sets of parents (Model A1) predicts a 1.7% increase and the multinomial logit model with multiple children (Model A2) predicts a 1.6% increase. For highly educated children, the matching model predicts that a 20% increase in the housing price leads to a 3.7% increase in their co-residence rate, but Model A1 predicts a 1.8% increase and Model A2 predicts a 3.1% increase. In addition, both multinomial logit models underestimate the responses of both less educated and more educated parents when the housing price increases compared to the matching model. The matching model predicts that a 20% increase in the housing price increases the co-residence rates of less educated and more educated parents by 3.2% and 4.3%, respectively, while Model A1 predicts 1.4% and 1.6% co-residence rate increases and Model A2 predicts 2.7% and 3.4% co-residence rate increases.

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 sets of parents. In this counterfactual experiment, we keep only one married child from each family in our current sample and re-simulate the marriage and co-residence decisions for these adult children. We need to re-simulate the marriage matching because doing so is necessary to predict the social network and co-residence matching.

To conduct this counterfactual analysis, we consider two ways of sampling adult children: keep the oldest child and randomly sample a child in the family. Each approach has its own advantage. The only child is likely to have a larger parent–child age gap than the oldest child and a similar age gap as the random child. Meanwhile, the only child may have higher education attainment than the random child and similar or slightly higher education attainment than the oldest child due to the quantity–quality tradeoff and birth order effect.[36] Moreover, under the one-child policy, every child is the oldest/youngest child and the birth order effect may not exist. Therefore, we consider four possible cases in total: 1) keep the oldest child in the family and keep the birth order effect, 2) keep the oldest child in the family and remove the birth order effect, 3) randomly keep a child in the family and keep the birth order effect, and 4) randomly keep a child in the family and remove the birth order effect. The results provide us with an upper and a lower bound of the effect of one-child policy.

The bottom panel of Table 9 presents the results of these counterfactual experiments. We predict an increase in the co-residence likelihood of adult children, from 20.0% to a range of 21.1%–26.2%. We find that randomly keeping a child generates a larger response than keeping the oldest child, mainly because the former has a larger parent–child gap. Not surprisingly, keeping the birth order effect also leads to a higher co-residence rate. Now we focus on the most conservative prediction where we keep the oldest child and remove the birth order effect. Among different subgroups, a 7.4% (2.4 ppt) increase occurs in the probability of male children living with their own parents, and a 14.5% (1.0 ppt) increase occurs in the probability of female children living with their own parents. We also find that less educated children and children with a larger 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 a range of 21.1%–26.2% when each family has one child instead of 2.2 children on average. In addition, we find that less educated parents experience a larger decline in the co-residence likelihood compared to highly educated parents in all four cases. 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 eldercare in China.[37]

## 8 Conclusions

We develop a TU model to analyse 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 the co-residence context. The model captures several reasons for co-residence: social norms that are reflected in congestion costs 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

[36] Guo *et al.* (2021) use the CFPS data and show that children with higher birth order have worse education outcomes.

[37] We do not show the counterfactual results of the one-child policy of the two multinomial models because the matching model is degenerate to the multinomial logit model with both sets of parents (Model A1) when all adult children have no siblings. In other words, the predictions of the matching model and Model A1 would be the same when they have the same parameters.

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 that are not directly observed in the data. We estimate our co-residence model using the CFPS data and conduct counterfactual experiments to analyse 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, birth order, education, parental education, the age gap between them and their parents, and local housing prices. 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 parental education, and the spouse's number of siblings affect the co-residence likelihood of an adult child. We find that our matching model dominates the alternative multinomial logit models that exclude child-side or parent-side competition in terms of fitting these patterns. 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 adult children who are female, first-born, or highly educated and parents who 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, whereas 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. A few complications 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 consider marriage matching and co-residence matching as sequential decisions. We allow the marriage matching model to account for the expected utility from co-residence 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, people may make marriage and co-residence decisions simultaneously and if we want to explicitly model such simultaneous decisions, 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. As this analysis is beyond the scope of this study, we leave it for future research.

## Affiliations

[1]Department of Economics, the Chinese University of Hong Kong

[2]Department of Economics, University of Essex

[3]School of Economics, Zhejiang University

## References

Barczyk, D. and Kredler, M. (2018). 'Evaluating long-term-care policy options, taking the family seriously', *The Review of Economic Studies*, vol. 85(2), pp. 766–809.

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

Byrne, D., Goeree, M.S., Hiedemann, B. and Stern, S. (2009). 'Formal home health care, informal care, and family decision making', *International Economic Review*, vol. 50(4), pp. 1205–1242.

Card, D. and Lemieux, T. (1997). *Adapting to Circumstances: The Evolution of Work, School, and Living Arrangements Among North American Youth*, University of Chicago Press.

Chiappori, P.A. (1992). 'Collective labor supply and welfare', *Journal of Political Economy*, vol. 100(3), pp. 437–467.

Chiappori, P.A., Dias, M.C. and Meghir, C. (2018). 'The marriage market, labor supply, and education choice', *Journal of Political Economy*, vol. 126(1), pp. S26–S72.

Chiappori, P.A. and Gugl, E. (2015). 'Transferable utility and demand functions: A complete characterization', Working paper.

Chiappori, P.A. and Mazzocco, M. (2017). 'Static and intertemporal household decisions', *Journal of Economic Literature*, vol. 55(3), pp. 985–1045.

