A House for a Bride: Marriage and Homeownership in China*

Scarlet Chen
Stanford University

Adam Zhang
Stanford University

May 2023

Link to the Latest Draft

Abstract

This paper studies the phenomenon of marriage house in China and its effects on demographics and homeownership. We first show empirical evidence: single males with a marriage house (a house where the newlywed can move into) have 70% higher odds of getting married compared to their counterparts who do not have a marriage house. In addition, the timing of home purchase exhibits a clear cut-off around the time of marriage, with the probability of purchasing a house peaking 0-2 years before marriage and slumping immediately after the time of marriage. Moreover, in the cross section, county house prices and average age at marriage are highly correlated in both level and in growth rate. We then quantify the marriage related incentives for homeownership using a lifecycle consumption-savings model with housing demand and ownership-dependent marriage shocks. The model captures both the marriage market premium of homeownership for single males and the convenience that a marriage house provides for a newly-wed couple. Without the marriage market premium, young households generally delay their home purchases by about 5 years. Eliminating the convenience value of a marriage house reduces the overall ownership by 35% for families with a male child between age 15 and 45. Furthermore, the marriage house friction accounts for 40% of the rise in marriage age between 1995 and 2010. Our results suggest that policies directed at either housing affordability or demographics can have significant consequences for both marriage and housing markets.

*We are grateful to Monika Piazzesi, Martin Schneider, Chris Tonetti, Adrien Auclert, Liran Einav, Luigi Pistaferri, Pete Klenow, Nick Bloom, and seminar participants at AREUEA National Conference, Stanford GSB, and Stanford Economics for helpful comments and feedback. All errors are our own. Email: sijiac@stanford.edu and adamjcz@stanford.edu.

1 Introduction

Homeownership and marriage are both significant milestones in life. The decision and timing for one event is not independent from the other. This close relationship is especially true in China, where the tradition of marriage house has persisted for generations: males (and their families) are expected to provide a marriage house (*Hunfang* in Chinese) - a house where the newlywed can move into upon their marriage. Not yet fully understood is how this tradition affects China's demographics and homeownership.

This paper studies the interaction between marriage and homeownership and its effects on China's housing and demographic outcomes. We first provide novel empirical evidence supporting the complementarity between the two lifetime events. Specifically, marriage house ownership carries a premium on the marriage market for single males. Moreover, the timing of home purchase exhibits a clear cutoff around the time of marriage, with the vast majority of families buying houses before marriage but not after. In addition, the average age at first marriage is strongly correlated with both the level and the growth rate of house price across China's counties. Then we use a unitary household consumption savings model with housing demand and family structure to quantify the impact of the marriage house phenomenon on homeownership decision and demographics. The model suggests that the marriage-house complementarity is crucial to understand the high and early homeownership among Chinese households. Without the marriage market premium, households would delay their home purchases by around 5 years. Eliminating the convenience value of a marriage house reduces the overall ownership by 35% for families with a male child between age 15 and 45. Removing this homeownership friction from the marriage market would have reduced China's rise in the first-marriage age by 40% from 1995 to 2010.

We start our analysis by showing micro-data evidence supporting the marriage house tradition in China. Using panel survey data, we conduct hazard analysis on the effect of marriage house ownership on single males' odds of getting married. We find that marriage house ownership significantly increases a single male's chances of getting married. Specifically, a single male with a marriage house ready has a 1.7 higher odds compared to their counterparts without one, controlling for family wealth, income, and other covariates. Furthermore, this positive relationship is only observed for single males but not for single females, which corroborates to the anecdotal wisdom that single males and their families are generally expected to provide a marriage house before marriage.

We further substantiate the complementarity between marriage and homeownership by showing a clear cut-off of home purchases at the time of marriage. Specifically, the probability of purchasing a house peaks 0-2 years before marriage and slumps immediately after the time of marriage. This cutoff in the timing of homeownership indicates that endogeneity concerns about our hazard analysis are unlikely. The primary threat to our hazard analysis is reverse causality: the benefits of owning a house are larger when family size is expected to increase because of marriage. Therefore, single males purchase marriage houses when they know marriage is on the horizon. This reverse causality argument can also produce a positive correlation between the timing of marriage and the timing of homeownership. However, this argument also implies more people should buy houses immediately after marriage as well since family size has indeed increased, which is hard to be reconciled with the fact that 90% of families would buy houses before marriage and that only 10% would buy houses afterwards.

Our empirical results from micro-data point to marital home ownership as a friction in the Chinese marriage market. One interpretation is that owning a marriage house helps single males attract partners. The other interpretation is that, conditional on finding a partner, owning a marriage house increases the chances of converting the relationship into a marriage. Both interpretations would result in the observed positive correlation between marriage house ownership and marriage probabilities for single males only and generate the observed cut-off in the timing of homeownership and marriage.[1]

The implication of this marriage house friction is that delays in martial home purchases lead to delays in marriage. In the cross section, counties with higher house prices also have higher average age at marriage. Locations with twice as high house price than average also see 0.5 higher age at marriage. We rule out the likelihood of other covariates driving the observed patterns by residualizing on income, wealth and by including a rich set of fixed effects. Furthermore, the robust positive relationship in levels also manifests in growth rates: counties with 1% higher house price growth rates also saw 1% higher growth rate in marriage age between 1990 and 2012. The cross sectional evidences are consistent with our micro-data findings which point to marriage house ownership as a friction in the marriage market.

We then study the effect of marriage-house complementarity on ownership and demographics using a unitary lifecycle consumption-savings model for single males and their families with housing demand. Exogenous marriage shocks arrive according to marriage house ownership status, and marriage is desirable. Married couples move into the marriage house and enjoy housing consumption from their marital home. Therefore, households in our model have three different incentives for ownership: to save for future consumption, to shorten the waiting time for marriage, and to derive convenience provided by a marriage house. The flip side of this marriage-house complementarity is that rising house prices delay marriage by deferring marriage house ownership, as households take longer to save up for home purchases.

After quantification of the benchmark model using aggregate moments, the model matches

[1] Micro-foundations of the marriage house phenomenon are beyond the scope of this paper. Possible candidates are: marriage house as a commitment device, convenience value a marriage house provides for the bride and her child, etc.

the joint distribution of marital status and homeownership in the cross section, for males between age 15 and 45. The success of the model comes from its ability to produce not only realistic age profile of ownership but also the correct timing of ownership. People in different stages of the lifecycle purchase marital homes for different reasons. Single agents trade off the cost of ownership with marriage market premium, convenience value of marriage house, and investment potential as an asset. Young, married non-owners, on the other hand, care only about the latter two incentives. Older, married agents would choose to be owners only for the convenience value of a marriage house. The success of the model in matching the joint distribution of marriage and ownership is a testimony to the close relationship between marriage and homeownership throughout different stages of the lifecycle.

We then show counterfactual ownership rate and demographics when the marriage-house complementarity is absent. The two marriage-related benefits of homeownership, marriage market premium and convenience value, affect homeownership deferentially. The marriage market premium changes the timing of home purchases, while the convenience value alters the overall ownership. We find that the marriage market premium primarily explains the timing of home purchases and less so about the ownership rate. Specifically, removing the marriage market premium delays home purchases by about 5 years for males between age 15 and 45, and overall ownership rate When the convenience value of a marriage house is absent, the overall homeownership rate drops by 35%. Without the marriage house friction in the marriage market, the rise in marriage age would have been 60% of what China experienced between 1995 and 2010.

Related Literature. We contribute to the existing literature on three fronts. First and foremost, this paper highlights a new aspect in which families and the macroeconomy are interconnected and how family decisions and economic decisions affect each other. In addition, we are the first to structurally model the marriage house phenomenon and study the implications for lifecycle portfolio choice and demographics. Last but not least, we show novel empirical evidence for the complementarity between marriage and homeownership in China.

A large literature has shown the close relationship between economic and family decisions. One strand of this literature has combined family economics into macroeconomics and examined classical macroeconomic questions, such as the national savings rate, economic inequality, employment, etc (Benhabib, Rogerson and Wright, 1991, Greenwood and Hercowitz, 1991, De La Croix and Doepke, 2003, Fernandez, Guner and Knowles, 2005, Guler, Guvenen and Violante, 2012, Jaimovich, Pruitt and Siu, 2013, Jones, Manuelli and McGrattan, 2015, Doepke and Tertilt, 2016, Dyrda, Kaplan and Ríos-Rull, 2019). Particularly related to this paper are works that examine family structure and financial decisions (Cubeddu and Ríos-Rull, 2003, Love, 2010, Fernández and Wong, 2014a,b, Voena, 2015, Guvenen and Rendall, 2015, Piazzesi and Schneider, 2016, Peter, Schneider and Piazzesi, 2020). The other strand of this literature has adapted macroeconomic models and answered questions related family decisions, such as fertility, marriage rate, and so on (Mortensen, 1988, Galor and Weil, 2000, Regalia and Rios-Rull, 2001, Greenwood and Seshadri, 2002, Boldrin and Jones, 2002, Kalemli-Ozcan, 2003, Doepke, 2005, Greenwood, Seshadri and Vandenbroucke, 2005, Greenwood and Guner, 2008, Bar and Leukhina, 2010, Doepke, Hazan and Maoz, 2015, Goussé, Jacquemet and Robin, 2017, Greenwood, Guner and Vandenbroucke, 2017). This paper structurally models homeownership decisions in the presence of marriage-house complementarity and investigates the implications for age at marriage.

We also add to the broad literature that studies portfolio choice with housing in the lifecycle. Fernandez-Villaverde and Krueger (2011) and Yang (2009) emphasize on the collateral value of housing and the hump-shaped age profile of housing consumption Lustig and Van Nieuwerburgh (2005, 2010). Li and Yao (2007) and Díaz and Luengo-Prado (2008) bring tenure decision into the lifecycle and examine tax benefits of ownership. Chambers, Garriga and Schlagenhauf (2009a,b) analyze mortgage choice of the lifecycle in equilibrium. Attanasio, Bottazzi, Low, Nesheim and Wakefield (2012) studies the lifecycle choice of house size when there are transaction costs. Home equity has been argued to function like buffer stock for consumption smoothing Hurst and Stafford (2004), Mian and Sufi (2011), Chen et al. (2020). Particularly related to this paper are previous works studying housing as an asset in the life cycle (Flavin and Yamashita, 2002, Cocco, 2005, Cocco, Gomes and Maenhout, 2005, Yao and Zhang, 2005, Davidoff, 2010, Attanasio, Leicester and Wakefield, 2011, Gorea and Midrigin, 2017, Kaplan, Mitman and Violante, 2017, Berger, Guerrieri, Lorenzoni and Vavra, 2018, Landvoigt, 2017). In this paper, we study ownership decision in the life cycle when housing provides non-pecuniary benefits in the marriage market, in addition to being an asset.

In addition, we complement the empirical literature studying homeownership and demographics. Fisher and Gervais (2011) documents a secular decline in homeownership among the young in the US from 1980s to 2000s and argue that one of the key explanations is the trend towards marrying later. Chang et al. (2019) argues that the decrease in marriage and rise in divorce probability in US from 1970 to 1995 can explain 29% of the increase in homeownership rate among young singles. Andrews and Sánchez (2011) argues that deferral in marriage and child-bearing and rise in single-parent household can explain the rise in overall homeownership rate in Europe in past decades. Dettling and Kearney (2014) found that increase in house price leads to increase in fertility for home-owners in US. Particularly related to this paper are previous empirical works documenting evidence for the marriage house phenomenon in China (Fisher and Gervais, 2011, Choukhmane, Coeurdacier and Jin, 2013, Wei, Zhang and Liu, 2017, Chang et al., 2019). For example, Wei and Zhang (2011) document that, due to rising sex ratio in China, marriage market competition increased for males, and as a result, single males' families save more for their son to be attractive on the marriage market. In interviews they conducted for the paper, many parents said 'having a house for a son is essential for their son's prospects in successfully finding a wife.' In addition, many scholars have provided empirical evidence for rising house prices leading to delays in marriage (Wrenn, Yi and Zhang, 2019, Su, Khan, Hao, Tao and Peculea, 2020, Ma, Ma, Yang, Wang and Wang, 2023) using national data, micro household level data. In this paper, we contribute to this empirical literature in two ways. Firstly, we provide novel micro-data evidence for the marriage house custom. Secondly, we supplement existing empirical literature on rising house prices and delays in marriage by adding cross-county level analysis. In addition, we take a structural approach to study the homeownership decision and demographic effects, quantifying the strength of these channels documented empirically.

The rest of the paper is organized as follows: in section 3, we discuss institutional background that are important for understanding the key modeling choices; in section 4, we introduce the data; in section 5, we show empirical evidence that supports key mechanism in the model; in section 6, we introduce the model; in section 7, we calibrate the model; in section 8, we present model fit and result from a quantitative exercise; section 9 concludes.

2 The marriage house tradition and China's housing boom

This section provides the institutional background of China's marriage house tradition and its recent housing boom. The rental market in China is relatively small, and the male partner and his family are expected to provide a marriage house for the newly wed couple. These institutional features in China's marriage and housing markets inform our modeling choice and quantitative exercise in later sections.

Small rental market. In China, renting is rare and young people tend to live in their parents' house until marriage. In figure 1 we show the share of households and share of married individuals by housing arrangements. As we can see, 85% of households live in an owned housing and less than 10% of households rent. This is even more so for married people: merely 5% of individuals who are married live in rental housing. This is consistent with a world in which young individuals move from living in parents' house directly to living in their own house, without an (or with a very short) intermediate stage where they rent. Figure 2 confirms this hypothesis: the share of people who are single and living without their parent is lower than 5% at any given age. This explains why a marriage house is 'needed' for marriage: since all young people are living in their parents' house before their marriage, if a marriage house is not available at the time of their marriage, the newlywed would need to move into their parent's house after marriage, which is not ideal. In other words, with a under-developed rental market, a marriage house provides convenience for the newly-weds.

Males' family buying marriage house. In 1978 China enacted the famous 'One-Child

Policy'. Due to preference for male child and availability of selective abortion, the sex ratio (number of males over females) at birth rose dramatically in the following decades(Ebenstein, 2010, Ding and Hesketh, 2006). In China, the legal marriage age is 22 for males and 20 for females. Thus, the cohort of individuals born in the One-Child Policy era started to enter the marriage market in the early 2000s.[2] The marriage market has become much more competitive for males since then. Many families with sons view ownership of a marriage house as a key factor in their sons' marriage prospects Wei and Zhang (2011).

Parents' subsidy in house purchase In figure 3, we plot the ratio of net house value over annual household income spent on housing against total years worked for young couples, i.e. those who got married in the past 5 years.[3] What we plot on the x-axis measures how many years would the household need to work in order to finance the housing stock they own, and the y-axis plots the maximum number of years they could have worked. The red line is the 45 degree line. If a household lies below the line, it means they could not have afforded the total housing stock they own all by themselves using only their own income. 66% of the households lie below the line, suggesting that they received subsidy in financing their house. In figure 4 we show the sources of down-payment financing for young households. Only 50% paid for the down-payment fully by themselves, and most of the rest received subsidy fully or partially from parents. Also, given the very high down-payment ratio in China (40% if it is a second house) and how early marriage happens (see figure 2), it is unrealistic for many young people to fully finance the down-payment of their marriage house.

Housing boom Since its housing reform, China experienced a housing boom that lasted for almost two decades.[4]. Figure F.21 shows real house price index by tier of cities in US and China for their most recent boom: in China, house prices grew at 13% a year for 2003-2013; compared to which even the US housing boom before 08 is dwarfed. What's more, income growth has not caught up. As shown in figure F.22, the annual growth rate of per capita real disposable income was only 8.9% a year over the same period for China, which leaves house-price-to-income ratio growing at roughly 4% a year, resulting in a 50% increase in this ratio by the end of this period. Importantly, this housing boom is largely due to reasons exogenous to the marriage market. The key drivers of the boom include massive rural-to-urban migration, government's manipulation of land prices, and high return of housing as an investment compared to other assets in China (Yanyun Man, 2011).[5]

[2]One-Child Policy was abolished in 2015, so all the way until 2030s, the marriage market will be dominated by children born under this policy.

[3]Assuming 30% of annual household income is spent on housing

[4]See brief timeline of China's housing reform in table 1

[5]Urban population grew from 20.6% in 1982 to 56.1% in 2017 (National Bureau of Statistics). Urban land is owned by government in China and proceeds from selling land leases is the most important source of revenue for local governments Yanyun Man (2011). The average annual growth of Shanghai Stock Exchange Index is only 3.5% over 2003-2013 while that of housing is around 13% over the same period.

Accompanying this massive housing boom is a steady rise in the age at first marriage. From figure 5 we can see that average marriage age increased for around 2 year from 1990 to 2010, and importantly, the age gap at marriage between male and female has widened (see table 2) by 0.5 years of age from the 1990s to 2000s. These facts show that it has become harder for males to find spouses in the 2000s compared to before, which is consistent with a world in which males' family need to pay some form of 'bride price', i.e. buying a marriage house, in order to attract spouse.

3 Data

Our empirical analysis utilizes three standard and well-known household surveys in China: the China Family Panel Study, the Chinese Household Income Project, and the China Household Finance Survey. This section focuses on how we analyze each data source, and we leave the description of these datasets to the Appendix.

3.1 Data - China Family Panel Study

CFPS is a nationally representative panel survey launched in 2010 by the Institute of Social Science Survey (ISSS) of Peking University. The survey is biennial and 4 waves (2010,12,14,16) have been released by the time of this draft. The survey collects both individual and household level data including age, sex, education attainment, income, marital status, etc. of the individual, and ownership status of the resident and other house, consumption, asset, and debt, etc. of the household. In the baseline survey there are 14k households and 33k individuals. Following waves follow the original individuals and households and new households created by them.

Given that CFPS is the longest publicly available panel with both marriage and homeownership information, it naturally becomes the main dataset for our analysis. For married individuals, CFPS asks for marriage history, importantly, year of marriage for each marriage. For households who own their residential house, CFPS asks for the year in which the house was purchased, the total area of the house, and the market value of the house at purchase. From these information, we construct a county-year level dataset of average house price and average marriage age, and conduct analysis of whether in county-years where house price is higher/grows faster, marriage age is higher/grows faster; construct a dataset of married couples' residential house purchase year and marriage year, and conduct analysis on whether marriage tends to happen before or after marriage. In addition, CFPS asks for each individual's marital status in each wave. Thus, for the individuals who are single in 2010, we can construct an individual level dataset of marital status and their household's ownership status of other house for 2010-2016, and conduct hazard analysis on whether males whose family own other house get married faster. Description of data construction can be found in the appendix.

3.2 Chinese Household Income Project

CHIP is a nationally representative survey launched by China Institute of Income Distribution and has been conducted 5 waves so far for 1989, 1996, 2003, 2008 and 2013. The size of the sample has changed across years; for the most recent 3 waves, the number of households is around 5k and the number of individuals is around 20k. The dataset provides individual level data on sex, age, education, income, etc. and household level data on expenditure, debt, and asset, etc.

Since the first year the house price data is available is 2003, we also use 2003 as the baseline year to calibrate the model. Thus, we use the 2003 CHIP for constructing data moments for use in the calibration. More on how the data moments are constructed can be found in the Calibration section.

3.3 China Household Finance Survey

CHFS is a nationally representative panel launched in 2011 by the Survey and Research Center for China Household Finance in Southwestern University of Finance and Economics in China. It has been conducted for two waves so far, in 2011 and 2013 respectively. The baseline covers 9k households with 30k individuals and the follow-up expanded to 28k households with 98k individuals. The survey again provides both individual level characteristics and household level financial variables as mentioned above. We benefit from CHFS for its larger sample size. But given that there are only 2 years, the panel dimension is not as useful as that of CFPS. The CHFS asks 'who paid for the down payment on your house', which we use to gauge how important parent's role is in financing the marriage house.

4 Empirical Evidence

In this section, we show a set of empirical evidence that support the key mechanism of our model. Single males whose families own marriage house are more likely to get married than those whose family do not. House purchases tens to happen before marriage, and marriage probability is much higher right after house purchase. In the cross section, both the level and growth rate of marriage age and house prices are strongly correlated with each other, even controlling for time and place fixed effects.

4.1 Marriage and Homeownership at Individual Level

This section starts with an event study that shows marriage house ownership increases a single male's marriage probability. We then formally estimate this effect using hazard analysis in preparation for our quantitative exercise later. We finally address the possible endogeneity issues and show it is unlikely the observed patterns are driven by reverse causality.

4.1.1 Event Study: Marriage Rate Before and After House Purchase

In this section, we conduct an event study around time zero to investigate whether the probability of marriage disproportionately increases after the marriage house purchase.

In Figure 8, we present the cumulative density function plots of marriage over years since house purchase for single males.[6] Each subplot corresponds to people born in a given year. Take the subplot for birth year 1982 as an example: 10 years before house purchase, no one is married; 1 year before house purchase, 20% of them have been married; in the year of house purchase, 30% have been married; 10 years after house purchase, everyone is married.

As Figure 6 shows, in many subplots, there is a relatively obvious kink in the slope around time zero, the year of the house purchase. The area under the dotted line reflects the share of individuals who have been married by each year, and the period-by-period increment, or the slope of the line, reflects the rate at which those who are still single transition into marriage, i.e. the marriage probability of that year. A formal statistical test shows that the slope after time 0 is significantly different from that of before time 0 (see table 3). The significant change in marriage probability around the time of home purchase corroborates to the anecdotal story that marriage house ownership increases a single male's marriage outlook.

4.1.2 Marriage House and Marriage Probability: a Hazard Analysis

We formalize the above patterns with a hazard analysis. In the hazard analysis, the failure event is getting married, and we are interested in whether being in a household that owns houses other than the residential house is associated with higher hazard rate of getting married compared to those whose family owns only residential house. We merge the 4 waves of CFPS together so that we can trace the marital status of male individuals and ownership status of his family across years. We keep only males who are single and have never married in 2010. All control variables are time varying. We categorize the population of households into 4 groups:

[6] We construct the data in the same way as we did for the one for Figure 6, with one additional step. After getting an individual level dataset on marriage and house purchase year, we can construct an individual level panel dataset of marital status and homeownership status. For example, if a person got married in 2010 and bought house in 2008, we know that his marital status is 0 for all the years until time -2, and is 1 for all the years from -2 onwards. Thus, for each person, we know his marital status in each year since house purchase (including in each year before house purchase). From such a dataset, we can calculate the share of individuals who have been married by each year since house purchase and use this to plot figure 8.

own resident or not by own other house or not, and create a dummy variable for each of them.

