Understanding household savings in China: the role of the housing market and borrowing constraints

Matthieu Bussière and Yannick Kalantzis and Romain Lafarguette and Terry Sicular

Banque de France, Banque de France, Peking University (CCER) and Paris School of Economics, University of Western Ontario

26. February 2013

Understanding Household Savings in China: the Role of the Housing Market and Borrowing Constraints*

Matthieu Bussière[†] Yannick Kalantzis[‡] Romain Lafarguette[§] Terry Sicular[¶]

February 25, 2013

Abstract

This paper examines the role of rising housing prices and borrowing constraints as determinants of China's high household saving rate, especially among young households. Using a life-cycle model of saving behavior in the presence of borrowing constraints, we show that the relationship between housing prices and saving exists only under certain conditions and for certain groups of households. Specifically, when the return on financial instruments is low (which is the case in China), the saving rate of young households may increase with housing prices. This relationship, moreover, is non-linear and depends on the level of wealth. Employing an empirical strategy motivated by the theoretical model, we analyze a dataset of over six thousand Chinese urban households spanning the years 1995, 2002 and 2007. We find evidence that higher housing prices do in fact increase the saving rates of young households. We also find evidence for the predicted non-linearity.

Keywords: saving rates, housing market, China, global imbalances, borrowing constraints.

JEL Classification Numbers: D91, E21, C81, O16.

*The paper has benefited from discussions with Nicolas Coeurdacier, Meredith Crowley, Maurice Obstfeld, Stephanie Schmitt-Grohe, Yang Yao, Yi Wen, and from comments by participants at the 2012 "Columbia-Tsinghua Conference in International Economics" in Beijing. The views presented in this paper are those of the authors and do no necessarily reflect those of the Banque de France or the Eurosystem.

†Corresponding author. Banque de France, 31 rue Croix des Petits Champs, 75001 Paris, France; email: matthieu.bussiere.at.banque-france.fr. Phone: +33 1 42 92 29 21. Fax: +33 1 42 92 27 66.

‡Banque de France; email: yannick.kalantzis.at.banque-france.fr

§Peking University (CCER) and Paris School of Economics; email: romain.lafarguette.at.ens.fr

¶University of Western Ontario; email: sicular.at.uwo.ca

1 Introduction

The particularly high household saving rates in China have attracted much attention among policy makers and academics. In policy circles, high saving rates in emerging Asia are seen as a key factor behind global imbalances, in which China appears to be a major player (Bernanke, 2005, 2006). Meanwhile, concerns have been expressed about China's high saving rates (and correspondingly weak consumption), which can represent a potential threat to the Chinese economy itself by putting China's growth model off-balance (see in particular Yellen, 2011, Zhou, 2011, and repeated G20 communiqués).

Data from household surveys suggest that saving rates are especially high for young households (Chamon and Prasad, 2010). Along with several other factors, high housing prices, together with strong borrowing constraints, have been proposed as a possible explanation. First, house purchases are part of saving so that, other things equal, higher prices could translate into larger saving. Second, with strong borrowing constraints, young households need to save more prior to the purchase in order to make the required down payment (Chamon and Prasad 2010, Wang and Wen 2012). At first glance, this mechanism seems a promising explanation given the developments on the Chinese housing market in the last twenty years. The privatization of urban housing started in the early 1990's and was close to completion in the early 2000's. A sizable appreciation of housing prices took place during the same period and accelerated even further afterward. While appealing, this potential link between housing prices and household savings is not as straightforward as it may seem: if households have the choice between buying and renting a home, it is not obvious why they should distort their saving behavior to become owners.

This paper aims at carefully studying the relationship between housing prices and the savings behavior of young households in a context of strong borrowing constraints. Our contribution is twofold. First, we build a life-cycle model and identify conditions under which there can indeed be a positive relationship between housing prices and saving. Second, we bring the model to the data, using the China Household Income Project (CHIP) surveys, a unique dataset of over six thousand urban households. The CHIP survey provides data for 1995, 2002, and 2007, years that span the period during which China's urban housing privatization was fully implemented.