Chiappori, P.A., Oreffice, S. and Quintana-Domeque, C. (2012). 'Fatter attraction: Anthropometric and socioeconomic matching on the marriage market', *Journal of Political Economy*, vol. 120(4), pp. 659–695.

Chiappori, P.A., Salanié, B. and Weiss, Y. (2017). 'Partner choice, investment in children, and the marital college premium', *American Economic Review*, vol. 107(8), pp. 2109–67.

Choo, E. and Siow, A. (2006). 'Who marries whom and why', *Journal of Political Economy*, vol. 114(1), pp. 175–201.

Chu, C.C., Xie, Y. and Yu, R.R. (2011). 'Coresidence with elderly parents: A comparative study of southeast China and Taiwan', *Journal of Marriage and Family*, vol. 73(1), pp. 120–135.

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

Dostie, B. and Léger, P.T. (2005). 'The living arrangement dynamics of sick, elderly individuals', *Journal of Human Resources*, vol. 40(4), pp. 989–1014.

Dupuy, A. (2018). 'Migration in China: To work or to wed?', Working paper.

Engers, M. and Stern, S. (2002). 'Long-term care and family bargaining', *International Economic Review*, vol. 43(1), pp. 73–114.

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

Fox, J.T. (2010). 'Identification in matching games', *Quantitative Economics*, vol. 1(2), pp. 203–254.

Fox, J.T. (2018). 'Estimating matching games with transfers', *Quantitative Economics*, vol. 9(1), pp. 1–38.

Galichon, A., Kominers, S.D. and Weber, S. (2019). 'Costly concessions: An empirical framework for matching with imperfectly transferable utility', *Journal of Political Economy*, vol. 127(6), pp. 2875–2925.

Galichon, A. and Salanié, B. (2021). 'Cupid's invisible hand: Social surplus and identification in matching models', *Review of Economic Studies*.

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

Gemici, A. (2007). 'Family migration and labor market outcomes', Working paper.

Giuliano, P. (2007). 'Living arrangements in western europe: Does cultural origin matter?', *Journal of the European Economic Association*, vol. 5(5), pp. 927–952.

Guo, N., Zhang, J. and Zhang, S. (2021). 'Birth order and human capital: Evidence from china', Working paper.

Haurin, D.R., Hendershott, P.H. and Kim, D. (1993). 'The impact of real rents and wages on household formation', *Review of Economics and Statistics*, vol. 75(2), pp. 284–293.

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

Hoerger, T.J., Picone, G.A. and Sloan, F.A. (1996). 'Public subsidies, private provision of care and living arrangements of the elderly', *Review of Economics and Statistics*, vol. 78(3), pp. 428–440.

Hu, W.Y. (2001). 'Welfare and family stability: Do benefits affect when children leave the nest?', *Journal of Human Resources*, vol. 36(2), pp. 274–303.

Institute of Social Science Survey, P.U. (2015). 'China family panel studies', *https://doi.org/10.18170/DVN/45LCSO, Peking University Open Research Data Platform, V42.*

Johar, M. and Maruyama, S. (2014). 'Does coresidence improve an elderly parent's health?', *Journal of Applied Econometrics*, vol. 29(6), pp. 965–983.

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

Konrad, K.A., Künemund, H., Lommerud, K.E. and Robledo, J.R. (2002). 'Geography of the family', *American Economic Review*, vol. 92(4), pp. 981–998.

Ma, S. and Wen, F. (2016). 'Who coresides with parents? An analysis based on sibling comparative advantage', *Demography*, vol. 53(3), pp. 623–647.

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

Maruyama, S. and Johar, M. (2017). 'Do siblings free-ride in 'being there' for parents?', *Quantitative Economics*, vol. 8(1), pp. 277–316.

Maurer-Fazio, M., Connelly, R., Chen, L. and Tang, L. (2011). 'Childcare, eldercare, and labor force participation of married women in urban China, 1982–2000', *Journal of Human Resources*, vol. 46(2), pp. 261–294.

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

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

Pezzin, L.E. and Schone, B.S. (1999). 'Intergenerational household formation, female labor supply and informal caregiving: A bargaining approach', *Journal of Human Resources*, vol. 34(3), pp. 475–503.

Rosenzweig, M. and Zhang, J. (2014). 'Co-residence, life-cycle savings and inter-generational support in urban China', Working paper.

Rosenzweig, M.R. and Wolpin, K.I. (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*, vol. 11(1), pp. 84–112.

Rosenzweig, M.R. and Wolpin, K.I. (1994). 'Parental and public transfers to young women and their children', *American Economic Review*, vol. 84(5), pp. 1195–1212.

Sakudo, M. (2007). 'Strategic interactions between parents and daughters: Co-residence, marriage and intergenerational transfers in Japan', Working paper.

Shapley, L.S. and Shubik, M. (1971). 'The assignment game I: The core', *International Journal of Game Theory*, vol. 1(1), pp. 111–130.

Siow, A. (2015). 'Testing Becker's theory of positive assortative matching', *Journal of Labor Economics*, vol. 33(2), pp. 409–441.

Stern, S. (2014). 'O brother, where art thou? we need your help', Working paper.

Zorlu, A. and Mulder, C.H. (2011). 'Ethnic differences in leaving home: Timing and pathways', *Demography*, vol. 48(1), pp. 49–72.