We use the Cox proportional hazard model:

$$H(t) = H_0(t)e^{X\beta}$$

to estimate the effect of being in a family who owns other house on marriage hazard.[7] The age-dependent hazard $H(t)$ depends on a baseline hazard $H_0(t)$ and the value of the covariates, i.e. X. In the regression table we report both the coefficients, β, and the hazard ratios, $e^{X\beta}$. The hazard ratio reflects how many times more likely is the marriage event for someone with certain characteristics. For example, the bolded coefficient in table 4 can be interpreted as, in any given age, the marriage hazard of single males whose family owns resident and other house is 1.707 times that of the single males whose family owns only the residential house. Since the age profile of marriage for female is significantly different from that for males, we run the regression for male and female separately to allow for different shape of the baseline hazard for the two genders.[8]

As Table 4 suggests, the marriage probability is not significantly higher for either male or female whose family own the resident but not other house. However, it is significantly so for males whose family own other house, but not for females. This is consistent with the story in which marriage house increases marriage probability for males: as the marriage market got more competitive for males (due to higher sex ratio, as introduced in the introduction), males whose family owns other house, which can be used as 'marriage house', is more likely to successfully find spouse and get married.

There may be two reasons why this second pattern is true. First reason is that it might be males with marriage houses are more attractive to females, houses help one to find a spouse. A second reason might be that singles males might have found a partner, but the house is a prerequisite for the marriage (marriage cannot happen until a house is ready). Both reasons suggest that marriage house is a friction in the marriage market. We do not distinguish between these two stories because for the purpose of this study, whether increase in house price delays marriage, these two have the same effect on the outcome.

A more serious threat to our interpretation of the hazard analysis is that a marriage house provides convenience to the newly-weds, and the husband would purchase a marriage house for the convenience value. In other words, the marriage house is not a friction but an endogenous choice. Under this conjecture, we would expect more frequent house purchase activities among males as they enter marriage. So we test whether our hazard analysis is driven by this

[7] The Cox proportional hazard model impose no restriction on the shape of baseline hazard.

[8] See appendix for overall pattern of marriage by gender

alternative conjecture in the following section.

4.1.3 Timing of Marriage and House Purchase

In figure 6, we plot the density of marriage year minus house purchase year. The x-axis corresponds to the number of years since house purchase. Note that each bar represents 2 years. For example, 2 means 2 or 3 years after house purchase and -2 means 1 or 2 years before house purchase. The y-axis corresponds to share of households who fall into each category. For example, the bar at time 0 means 'around 20% of households got married in the year of house purchase or the year after'; the bar at time 2 means 'around 13% of households got married 2 or 3 years after house purchase'. As we can see, the vast majority (more than 90%) of households got married after house purchase.[9]

One might be worried that as long as the age profile of the two events is such that marriage tends to happen after house purchase, even if there is no linkage between the two, one may still get a graph in the above shape. To address this issue, we conduct an independence test. We use the data we used to plot figure 6 to calculate the probability of marriage and the probability of house purchase at each age *assuming these two events are independent*. Then, we simulate a large number of people whose marriage and house purchase follow the above calculated independent age profiles. We record the age of house purchase and age of marriage for each individual in the simulation. We plot the density of marriage year minus house purchase year using the generated data in figure 7. As we can see, the graph looks very different from the one plotted using the empirical distribution: there is no spike right after time 0, and there are less households who get married after house purchase (less than 80%). This suggests that marriage and house purchase are not independent of each other, and marriage tends to happen after house purchase.

Our analysis on the timing of marriage and house purchases confirm that the hazard analysis results are unlikely to be driven by the alternative story. In other words, the boost in marriage probability after house purchase is unlikely to come from males choosing homeownership for its convenience value.

4.2 Marriage Age and House Prices in the Cross Section

To supplement micro-data analysis above, we provide additional evidence supporting the marriage-house complementarity on a macro scale. This section shows that the cross-sectional house price and age at first marriage are strongly correlated in both level and in growth rate.

[9] Apparently a household may switch to a bigger houses after marriage, especially after they have children. Also, if the couple lives with either parent, the house might be the parent's, and can be bought long before their marriage. To address these issues, we keep only couples who are married and do not have children yet and whose parents are not present in the household. Thus, each household corresponds to a married couple living in an owned housing with no parent or child.

We start by examining the correlation between the level of marriage age and the level of house price in the cross section. Figure 9 plots the average age at marriage against the average house price at county-year level. Across the county-year pairs, a 300% increase in the level of house price, which is the same order of magnitude of housing price boom suggested in Figure F.21, is associated with a year increase in the average age at marriage. The cross sectional pattern points to the possibility that rising house prices can deter marriage in the presence of marriage-house complementarity.

One concern is that the above pattern is confounded by the fact that more developed areas tend to have both higher house prices and higher marriage age. To cope with the issue, we formalize the above pattern in a regression table, as shown in Table 5.[10] Even after controlling for county and year fixed effects, marriage age and house price are strongly positively correlated in the cross section. Tripling house price is associated with a more than 0.5 age increase in marriage age in the cross section.

Another way to address the above concern - that when a place gets richer both house price and marriage age increase - is to directly control for average income in a county-year. However, due to data limitation we can only implement a variant. we residualize marriage age and house price at purchase on person or household level income or asset before aggregating these variables to county-year level. We present the result in figure 10 and table 6.[11][12] As we can see, the results are qualitatively and quantitatively similar.

We next investigate whether the positive correlation between house price and marriage age also holds for the growth rate. This exercise speaks directly to whether we should expect to see age at marriage rise as house price grows in a region.

The exact relationship between marriage age and house price growth rates is tricky to show, for the following two reasons. First, the standard growth rate $(y_t - y_{t-1})/y_{t-1}$ may capture high frequency noise in the data instead of longer-term growth rate in house price or marriage age. To address this, we use the regression coefficient of house price or marriage age on time as the measure for how fast the variable grew in a period leading up to a certain year in a

[10] The county-year level dataset on house price and marriage age is constructed from taking average from individual or household level recall data on marriage age and house price at purchase. As a result, there are only 1-30 observations per county-year and the constructed dataset contains quite a lot of noise. Thus, we take 3 or 5 year moving averages (with various weights) to smooth out high frequency noises. Results from here onward are all on the smoothed version of the dataset.

[11] We construced the county-year level dataset on house price and marriage age from re-call data on individual or household level marriage age and house price at purchase. Unfortunately CFPS does not ask for one's or one's family's income level at the time of marriage or house purchase. As a result, there is no information for us to construct county level income measure back in time (i.e. for earlier than 2010).

[12] We only have information on income and/or asset for the individual or household in the survey year.

certain county.[13]

Second, the relationship between increase in house price and increase in marriage age may not be concurrent, for two reasons: marriage tends to happen *after* house purchase, leading to a mechanical gap between the two events; and, even if marriage happens in the same year as house purchase, if high house price *delays* house purchase, higher price in a certain period will mechanically lead to higher marriage age in future period.[14] For example, consider a simple model in which there are two levels of house price: low and high. Every period a measure 1 of single males of the same age come into the model. If house price is low, they buy house in the current period. If house price is high, they buy house the next period (because he needs to save for longer to afford the house). Marriage happens in the same period as house purchase. Apparently, if they get married in the period they come in the model, they get married *young*; if they get married one period after, they get married *old*. Apparently, high house price in period t means agents have to get married in period $t+1$, which means high marriage age in $t+1$ instead of t. Thus, high price growth from $t-1$ to t means high marriage age growth from t to $t+1$.

To address the above two issues, we run regression of smoothed growth rate of marriage age on lagged smoothed growth rate of house price, using the measure of smoothed growth rate as defined in the above paragraph. The results are shown in table 7. As we can see, growth in house price is strongly correlated with growth in marriage age in the period after.[15] The interpretation of the coefficient is also harder due to the more complicated data construction. An example would be, if one moves from a scenario where house price triples every 2 years to a scenario where house price triples every year, the annual increase in marriage age should go up by roughly 0.5 year.[16] As a robustness check, we carry out the same analysis using the residualized marriage age and house price is in table 8. The coefficients are qualitatively and quantitatively similar.

[13] For example, to measure how fast house price grew for the 3- or 5- year period leading up to year t in county i we run a regression of house prices (the smoothed version, as explained before) in years $t-2$ to t (or $t-4$ to t) in county i on time (measured in years). We do this for all years and all counties for both house price and marriage age.

[14] As shown in figure 6, although only less than 10% of marriage happen in the year of house purchase, 50% of them happen within 5 years of house purchase.

[15] A period is defined as 3 years, because both the growth rate and lag is taken for 3-year periods. We construct the measure of 'growth rate' by running regression of house price or marriage age in year $t-2$, $t-1$, t (i.e. previous 3 years) in county i on time (measured in years). Then we regress this generated growth rate of marriage age on the 3-year lagged generated growth rate of house price.

[16] Here 'scenario' refers to a certain county in a certain period, as the regression is at county-year level. The coefficients of the regressions in table 7 is roughly 1. Suppose there are two scenarios, a and b. Let y denote the generated growth rate of marriage age and x denote the generated growth rate of house price. Recall that x_a means 'every year, log house price in place a increase by x_a'. If $x_a = 1$, house price almost triples every year; if $x_a = 0.5$, it almost triples every 2 years. Then if one moves from a place where house price almost triples every 2 years to a place where house price almost triples every year, i.e. $x_b - x_a = 0.5$, then the annual increase in marriage age also increase by 0.5 year, i.e. $y_b - y_a = 0.5$.

To summarize, house price and marriage age positively correlates in the cross section (across county-years), both in levels and in growth rate, even controlling for place and time fixed effects, i.e. having considered that more developed places have both higher/faster-growing house price and higher/faster-growing marriage age.

5 Model

We build a life-cycle consumption-savings model with housing demand for households that consist of parents and a male child to capture the marriage incentive of house purchase described above. It is a unitary model where parents make decision to maximize per person utility in the household.[17] They take marriage probability and the path of house prices as given and make consumption, saving and house purchase decisions.[18] There are several key elements of the model. First, marriage is modeled as a shock and its probability depends on homeownership (Cubeddu and Ríos-Rull, 2003, Hong and Rios-Rull, 2012). Second, marriage is desirable: there is an additive utility premium for marriage; and there is complementarity between marriage and ownership: there is another additive utility premium for owning while married. Thus, parents have incentive to buy a marriage house such that their male child can get married faster and they can enjoy higher utility induced by marriage. Even if marriage happens before house purchase, they still has incentive to buy because there is additional utility gain from owning while married.[19]

5.1 Model Set-up

Environment. The economy is populated by a continuum of households. A household consists of two generations. The household enters the model when the first generation ('the parents') gets married and exits the model when they die at age t_T; the second generation ('the male child' or 'the child') is born in the same period of when the first generation gets married. Each couple has one and only one child and the child is male.[20][21] Time is discrete and each period is a year. In each period, the age of the first generation is t and the age of the second generation is t_s.

Preferences. Agents enjoy consumption of a numeraire good c, additive utility premium ω when married, and additive utility premium η when owning while married, and bequest d

[17]In this paper we do not consider the decision problem of the families with female child. For the period of interest, most families with male children has one male child - see appendix Fertility.

[18]We assume rational expectations, i.e. agents in the model can perceive the future path of house prices over their life.

[19]For example, it can be because it is crowded to have two generations living in one house.

[20]Before age t_T, the probability of death is zero.

[21]We only consider the sub-population of households with one male child.

in their last period of life:

$$u(c|m,o,t) = \frac{c^{1-\gamma}}{1-\gamma} + \mathbb{1}_{m=1} \cdot \omega + \mathbb{1}_{m=1} \cdot \mathbb{1}_{o=1} \cdot \eta + \mathbb{1}_{t=t_T} \kappa \frac{d^{1-\gamma}}{1-\gamma}$$

where γ governs the intertemporal elasticity of substitution. At the household level, the older generation of the household (i.e. 'parents') maximizes per capita utility in the above form subject to household budget constraint.[22]

The reason why there is no housing service in the utility function is that all households are endowed with a residential house when they enter the model, which cannot be sold. The only margin of decision is on the 'marriage house'. The 'marriage house' does not provide any utility benefit before marriage happens - because the household still lives in the residential house endowed to them. However, upon marriage, the marriage house brings additional utility benefit - this can capture the fact that, instead of having two generations crowded in the same residential house, the newlywed can move out into the marriage house.

Income. Individuals receive income every period from age t_{L_1} to age t_{L_T} according to a deterministic age profile.[23] A household's total income is the sum of the income of all its members in a given period. There is growth in income over time, i.e. individuals born in a later generation receives a higher income path than those born in an earlier generation.[24]

Assets. There are two assets: one period bond and housing. Bonds pay a constant gross return R each period. There is no rental market. Houses are homogeneous each with price P_t in a given period. Agents can observe the house price in each period of their life time.

A household can own up to two houses, a 'residential house', i.e. where the household lives in, and a 'marriage house', i.e. where the child can move into after he gets married. The parents are endowed with a residential house when they come into the model. The residential house cannot be sold. The parents can buy and sell a marriage house subject to transaction cost τ as long as the child is single.[25] Households can borrow against the marriage house with a maximum loan-to-house-value ratio ϕ using one-period bond introduced above.[26] Without a marriage house, households cannot borrow.

Marriage and Family Structure. Agents are subject to marriage shock from age t_{M_1} to t_{M_T} as long as they are single. The probability of getting married depends on age and whether his

[22] In other words, it is a unitary model.

[23] L_1 and L_T are the first and last periods when an agent is in the labor market

[24] Both of the parents and the male child.

[25] If a marriage house is available when the child gets married, an event called 'split-off' happens (which will be explained in the following sub-section), and the marriage house becomes the child's family's residential house, and thus cannot be sold by the parents anymore.

[26] Including in the period when they buy the marriage house.

family owns a marriage house.[27] There is no divorce.

We summarize the 4 states of a household can be in in figure 11. In the figure, each dot is an agent: the blacks the parents, the blues the child, and the orange the grandchild (i.e. darker the color, more senior the generation). The residential house is in black, and the marriage house is in red. 'Single' and 'Married' refers to the marital status of the second generation; 'No house' and 'With house' refers to whether the household owns a marriage house.

All households start from the left: when they enter the model, they are endowed with a residential house, and the child is single, and they do not own a marriage house. When they buy a marriage house, they come to the state at the bottom. While they are at the bottom state they can always sell the marriage house and go back to the original state on the left. While they are at the state at the bottom, if the child gets a marriage shock, they move to the state on the right, i.e. the child's family moves into the marriage house, and the entire household starts enjoying both the additive marriage premium and premium for owning while married. Another possible path from the left to the right is through the top: if the child gets a marriage shock while the family does not own a marriage house, the family comes to the state at the top, i.e. the three generations live together in the residential house, and this household only enjoys the marriage premium. While they are at that state, if the family chooses to buy a marriage house, the family comes to the state on the right, and starts enjoying the additional premium for owning while married.

5.2 Recursive Formulation

We summarize the model in Bellman equations. The state variables are: the amount of bond the household holds, b; the age of the first generation, t; the age of the second generation, t_s; marital status of the second generation, m (1 if married, 0 if single); whether a marriage house exists, o (1 if yes, 0 if not).

The choice variables are: the amount of consumption c; whether to buy a marriage house (when the family does not own one) H (0 not buy, 1 buy); whether to sell the marriage house (when the family owns one and is allowed to sell) Q (0 not sell, 1 sell).[28]

Recall that t_{M_1} and t_{M_T} denote the first and last age at which agents receive marriage shock; t_{L_1} and t_{L_T} denote the first and last age at which agents receive income; let P_t denote house price in a given period; and ϕ denotes the maximum loan to value ratio. For simplicity, we omit the time subscript on state variables and use prime to indicate state variables of the next

[27] Note that we do not model the matching process in the marriage market. From the male individual's perspective, marriage is an exogenous shock.

[28] i.e. when the child is not married yet

period.

$$
V_t(t_s, m, o, b) = \max_{c,H,Q,b'} u(c) + \cdots
$$
$$
\begin{cases}
\pi_{o,t_s}\beta V_{t+1}(t_s', 1, o', b') + (1-\pi_{o,t_s})\beta V_{t+1}(t_s', 0, o', b') & \text{if } t \neq T \text{ and } m = 0 \text{ and } t_s' \in [t_{M_1}, t_{M_T}] \\
\beta V_{t+1}(t_s', m, o', b') & \text{if } t \neq T \text{ and } (m = 1 \text{ or } t_s' \notin [t_{M_1}, t_{M_T}]\ ) \\
v(d) & \text{if } t = T
\end{cases}
\tag{5.1}
$$

where

$$
u(c) = \frac{c^{1-\gamma}}{1-\gamma} \tag{5.2}
$$

$$
v(d) = \kappa \frac{d^{1-\gamma}}{1-\gamma} \tag{5.3}
$$

$$
o' = \begin{cases} 1 & \text{if} \quad (o = 1 \quad \text{and} \quad Q \neq 1) \quad \text{or} \quad (o = 0 \quad \text{and} \quad H = 1) \\ 0 & \text{if} \quad (o = 0 \quad \text{and} \quad H \neq 1) \quad \text{or} \quad (o = 1 \quad \text{and} \quad Q = 1) \end{cases} \tag{5.4}
$$

subject to budget constraint

$$
\begin{cases}
c_t + R^{-1} b_t \leq b_{t-1} + y_t & \text{if} \quad m = 1, h = 1 \\
c_t + d_t \mathbb{1}_{t=t_T} + R^{-1} b_t + P_t H_t \mathbb{1}_{o=0} \leq b_{t-1} + y_t + P_t Q_t (1-\tau) \mathbb{1}_{o=1} & \text{otherwise}
\end{cases}
\tag{5.5}
$$

where

$$
b' \geq -\phi P \mathbb{1}_{o'=1} \tag{5.6}
$$

6 Calibration

We estimate 3 parameters: utility premium for marriage ω, utility premium for owning while married η and utility multiplier on the bequest κ, to match 3 moments: male marriage rate, male ownership rate, and net asset to income ratio, from 2010 CFPS. We summarize the result of the calibration in Table 9. We calibrate the rest of the parameters according to common practice in the literature or directly from the data, which we summarize in Table 10.

6.1 Calibrated from the Literature or the Data

We pick most parameters using data and the existing literature. Table 10 shows such parameters and their sources. The remaining of this section discusses these picked parameters in detail.

Key ages in the model. Agents in the model receive income from age $t_{L_1} = 16$ to age

$t_{L_T} = 70$, have marriage probability from age $t_{M_1} = 16$ to age $t_{M_T} = 45$, die at age $t_T = 70$. When entering the model, both parents are age 25. Recall that the male child is born in the same year as when the parent generation enters the model. This suggests that the age gap between the parent and the male child is also 25.

Parameters from the literature. For most parameters that are not specific to the model, we pick according to common practices in the literature. We pick time preference parameter β to be 0.98 (Frederick, Loewenstein and O'donoghue, 2002); we pick the constant relative risk aversion parameter to be 5 (Landvoigt, Piazzesi and Schneider, 2015).

Income Age Profile. We estimate the income age profile from the cross section of CHIP 1995-2008 (as plotted in figure 12).[29] We run a regression of log of annual personal income (in 2010 Chinese Yuan) on age polynomials up to the 4th order, educational group fixed effects, and year fixed effects, for males and for females separately.[30]

6.1.1 Marriage Probability

We want to estimate the probability of getting married at each age for agents whose family owns and does not own a marriage house. Arguably, the age-specific marriage probabilities have shifted throughout the years we study. Therefore, we estimate the marriage probabilities in two steps.

In the first step, we estimate age-dependent marriage probabilities for non-owners in 1995 and 2010 separately. Then we obtain such probabilities for 1996-2009 by linearly interpolating these two series for each age. To construct age-specific marriage probabilities, we apply the Kaplan-Meier estimator on a smoothed version of the empirical cdf of marriage.[31][32] The estimated pdf and cdf are plotted in figure 13.[33] We report the estimated marriage probability for non-owner males in table 11.

We then obtain age-specific marriage probabilities for marriage house owners using the

[29] As we can see, the shape of the income profile is not exactly the usual hump-shape we are familiar with: instead of peaking at around 45, it peaks at a much earlier age. We interpret this as China not being in its steady state: China experienced rapid growth since its Reform and Re-opening in 1978; real income has more than doubled over the past decade (see figure F.22). Thus, younger generations have higher life-time real income than the old. As a result, the peak of the income profile in the cross-section is shifted towards younger ages.

[30] Note that we are not considering any heterogeneity, e.g. education, yet; in other words, the age profile is that of an average male in the economy.

[31] The idea of the Kaplan-Meier estimator is as follows: if the share of married individuals in year 1 and 2 are s_1 and s_2 respectively, then the probability of marriage in year 2 is $(s_2 - s_1)/(1 - s_1)$, i.e. the increment from year 2 to year 1 divided by the share of single individuals in year 1.

[32] The smoother we use is a 3-year moving average and we apply the smoother three times to the empirical cdf to obtain the estimated cdf.

[33] As one can see, since the empirical cdf is simply from the cross-section of the share of married individual by age, there can be a lower share of married individuals in age x than age (x-1), resulting in negative empirical pdf. This is the reason why we need to apply the smoother to the empirical cdf before estimating the marriage probability.

estimated proportional hazard parameter from the Cox model estimated earlier. From the hazard regression presented in the empirical section in table 4, we can read that, on average, in a given age, the marriage probability of those whose family owns marriage and residential house is 1.707 times the probability of those whose family owns residential house only. Then once we get the marriage probability at each age for those whose family owns residential house only, we know marriage probabilities for all ages and ownership statuses for all years studied.