In the model, households live for three periods (young, middle age, and old). They receive labor income in the first two periods but not the last one. We assume they can use two alternative assets as saving instruments: financial assets (deposits) and real assets. While deposits have a potentially lower rate of return, real assets are subject to a minimum investment requirement. A typical example of such a real asset is an equity share in housing when homes have a minimum size and the debt-to-value ratio is constrained. Using this framework, we show that the relationship between housing prices and saving exists only under certain conditions and for certain groups of households.

We first assume that the two assets have the same rate of return. Then, there is no systematic relationship between household saving and housing prices, even with strong borrowing constraints. The reason is that households prefer to rent housing, rather than having to distort their saving profile to buy a home. Next, we assume that deposits have a lower rate of return than real assets, a feature that describes the Chinese economy well. With this assumption, households can be willing to increase their saving rate when young, in order to accumulate enough wealth to purchase a home in their middle age and enjoy the larger rate of return when old. In that case, saving rates of young households are shown to be an increasing function of the housing price. Importantly, the relationship between housing prices and saving rates of young households is found to be non-linear: it only holds for households with intermediate levels of wealth, but not for very poor or very rich households.

The model's testable implication is that the saving rate of young households should be negatively correlated with the ratio of income to housing prices for certain levels of wealth. In the empirical exercise, we use the geographical variation of housing prices at the province level to test for this mechanism. Our identification strategy consists of comparing households with different shares of young members facing different real estate prices at the provincial level. Specifically, we interact the income-to-housing price ratio with the share of young members in the household.[1] If the mechanism is present in the data, this interacted term should be negative. This identification strategy addresses potential correlations between provincial housing prices and other variables likely to affect saving behaviors.

Our result provides support for the mechanism described in the model. We do indeed find that the interacted coefficient is negative, with estimated values statistically significant at conventional levels for 1995 and 2007 (for 2002, the estimation is less accurate but the p-value is close to 10%). This result is robust to a variety of alternative specifications. We also find support for the predicted non-linearity. When restricting the samples to wealth quartiles, we find that the relationship only holds for lower levels of wealth, with statistically significant coefficients for all three years. According to our results, a 50% increase in housing prices should be associated ceteris paribus with a 10 percentage point increase of the saving rate of young individuals belonging to the first quartile of the wealth distribution. We therefore conclude that house prices play a key role in understanding the high Chinese saving rate.

The rest of this paper is organized as follows. Section 2 reviews the relevant literature. Section 3 presents key stylized facts on household saving in China, with reference to Chamon and Prasad (2010). Section 4 presents our theoretical model of household saving. The empirical exercises are carried out in Section 5. Section 6 concludes.

[1] A large proportion of Chinese households are multi-generational.

2 Review of the literature

The literature on savings in China has grown rapidly in recent years, much of it prompted by China's unusually high savings rates. Estimates of the average propensity to save by households vary depending on the data used. National accounts flow of funds data yield aggregate savings rates in the range of 20 to 25 percent in the early 1990s, rising to 30 percent in the late 1990s and early 2000s and then increasing further to nearly 40 percent by 2008 (Ma and Yi 2010). Households have contributed from 40 percent to 50 percent or more of aggregate savings, and studies based on national accounts data report shares of household savings in GDP generally in the range of 17 to 23 percent (Ma and Yi 2010; Chamon and Prasad 2010, Yang, Zhang and Zhou 2011). Household survey data yield lower average savings rates, but with similar upward trends. For urban households, the focus of this paper, average savings rates calculated using household survey data from the National Bureau of Statistics (NBS) rose from 15-18 percent in the 1990s to 20 percent in the early 2000s, and further to over 25 percent in the late 2000s (Chamon and Prasad 2010; Yang, Zhang and Zhou 2011).