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

| | Baseline | Testing child-side competition | Testing parent-side competition |
|---|---|---|---|
| | (1) | (2) | (3) |
| Male | 0.2431*** | 0.2581*** | 0.6120*** |
| | (0.0154) | (0.0133) | (0.0385) |
| First born | 0.0363* | 0.0250 | 0.0266 |
| | (0.0191) | (0.0180) | (0.0385) |
| Years of education | -0.0042* | -0.0119*** | -0.0013 |
| | (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 | 0.0051** | 0.0059*** | -0.0018 |
| | (0.0021) | (0.0018) | (0.0040) |
| Number of siblings | -0.0434*** | -0.0365*** | -0.0735*** |
| | (0.0067) | (0.0061) | (0.0176) |
| Logged housing price | 0.0463** | 0.0408** | 0.0116 |
| | (0.0193) | (0.0165) | (0.0402) |
| Average education of siblings | | 0.0073* | |
| | | (0.0038) | |
| Spouse' parental education | | | 0.0198*** |
| | | | (0.0059) |
| Spouse' number of siblings | | | 0.0867*** |
| | | | (0.0133) |
| Constant | -0.4131* | -0.3995** | 0.0983 |
| | (0.2114) | (0.1830) | (0.4605) |
| Observations | 3,526 | 3,328 | 437 |
| R-squared | 0.1270 | 0.1375 | 0.5205 |

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 the 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. In column (1), we run baseline regressions, which only include the characteristics of parents and adult children. We add siblings' attributes in column (2) to test the child-side competition and add the attributes of spouse's parents in column (3) to test the parent-side competition.

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 |
| First-born child | 35.4% | 45.5% | 0.5 | 0 | 1 |
| 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 |
| Having a child under age six | 30.0% | 30.0% | 0.5 | 0 | 1 |
| City Level, N=104 | | | | | |
| Raw housing price (RMB per square meter) | 3,989 | 3,989 | 2,528 | 1,454 | 17,315 |
| Adjusted housing price (via GDP per capita) | 10.5 | 10.5 | 0.5 | 9.5 | 11.8 |
| 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 |

This table presents the descriptive statistics of key variables in our analysis. The first panel presents the summary statistics of adult children in the child sample; the second panel presents the statistics of parents in the family sample; the third panel presents the statistics of child couples in the couple sample; and the last panel presents the statistics of city-level characteristics. We present the raw and weighted average of key variables, where the weighted average uses the inverse of the number of adult children in a family as the weight to adjust the oversampling bias for large-size families. However, the statistics of child couples sample is not affected by this weight.

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 |
| | (0.024) | (0.091) | (0.020) | (0.026) | (0.012) |
| Age | 0.073 | 1.174 | 0.115 | 0.037 | 0.315 |
| | (0.010) | (0.012) | (0.012) | (0.048) | (0.010) |
| Number of siblings | -0.048 | 0.176 | 0.124 | -0.061 | 0.016 |
| | (0.017) | (0.017) | (0.025) | (0.003) | (0.008) |
| Parental education | 0.182 | -0.103 | -0.052 | 0.291 | 0.087 |
| | (0.048) | (0.055) | (0.011) | (0.022) | (0.015) |
| Parental age | 0.009 | 0.370 | 0.021 | 0.084 | 0.106 |
| | (0.025) | (0.038) | (0.007) | (0.032) | (0.006) |

This table presents the estimation results of the marriage matching surplus in Equation (13). 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 | | Child couple's income | | Parents in poor health | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Age | -0.059*** | -0.060 | -0.007** | 0.080*** | 0.003* | 0.019 |
| | (0.013) | (0.221) | (0.003) | (0.030) | (0.002) | (0.046) |
| Age square | | 0.000 | | -0.001*** | | -0.000 |
| | | (0.002) | | (0.000) | | (0.000) |
| Years of education | 0.051** | 0.051** | 0.057*** | 0.057*** | -0.007*** | -0.007*** |
| | (0.021) | (0.021) | (0.006) | (0.006) | (0.003) | (0.003) |
| Constant | 3.296*** | 3.311 | -0.314** | -1.946*** | 0.086 | -0.422 |
| | (0.876) | (6.973) | (0.151) | (0.588) | (0.123) | (1.478) |
| City FE | Yes | Yes | Yes | Yes | No | No |
| Observations | 545 | 545 | 1,026 | 1,026 | 1,438 | 1,438 |
| R-squared | 0.097 | 0.097 | 0.097 | 0.105 | 0.008 | 0.008 |

The dependent variables are the logarithm value of parent's total income in columns (1) and (2), logarithm value of child couple's total income in columns (3) and (4), and an indicator of whether parents are in poor health in columns (5) and (6). We use ordinary least squares to estimate columns (1) to (4) and linear probability models to estimate columns (5) to (6). Columns (1) to (4) 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) and (4), the sample is restricted to urban local residents whose parents are aged 55 to 75.

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

Table 5. *Out-of-model estimation (cont'd) – effects on fertility*

| | (1) OLS | (2) 2SLS | (3) First-stage |
|---|---|---|---|
| Co-residence | 0.043* | 0.440** | |
| | (0.025) | (0.199) | |
| Age | -0.137*** | -0.106*** | -0.066*** |
| | (0.015) | (0.024) | (0.017) |
| Age square | 0.001*** | 0.001*** | 0.001*** |
| | (0.000) | (0.000) | (0.000) |
| Years of education | 0.000 | 0.002 | -0.011** |
| | (0.003) | (0.004) | (0.004) |
| Number of siblings | | | -0.052*** |
| | | | (0.013) |
| Constant | 3.220*** | 2.289*** | 2.193*** |
| | (0.307) | (0.598) | (0.335) |
| City FE | Yes | Yes | Yes |
| Observations | 1064 | 1056 | 1056 |
| R-square | 0.325 | 0.140 | 0.112 |
| Kleibergen-Paap F statistic | | 17.249 | |
| Underidentification LM-Test | | 0.000 | |