6.1.2 House Prices

In the model, the agents can observe the full path of house price to income ratio for their life time. Since China has only had a housing market for around 20 years, but agents in the model live for as long as 50 years, we need to make assumptions about house price and income growth paths in the future.[34]

The house price path we use can be decomposed into four time periods: 1987-2003, 2004-2013, 204-2017, and 2017-2058. The challenge with constructing a reliable and consistent house price path is two-fold. Firstly, as noted in Fang et al. (2016), the official house price data from China's national Bureau of Statistics underestimate China's house price boom between 2004 and 2013. Secondly, we need to predict house price growth into the future. To circumvent these challenges, we adopt the following strategy. House prices between 1987 and 2003 are taken from from China's National Bureau of Statistics. For house prices between 2004 and 2013, we use real house price growth rates calculated in Fang et al. (2016). As Liu and Xiong (2018) points out, NBS house price data are more reliable after their change in methodology. We then calculate house prices between 2014 and 2017 using house price growth rates reported by the NBS. For years beyond 2017, we predict future house prices by estimating an ARMA(1,2) process with year trend on the logarithm of 1987-2017 house prices.[35]

The last step in constructing the price-to-income ratio series is to predict an income growth path into the future. For this purpose, we use household real disposable income per capita published by the NBS for 1987-2017. Beyond 2017, we predict an income growth path by estimating an ARMA(2,1) process with year trend on log real household disposable income per capita.[36] The final price-to-income ratio is then defined as the ratio between the house price series and the income series.

[34]If the first generations got married when they are 16 (the earliest age at which agents can get married), they live for 50 years in the model (as they die at age 65). For those who got married later than 16, they live fewer number of periods in the model. For example, for the first generations that got married in year 2000 (when they are age 16), they would have expectations about house price for year 2000 to 2050 but we apparently do not know the empirical house price for 2019 to 2050 yet.

[35]The number of lags for AR and MA components is determined by ACF and PACF.

[36]The number of lags for AR and MA components is determined by ACF and PACF.

6.1.3 Other parameters from the Data

For interest rate R: in China households keep majority of their liquid asset in current or fixed deposits at commercial banks (see appendix Portfolio Share); it is reasonable to use the interest rate on current and fixed deposit in China as benchmark for setting R in the model. In China, the central bank, i.e. People's Bank of China (PBoC) sets a 'benchmark deposit rate' and commercial banks set their own rate according to the benchmark with no more than 10% of deviation. From 2006 to 2015, PBoC benchmark deposit rate varied from 2.50% to 1.50%. We simply take the midpoint, i.e. 2.6% as $(R-1)$ in the model.

For loan-to-value ratio ϕ: in China, the central government sets an explicit lower bound for down-payment ratio which is different for the first and second house a household purchases in order to control speculation in the housing market. The exact rule changes very frequently because this is a major policy tool to either combat overheating or to stimulate the housing market.[37] Fang, Gu, Xiong and Zhou (2016) documents that the down payment ratio was typically 40% for first-time homeowners but higher for second house purchases between 2003-2012. For simplicity, we pick loan-to-value ratio ϕ to be 0.55 to reflect that, in addition to the family's resident house, the marriage house is usually a second house of the male child's family.

For transaction cost in our model captures pecuniary costs for the seller in a house sale. In reality, house buyers and sellers pay VAT taxes, deed taxes, business taxes, and income taxes in a housing transaction. The exact marginal tax rate may vary depending on tenure period, value added, and total number of houses owned. According to Cruz et al. (2008), the total sales taxes in a transaction paid by both the seller and the buyer can exceed 10% of housing value in China. We pick middle ground 5% for sellers in our model to abstract away from how these taxes are shared between buyers and sellers.

6.2 Estimated by Target Data Moments

We calibrate our model using simulated method of moments. The three parameters remaining to be set are the additive utility premium for marriage ω, the additive utility premium for owning while married η, the utility multiplier on bequest. κ.

The model is initialized with four distributions among males aged 26-70 in 1995 taken from the 1995 CHIP. These four distributions are age distribution, marital status by age, house ownership status by age, and household net non-housing wealth to house price ratio by age.

[37] In 1999, the down-payment ratio for the first house was adjusted to 20% from 30% and the guideline says 'that of the second house should be higher than the first' with no explicit rule; in 2007, it was raised to 30% for the first and 40% for the second; in 2008, in face of the financial crisis, it was relaxed to 20% and 30%; in order to control overheating of the housing market, it was raised to 30% and 50% in 2010 and to 30% and 60% in 2011; it was again relaxed in 2014 to 30% and 30%, and then in 2015 to 30% and 40% Yanyun Man (2011)

We then simulate forward to 2010 to match the 2010 CFPS cross section in the following three moments: 73% of males aged 16-45 are married; 62% of households with a male child own a marriage house; household net wealth to annual household income ratio 4.6.

The three targeted moments exactly identify the three calibrated parameters. To illustrate the identification strategy, notice first the three calibrated parameters govern three motives for house purchase in the model. The strength of bequest motive influences agents' motive to save. Therefore, a stronger bequest parameter κ will lead to higher household wealth positions in 2010.

Marriage premium ω and marriage house premium η together prompt agents to purchase a marriage house in exchange for higher age-specific marriage shock probabilities. That is to say, ω and η jointly affect ownership rate in the 2010 cross section. Distinct from ω, however, marriage house premium η further awards marriage house purchases for married households. In other words, holding home ownership rate fixed, the relative strength of marriage house premium η to marriage premium ω determines the relative share of pre-marriage buyers to post-marriage buyers.

The relative share of pre/post marriage buyers lead to different levels of marriage rate in the 2010 cross section, as pre-marriage purchases result in higher marriage shock probabilities whilst post-marriage purchases do not. Therefore, the targeted 2010 ownership rate and marriage rate together determine the value of marriage premium and marriage house premium.

Table 9 displays results from the calibration exercise. The model is able to match the targeted moments relatively well. The calibrated marriage premium and marriage house premium are roughly 11 and 8 years of consumption for an average-asset single non-owner household in 1995.

7 Quantitative Exercise

In this section, we show that the model captures well the timing of marriage and homeownership. Then in counterfactual exercises, we isolate out the effect of marriage probability booster and marriage house convenience on age profile of homeownership. Last but not least, we examine counterfactual marriage age when the marriage house friction is absent.

7.1 Non-targeted Moments

To assess how well the model captures the relationship between marriage and house purchase, we compare the joint distribution of marital status and marriage house ownership between the model and the data.

We focus on the families with a male child whose age is between 16 and 45. Thus, there are a total of $3 \times 30 = 90$ untargeted moments.[38] We show the fit between the model and data moments in figure 14. See table 12 for how households in the data are classified into the above 4 states so that they are comparable to that in the model.

As males age, they tend to leave the single, non-owner state, as marriage shock arrives and as their family accumulates enough money for purchasing a marriage house. Both the estimated marriage probability for non-owners and the endogenous decision of purchasing a marriage house determine how fast agents leave the single, non-ownership state.

Conditioning on leaving the single non-owner state, the pre-marriage homeownership decision affects how many agents transition into the single-owner state. The marriage probability multiplier effect then affects how soon agents transition from single-owners to married-owners. In addition, post-marriage house purchase decisions determine the rate at which agents leave the married, non-owner state into the married-owner state.

Therefore, matching the join distribution of marital status and homeownership is a testament to the validity of exogenous input of non-owner marriage probability, marriage probability multiplier estimates, and the endogenous homeownership decisions before and after marriage in the model.

The model is also able to reproduce the gradual rise in marriage age in China over the past few decades as shown in figure 15. Notice that the *levels* is higher in the model than in the data: this is normal because we used 2010 data to estimate the by age hazard rate of marriage; since there has been a delay in marriage, the hazard rate estimated from 2010 is necessarily lower (at each age on average) than that in earlier years. In the model, two forces generate this gradual rise in marriage age. Firstly, the underlying marriage probability estimated in time-varying and captures social and economic transitions that have caused delay in marriage and that are not explicitly modeled in the benchmark. Secondly, house price is rising in the model which means households take longer time to save up for home purchases. We isolate out how much each force contributes to the delay in marriage in the following section.

7.2 Counterfactual Exercises

In this section, we quantify the effect of marriage market benefits of homeownership. The marriage probability boost effect is responsible for why homeownership happens so early in the lifecycle for Chinese households. The marriage house convenience value, on the other hand, affects mostly the overall level of homeownership. Removing the marriage house friction from the marriage market implies that the age at marriage would have only increased by 60% of what was observed for China between 1995 and 2010.

[38] The reason why it is 3 instead of 4 per age is that since the shares add up to 1, there is one less degree of freedom.

The marriage probability booster effect of marriage house ownership affects mostly the timing of home purchase but not the level of ownership. In Figure 16, we re-plot the joint distribution of marital status and homeownership when marriage house ownership does not bring advantages on the marriage market. Immediately stands out is that young agents who leave the single, non-owner state do so at a much slower pace. As the married, non-owner subplot shows, this slow pace of transitioning out of the single, non-owner state is not a result of changed underlying marriage probability for non-owners. Instead, the removal of marriage probability booster effect delays marriage house ownership by about 5 years. Some single agents still choose to buy a marriage house, in anticipation of future marriage shock. As the single-owner subplot demonstrates, the ownership profile for single males is compressed to the right. The implication is that a lot less people transition from the single-owner state into the married-owner state. In fact, at age 30, only 40% of households are married and owners, as compared to 60% in the benchmark case. This difference, however, starts to diminish as agents age past 30. This is a consequence of people leaving the married, non-owner state by purchase houses post-marriage. By age 45, the difference in the share of households who are married has disappeared. The overall ownership rate among households age 15-45 goes down slightly from 61.5% to 58.4%.

The marriage house convenience, however, dramatically changes the level of ownership rate. Figure 17 plots the joint distribution of marital status and homeownership when a marriage house does not provide any convenience for the newly-weds. Similar to the case above where a marriage house does not increase a single males' marriage prospects, agents leave the single, non-owner state at a much slower rate. The reason for this change, however, is totally different. As the single-owner subplot suggests, the majority of young people choose to be non-owners rather than delaying home purchases. Some young agents, yet, still choose to buy a marriage house before marriage to take advantage of the marriage probability booster effect, particularly between age 21 and age 30. In fact, the married, non-owner subplot indicates that there is substantial reduction in home purchases post-marriage as well. Across all ages, substantially more people choose to be non-owners. The overall ownership rate plummets from 61.5% to 25.1%.

The marriage probability boost and marriage house convenience also interact in the model. Figure 18 demonstrates the drastically different joint distribution of marital status and homeownership when both marriage-related incentives for home purchase are gone. Compared to the previous case where only the marriage house convenience is removed, single males no longer purchase houses before marriage. All home purchases happen after marriage for the purpose of investment. The overall ownership further drops to 16.9%. Notice this 8 percentage points reduction in ownership more than doubles the reduction of 3% if we compare the benchmark results to the case where only the marriage probability boost effect is removed. The interpretation is that either effect is sufficient to prompt about 5% of households, particularly young singles males, to purchase a marriage house.

Last but not least, removing the marriage house friction from the marriage market slows down the rise in age at marriage. To quantify the the contribution of marriage probability premium, we boost the marriage shock probabilities for single non-owners to the same level as marriage house owners. In other words, single males without a marriage house are not penalized on the marriage market. Note that factors not explicitly modeled are still present in the underlying marriage probabilities since our estimates are time varying (see Section 6.1.1). Thus, the interpretation for our decomposition exercise is that structural changes like the increasing female labor participation, urbanization, growing skewness of sex ratio, etc are taking place in the background, but a marriage house is no longer needed for attracting a spouse or for tying the knot conditional on being a relationship. Figure 19 plots the counterfactual age at marriage when the marriage probability premium is absent. Without the marriage house friction, age at marriage would have risen by one year instead of 1.8 years between 1996 and 2010, which marks a 40% reduction of the delay in marriage.

8 Conclusion

In this paper we study the relationship between marriage and house purchase in China. China has a 'house for bride' tradition where the males households provide a 'marriage house'. This non-pecuniary benefit of ownership can lead to a high ownership rate that also happens early in the life cycle, and changes in the housing market can affect the marriage market. We first provide empirical evidence supporting the existence of the 'house for bride' tradition, and use a quantitative model that includes this mechanism to estimate its effect on ownership and age at marriage.

In the empirical part we show that: house purchase tends to happen before marriage; marriage is more likely to happen right after house purchase; single males whose family owns marriage house are more likely to get married than those whose family do not; both the level and growth of marriage age and house price are strongly positively correlated in the cross section, even controlling for time and place fixed effects. These patterns are consistent with the 'house for bride' tradition: marriage house is required on or before marriage and marriage house increases single males' marriage probability.

We then build a model of household choice problem with housing demand to capture the 'house for bride' force: parents of single male child choose whether and when to buy a marriage house for their son, taking house price and marriage probability as given, knowing that the probability of getting married will be higher if the own a marriage house, and that their utility will be higher after their son is married.

The key force in the model is as follows: parents of male child wants their son to get married (because after their son is married, their utility is higher); in order to make marriage happen faster, they need to save up to buy a marriage house; this leads to a higher homeownership rate than if housing only had financial returns. Also, if the price of marriage house is higher, they need more years to save enough to afford the house; given that marriage is more likely to happen with a marriage house, if house price is higher, less number of people in a given age owns marriage house, marriage is less likely to happen, and average age at marriage will be higher.

We calibrate 3 parameters of the model to match 3 key patterns of marriage and ownership in the 2010 cross section. We conduct 2 tests of model fit: the share of households by ownership status and marriage status over age, and the path for age at marriage for males over 1995-2010. The model-generated series match the data series very well.

At the end, we use the model to answer the question: how much of the ownership in China can be explained by the complementarity between ownership and marriage, and how much of the rise in marriage age in China in the past two decades can be explained by the 'house for bride' channel combined with the rising house price. We simulate a large number of cohorts of households, each starting their life in a different year, so that we have a cross section of households whose heads are from age 16 to 65 for each year from 1995 to 2010. We found that 60% of the ownership can be explained by the complementarity between ownership and marriage, and 30% of the rise in marriage age can be explained by the 'house for bride' tradition combined with rise in house prices.

There are many steps we would like to take to improve the paper. On the empirical side, we would like to find instruments for house prices so that we can identify the causal effect of having a marriage house on marriage. For example, since the proceeds from selling land leases is one of the most important sources of local government revenue, shocks to local government's budget that induces them to sell more or less land in certain years can be used as instrument for house prices. Other than instruments for house prices, we can find instruments for relaxing agents' budget constraints during house purchase: central government of China changes regulations on mortgage interest rate, down-payment ratio, and taxation on house sales very frequently, which provides exogenous variation to how hard it is to buy a house.

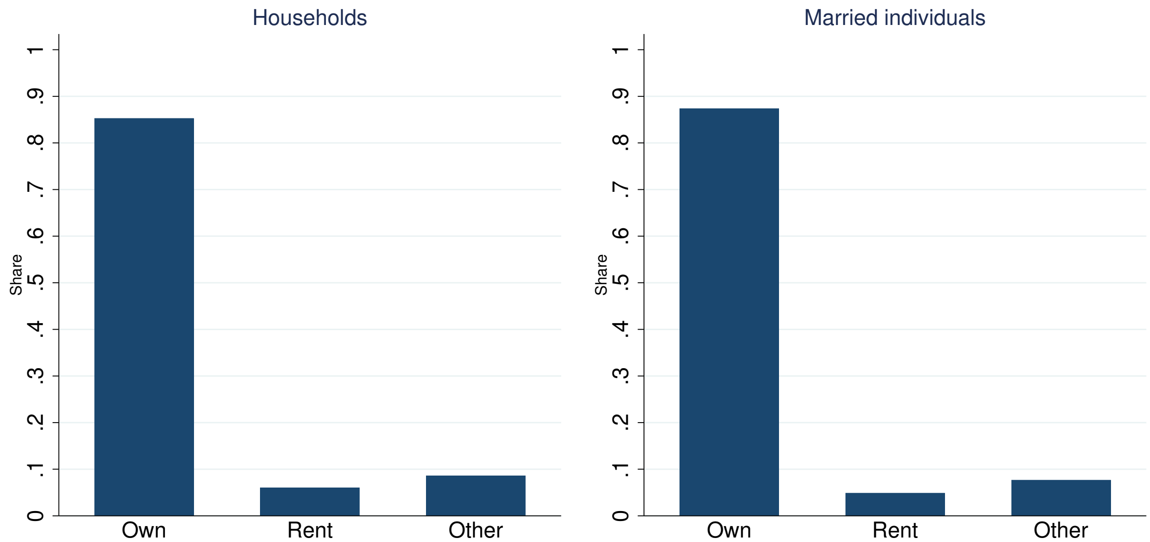
On the modeling side, we are considering adding in heterogeneity in terms of income or growth rate of house prices. For example, house price grew faster in top tier cities in China. We could conduct the calibration and simulation separately for different cities, and plot the model generated house price vs marriage age, and see how much of the variation in the data between the two variables can the model explain. We can also experiment with alternative formulations of expectations of house prices. For example, we can allow different scenarios for future price paths (boom, bust, plateau) or allow different agents to have different beliefs.

Table 1: China's Housing Reform - Brief Timeline

1949:	Establishment of People's Republic of China
1949-1976:	Mao's era; planned economy; socialist housing system
1978:	Start of Deng's era; Reform and Reopening; start of housing reform
1978-1993:	Experiments in selected cities (e.g. Shenzhen)
1988:	Start of nationwide reform to privatize and commercialize urban housing
1993-1997:	Restructuring of the construction, finance, and provision system
1998:	Completely ended linkage between housing and employers

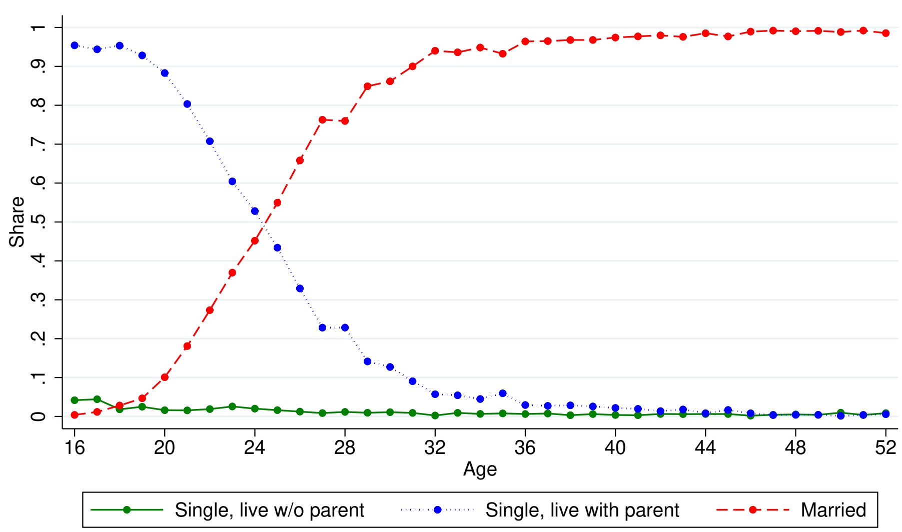
Notes. Taken from Yanyun Man (2011); in 1949-1976, in rural area, land is collectively owned; in urban area, rental housing is distributed to employees by their employers.

Figure 1: Share of People by Housing Arrangement



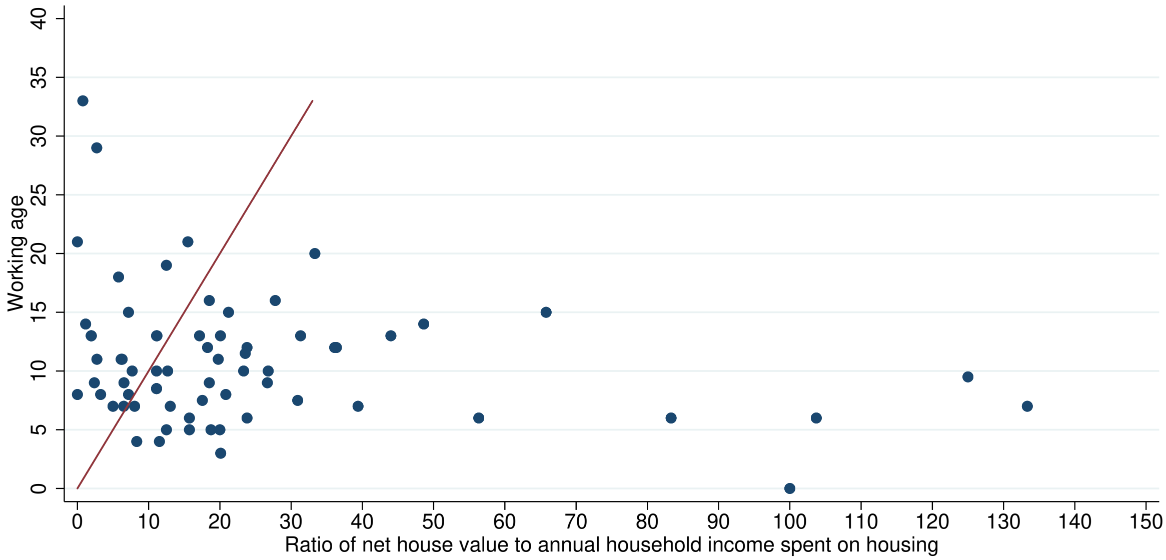
Notes. CFPS 10. 'Other' includes partial ownership, living in employer or government provided free housing, living in friends or relatives provided free housing, living in public housing, etc.