Explanations for China's high household saving rates generally take the lifecycle hypothesis (LCH) as a starting point. Several studies explicitly test the applicability of the LCH to China, in some cases making comparisons with Keynesian models (Ang 2009, Horioka and Wan 2007, Modigliani and Cao 2004). The LCH implies that household saving rates may be affected by (anticipated) growth in per capita income, age or stage in the lifecycle, interest rates, and inflation. Uncertainty about future income growth can create a motive for precautionary savings, and so may also be relevant (Kraay 2000, Nabar 2011). Modigliani and Cao (2004), using estimates of national savings based on aggregate asset data, and Horioka and Wan (2007), using provincial-level aggregated household survey data, find that China's high and increasing household savings rates are largely explained by income growth and changes in demographic structure. Horioka and Wan (2007) also reports evidence of behavioral inertia/persistence and identifies differences between urban and rural savings behavior.

Kraay (2000) and Nabar (2011), both of which use provincial-level aggregated household survey data, come to somewhat different conclusions. Kraay's empirical results suggest that a standard LCH model does not adequately explain Chinese household savings behavior. He suggests that other factors, some unique to China, should be considered. In this regard he raises the importance of measurement issues; credit constraints in conjunction with the need to save for purchases of durables and lifecycle events; and precautionary saving associated with institutional changes and the expansion of new and unpredictable markets in the process of China's economic transition. Nabar (2011) finds that the coefficients on income growth and income volatility are significant in simple models, but that when controls are introduced for lagged savings rates (to capture behavioral persistence), demographics, employment structure, and other variables, income growth and volatility are no longer significant. In almost all cases, however, interest rates have a significant, negative coefficient, which he attributes to target savings behavior.

Recent studies have benefited from access to household-level data, which makes possible empirical analysis of the micro foundations of savings in China using a wider range of explanatory variables. Household-level data display not only high savings rates, but also a U-shaped household age-savings profile. Savings rates are higher for younger and older households, and lower for middle-aged households (Chamon and Prasad 2010, Handelshøyskole 2011, Yang, Zhang and Zhou 2011). This contrasts with the standard hump-shaped age-savings profile implied by the LCH.

Explanations for the unusual pattern of household savings in China generally hinge on the presence of imperfect or incomplete markets reflecting credit constraints, limited insurance options, or specific policies and institutions. Some studies investigate the role of saving for retirement, which has been affected by changes in pension policies and family structure (e.g., Ang 2009, Chamon, Liu and Prasad 2010, Chamon and Prasad 2010, Feng, He and Sato 2011). Some look at the role of precautionary savings, which may have increased following employment reforms and changes in social safety net programs such as health and unemployment insurance (e.g., Baldacci et al. 2010, Barnett and Brooks 2010, Chamon and Prasad 2010, Meng 2003). Banerjee, Meng and Qian (2010) and Wei and Zhang (2010) analyze the impact of lifecycle events such as marriage, as influenced by China's changing demographics and sex ratios in the wake of China's family planning policies. Several recent papers analyze the impact of housing wealth and homeownership on household savings, especially in urban areas after the urban housing privatization and during a period characterized by the rapid appreciation of housing prices (e.g., Chen, Funke and Mehrota 2011, Handelshøyskole 2011, Jin 2011, and Wang and Wen 2012). Finally, Coeurdacier et al. (2012) argue that the abnormal U-shape pattern results from a measurement bias: young households are likely to be individuals living on their own rather than staying with their parents, earning a higher income, and therefore having a higher saving rate than average.

The relationship between housing and savings, the focus of this paper, is the topic of some studies in the general economics literature. A simple LCH model predicts that an unanticipated increase in housing prices will increase the consumption and thus reduce the savings of homeowners. Studies using micro data generally find evidence in support of such a housing wealth effect. Engelhardt (1996), for example, finds a significant effect for households that experience capital losses, but not for households that experience capital gains, in housing. In a study of U.S. states and also across countries, Case, Quigley and Shiller (2004) find evidence of a fairly large housing wealth effect on consumption. Bostic et al. (2009), Campbell and Cocco (2007) and Chen, et al. (2007) report evidence of differences in the housing wealth effect among wealth quartiles and also between younger and older homeowners, households that are and are not credit constrained, and homeowners and renters. These studies highlight the potential importance of heterogeneity among groups in the population.