The dependent variables are an indicator of whether the child couple has a grandchild aged 0–6 in columns (1) and (2) and an indicator of whether the child couple lives with any side of the parents in column (3). Column (1) reports the result of a linear probability model, column (2) reports the second stage result of the IV model where we instrument co-residence with the number of siblings, and column (3) reports the first stage of the IV regression. All columns use the CFPS data, and the sample is restricted to urban local child couples whose parents are aged 55 to 75.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6. *Estimation results of the co-residence matching model and two alternative multinomial logit models*

| Parameter | Explanation | Matching model | Mlogit model A1 | Mlogit model A2 |
|---|---|---|---|---|
| $\frac{0.71}{\kappa_0+\kappa_1+\kappa_2} - 0.5$ | Congestion costs for first-born male children | -0.078 $e{-6}$ | 1.540 $e{-6}$ | 0.748 $e{-6}$ |
| | | (0.012 $e{-6}$) | (0.356 $e{-6}$) | (0.180 $e{-6}$) |
| $\frac{0.71}{\kappa_0+\kappa_1} - 0.5$ | Congestion costs for non-first-born male children | -0.371 $e{-6}$ | -0.008 $e{-6}$ | 0.271 $e{-6}$ |
| | | (0.124 $e{-6}$) | (0.002 $e{-6}$) | (0.075 $e{-6}$) |
| $\frac{0.71}{\kappa_0+\kappa_2} - 0.5$ | Congestion costs for first-born female children | -2.239 $e{-6}$ | -0.412 $e{-6}$ | -2.380 $e{-6}$ |
| | | (0.665 $e{-6}$) | (0.123 $e{-6}$) | (0.084 $e{-6}$) |
| $\frac{0.71}{\kappa_0} - 0.5$ | Congestion costs for non-first-born female children | -2.519 $e{-6}$ | -1.960 $e{-6}$ | -2.857 $e{-6}$ |
| | | (0.724 $e{-6}$) | (0.429 $e{-6}$) | (0.533 $e{-6}$) |
| $\delta_0$ | Constant in living costs | -0.169 | -0.789 | -0.370 |
| | | (0.045) | (0.219) | (0.110) |
| $\delta_1$ | Effect of log housing price on housing costs | 0.034 | 0.019 | 0.036 |
| | | (0.007) | (0.005) | (0.006) |
| $Z_{ec}$ | Eldercare costs | 0.548 | 2.272 | 0.819 |
| | | (0.084) | (0.468) | (0.251) |
| $Z_{cc}$ | Childcare costs | 0.133 | 0.178 | 0.277 |
| | | (0.046) | (0.032) | (0.055) |

We present the estimation result of the co-residence matching model, a multinomial logit model with a child couple and both side of parents (A1), and another multinomial logit model with a parent couple and all of their adult children (A2). Standard errors are in parentheses.

Table 7. *Model fit: A child's co-residence probability with his/her own parents*

| | Baseline | | Testing child-side competition | | Testing parent-side competition | |
|---|---|---|---|---|---|---|
| | Data | Model | Data | Model | Data | Model |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Male | 0.2431*** | 0.2419*** | 0.2581*** | 0.2297*** | 0.6120*** | 0.6103*** |
| | (0.0154) | (0.0016) | (0.0133) | (0.0017) | (0.0385) | (0.0027) |
| First-born | 0.0363* | 0.0366*** | 0.0250 | 0.0341*** | 0.0266 | 0.0328*** |
| | (0.0191) | (0.0019) | (0.0180) | (0.0021) | (0.0385) | (0.0032) |
| Years of education | -0.0042* | -0.0023*** | -0.0113*** | -0.0180*** | -0.0013 | -0.0008* |
| | (0.0025) | (0.0003) | (0.0031) | (0.0003) | (0.0058) | (0.0004) |
| Parent's education | -0.0044* | -0.0053*** | -0.0045* | -0.0046*** | -0.0184*** | -0.0101*** |
| | (0.0026) | (0.0003) | (0.0023) | (0.0003) | (0.0064) | (0.0005) |
| Parent–child age gap | 0.0051** | 0.0046*** | 0.0059*** | 0.0045*** | -0.0018 | -0.0040*** |
| | (0.0021) | (0.0002) | (0.0018) | (0.0002) | (0.0040) | (0.0003) |
| Number of siblings | -0.0434*** | -0.0396*** | -0.0365*** | -0.0383*** | -0.0735*** | -0.0674*** |
| | (0.0067) | (0.0007) | -0.0061 | (0.0008) | (0.0176) | (0.0012) |
| Logged housing price | 0.0463** | 0.0453*** | 0.0408** | 0.0405*** | 0.0116 | 0.0001 |
| | (0.0193) | (0.0018) | (0.0165) | (0.0020) | (0.0402) | (0.0031) |
| Average education of siblings | | | 0.0073* | 0.0070*** | | |
| | | | (0.0038) | (0.0004) | | |
| Spouse's parental education | | | | | 0.0198*** | 0.0172*** |
| | | | | | (0.0059) | (0.0005) |
| Spouse's number of siblings | | | | | 0.0867*** | 0.0615*** |
| | | | | | (0.0133) | (0.0011) |
| Constant | -0.4131* | -0.4505*** | -0.3995** | -0.3991*** | 0.0983 | 0.0684** |
| | (0.2114) | (0.0204) | (0.1830) | (0.0219) | (0.4605) | (0.0345) |
| Observations | 3526 | 23,663,900 | 3328 | 20,015,400 | 437 | 9,016,000 |
| R-squared | 0.1270 | 0.1044 | 0.1375 | 0.0951 | 0.5205 | 0.3890 |
| Root MSE | 0.3764 | 0.3785 | | | | |

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.