Figure 2: Share of Individuals by Cohabitation Status



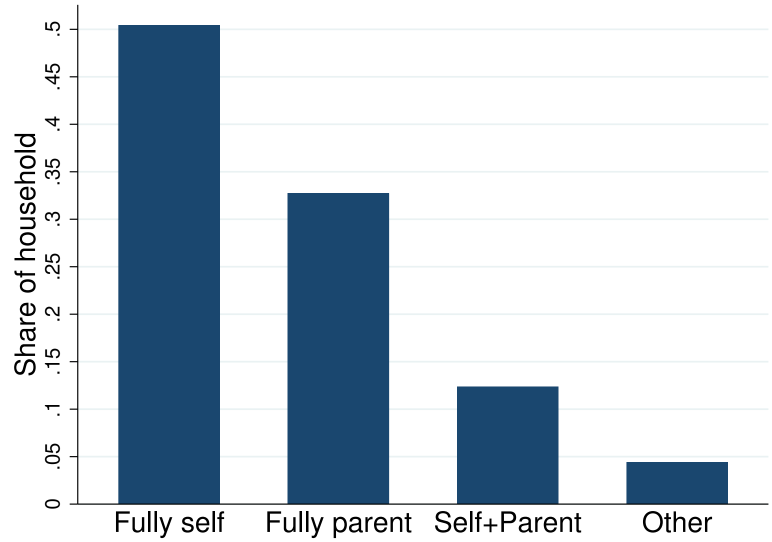
Notes. CFPS 10. Only those who are currently married or never married plotted (only 6.5% of individuals do not fall into these two categories). CFPS provides a family roster file in which it reports whether an individual's father or mother lives in the same household. Living with parent is defined as either parent being in the same household.

Figure 3: Housing Affordability



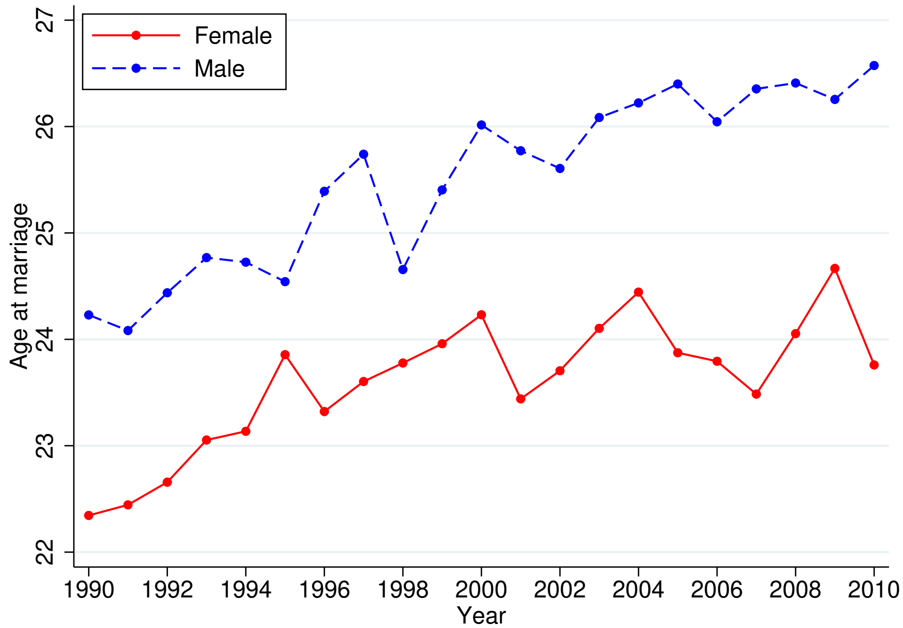
Notes. CFPS 10. We keep couples married in the last 5 years, live in an owned house, do not live with parents, and do not have kids yet. Net house value is the current market value of the residential house minus debt on the house. We use the variable total household income provided by CFPS. We assume 30% of household income is spent on housing. We calculate working age as the husband's age minus his years of education minus 6.

Figure 4: Who Paid for Downpayment?



Notes. China Household Finance Survey 2013. Households are asked who paid for the downpayment of the residential house. Options include self and spouse, parents, children, relatives, friends, banks, etc. We keep only households with only a head and his/her spouse, and who married in the past 5 years.

Figure 5: Age at First Marriage



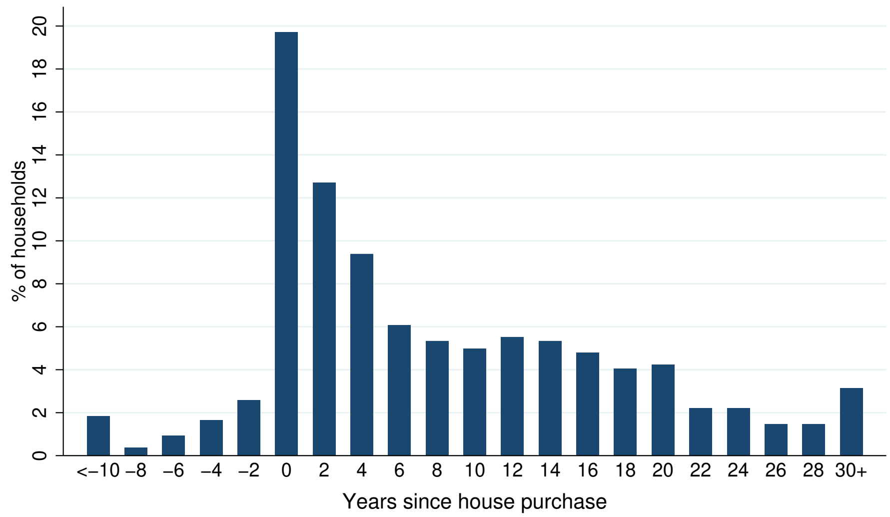
Notes. CFPS 10-16; we calculate age-at-marriage from individuals' birth year and year of marriage; only urban population and those who are married for the first time included.

Table 2: Gender Age Gap at Marriage

Dep. Var: Age at Marriage				
Male dummy	1.905***		1.929***	1.617***
	(0.0951)		(0.0944)	(0.130)
POST 2000 dummy		0.901***	0.949***	0.645***
		(0.0966)	(0.0938)	(0.120)
Male X POST				0.631**
				(0.189)
Observations	6,766	6,766	6,766	6,766
R-squared	0.057	0.013	0.071	0.072

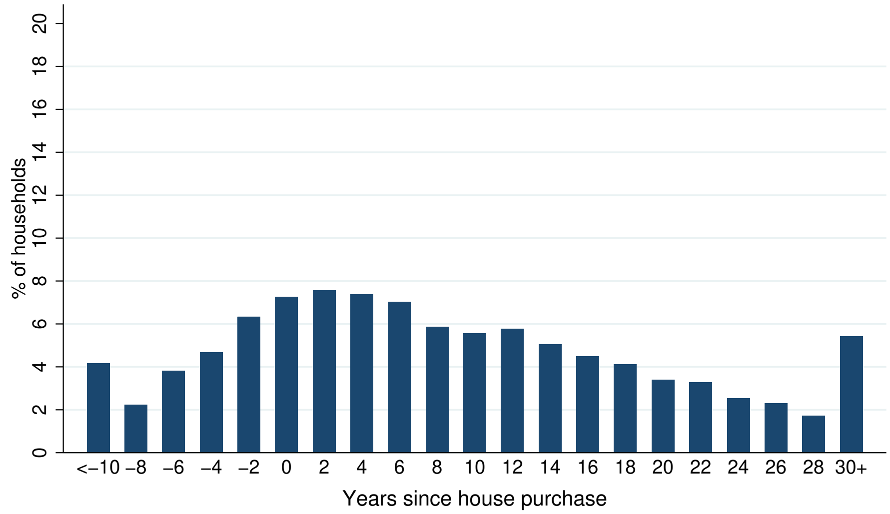
Notes. CFPS 10-16; we calculate age-at-marriage from individuals' birth year and year of marriage; only urban population and those who are married for the first time included. Years 1990 to 2012 included; POST is a dummy for greater than 2000; robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Figure 6: Density of Marriage Year Minus House Purchase Year



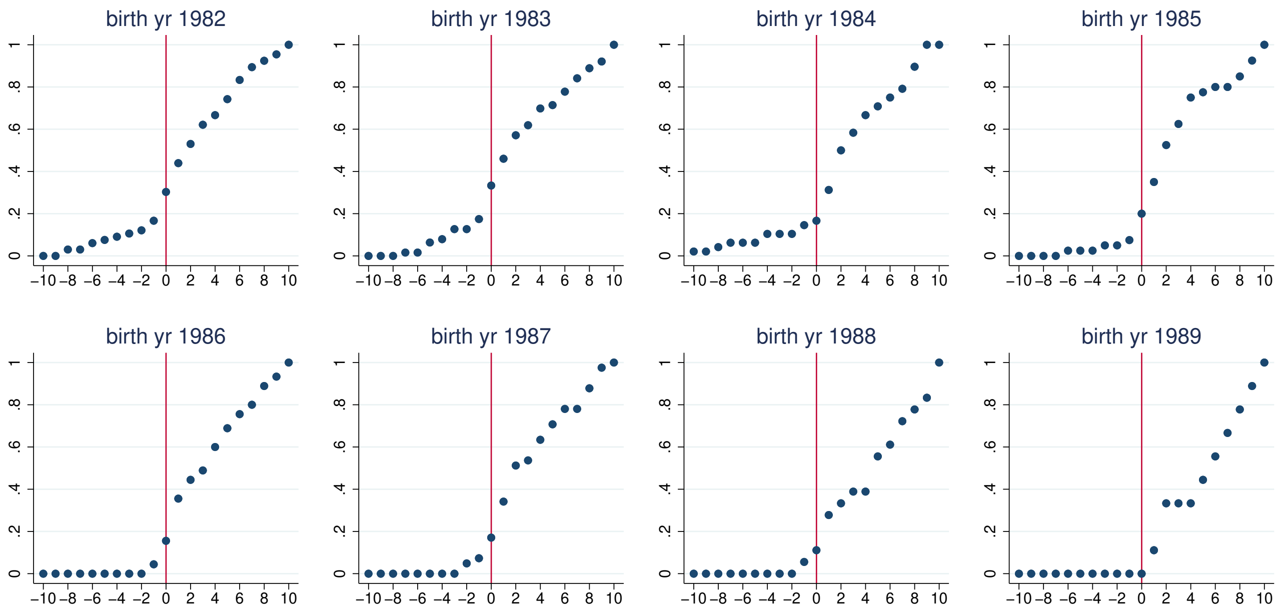
Notes. CFPS 10-16. We keep individuals who live in an owned housing, are married for the first time, do not live with parents, do not have kids yet, between age 16 and 51 (consistent with ages that have positive marriage probability in the model), then we keep one observation per household. We observe the year in which the residential house is purchased and the year in which the individual got married. From that we can calculate the difference between the two. Each bar represents 2 years, i.e. bar -1 represents 2 years and 1 year before house purchase; bar 0 represents the year of house purchase and 1 year after house purchase, etc.

Figure 7: Density of Marriage Year Minus House Purchase Year Assuming Independence



Notes. We use the data we used to plot figure 6 to calculate the probability of marriage and the probability of house purchase at each age assuming these two processes are independent. Then we simulate a large number of individuals whose life follow these two independent age profile and record the age at which they buy house and get married. Then we plot the density of marriage year minus house purchase year from this generated dataset. Each bar represents 2 years, i.e. bar -1 represents 2 years and 1 year before house purchase; bar 0 represents the year of house purchase and 1 year after house purchase, etc.

Figure 8: Event Study: the Share of Married Before and After House Purchase



Notes. CFPS 10-16. Each subplot is on individuals born in a given year. Refer to note of figure 6 for construction of the raw data. From that we can construct a panel dataset on marital status for each individual. For example, if an individual purchased the house in 1998 and got married in 2002, his marital status is 0 for all the years until year -4, and will be 1 for all the year onwards. Then we can calculate the cross-sectional share of individuals who have been married by each year since house purchase.

Table 3: Event Study: Test on Break in Slope

Dep. Var: Married dummy	1982	1983	1984	1985	1986	1987	1988	1989
Time X Post	5.12	4.13	6.39	6.01	7.62	6.57	7.78	9.39
	(0.450)	(0.414)	(0.666)	(0.874)	(0.483)	(0.630)	(0.463)	(0.569)
Observations	21	21	21	21	21	21	21	21
R-squared	0.994	0.995	0.986	0.979	0.993	0.989	0.992	0.989

Notes. CFPS 10-16. The table shows the coefficient on the interaction between time and POST in the regression of marital status on time, POST and time X POST, where time is from -10 to 10 (i.e. years since house purchase), and POST is 1 if time$\geq$0. All coefficients significant with p-value < 0.01.

Table 4: House Ownership Status and Marriage Hazards

	Coefficient		Hazard Ratio	
	female	male	female	male
Own neither resident nor other house	-0.196	0.047	0.822	1.048
	(0.147)	(0.190)	(0.121)	(0.199)
Own resident and other house	-0.310	0.535***	0.733	**1.707***
	(0.199)	(0.141)	(0.146)	(0.241)
Own other house but not residential house	-0.267	-0.127	0.766	0.881
	(0.300)	(0.377)	(0.230)	(0.332)
Years of education	-0.152***	-0.039**	0.859***	0.962**
	(0.017)	(0.015)	(0.015)	(0.015)
Log personal income	-0.004	0.048***	0.996	1.049***
	(0.027)	(0.015)	(0.027)	(0.016)
Log family asset	0.001	0.001	1.001	1.001
	(0.005)	(0.005)	(0.005)	(0.005)
Urban	0.466***	-0.229*	1.593***	0.795*
	(0.128)	(0.125)	(0.203)	(0.099)
N	1,356	1,933	1,356	1,933
Province FE	Y	Y	Y	Y
Age Polynomials	Y	Y	Y	Y

Notes. CFPS 10-16. We categorize single males by ownership status of his household: owns neither resident nor other house; owns residential house only; owns other house only; owns both resident and other house. The left-out group is those whose family owns residential house only. Education is years of completed education; income is log of individual annual income; family asset is log of net asset; income and asset in 10,000 of 2010 RMB. We use the Cox proportional hazard model for the baseline hazard. Individuals enter the model at age 16, i.e. start to have marriage hazard starting from age 16.

Figure 9: Marriage Age and House Price in the Cross Section

Notes. CFPS 10-16. Plotted using binscatter with 50 bins. Each underlying dot is a county-year. There are 162 counties; years span 1990-2012. For data on marriage: we keep individuals who are married for the first time, calculate their age at marriage from their marriage year and birth year, and calculate the average age at marriage for each county in each year. For data on house price: we keep households living in owned housing; we can observe the year in which the house is purchased, the total area, and the value at purchase; from this we can calculate the house price per square meter for that house; we deflate house price to 2010 RMB and then take logs; then we can calculate the average house price in each county in each year.

Table 5: Marriage Age and House Price in the Cross Section

Dep. Var: Age at Marriage (at county-year level)				
Log of house price	1.176***	0.695***	1.196***	0.537**
(2010 RMB/square meter)	(0.114)	(0.196)	(0.117)	(0.230)
Year FE	N	N	Y	Y
County FE	N	Y	N	Y
Observations	394	373	393	372
R-squared	0.214	0.834	0.231	0.847

Notes. Refer to the note for figure 9 for data construction. After those steps, we take 5 year moving average for both log of house price and marriage age with weights 1/9, 2/9, 3/9, 2/9, 1/9 for $y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}$ respectively. The same regression using alternative weights (e.g. uniform weights; only middle three years) give similar results. *** p<0.01, ** p<0.05, * p<0.1.

Figure 10: Marriage Age and House Price in the Cross Section - Residualized

Notes. CFPS 10-16. Plotted using binscatter with 50 bins. Each underlying dot is a county-year. There are 162 counties; years span 1990-2012. Refer to the note for figure 9 for data construction. The only variant is that before taking county-year level average on marriage age and house price, we residualize them on individual level income or household level net asset.

Table 6: Marriage Age and House Price in the Cross Section - Residualized

Dep. Var: Age at Marriage (at county-year level, residualized)				
Log of house price	1.122***	0.579***	1.150***	0.457*
(2010 RMB/square meter, residualized)	(0.134)	(0.205)	(0.138)	(0.241)
Year FE	N	N	Y	Y
County FE	N	Y	N	Y
Observations	384	363	384	363
R-squared	0.155	0.821	0.169	0.833

Notes. Refer to the note for figure 9 for data construction. The only variant is that before taking county-year level average on marriage age and house price, we residualize them on individual level income or household level net asset. After those steps, we take 5 year moving average for both log of house price and marriage age with weights 1/9, 2/9, 3/9, 2/9, 1/9 for $y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2}$ respectively. The same regression using alternative weights (e.g. uniform weights; only middle three years) give similar results. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: Growth of Marriage Age and House Price in the Cross Section

Dep. Var: Growth of Age at Marriage (at county-year level)				
Growth of house price	0.721***	1.017**	0.726***	0.959**
(2010 RMB/square meter)	(0.275)	(0.392)	(0.277)	(0.423)
Year FE	N	N	Y	Y
County FE	N	Y	N	Y
Observations	147	132	147	132
R-squared	0.045	0.302	0.099	0.355

Notes. We construct the raw data according to footnote of figure 9. After those steps, we smooth marriage age and log house price using a 5-year moving average. Then we construct the measure of 'growth rate' as follows: run a regression of house price or marriage age in year $t-2$, $t-1$, t in county i on time (measured in year) and use the regression coefficient as the measure for growth rate of the variable leading up to year t. Then we run regression of the generated growth rate of marriage age on the 3-year lagged generated growth rate of house price. The same regression using alternative smoother or number of periods to construct growth rate or number of periods for lags give similar results. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Growth of Marriage Age and House Price in the Cross Section - Residualized

Dep. Var: Growth of Age at Marriage (at county-year level, residualized)				
Growth of house price	0.919***	1.149**	0.888***	1.068**
(2010 RMB/square meter, residualized)	(0.286)	(0.397)	(0.288)	(0.430)
Year FE	N	N	Y	Y
County FE	N	Y	N	Y
Observations	140	126	139	126
R-squared	0.070	0.338	0.111	0.376

Notes. We construct the raw data according to footnote of figure 9. But before taking the county-year level average of marriage age and house price, we residualize them on individual level income or household level net asset. After those steps, we smooth marriage age and log house price using a 5-year moving average. Then we construct the measure of 'growth rate' as follows: run a regression of house price or marriage age in year $t-2$, $t-1$, t in county i on time (measured in year) and use the regression coefficient as the measure for growth rate of the variable leading up to year t. Then we run regression of the generated growth rate of marriage age on the 3-year lagged generated growth rate of house price. The same regression using alternative smoother or number of periods to construct growth rate or number of periods for lags give similar results. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Summary of Estimated Parameters

Parameter		Target	Data	Model	Value
μ	Marriage premium	2010 male marriage rate	73%	74.9%	0.06
η	Marriage house premium	2010 male ownership rate	62%	61.4%	0.04
κ	Bequest motive	2010 HH net wealth/income	4.6	4.64	0.89

Table 10: Summary of Calibrated Parameters

		Parameter	Value
Preferences	β	Discount factor	0.98
	γ	CRRA	5
Assets	R	Real interest rate	1.026
	ϕ	Maximum loan-to-value ratio	0.55
	τ	Transaction cost	0.05
Marriage	λ	Marriage probability multiplier	1.7

Notes. NBS refers to National Bureau of Statistics; for discount factor we follow the common practice from the literature (Frederick et al., 2002). Age profile of income can be found in footnotes in subsubsection 'Income Profile'.