To date few papers have analyzed the relationship between housing wealth and savings (or consumption) in China. Most look at the urban sector since the mid- or late 1990s, the period during which China's urban housing privatization was completed and real estate markets became more developed. Jin (2011) uses household data from Beijing to analyze the housing wealth effect on consumption and finds a weak relationship. His study, however, is based on a small, nonrandom sample from a single city. Chen, Funke and Mehrota (2011) examine the housing wealth effect using a larger, city-level dataset and panel techniques. They report a positive and significant effect of housing prices on consumption, i.e., housing price appreciation reduces savings. The effect is stronger in cities where house prices have risen more quickly and have been more volatile. These findings suggest that housing wealth effects in China resemble those elsewhere, but they imply that China's urban savings rates should be declining (because housing prices have been increasing). Consequently, they do not solve the puzzle of China's high and rising savings rates.

Credit constraints provide a key to the puzzle. In China credit markets are underdeveloped and households' ability to borrow to finance housing purchases is limited. Households therefore must save to purchase housing. Moreover, the amount that households save will depend on the cost of housing (Chamon and Prasad 2010, Handelshøyskole 2011, Wang and Wen 2012). If housing prices increase relative to income, then households that wish to purchase a house (or help their children purchase a house) will increase their savings rates. This mechanism may lead to heterogeneous savings behavior between households that are and are not saving for the purchase of a home, thus providing a possible explanation for China's U shaped age-savings profile.

Analysis of the impact of credit constraints and housing purchases on household savings behavior in China is as yet limited. A few papers that estimate LCH-type models of savings add control variables for housing. Chamon and Prasad (2010), for example, include a dummy variable for homeownership. Contrary to expectations, the coefficient on this dummy variable is positive or not significantly different from zero. When they include interactions between the homeownership dummy and five-year age ranges for the household head, however, they find the expected negative relationship for younger age groups in recent years. Nabar (2011) uses city-level data to estimate an LCH model and includes the city housing price level and housing price growth as control variables. He finds that the level of the housing price has a positive and significant effect on household savings, although growth in housing prices is not significant.[2]

While some progress has been made in the quest to better understand the relationship between homeownership and savings in China, more work remains to be done. First, while the positive relation between housing prices and household saving in the presence of credit constraints seems appealing, theoretical work is needed to determine the conditions under which such a mechanism really applies. Second, both the theoretical and empirical analysis of savings would benefit from deliberate and explicit attention to heterogeneity in savings behavior. Third, access to new household-level datasets, such as the CHIP survey data, makes possible a more complete, nuanced

[2]The issue of high house prices in China itself has attracted attention. For a recent analysis see Wei, Zhang and Liu (2012), as well as the references therein, who focus on the role of status competition in explaining house prices.

empirical analysis. Our aim in this paper is to make contributions in all these regards.

3 Stylized facts

In this section, we show that our data display patterns similar to those documented by earlier works, *e.g.* Chamon and Prasad (2010), and explore how these patterns are affected by different assumptions regarding the definition of the saving rate. We compute the age-profiles of the saving rates using all three surveys, 1995, 2002, and 2007. Age profiles are built by averaging the saving rate over households of a given age using a three-year moving average to smooth the resulting curve.

We start by replicating the methodology used by Chamon and Prasad (2010) to reproduce their main finding (Figure 5. p. 103). As Chamon and Prasad (2010), we include durable goods and education expenses in the definition of consumption when computing the saving rate. Our methodology differs from theirs in two ways. First, our measures of income and consumption include an estimation of imputed rents. Second, we define the age of the household as the age of its member with the highest income, instead of the head of household. "Household head" is a self-reported variable. In multi-generation households, an older parent may be listed as the household head even though an adult child is the main income earner. In multi-generation households in the CHIP data, the self-reported household heads are often older than the household member with the highest income. We conclude that the age of the household head may not reflect well the stage in the household's lifecycle.