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

Table 8. *Model fit: Multinomial logit models*

| | Baseline | | | Testing child-side competition | | | Testing parent-side competition | | |
|---|---|---|---|---|---|---|---|---|---|
| | Data | Model A1 | Model A2 | Data | Model A1 | Model A2 | Data | Model A1 | Model A2 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Male | 0.2431*** | 0.1630*** | 0.2201*** | 0.2581*** | 0.1528*** | 0.2057*** | 0.6120*** | 0.3972*** | 0.5765*** |
| | (0.0154) | (0.0016) | (0.0016) | (0.0133) | (0.0017) | (0.0017) | (0.0385) | (0.0030) | (0.0028) |
| First-born | 0.0363* | 0.1252*** | 0.0492*** | 0.0250 | 0.1247*** | 0.0428*** | 0.0266 | 0.1474*** | 0.0457*** |
| | (0.0191) | (0.0020) | (0.0020) | (0.0180) | (0.0021) | (0.0021) | (0.0385) | (0.0037) | (0.0034) |
| Years of education | -0.0042* | 0.0009*** | -0.0008*** | -0.0113*** | 0.0008*** | -0.0005* | -0.0013 | -0.0018*** | -0.0027*** |
| | (0.0025) | (0.0003) | (0.0003) | (0.0031) | (0.0003) | (0.0003) | (0.0058) | (0.0005) | (0.0005) |
| Parent's education | -0.0044* | -0.0171*** | -0.0061*** | -0.0045* | -0.0165*** | -0.0050*** | -0.0184*** | -0.0224*** | -0.0082*** |
| | (0.0026) | (0.0003) | (0.0003) | (0.0023) | (0.0003) | (0.0003) | (0.0064) | (0.0005) | (0.0005) |
| Parent–child age gap | 0.0051** | 0.0051*** | 0.0083*** | 0.0059*** | 0.0051*** | 0.0076*** | -0.0018 | 0.0060*** | 0.0081*** |
| | (0.0021) | (0.0002) | (0.0002) | (0.0018) | (0.0002) | (0.0002) | (0.0040) | (0.0004) | (0.0004) |
| Number of siblings | -0.0434*** | -0.0004 | -0.0223*** | -0.0365*** | -0.0007 | -0.0205*** | -0.0735*** | 0.0022* | -0.0218*** |
| | (0.0067) | (0.0007) | (0.0007) | (0.0061) | (0.0008) | (0.0008) | (0.0176) | (0.0013) | (0.0013) |
| Logged housing price | 0.0463** | 0.0144*** | 0.0344*** | 0.0408** | 0.0157*** | 0.0302*** | 0.0116 | 0.0011 | 0.0125*** |
| | (0.0193) | (0.0019) | (0.0019) | (0.0165) | (0.0020) | (0.0020) | (0.0402) | (0.0035) | (0.0032) |
| Average education of siblings | | | | 0.0073* | -0.0004 | 0.0080*** | | | |
| | | | | (0.0038) | (0.0004) | (0.0004) | | | |
| Spouse's parental education | | | | | | | 0.0198*** | 0.0226*** | 0.0064*** |
| | | | | | | | (0.0059) | (0.0005) | (0.0005) |
| Spouse's number of siblings | | | | | | | 0.0867*** | 0.0570*** | 0.0212*** |
| | | | | | | | (0.0133) | (0.0012) | (0.0011) |
| Constant | -0.4131* | -0.1399*** | -0.4301*** | -0.3995** | -0.1450*** | -0.3641*** | 0.0983 | 0.0352 | -0.0882 |
| | (0.2114) | (0.0207) | (0.0208) | (0.1830) | (0.0222) | (0.0223) | (0.4605) | (0.0390) | (0.0356) |
| Observations | 3526 | 23,663,900 | 23,663,900 | 3328 | 20,015,400 | 20,015,400 | 437 | 8,669,800 | 8,669,800 |
| R-squared | 0.1270 | 0.0792 | 0.0919 | 0.1375 | 0.0733 | 0.0800 | 0.5205 | 0.2071 | 0.3557 |

Data columns use the CFPS data to run OLS regressions, the same as Table 1. Model A1 columns use a multinomial logit model with one child couple and both sets of parents to predict the living arrangement. Model A2 columns use a multinomial logit model with one parent and multiple children to predict the living arrangement. Both models use simulated co-residence decisions in the bootstrap sample to run the same regressions.