Figure 11: 4 States of a Household

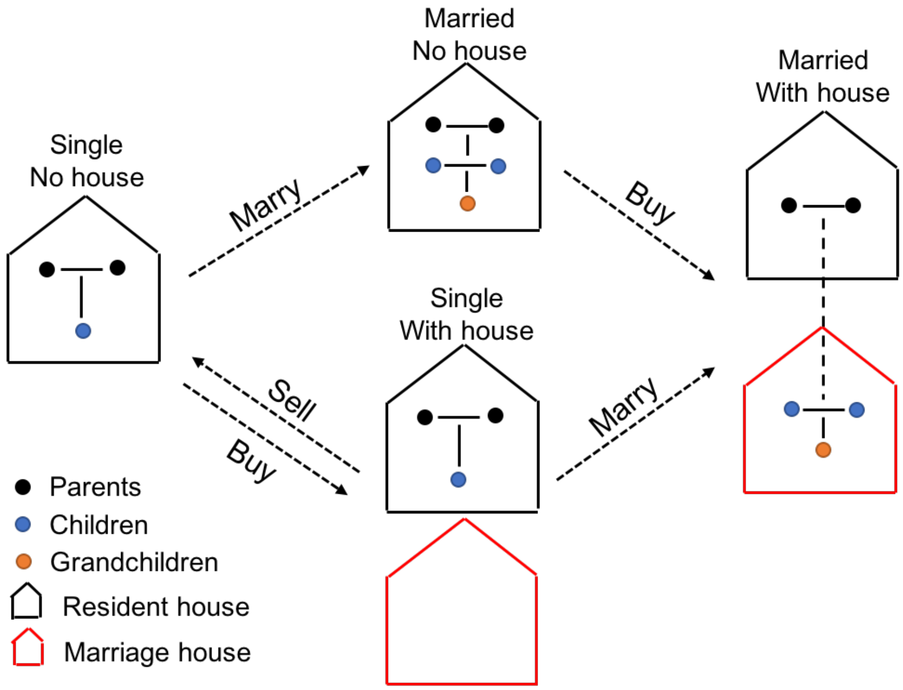
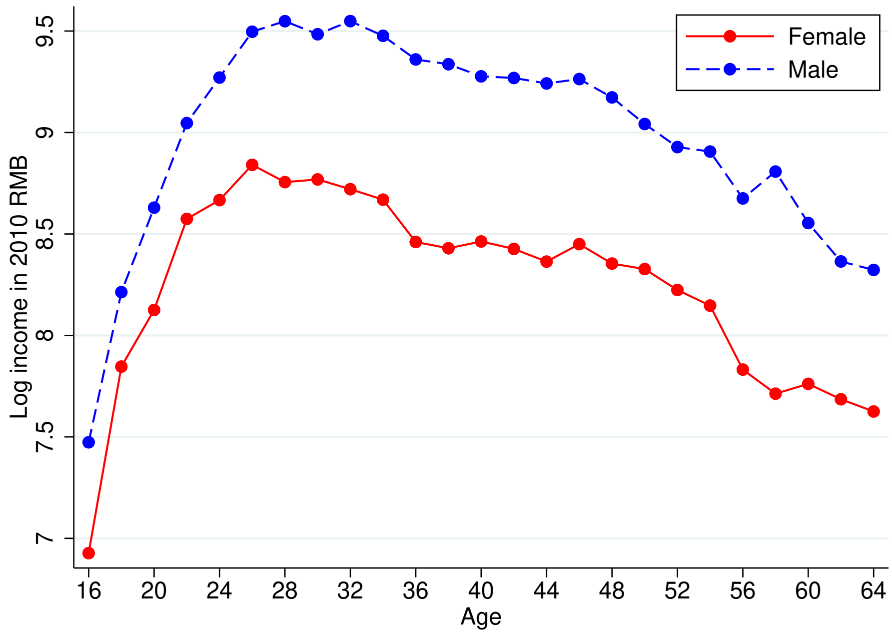
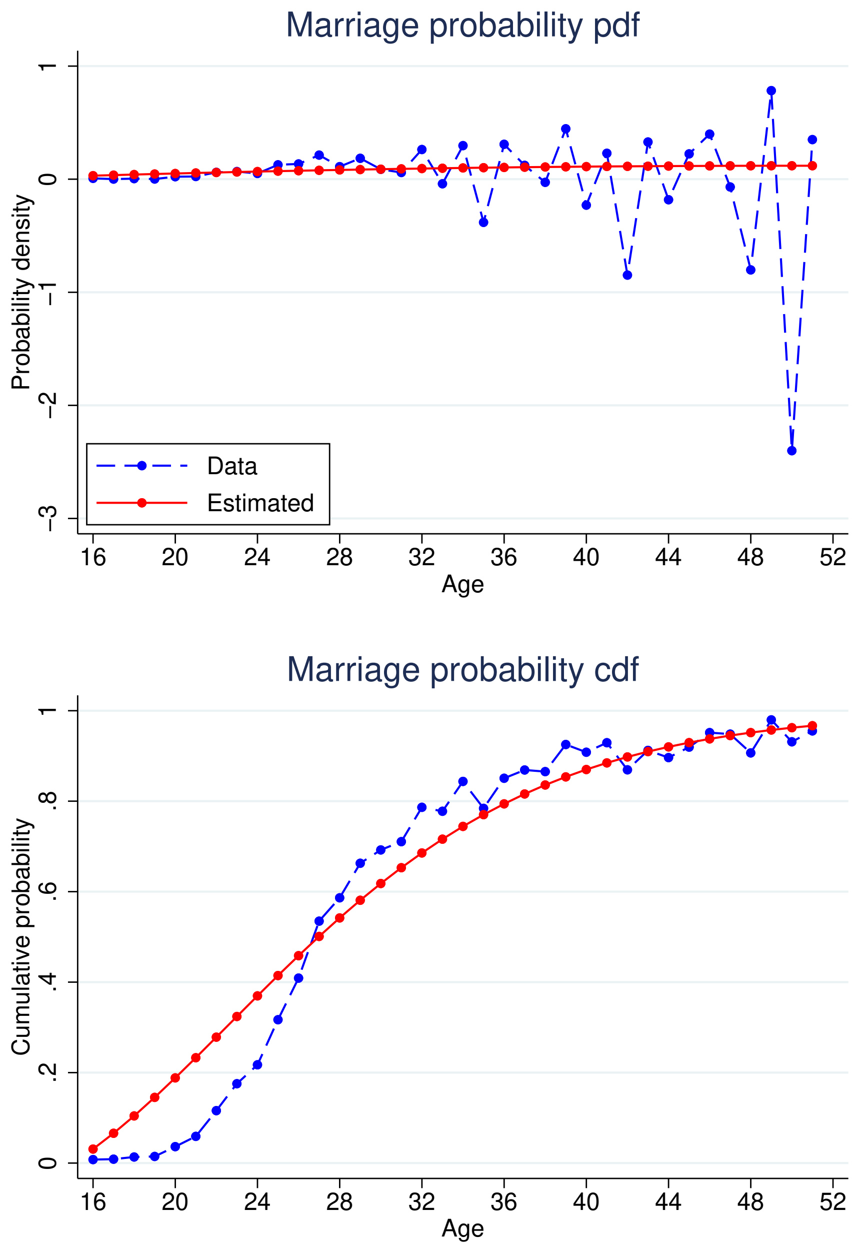


Figure 12: Age Profile of Income - by Gender



Notes. CFPS 2010. The graph plots the log of annual income in 2010 RMB for individuals of a given gender in a given age. Age 16 refers to the average of age 16 and 17; age 20 refers to the average of age 20 and 21, etc.

Figure 13: Marriage Probability by Age, Non-owners, Data vs Estimated



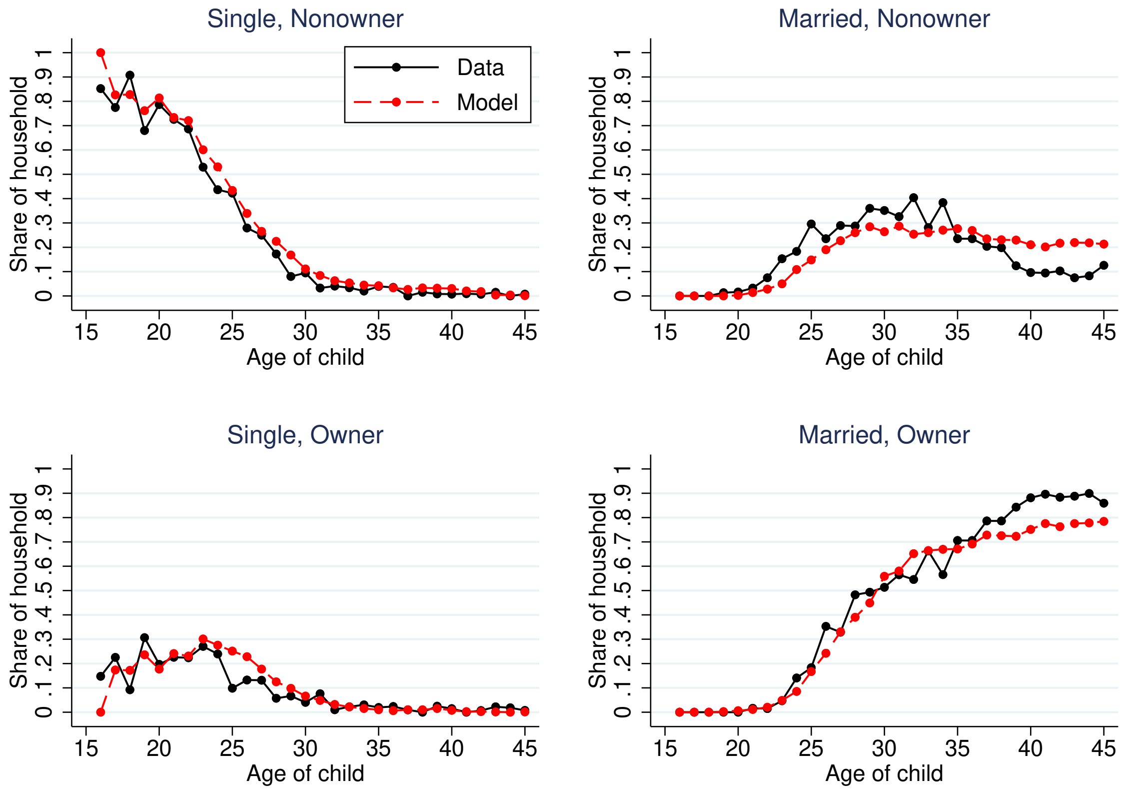
Notes. The graph shows the empirical and the estimated pdf and cdf of marriage probability by age for non-owners, i.e. those whose family does not own marriage house but owns residential house. The data is from CFPS 2010, i.e. it is a cross section. The empirical cdf is simply the share of non-owner males who are married by each age in the data. The empirical pdf is calculated by the Kaplan-Meier estimator, i.e. increment of the share of married individuals in a given year divided by the share of single individuals in the last year. The estimated cdf is from applying a 3-year moving average three times to the empirical cdf. The estimated pdf is from applying the Kaplan-Meier estimator to the estimated cdf.

Table 11: Estimated Marriage Probability by Age

Age	Marriage probability	Age	Marriage probability
16	0.030999	34	0.099321
17	0.03608	35	0.10168
18	0.04101	36	0.10389
19	0.04579	37	0.10594
20	0.050417	38	0.10785
21	0.054894	39	0.1096
22	0.059219	40	0.1112
23	0.063393	41	0.11265
24	0.067415	42	0.11395
25	0.071287	43	0.1151
26	0.075007	44	0.1161
27	0.078576	45	0.11695
28	0.081993	46	0.11764
29	0.085259	47	0.11818
30	0.088374	48	0.11858
31	0.091338	49	0.11882
32	0.09415	50	0.11891
33	0.096811	51	0.11884

Notes.The table reports estimated marriage probability for non-owner males, i.e. those whose family does not own marriage house but owns residential house, in each age. The estimation procedure is described in section Calibration, subsection From the literature or the data, subsubsection Marriage probability.

Figure 14: Age Profile of Marriage and Ownership



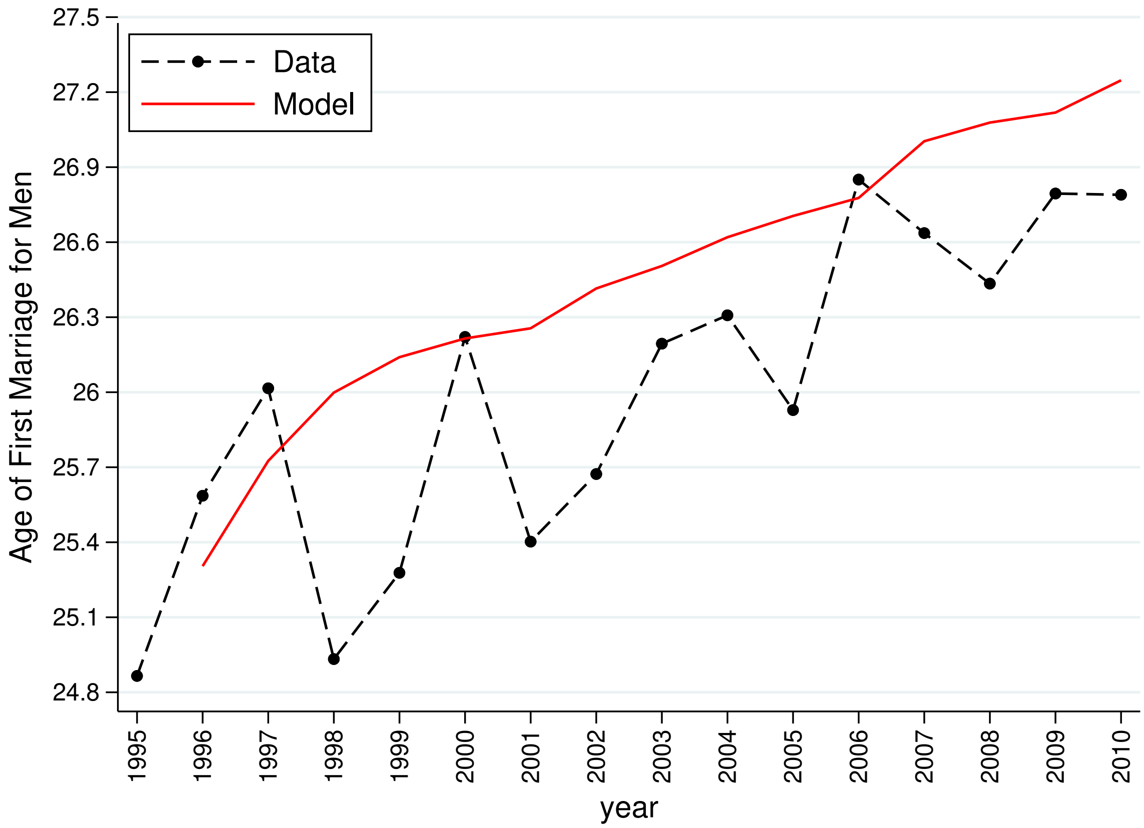
Notes. The 'married' and 'single' refers to that of the child and 'no house' and 'with house' refers to whether his family owns a marriage house in the model. For how the data series are constructed, see section Quantitative Exercise.

Table 12: Categorizing individuals in the data into 4 states

State	Criteria
Single and no house	Individual is single and his parents are present in the household and his household does not own house other than the residential house
Single and with house	Individual is single and his parents are present in the household and his household owns house other than the residential house
Married and no house	Individual is married and his parents are present in the household and his household does not own house other than the residential house
Married and with house	Individual is married and ((his parents are not present in the household and his household owns the residential house) or (his parents are present in the household and his household owns house other than the residential house))

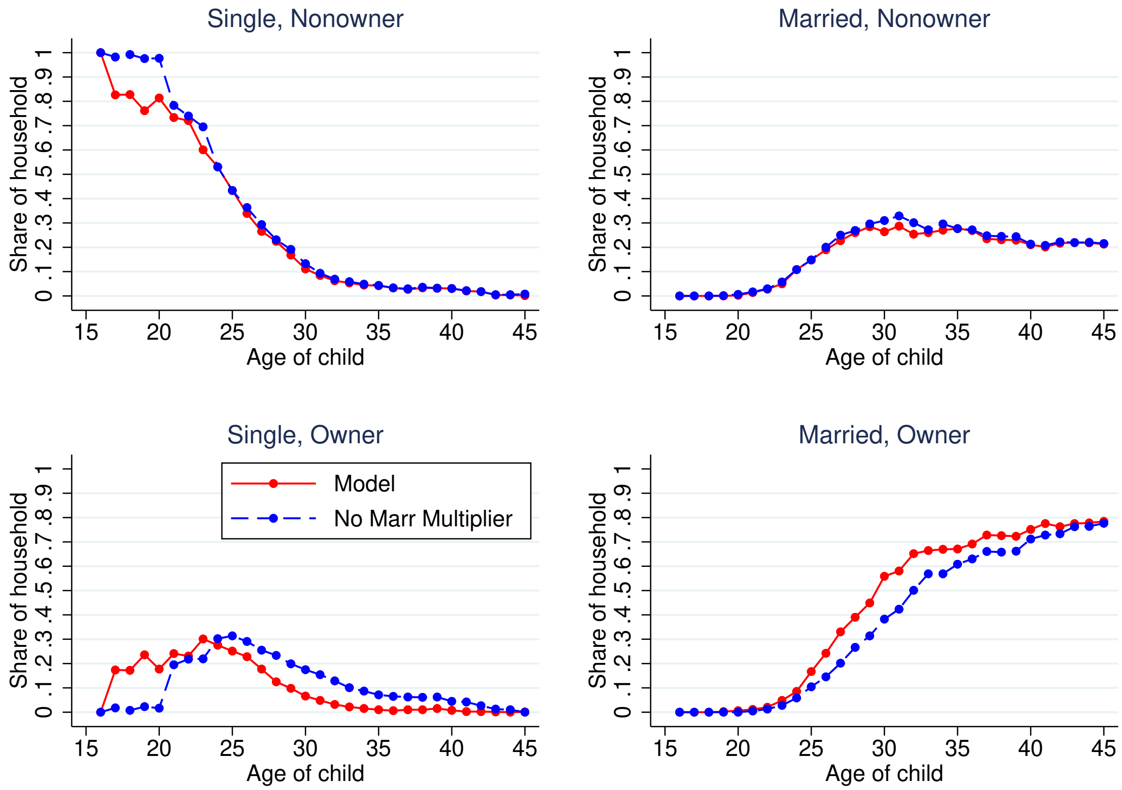
Notes. The table presents the criteria for categorizing individuals in CFPS 2010 into the 4 states that are comparable to those in the model. Explanation can be found in section Quantitative Exercise, subsection Model and Data Moments.

Figure 15: Rise in Marriage Age Over Time (Males)



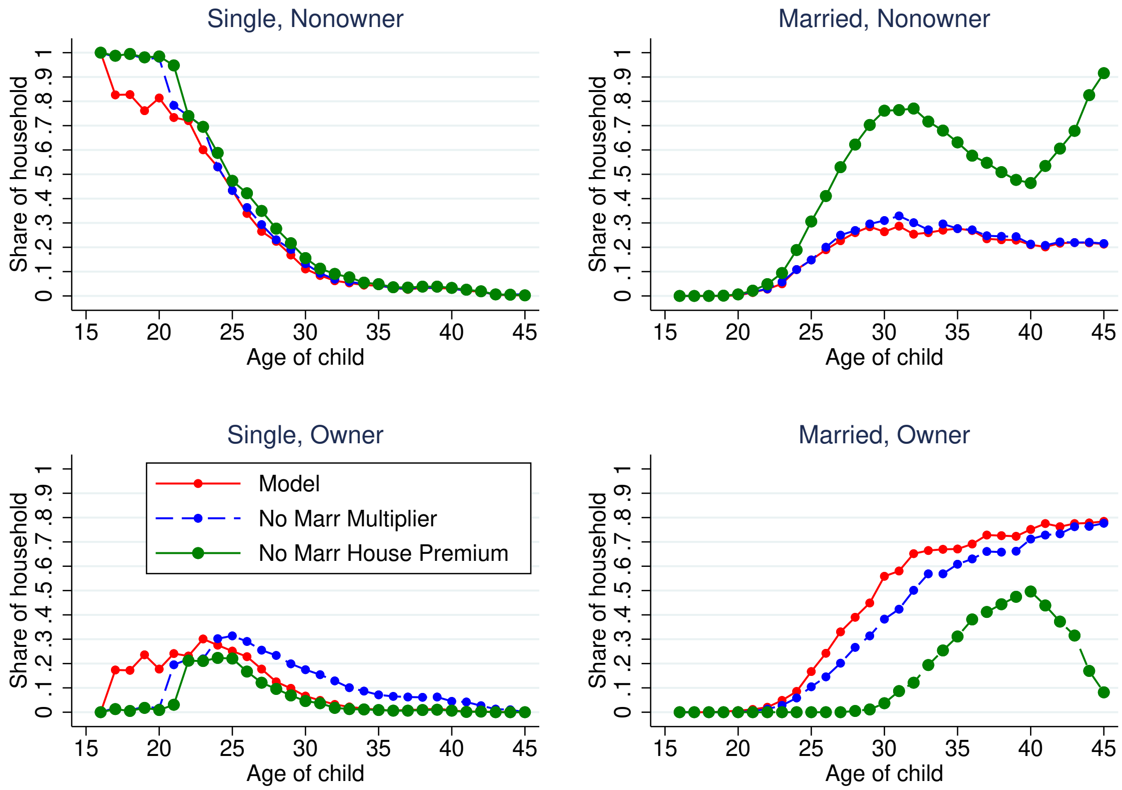
Notes. average age at marriage for males who got married in that year (first marriage only) calculated from CFPS (CFPS asks married individuals which year they got married; combined with their age, we can back out the age at marriage).

Figure 16: Counterfactual: turn off marriage probability multiplier



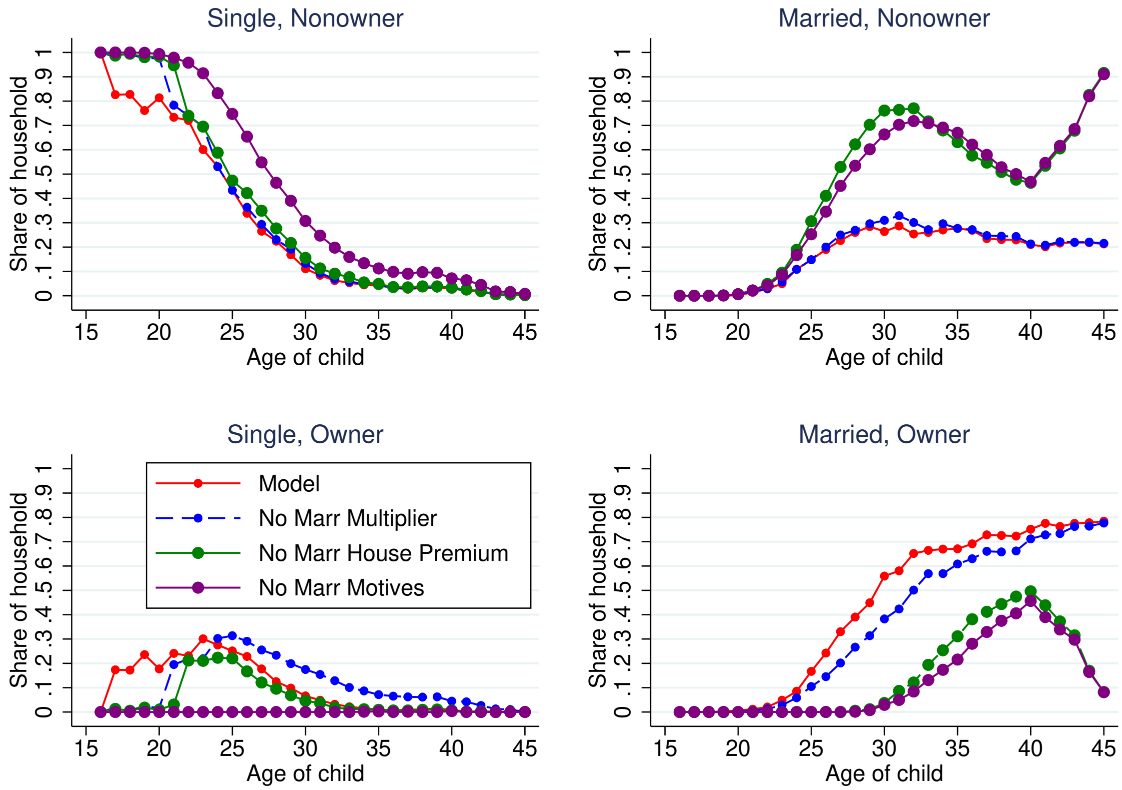
Notes. The baseline model feature both marriage multiplier and marriage premium of homeownership. The counterfactual exercise mutes the marriage multiplier, in the sense that homeownership does not increase the probability of marriage. The marriage premium of homeownership, which refers to convenience and consumption value of the marriage house after marriage, stays in the counterfactual case.