The resulting age profiles are plotted in Figure 4. Two observations stand out. First, the figure clearly shows the large increase of saving rates between 1995 and 2007. Second, the age-profile does not follow the inverted U-shape corresponding to the standard life-cycle. Of particular note, the saving rates of young households continued to rise from 1995 to 2002 and again from 2002 to 2007, by which time their saving rates were in the range of 20 to 30 percent and the highest of among all age groups.

Next, we explore alternative definitions of the saving rate. In Figure 5, we now exclude durables from the definition of consumption. Durable goods can indeed be considered as investments rather than consumption expenditures and methodological papers sometimes recommend excluding them from consumption when computed saving rates. As the Figure shows, the resulting age-profile are similar to the benchmark case.

The next section builds a model to account for this high saving rate of young households.

4 Theoretical model

This section uses a simple life-cycle model to study how the low return on saving instruments in a financially undeveloped economy can lead to large saving by young households. When the return on saving instruments available in the financial system is low and borrowing constraints are strong, young households have an incentive to accumulate enough wealth so that they can buy real assets with larger returns such as housing.

4.1 A simple framework

We start with a simple framework where we can study the mechanism in a transparent way. We consider a life-cycle model where households live for three periods: young, middle age, and old, indexed by $t = 1, 2, 3$. Households receive a labor income in their young and middle age, but not in their old age. We assume an institutional setting characterized by an undeveloped financial system, with a lack of saving instruments providing a good rate of return, so that households use real assets as substitutes for financial assets. We model this institutional setting with the three following assumptions:

1. There are strong borrowing constraints: households cannot borrow.
2. The financial system provides saving instruments D (henceforth *deposits*), with a low gross rate of return r.
3. At $t = 2$, households can also buy a real asset A with a larger rate of return $R > r$, but subject to a minimum investment requirement $\underline{A}$.

Our approach is partial equilibrium and we take the rate of returns r and R as exogenous.

As we will see, these assumptions are enough to generate large saving when young. As in a standard life-cycle model, households want to transfer consumption from their middle age to their old age, but are limited by the low rate of return on deposits. To benefit from the higher rate of return on the real asset, households have to save enough when young to meet the minimum investment requirement in their middle age.

Section 4.2 below shows that this simple framework can be derived from a more realistic set-up with housing and mortgages.

Consider a household that derives utility from consumption:

$$U = u(C_1) + \beta u(C_2) + \beta^2 u(C_3). \tag{1}$$

To get closed-form solutions, we study the case of a logarithmic utility function: $u(C_t) = \log(C_t)$.

The household enters the economy with wealth rD_0. It receives labor income Y at $t = 1$ and gY at $t = 2$. Denote D_t the deposits bought in period t and A_2 the real assets bought at $t = 2$. The budget constraints at $t = 1, 2, 3$ are:

$$rD_0 + Y = C_1 + D_1, \tag{2}$$

$$rD_1 + gY = C_2 + D_2 + \zeta_2 A_2, \tag{3}$$

$$rD_2 + R\zeta_2 A_2 = C_3, \tag{4}$$

where $\zeta_t = 1$ if the household buys the real asset at t and 0 otherwise.

The household faces borrowing constraints and a minimum investment requirement for the real asset:

$$D_1 \geq 0, \tag{5}$$

$$D_2 \geq 0, \tag{6}$$

$$A_2 \geq \underline{A}. \tag{7}$$

An optimal plan is a set of $\{C_t\}_{t=1}^3$, $\{D_t\}_{t=1}^2$, A_2, and ζ_2 that maximizes the utility (1) subject to the budget constraints (2), (3), and (4), the borrowing constraints (5) and (6) and the minimum investment requirement (7).