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

Table 9. *Counterfactual: Effects of the housing price and fertility policy on co-residence probability (%)*

| | Children's probability of co-residing with own parents | | | | | | | | | Parent's co-residence probability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Child total | Child's gender | | Child's birth order | | Child's education | | Parent–child age gap | | Parent total | Parental education | |
| | | male | female | first-born | non-first-born | $<12$ | $\geq 12$ | $<27$ | $\geq 27$ | | $<9$ | $\geq 9$ |
| 0. Baseline | 0.200 | 0.318 | 0.072 | 0.221 | 0.182 | 0.201 | 0.198 | 0.198 | 0.201 | 0.439 | 0.476 | 0.343 |
| 1. Changes in housing price | | | | | | | | | | | | |
| -20% | 0.192 | 0.307 | 0.069 | 0.211 | 0.176 | 0.193 | 0.190 | 0.191 | 0.193 | 0.422 | 0.458 | 0.329 |
| 20% | 0.206 | 0.328 | 0.076 | 0.229 | 0.188 | 0.207 | 0.206 | 0.205 | 0.208 | 0.455 | 0.491 | 0.358 |
| 2. Each family has only one child | | | | | | | | | | | | |
| Oldest child, with birth order effect | 0.242 | 0.393 | 0.093 | n/a | n/a | 0.246 | 0.236 | 0.239 | 0.251 | 0.242 | 0.255 | 0.199 |
| Oldest child, without birth order effect | 0.211 | 0.342 | 0.083 | n/a | n/a | 0.219 | 0.200 | 0.210 | 0.215 | 0.211 | 0.226 | 0.165 |
| Any child, with birth order effect | 0.262 | 0.413 | 0.094 | n/a | n/a | 0.273 | 0.250 | 0.258 | 0.265 | 0.262 | 0.278 | 0.219 |
| Any child, without birth order effect | 0.229 | 0.358 | 0.083 | n/a | n/a | 0.243 | 0.212 | 0.225 | 0.232 | 0.229 | 0.246 | 0.183 |

The first row presents the predicted co-residence rates for adult children and parents in the baseline model. The next two rows present the counterfactual results when we reduce or increase the housing price by 20%. The last four rows present the counterfactual results on the one-child policy, where we consider four possible cases: 1) keep the oldest child and allow for the birth order effect, 2) keep the oldest child and ignore the birth order effect, 3) randomly draw a child and allow for the birth order effect, and 4) randomly draw a child and ignore the birth order effect. The first nine columns present adult children's co-residence probability of living with the child's own parents (not the spouse's). The last three columns present parents' co-residence probability of living with any of their own adult children.

Table 10. *Counterfactual: Effects of the housing price on the co-residence rate predicted by the three models*

| | Children's probability of co-residing with own parents | | | | | | | | | Parent's co-residence probability | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Child | Child's gender | | Child's birth order | | Child's education | | Parent–child age gap | | Parent | Parental education | |
| | total | male | female | first-born | non-first-born | <12 | ≥12 | <27 | ≥27 | total | <9 | ≥9 |
| 1. Matching model | | | | | | | | | | | | |
| Baseline (%) | 20.0 | 31.8 | 7.2 | 22.1 | 18.2 | 20.1 | 19.8 | 19.8 | 20.1 | 43.9 | 47.6 | 34.3 |
| 20% decline (%) | -3.98 | -3.61 | -5.08 | -4.42 | -3.54 | -3.81 | -4.22 | -3.80 | -4.13 | -3.98 | -3.88 | -4.34 |
| 20% increase (%) | 3.45 | 3.09 | 4.85 | 3.71 | 3.18 | 3.26 | 3.72 | 3.45 | 3.45 | 3.45 | 3.20 | 4.33 |
| 2. Multinomial logit model with both sets of parents (A1) | | | | | | | | | | | | |
| Baseline (%) | 20.3 | 28.5 | 11.3 | 26.1 | 15.5 | 21.0 | 19.3 | 21.9 | 19.0 | 38.2 | 43.2 | 25.2 |
| 20% decline (%) | -2.19 | -2.22 | -2.17 | -2.05 | -2.39 | -2.14 | -2.27 | -2.10 | -2.28 | -1.88 | -1.72 | -2.60 |
| 20% increase (%) | 1.79 | 1.82 | 1.69 | 1.71 | 1.91 | 1.80 | 1.78 | 1.78 | 1.80 | 1.48 | 1.44 | 1.64 |
| 3. Multinomial logit model with multiple children (A2) | | | | | | | | | | | | |
| Baseline (%) | 20.8 | 31.8 | 6.2 | 23.0 | 18.9 | 20.3 | 21.4 | 19.7 | 21.7 | 45.8 | 49.5 | 36.1 |
| 20% decline (%) | -3.57 | -3.49 | -2.40 | -3.74 | -3.41 | -3.53 | -3.63 | -3.51 | -3.62 | -3.58 | -3.46 | -4.00 |
| 20% increase (%) | 2.85 | 2.72 | 1.54 | 3.24 | 2.45 | 2.69 | 3.06 | 3.13 | 2.63 | 2.85 | 2.70 | 3.37 |

We compare the counterfactual results on changes in housing price between the matching model and the two multinomial models. In each panel, we present model's predicted co-residence rate under the current housing price (referred to as "baseline") in the first row. In the next two rows, we present the percentage changes in the co-residence rate for an increase and a decline in housing prices by 20% relative to the baseline. The first nine columns present adult children's co-residence probability of living with the child's own parents (not the spouse's). The last three columns present parents' co-residence probability of living with any of their own adult children.