Figure 17: Counterfactual: turn off utility complementarity



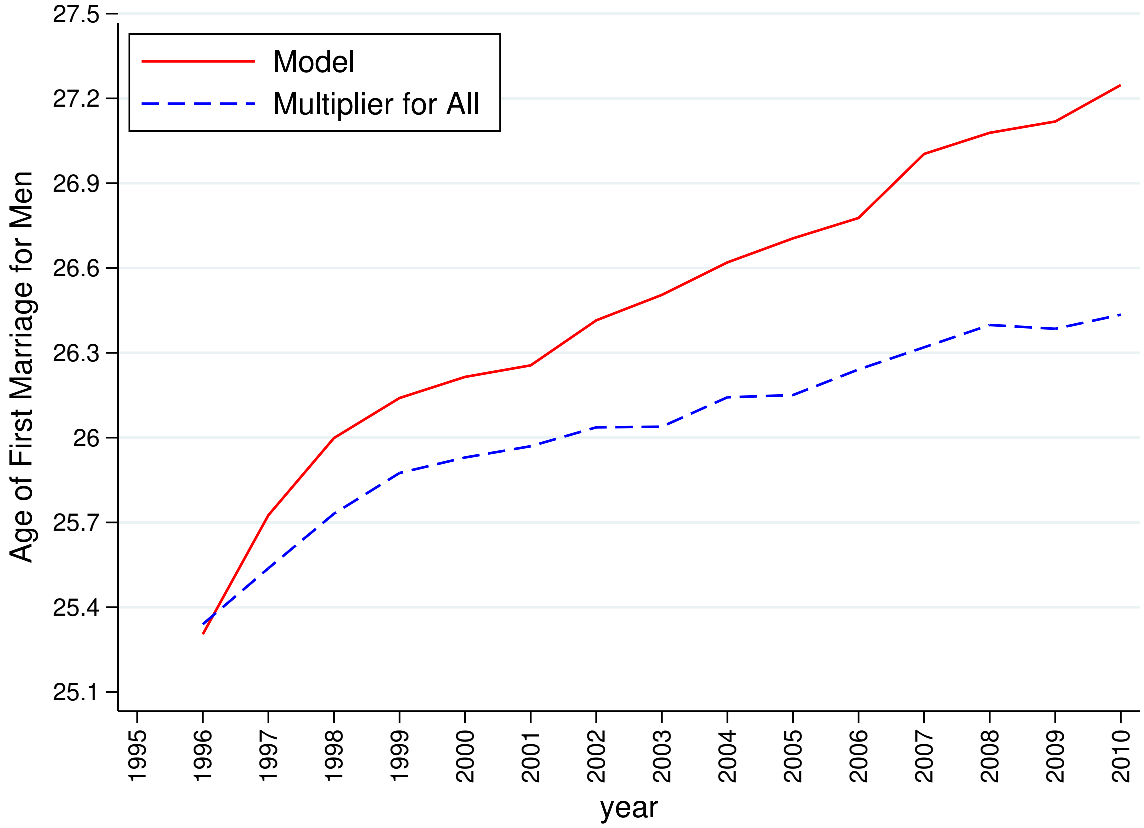
Notes. The baseline model feature both marriage multiplier and marriage premium of homeownership. Two counterfactual cases consider either the marriage multiplier effect or the marriage premium effect of homeownership is muted. The marriage premium of homeownership refers to convenience and consumption value of the marriage house after marriage. Marriage multiplier captures the fact that homeownership increases the probability of marriage.

Figure 18: Counterfactual: turn off both the probability multiplier and the utility complementarity



Notes. The baseline model feature both marriage multiplier and marriage premium of homeownership. Two counterfactual cases then consider either the marriage multiplier effect or the marriage premium effect of homeownership is muted. The marriage premium of homeownership refers to convenience and consumption value of the marriage house after marriage. Marriage multiplier captures the fact that homeownership increases the probability of marriage. In a third counterfactual, both non-pecuniary benefits of homeownership are absent.

Figure 19: Counterfactual: rise in male marriage age



Notes. In the baseline model, marriage probability depends on the homeownership status. Owning a marriage house increases the probability of marriage by 1.7 times, which comes from hazard analysis. In a counterfactual, we remove the marriage-probability multiplier effect of homeownership. The plot illustrates the counterfactual average age at marriage.

References

Andrews, Dan and Aida Caldera Sánchez (2011) "The Evolution of Homeownership Rates in Selected OECD Countries: Demographic and Public Policy Influences," *OECD Journal: Economic Studies*, 2011, 10.1787/19952856.

Attanasio, Orazio, Andrew Leicester, and Matthew Wakefield (2011) "Do house prices drive consumption growth? The coincident cycles of house prices and consumption in the UK," *Journal of the European Economic Association*, 9 (3), 399–435.

Attanasio, Orazio P, Renata Bottazzi, Hamish W Low, Lars Nesheim, and Matthew Wakefield (2012) "Modelling the demand for housing over the life cycle," *Review of Economic Dynamics*, 15 (1), 1–18.

Bar, Michael and Oksana Leukhina (2010) "Demographic transition and industrial revolution: A macroeconomic investigation," *Review of Economic Dynamics*, 13 (2), 424–451.

Benhabib, Jess, Richard Rogerson, and Randall Wright (1991) "Homework in macroeconomics: Household production and aggregate fluctuations," *Journal of Political economy*, 99 (6), 1166–1187.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra (2018) "House prices and consumer spending," *The Review of Economic Studies*, 85 (3), 1502–1542.

Boldrin, Michele and Larry E Jones (2002) "Mortality, fertility, and saving in a Malthusian economy," *Review of Economic Dynamics*, 5 (4), 775–814.

Chambers, Matthew, Carlos Garriga, and Don Schlagenhauf (2009a) "The loan structure and housing tenure decisions in an equilibrium model of mortgage choice," *Review of Economic Dynamics*, 12 (3), 444–468.

Chambers, Matthew, Carlos Garriga, and Don E Schlagenhauf (2009b) "Accounting for changes in the homeownership rate," *International Economic Review*, 50 (3), 677–726.

Chang, Minsu et al. (2019) "A house without a ring: The role of changing marital transitions for housing decisions," in *2019 Meeting Papers* (514), Society for Economic Dynamics.

Chen, Hui, Michael Michaux, and Nikolai Roussanov (2020) "Houses as ATMs: mortgage refinancing and macroeconomic uncertainty," *The Journal of Finance*, 75 (1), 323–375.

Choukhmane, Taha, Nicholas Coeurdacier, and Keyu Jin (2013) "The One-Child Policy and Household Savings,"Technical Report ID 2341051, Social Science Research Network, Rochester, NY, https://papers.ssrn.com/abstract=2341051.

Cilluffo, Anthony, Abigail W Geiger, and Richard Fry (2017) "More US households are renting than at any point in 50 years."

Cocco, João F (2005) "Portfolio Choice in the Presence of Housing," *The Review of Financial Studies*, 18 (2), 535–567, 10.1093/rfs/hhi006.

Cocco, João F, Francisco J Gomes, and Pascal J Maenhout (2005) "Consumption and Portfolio Choice over the Life Cycle," *The Review of Financial Studies*, 18 (2), 491–533, 10.1093/rfs/hhi017.

Cruz, Prince Christian R et al. (2008) "Transaction costs and housing affordability in Asia," *International Real Estate Review*, 11 (1), 128–150.

Cubeddu, Luis and José-Víctor Ríos-Rull (2003) "Families as shocks," *Journal of the European Economic Association*, 1 (2-3), 671–682.

Davidoff, Thomas (2010) "Home equity commitment and long-term care insurance demand," *Journal of Public Economics*, 94 (1), 44–49, 10.1016/j.jpubeco.2009.09.006.

De La Croix, David and Matthias Doepke (2003) "Inequality and growth: why differential fertility matters," *American Economic Review*, 93 (4), 1091–1113.

Dettling, Lisa J and Melissa S Kearney (2014) "House prices and birth rates: The impact of the real estate market on the decision to have a baby," *Journal of Public Economics*, 110, 82–100.

Díaz, Antonia and María José Luengo-Prado (2008) "On the user cost and homeownership," *Review of Economic Dynamics*, 11 (3), 584–613.

Ding, Qu Jian and Therese Hesketh (2006) "Family size, fertility preferences, and sex ratio in China in the era of the one child family policy: results from national family planning and reproductive health survey."

Doepke, Matthias (2005) "Child mortality and fertility decline: Does the Barro-Becker model fit the facts?" *Journal of population Economics*, 18 (2), 337–366.

Doepke, Matthias, Moshe Hazan, and Yishay D Maoz (2015) "The baby boom and World War II: A macroeconomic analysis," *The Review of Economic Studies*, 82 (3), 1031–1073.

Doepke, Matthias and Michèle Tertilt (2016) "Families in macroeconomics," in *Handbook of macroeconomics*, 2, 1789–1891: Elsevier.

Dyrda, Sebastian, Greg Kaplan, and José-Víctor Ríos-Rull (2019) "Living Arrangements and Labor Market Volatility of Young Workers."

Ebenstein, Avraham (2010) "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy," *Journal of Human Resources*, 45 (1), 87–115, 10.3368/jhr.45.1.87.

Fang, Hanming, Quanlin Gu, Wei Xiong, and Li-An Zhou (2016) "Demystifying the Chinese housing boom," *NBER macroeconomics annual*, 30 (1), 105–166.

Fernandez, Raquel, Nezih Guner, and John Knowles (2005) "Love and money: A theoretical and empirical analysis of household sorting and inequality," *The Quarterly Journal of Economics*, 120 (1), 273–344.

Fernández, Raquel and Joyce Wong (2014a) "Divorce risk, wages and working wives: A quantitative life-cycle analysis of female labour force participation," *The Economic Journal*, 124 (576), 319–358.

——— (2014b) "Unilateral divorce, the decreasing gender gap, and married women's labor force participation," *American Economic Review*, 104 (5), 342–347.

Fernandez-Villaverde, Jesus and Dirk Krueger (2011) "Consumption and saving over the life cycle: How important are consumer durables?" *Macroeconomic dynamics*, 15 (5), 725–770.

Fisher, Jonas D. M. and Martin Gervais (2011) "Why Has Home Ownership Fallen Among The Young?," *International Economic Review*, 52 (3), 883–912, 10.1111/j.1468-2354.2011.00653.x.

Flavin, Marjorie and Takashi Yamashita (2002) "Owner-occupied housing and the composition of the household portfolio," *American Economic Review*, 92 (1), 345–362.

FRED, Federal Reserve Bank of St. Louis (2018) "U.S. Bureau of the Census, Homeownership Rate for the United States [RHORUSQ156N]," 7, `https://fred.stlouisfed.org/series/RHORUSQ156N`.

Frederick, Shane, George Loewenstein, and Ted O'donoghue (2002) "Time discounting and time preference: A critical review," *Journal of economic literature*, 40 (2), 351–401.

Galor, Oded and David N Weil (2000) "Population, technology, and growth: From Malthusian stagnation to the demographic transition and beyond," *American economic review*, 90 (4), 806–828.

Glaeser, Edward, Wei Huang, Yueran Ma, and Andrei Shleifer (2017) "A Real Estate Boom with Chinese Characteristics," *Journal of Economic Perspectives—Volume*, 31 (1—Winter), 93–116, 10.1257/jep.31.1.93.

Gorea, Denis and Virgiliu Midrigin (2017) "Liquidity constraints in the us housing market," *NBER Working Paper* (w23345).

Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin (2017) "Marriage, labor supply, and home production," *Econometrica*, 85 (6), 1873–1919.

Greenwood, Jeremy and Nezih Guner (2008) "Marriage and divorce since World War II: Analyzing the role of technological progress on the formation of households," *NBER macroeconomics annual*, 23 (1), 231–276.

Greenwood, Jeremy, Nezih Guner, and Guillaume Vandenbroucke (2017) "Family economics writ large," *Journal of Economic Literature*, 55 (4), 1346–1434.

Greenwood, Jeremy and Zvi Hercowitz (1991) "The allocation of capital and time over the business cycle," *Journal of political Economy*, 99 (6), 1188–1214.

Greenwood, Jeremy and Ananth Seshadri (2002) "The US demographic transition," *American Economic Review*, 92 (2), 153–159.

Greenwood, Jeremy, Ananth Seshadri, and Guillaume Vandenbroucke (2005) "The baby boom and baby bust," *American Economic Review*, 95 (1), 183–207.

Guler, Bulent, Fatih Guvenen, and Giovanni L Violante (2012) "Joint-search theory: New opportunities and new frictions," *Journal of Monetary Economics*, 59 (4), 352–369.

Guvenen, Fatih and Michelle Rendall (2015) "Women's emancipation through education: A macroeconomic analysis," *Review of Economic Dynamics*, 18 (4), 931–956.

Hong, Jay H and Jose-Victor Rios-Rull (2012) "Life insurance and household consumption," *American Economic Review*, 102 (7), 3701–3730.

Hurst, Erik and Frank Stafford (2004) "Home is where the equity is: Mortgage refinancing and household consumption," *Journal of Money, credit and Banking*, 985–1014.

Jaimovich, Nir, Seth Pruitt, and Henry E Siu (2013) "The demand for youth: Explaining age differences in the volatility of hours," *American Economic Review*, 103 (7), 3022–3044.

Jones, Larry E, Rodolfo E Manuelli, and Ellen R McGrattan (2015) "Why are married women working so much?" *Journal of Demographic Economics*, 81 (1), 75–114.

Kalemli-Ozcan, Sebnem (2003) "A stochastic model of mortality, fertility, and human capital investment," *Journal of Development Economics*, 70 (1), 103–118.

Kaplan, Greg, Kurt Mitman, and Giovanni L Violante (2017) "The housing boom and bust: Model meets evidence,"Technical report, National Bureau of Economic Research.

Landvoigt, Tim (2017) "Housing Demand During the Boom: The Role of Expectations and Credit Constraints," *The Review of Financial Studies*, 30 (6), 1865–1902, 10.1093/rfs/hhx026.

Landvoigt, Tim, Monika Piazzesi, and Martin Schneider (2015) "The housing market (s) of San Diego," *American Economic Review*, 105 (4), 1371–1407.

Li, Wenli and Rui Yao (2007) "The life-cycle effects of house price changes," *Journal of Money, Credit and banking*, 39 (6), 1375–1409.

Liu, Chang and Wei Xiong (2018) "China's Real Estate Market,"Technical report, National Bureau of Economic Research.

Love, David A (2010) "The effects of marital status and children on savings and portfolio choice," *The Review of Financial Studies*, 23 (1), 385–432.

Lustig, Hanno N and Stijn G Van Nieuwerburgh (2005) "Housing collateral, consumption insurance, and risk premia: An empirical perspective," *The Journal of Finance*, 60 (3), 1167–1219.

Lustig, Hanno and Stijn Van Nieuwerburgh (2010) "How much does household collateral constrain regional risk sharing?" *Review of Economic Dynamics*, 13 (2), 265–294.

Ma, Zunbo, Luyao Ma, Zhiyi Yang, Yibo Wang, and Zige Wang (2023) "The Influence of House Price on The Average Age of First Marriage in China," *Frontiers in Business, Economics and Management*, 8 (3), 126–133.

Mian, Atif and Amir Sufi (2011) "House prices, home equity–based borrowing, and the us household leverage crisis," *American Economic Review*, 101 (5), 2132–2156.

Mortensen, Dale T (1988) "Matching: Finding a partner for life or otherwise," *American journal of sociology*, 94, S215–S240.

Peter, Alessandra, Martin Schneider, and Monika Piazzesi (2020) "Houses and families across countries,"Technical report, Mimeo, columbia university.

Piazzesi, Monika and Martin Schneider (2016) "Housing and macroeconomics," *Handbook of macroeconomics*, 2, 1547–1640.

Regalia, Ferdinando and Jose-Victor Rios-Rull (2001) "What accounts for the increase in the number of single households?" *University of Pennsylvania, mimeo*.

Su, Chi-Wei, Khalid Khan, Lin-Na Hao, Ran Tao, and Adelina Dumitrescu Peculea (2020) "Do house prices squeeze marriages in China?" *Economic research-Ekonomska istraživanja*, 33 (1), 1419–1440.

Voena, Alessandra (2015) "Yours, mine, and ours: Do divorce laws affect the intertemporal behavior of married couples?" *American Economic Review*, 105 (8), 2295–2332.

Wei, Shang-Jin and Xiaobo Zhang (2011) "The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China," *Journal of Political Economy*, 119 (3), 511–564, 10.1086/660887.

Wei, Shang-Jin, Xiaobo Zhang, and Yin Liu (2017) "Home ownership as status competition: Some theory and evidence," *Journal of Development Economics*, 127, 169–186.

Wrenn, Douglas H, Junjian Yi, and Bo Zhang (2019) "House prices and marriage entry in China," *Regional Science and Urban Economics*, 74, 118–130.

Yang, Fang (2009) "Consumption over the life cycle: How different is housing?" *Review of Economic Dynamics*, 12 (3), 423–443.

Yanyun Man, Joyce (2011) *China's Housing Reform and Outcomes*, `https://www.lincolninst.edu/sites/default/files/pubfiles/china-housing-reform-and-outcomes-chp.pdf`.

Yao, Rui and Harold H Zhang (2005) "Optimal consumption and portfolio choices with risky housing and borrowing constraints," *The Review of Financial Studies*, 18 (1), 197–239.

Appendices

A Data Construction

A.1 Dataset for County-Year Level Analysis

The first piece of empirical evidence we show is that in county-years where house prices are higher/rising, marriage age is also higher/rising. Since CFPS is only available for 2010, 12, 14, 16, if we want to get data on marriage age and house price in the 1990s or 2000s, we need to rely on recall data. Conveniently, CFPS asks for the year in which individuals got married for each of their marriage, and ask for the year in which the house is purchased for the residential house owned by a household. From these, we can get historical data on house price and marriage age in a given county in a given year.

For each individual, we calculate the age at first marriage using the year of marriage and their birth year. For each house, we calculate the house price at purchase as the value at purchase divided by total area (in square meter). Then we deflate it to 2010 RMB using the CPI provided by National Bureau of Statistics. For the residualized version of the analysis, we residualize marriage age on individual annual income and residualize house price on household annual income (because homeownership can only be identified at household level), i.e. run a regression of marriage age and house price on the aforementioned right-hand-side variables, and retrieve the residual from the regression. Then we calculate the log of house price.

Then we calculate the average of log of house price in each county in each year (the year is the year in which the house was purchased) and the average of marriage age in each county in each year (the year is the year in which the marriage happened) (weighted by person and family national weight in CFPS) and merge the county-year level house price to the county-year level marriage age.

A.2 Dataset for Marriage Year vs House Purchase Year

The second piece of the empirical evidence is that marriage tends to happen after house purchase. We use the same variable as mentioned before: recalled marriage year, and recalled house purchase year.

We would like to ask the question whether a couple tends to first get married and then buy house, or first buy house and then get married. The issue with recall data is that we do not know whether the house a married couple lives in at the time of the survey is their house, or their parents' house, or their kids' house, or whether it is the house they were living in around the time of their marriage. To address these issues, we adopt the following sample selection criteria: we only keep married couples who do not live with ones' parents, do not have kids yet, and own their residential house [39]. We drop couples with children because people might switch to a bigger house or switch to houses closer to good school districts after they have children so the house they live in after children are born may not be the house they were living in at the time of marriage (or the first house they bought after marriage). We drop

[39] CFPS provides a family roster file in which there are variables indicating whether someone's parents live in the same unit as the person; it also shows the information on all of someone's children.

those cohabiting with parents because we cannot tell whether the house they live in belongs to them or the parents.

After the above selection, we have a dataset on married couples with their marriage year and their residential house purchase year. Then, we can calculate the difference between these two as the marriage year minus the house purchase year, and plot the density of this difference. For all the males in the above dataset, we can also construct a panel dataset of marital status and ownership status: for example, if person X got married in 2012, and the residential house he lived in at the time of the marriage was bought in 2010, I know that his marital status is 0 (i.e. single) for all the years before 2012 and 1 from 2012 onwards, and his ownership status is 0 (i.e. doesn't own) for all the years before 2010 and 1 from 2010 onwards. Thus, with this dataset, we can calculate the mean marital status across all individuals in each of the following year: 10 years before house purchase, 9 years before house purchase ... in the year of house purchase, 1 year after house purchase, 10 years after house purchase, etc.

A.3 Dataset for Hazard Analysis

For each individual, CFPS asks for the marital status in each survey year. For each household, it asks for ownership status of the residential house and other house(s). Thus, we can construct a panel dataset of male individuals' marital status and his family's ownership status of other house.

We use the following variables (directly provided by CFPS) as controls: urban dummy, years of completed education, individual annual income (all sources), and household net asset [40]. We deflate income and asset to 2010 RMB, divide by 10,000 [41], and then take logs. We also control for age polynomial up to the 4th order (taken over individual's age in 2010).

B Homeownership

B.1 Overall Homeownership Pattern

Homeownership rate is high in China but the exact statistic varies based on the source. According to Chinese Academy of Social Sciences, by the end of 2013, homeownership rate is 93.5% in China. According to Large-Sample Urban Household Survey, homeownership rate is 82.3% in 2007 and 84.3% in 2010 (Yanyun Man, 2011). This is high by international standards. The same statistic is 69.2% in US in the same quarter and 62.9% in q2 2016 (FRED, 2018). For OECD countries, in 2016, the median is Slovenia 75.1%, with 10th and 90th percentile being Austria 55% and Croatia 90%. If the statistic of China were to be taken seriously, it will rank among the highest worldwide.

Using the micro level data available to us, we cannot compute homeownership rate according to its official definition, which is the share of owner-occupied dwellings out of all dwellings, but we can provide a related statistic, which is the share of individuals or the share of households who live in a owner-occupied dwelling out of all individuals or households.

Figure F.23 shows this statistic by year from CFPS. The definition of all the categories from

[40] it includes all owned housing

[41] Unlike in the US where income and asset are often measured in thousands, 10,000 is the most common for Chinese Yuan.

top to bottom are: don't know; other; live in housing provided by and owned by one's employer; live in housing provided by and owned by the government; jointly own the housing with either employer or the government; live in housing provided by and owned by relatives or friends; rent; and own. As one can see, the share of people who live in owner-occupied housing is very high, around 85% throughout the 2010s, either in terms of individuals or households. Compared to owner-occupied housing, none of the other arrangements is comparable in terms of magnitude. In many OECD countries or US, renting is the biggest alternative to owning. For example, 36.6% of the population in US live in a rented dwelling in 2016 (Cilluffo, Geiger and Fry, 2017). In China, this number is around 5% around the same period. Thus, we abstract away from the rental market in modeling the Chinese housing market, and restrict agents to live in owner-occupied housing only.

B.2 Heterogeneity

The above statistics paint an overall picture of homeownership situation in China: either measured by share of total dwelling (taking the official statistic seriously) or share of people, living in one's owned housing is the most common arrangement compared to all other (including renting) by a large margin. Now we show some heterogeneity underneath this overall pattern: first, how common it is to own more than one house, and second, whether wealth, income, or education is systematically correlated with owning or owning more houses.

In figure F.24, we show the number of households by number of housing units owned from CFPS. CFPS provides the first three names on the deed of the residential house of a household, but for houses owned by the household other than the residential house, CFPS does not ask for which member of the household owns them. In CHFS, we do not even observe any house ownership information at the individual level. Thus, we cannot provide statistic on the share of individual by number of house owned, but can do that at the household level. As one can see from the figure, around 75% of households own 1 house; there are just above 10% of households who do not own any house; 10% of households own 2 houses, and those who own more than 2 are almost negligible. Thus, we group all households who own 2 or more and refer to them as 'own other house' or 'own 2+'.

C Marriage

C.1 Overall Marriage Pattern

Figure F.25 shows the share of individuals by marital status by age for the 4 waves in CFPS. The red solid line refers to people who are married and are married for the first time; the dashed blue line refers to people who are single and have never married; the dotted green line refers to all other (divorced, seperated, cohabitating, widowed, other). As one can see, from 16 to 40, almost everyone moves from being single and have never married to being married for the first time and have not divorced and there are very few people who fall into other categories. Two features stand out: first, the share of people who remain in their first marriage for the life time is very high (around 90%); second, marriage rate is very high and happens in a short period of the life time (16-40). Thus, in our model, we abstract away from divorce, assuming that agents can marry once and will stay in the same marriage for their life time, and also agents only have positive probability of getting married conditional on being single between 16 and 40. In empirical analysis, we keep only the group who are either single and have never married, or married for the first time and have never divorced. We refer to the first group as single (m=0) and the second group as married (m=1).

C.2 Heterogeneity

Does the above pattern look different for different demographic groups? We present it by rural vs urban, gender, education attainment, and income quartiles in figure F.26. There is a small rural-urban gap, a big gender gap, and a very big education gap. Rural, female, and lower education groups get married earlier than their counterparts. For income, individuals in the poorest quartile get married earlier, whereas effect of the other 3 quartiles are not significant. Thus, in our hazard analysis on whether home ownership affects marriage, we control for all the above factors (we use only urban population for our analysis since the housing arrangement in rural vs urban China is quite different).

D Fertility

D.1 Timing of Child Birth and Marriage

One of the assumptions in our model is that agents have a child in the same year as they get married. In figure F.27 we present a histogram over the difference between first child birth year and marriage year. We use CFPS 2010 data, and keep only agents who have never divorced and are in their first marriage at the time of the survey. From the family roster file, we are able to observe the birth year of each of the children an individual ever had. We keep only information on biological children, and use the year in which the first child was born for each agent. We merge this with the individual level dataset which contains information on marriage history, and generate a number which is the year of the first child birth minus the year of first marriage. This gives us 22k observations.

84% and 90% of individuals have their first child within 3 and 5 years of their first marriage respectively, which confirms that a model in which agents have kids soon after marriage is a good approximation of the real world. For the mass to the left of zero, we have two interpretations: 1) in China, one can only register one's child with a marriage certificate; in other words, technically out-of-the-wedlock fertility is allowed, but such children can only get *hukou*, i.e. offially registered as a citizen after the parents get married; 2) there are errors in either the marriage history or the birth year of the child or whether the child is biological; thus, there might be children who were born in previous marriages but are included in the graph.

D.2 Number of Kids Per Person

It is well known that China had a 'One-Child Policy' from 1979 to 2015: urban citizens can have one child; rural citizens can have one if the first is a male, and two if the first is a daughter; some ethnic-minority groups are exempted; if a pair of parents have more kids than the policy allows, they need to pay a fine, whose rate is pegged to the local living standards; government employees will lose their jobs if they violate this rule, etc.

Given that for the period of time under our interest (late 1990s to late 2010s), most households with children had been subject to this law, our assumption that each couple has one child is not a crazy assumption. However, it is still important to check how far this is from reality. Figure F.28 shows the share of individuals by number of kids they have by rural vs urban, using all years of CFPS. We use the CFPS family roster file to construct the data. For each individual covered by the survey, the family roster file provides information on all children they ever had in their life, biological or adopted; note that only 2010 data provides whether each child is biological or adopted to the parent; this is an important reason why individuals in 2010 CFPS have systematically more children than in other years.

Overall, the pattern agrees with what 'One-Child Policy' would predict: in rural areas, it is most common to have two children (45% of individuals) and in urban areas, it is most common to have one child (50% of individuals). There are 30% of urban individuals who have 2 children, and 15% have 3 or above.

However, this overall pattern conceals important heterogeneity. In figure F.29, we show number of children per person by the parents' birth cohort (CFPS 2010 urban individuals only). As we can see, there is a very strong cohort effect: for people born roughly before the establishment of the PRC (1949), they have significantly more kids than any other group; for people born roughly after the Chinese Economic Reform (1978), they have significantly less kids than any other group (of course this is also because for many people in this cohort they have not completed their fertility life yet). For this youngest cohort, 65% of them have one child, and for the rest, half have no child and have have 2; those who have 3 or more are almost negligible.

Given that in this model, the effect of a marriage house on probability of marriage is mostly on single males only, a more important metric is number of male children a couple has. In other words, even if many have two children, as long as the set of couples with two male children are low, our model is still a good approximation to their problem. In figure F.30, we show the share of individuals by the combination of kids they have: no kid; 1 male; 1 female; a male and a female; 2 males; 2 females; 3 and more. We again see a similar cohort effect as discussed above: the oldest group have significantly more children than any other group. For the youngest cohort, the the share of individuals who have 2 or more males is around 5%, almost negligible.

Given we are interested in the group of parents whose children are on the marriage market after the privatization of the housing market (late 1990s onwards), the fifth group (birth year 1976-94) who mostly have one male child are in fact exactly the group we are interested in. Thus, we confirm that the assumption of one child per parent is close enough to reality.

E Move-out, Marriage, and Marriage House

Other than making marriage more likely, another key feature of the model is that the existence of a marriage house enables the newlywed to move out: intuitively, absent of a rental market, if a single male gets married, he and his wife would need to move into his parents' house if his family does not own other house; however, if a marriage house exists at the time of the marriage, the newlywed can move into the marriage house, thus avoiding living with the male's parents.

E.1 The age profile of Cohabitation

How common is it to live with one's parents? When do children usually move out from parents' family? In figure F.31, we present the age profile of cohabitation pattern using CHFS data. 'prt' refers to parent; m and s refer to married and single respectively. Here cohabitation specifically refers to two generations living together. The reason why we use CHFS instead of CFPS is that CFPS does not identify household heads in a household, whereas CHFS does. Without information on who are the household heads in a dwelling, we cannot tell apart the following two types of families: old-age parents living in children's house so that they can get terminal care from their children; adult children living in their parents house because they have not moved out and formed new households yet. What we are interested in is the latter, i.e. during what life stage do youngsters live with their parents and when do they usually move out and form new households.

To plot the above two graphs, we first keep only individuals age 16 to 35, and then categorize them into 4 groups: single or married X whether they are children of household heads of the household that they belong to. Then we plot the share of each of the 4 categories by age. As we can see, throughout all ages, the share of people who are single and living alone is very low. In other words, in China, it is very rare to have moved out from parents' households before marriage. This supports one feature of our model, which is that even if a household owns both a residential house and a marriage house, the single child lives together with his parents until he gets married.

The overall pattern over age is that individuals transition from being single and living with parents to bring married and living with parents and then to being married and living without parents. In other words, if move-out happens at all, it happens after the marriage. The share of people who are married and living without parents is monotonically increasing over the life cycle, and the share of people who are married and living with parents first increase and then decline, indicating that there are many agents who choose to cohabit with their parents for a while after marriage. This supports another feature of our model: without a marriage house, the newlywed moves into the parents' residential house, and only move out when a marriage house is purchased.

E.2 Marriage House and New Household Formation

We test this against the data again using a hazard analysis. In table 13 we present the result. Again, we pool the 4 waves of CFPS together to create an individual level panel dataset. Thanks to the existence of a family roster file, we can observe, for each individual, whether his or her father and mother is part of his or her household in a given year. Thus, we can define a 'move-out' indicator for a given person in a given wave as his or her father or mother being in the household in the previous wave but neither the father nor the mother being in the household in the current wave. In other words, he or she used to be in the same household with his or her father or mother the last time we observe him or her, but in the next wave he or she is in a household with neither parent. If a parent passed away, it does not count as the child moving out. We also drop individuals who experienced multiple move-outs during the period we observe them.

Other controls are the same as the previous sub-section: individual level years of education, log of personal annual income, log of family asset including financial asset and housing; both income and asset are in 2010 Chinese Yuan. Note that we also control for a polynomial of age up to the 4th order, whose coefficients are not reported.

We regress the dummy 'move-out' on a dummy that indicates one's marital status (m=1 being married; m=0 being single; similar as above, only those never married or have married only once are kept), a dummy that indicates whether the individual's family owns other house, and an interaction term between the two. As we can see, marriage strongly predicts move-out; conditional on being single, the existence of a marriage house does not predict move-out; however, the interaction term between the two is significant for males, which we interpret as consistent with our hypothesis discussed at the beginning of this sub-section, that for males, the existence of a marriage house enable him and his wife to move out from his parents' house. The effect is also positive for females but not significant.

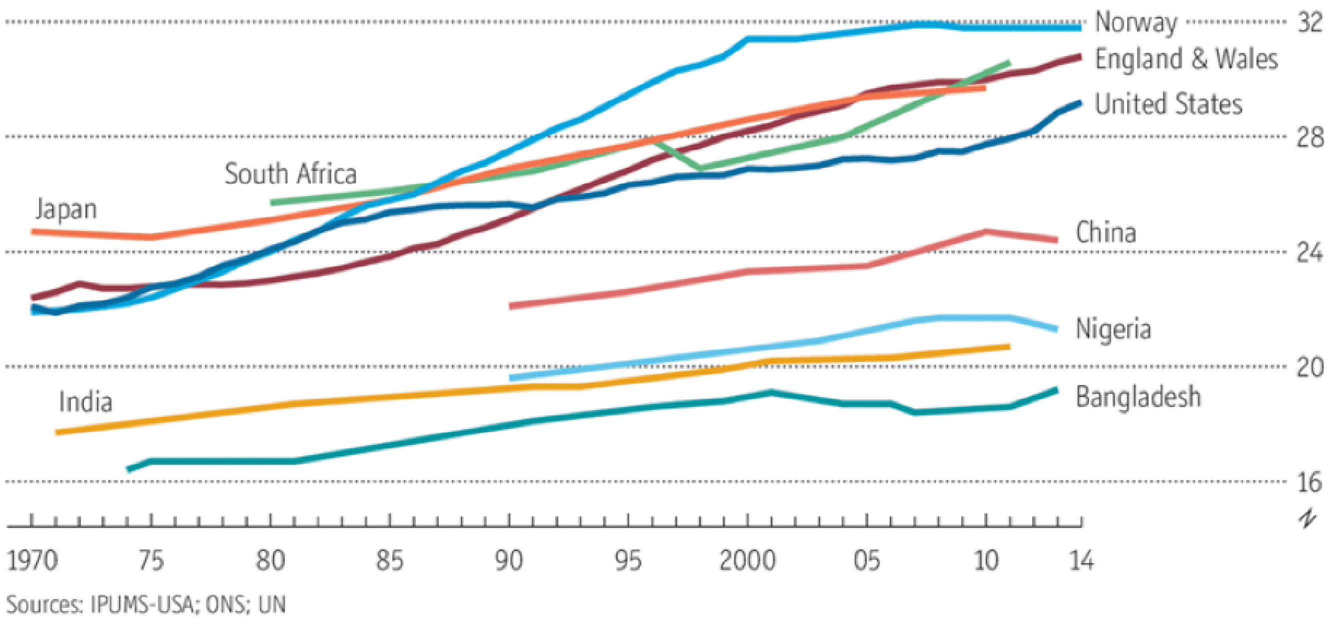
Other than the above key result, both higher education and income predict faster move-out, whereas higher family asset predicts slower move-out, consistent with the heuristics that move-out requires a certain level of financial independence.

F Portfolio Share

Another assumption in the model is that there are only two assets available: the risk free bond and the housing. Given the large amount of literature on housing in the portfolio share, especially on that housing crowds out stock holding, it is worthwhile to look at how important savings, housing, stocks and other financial assets are in the portfolio of Chinese households. For this exercise we use CHFS. CHFS is a survey specifically focused on household finance, and thus provides much more detailed data on household portfolio composition than all other datasets.

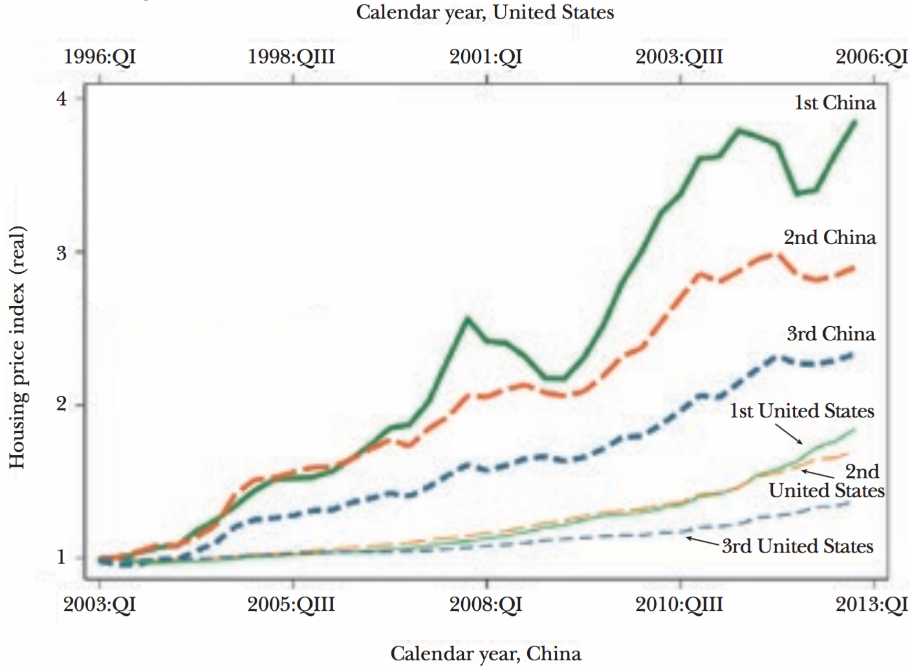
In figure F.32 we present the portfolio share of households by asset quintile in 2011 CHFS data. Note that total asset is defined as the sum of housing asset, car and other household appliances, lending to other households, and financial assets. We categorize different financial assets into two types: safe and risky. Safe asset includes cash, deposit (current and fixed), cash holdings in the household's stock market account, bonds (government and corporate), holdings in mutual funds, and gold. Risky asset includes stock, futures, warrants, other financial derivatives, financial product (provided by bank or other institutes), and asset in other currencies. As we can see, housing takes the largest share throughout all quintiles, and is weakly increasing by quintiles. The next two most important are safe assets and cars and other household appliances, whose share are both decreasing over asset quintiles. The share of risky asset is increasing over asset quintile, but its share, compared to the previous three categories, is almost negligible. Thus, we conclude that the abstraction from risky asset in the model other than housing is a reasonable depiction of the real world.

Figure F.20: Average Age at First Marriage - Global



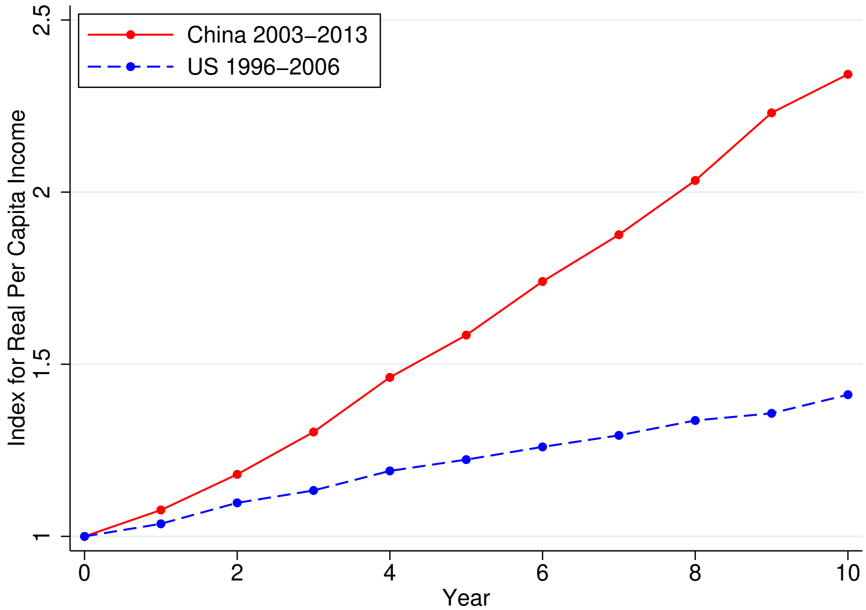
Notes. Special Report. (2017, Nov 25). The State of Marriage as an Institution. Retrieved October 23, 2018, from https://www.economist.com/special-report/2017/11/25/the-state-of-marriage-as-an-institution

Figure F.21: Real House Price Index, China vs US



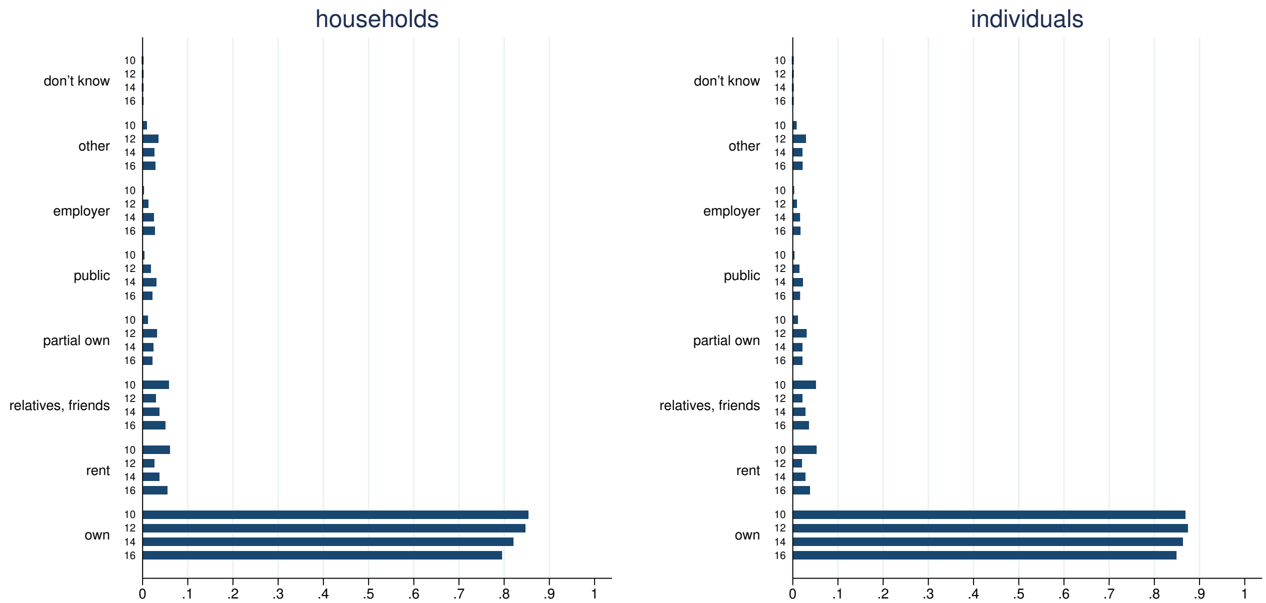
Notes. Figure 1 in Glaeser et al. (2017).

Figure F.22: Real Per Capita Income Index, China vs US



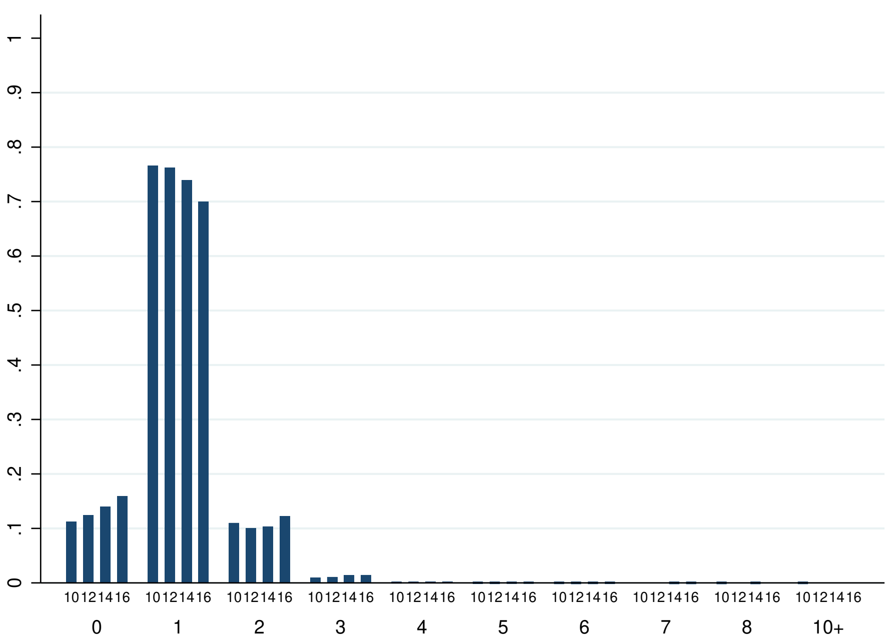
Notes. For the US time series, we retrieve this from FRED (U.S. Bureau of Economic Analysis, Real Disposable Personal Income [DSPIC96], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/DSPIC96, October 22, 2018). For the China time series, we retrieve urban household disposable income and urban CPI from the National Bureau of Statistics of China.