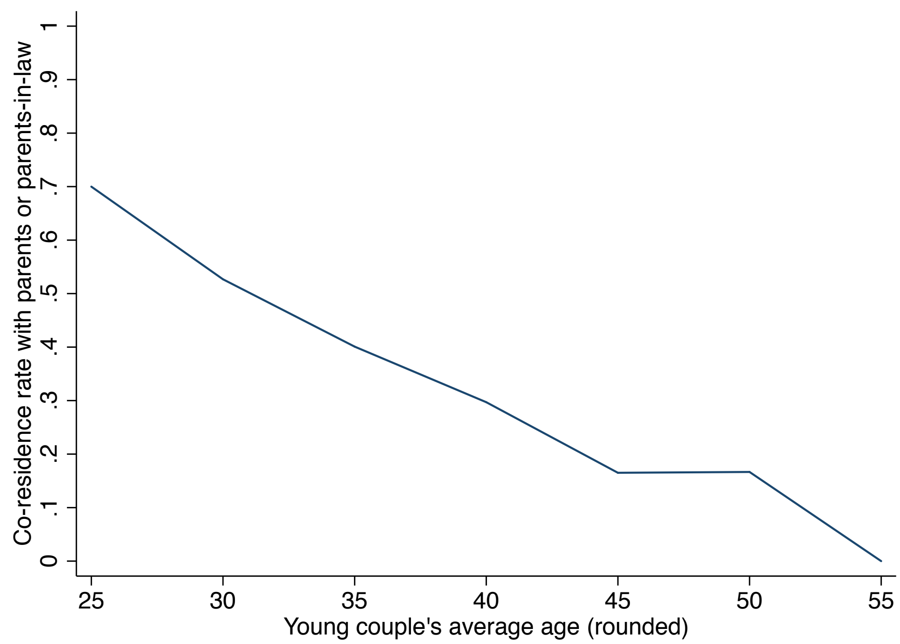


Figure 1. *Probability of co-residence with parents or parents-in-law for married couples by age. Data source: China Family Panel Study 2010*

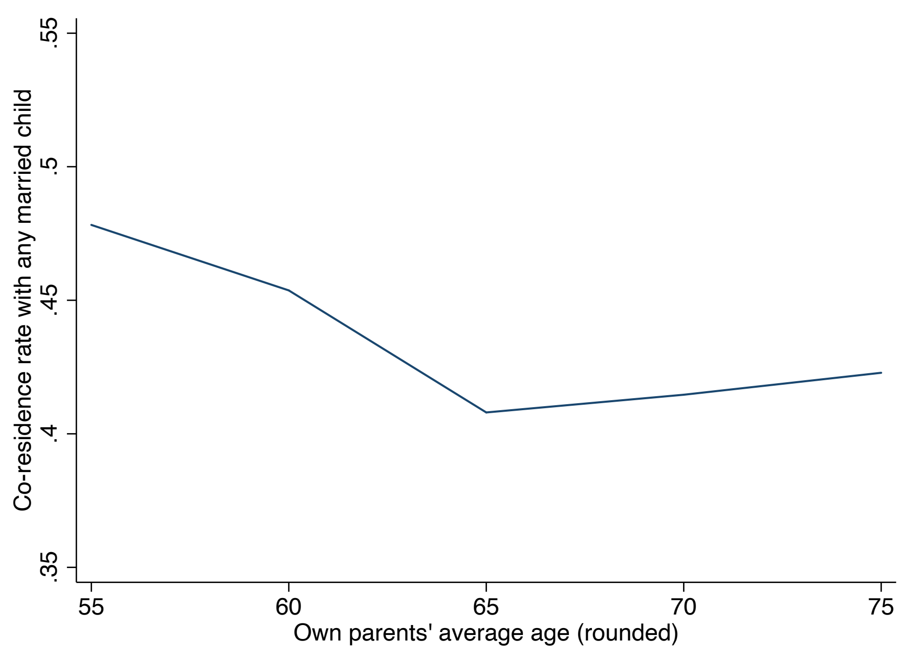


Figure 2. *Probability of co-residence with adult children for parents by age. Data source: China Family Panel Study 2010*