Figure F.23: Share of people by housing arrangement



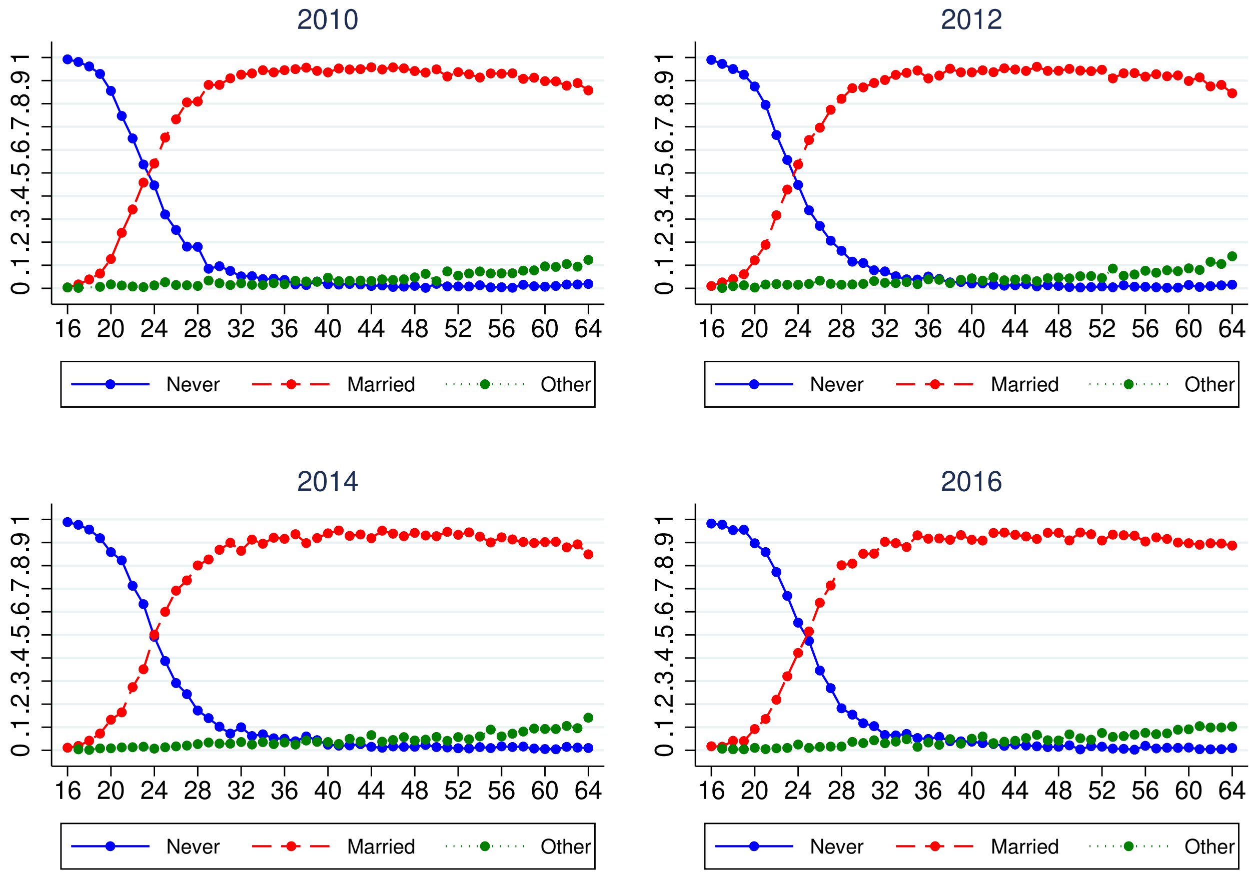
Notes. CFPS 10, 12, 13, 14 residential arrangement.

Figure F.24: Share of households by number of owned housing units



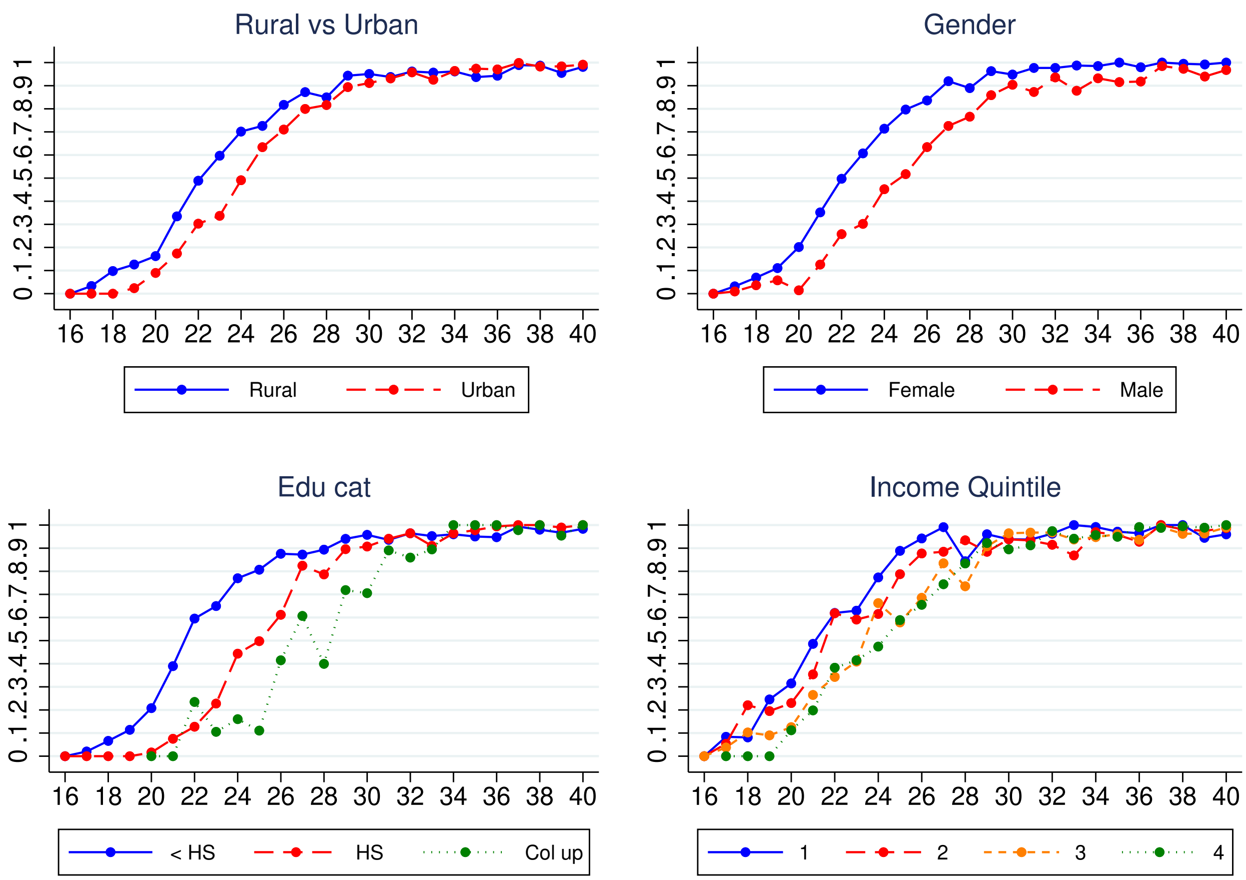
Notes. CFPS 10, 12, 13, 14.

Figure F.25: Share of individuals by marital status



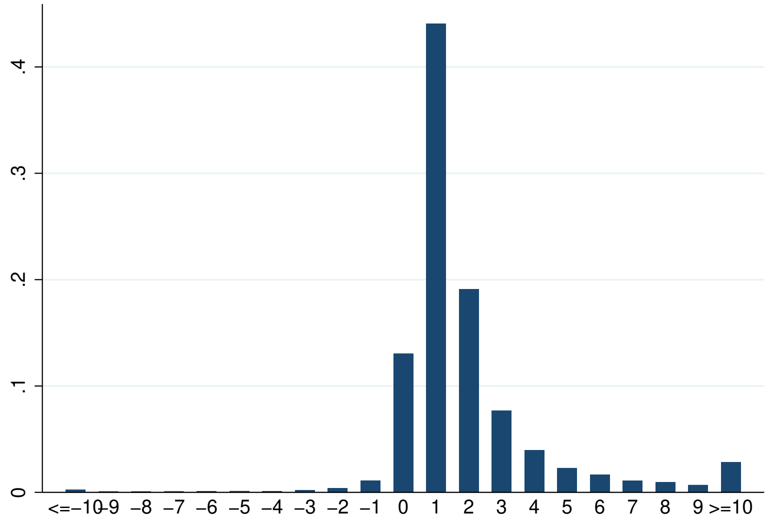
Notes. CFPS 10, 12, 13, 14. Never refers to never married; married refers to married for the first time; other refers to all other categories.

Figure F.26: Age Profile of Marital Status by Demographics



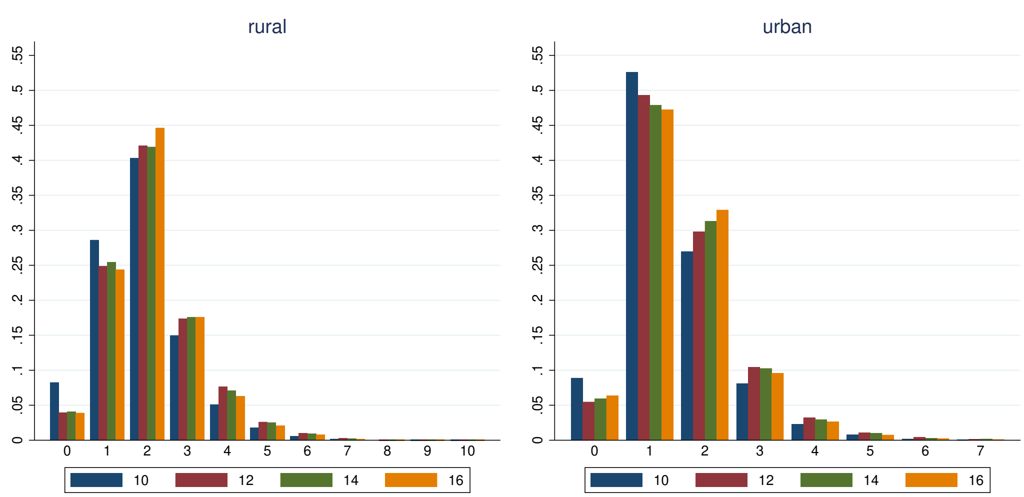
Notes. CFPS 10

Figure F.27: Year of First Child Birth Minus Year of Marriage



Notes. CFPS 10. The family roster file provides the year of birth for the first ten children for all individuals in CFPS. The CFPS individual dataset provides year of marriage. From these two we can calculate the difference between the year of marriage and the year of first child-birth for each individual.

Figure F.28: Number of Kids Per Person



Notes. CFPS 10, 12, 14, 16. The family roster file provides the year of birth for the first ten children for all individuals in CFPS.

Figure F.29: Number of Kids Per Person, by Birth Cohort

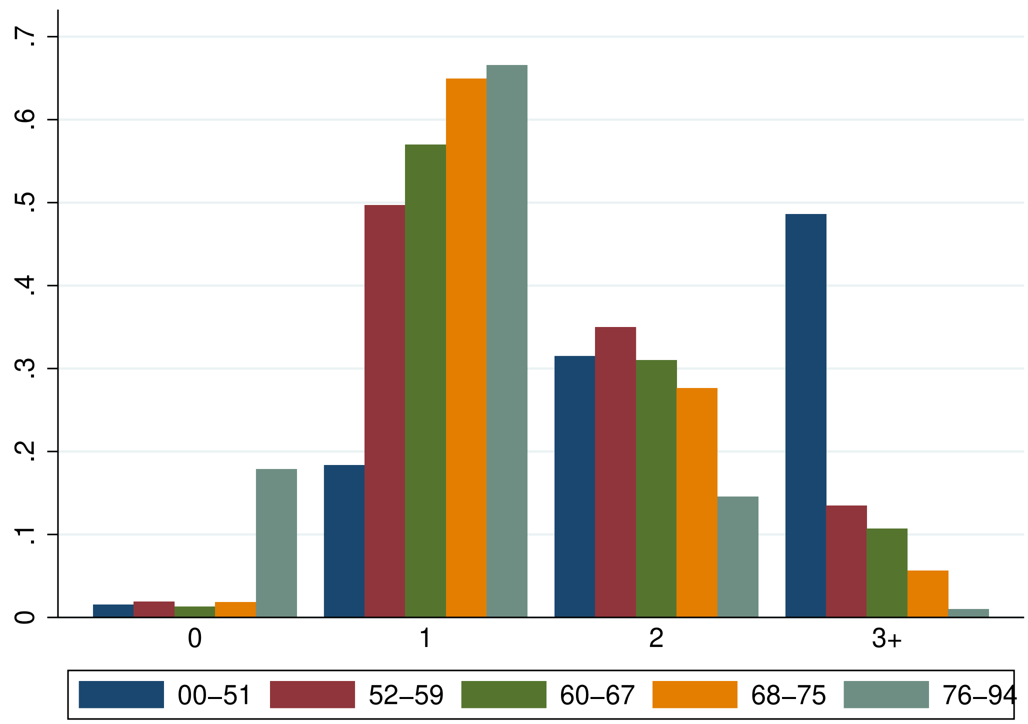


Figure F.30: Share of Individuals by Combination of Kids, by Birth Cohort

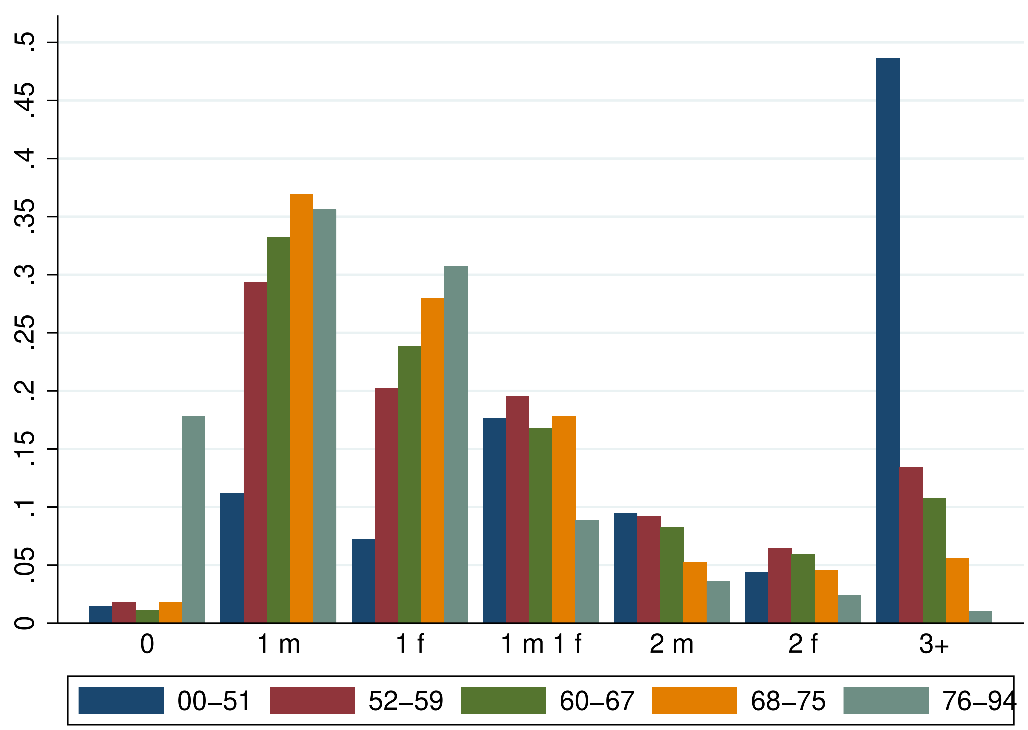
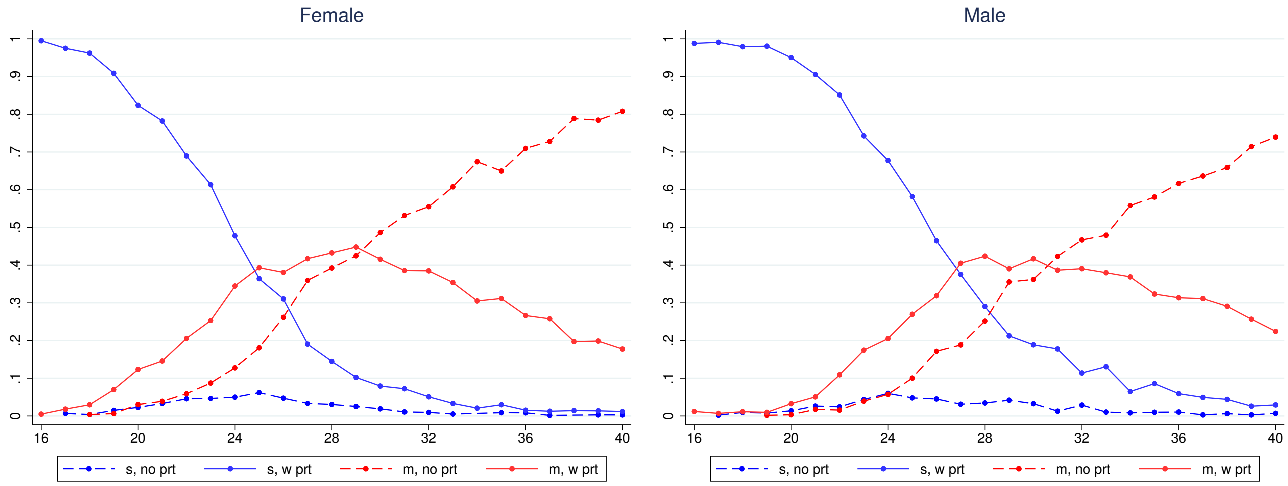


Figure F.31: Cohabitation Pattern by Gender



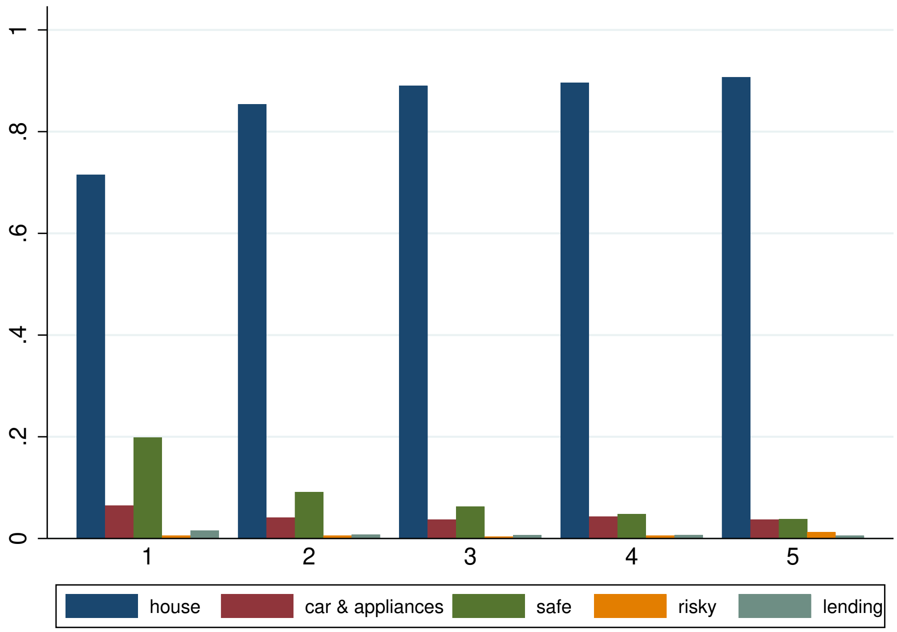
Notes. CHFS 13. Cohabitation refers to living with one's parent(s) or in-law(s). 's' and 'm' refers to single and married; 'prt' refers to parent living in the same dwelling. Here we plot the share of individuals in each age that fall into the 4 categories: single or married by cohabitation with parent(s). In CHFS, we observe each member's relationship to the head. We consider someone to be living with parent(s) only when he or she is a lower generation to the household head. For example, if a household has generation 1, 2, 3 (3 being youngest), generation 1 and 2 are not considered 'living with parents'; generation 3 is considered 'living with parents'.

Table 13: Marriage, Other House, and Move-out

	Coefficient		Hazard Ratio	
	f	m	f	m
Married	1.351***	0.325**	3.861***	1.385**
	(0.176)	(0.136)	(0.679)	(0.189)
Own other house	-0.780*	-0.156	0.458*	0.855
	(0.464)	(0.295)	(0.213)	(0.252)
Marr X Own oth	0.678	0.552*	1.969	1.737*
	(0.506)	(0.314)	(0.996)	(0.545)
Years of education	0.144***	0.0455***	1.155***	1.047***
	(0.0213)	(0.0113)	(0.025)	(0.012)
Log personal income	0.318***	0.164***	1.374***	1.179***
	(0.0641)	(0.0376)	(0.088)	(0.044)
Log family asset	-0.352***	-0.382***	0.703***	0.683***
	(0.0375)	(0.0247)	(0.026)	(0.017)
Observations	20,463	27,466	20,463	27,466
Province FE	Y	Y	Y	Y

Notes. CFPS 10, 12, 14, 16 merged. We keep only individuals whose family either own the residential house and no other house, or own the residential house and own other house, or does not own any house. We label individuals as either from a family that owns other house or one that does not. The left-out group is those whose family does not. Education is years of completed education; income is log of individual annual income; family asset is log of the sum of family's financial and housing asset. We use the Cox proportional hazard model for baseline hazard; controlled for province fixed effect.

Figure F.32: Portfolio Share by Asset Quintile



Notes. CHFS 13. We categorize different financial assets into two types: safe and risky. Safe asset includes cash, deposit (current and fixed), cash holdings in the household's stock market account, bonds (government and corporate), holdings in mutual funds, and gold. Risky asset includes stock, futures, warrants, other financial derivatives, financial product (provided by bank or other institutes), and asset in other currencies.