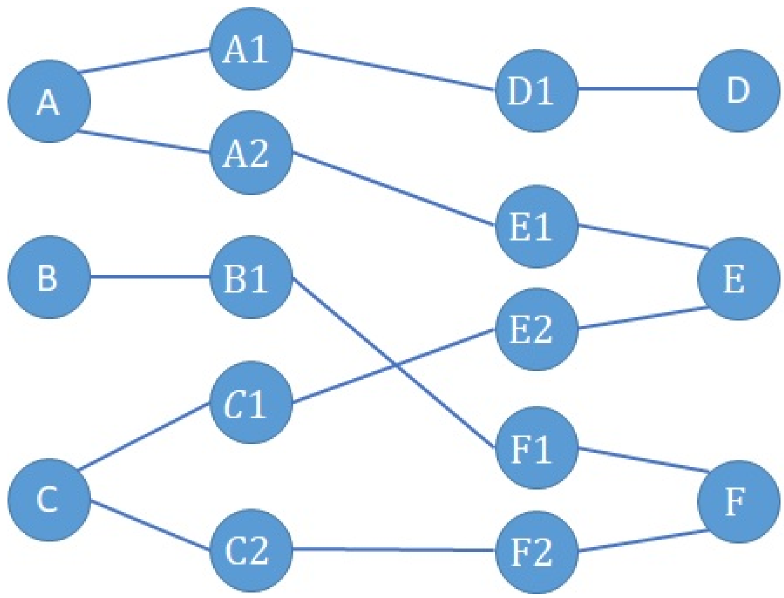


Figure 3. *Family network*

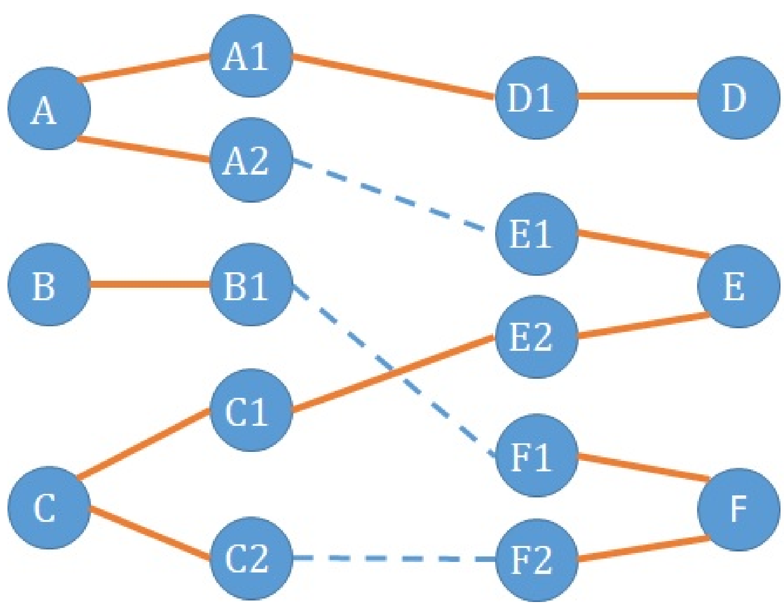


Figure 4. *Observed family network*

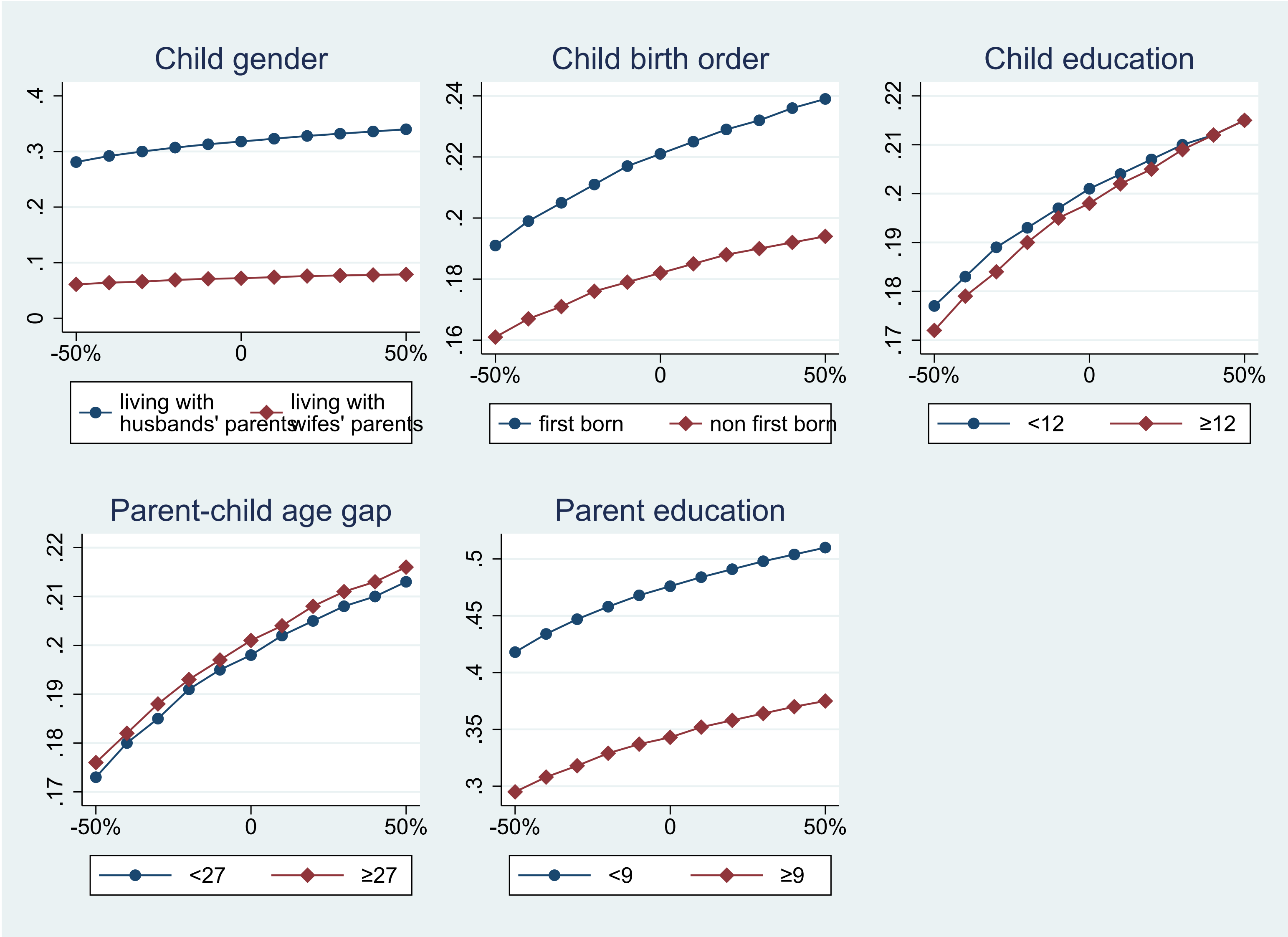


Figure 5. *Effects of changes in housing price on co-residence likelihood*

[a]

[a]The first four figures present the effects of changes in housing price on the co-residence likelihood of adult children by children's gender, birth order, education, and parent–child age gap. The last figure presents the effects of changes in housing price on the co-residence likelihood of parents by parents' education.