Define $W = Y(1 + rd_0 + g/r)$ the intertemporal wealth of the household. We use lower-case letters to denote variables normalized by first-period labor income: $d_0 = D_0/Y$, $d_1 = D_1/Y$, $w = W/Y$, $a_2 = A_2/Y$, $\underline{a} = \underline{A}/Y$, etc.

To restrict the set of possible solutions, we make the following assumption:

Assumption 1. $g \leq \beta r(1 + \beta)(1 + rd_0)$.

This assumption makes sure the household has non-negative net assets at the end of the first period.

4.2 Housing and mortgages

This subsection shows how the simple framework described above can be derived from a more realistic set-up with housing and mortgages.

Assume consumption consists of both goods C^G and housing services C^H, with a Cobb-Douglas aggregate: $C_t = \kappa(C_t^G)^{1-\eta}(C_t^H)^\eta$. The aggregate consumption basket C is chosen as the numeraire. Denote p^G the price of goods and ρ the rental price. We have $C_t = p_t^G C_t^G + \rho_t C_t^H$.

Households get housing services either by buying a home, with price p^H, or renting it: $C_t^H = \zeta_{t-1}H_{t-1} + X_t$, where H_{t-1} is the stock of housing bought in the previous period, X_t denotes housing services derived from renting, and $\zeta_t = 1$ if the household buys housing at t and 0 otherwise.[3]

We assume, as in Gervais (2002), that housing is not perfectly divisible: houses have a minimum size $\underline{H}$. This constrains a household wishing to buy a house. On the contrary, renting an accommodation is not subject to this indivisibility as several households can share the same house. Houses depreciate at rate δ.

Households can use their house as collateral to get a mortgage from a bank. Mortgages are subject to a maximum loan-to-value ratio γ: a household must make a down payment representing at least a fraction $1-\gamma$ of the value of the house. Denote R^m the gross rate of return on a mortgage.

[3] X_t can be negative if the household rents some of its house or apartment to another household.

Denote B_t the mortgage issued in period t. The budget constraint at time t is:

$$rD_{t-1} + p_t^H(1-\delta)\zeta_{t-1}H_{t-1} + B_t + Y_t = p_t^G C_t^G + \rho_t X_t + R^m B_{t-1} + p_t^H \zeta_t H_t + D_t. \tag{8}$$

In addition to the non-negativity constraints on D_t, (5) and (6), the household is now subject to borrowing constraints and the indivisibility of houses:

$$B_t \leq \gamma p_t H_t, \tag{9}$$

$$H_t \geq \underline{H}. \tag{10}$$

Define the rate of return on housing $R_t^H = [(1-\delta)p_t^H + \rho_t]/p_{t-1}^H$. Using the fact that $C_t^H = \zeta_{t-1}H_{t-1} + X_t$ and that $p_t^G C_t^G + \rho_t C_t^H = C_t$, we can rewrite the budget constraint (8) as follows:

$$rD_{t-1} + (R_t^H - R^m)p_{t-1}^H \zeta_{t-1}H_{t-1} + R^m(p_{t-1}^H \zeta_{t-1}H_{t-1} - B_{t-1}) + Y_t = C_t + (p_t^H \zeta_t H_t - B_t) + D_t.$$

We can consolidate the borrowing constraint (9) and the indivisibility constraint (10) in a minimum investment requirement in housing equity:

$$p_t^H H_t - B_t \geq (1-\gamma)p_t^H \underline{H}.$$

Assume banks arbitrage away any difference between the return on housing and the mortgage rate. Then, $R_t^H = R^m$, and the only housing-related term that matters in the budget constraint is housing equity $p_t^H H_t - B_t$. If, on the contrary, because of market imperfections, $R_t^H > R^m$, then any household that chooses to invest in housing will leverage up as much as possible so that the value of housing will be a multiple $1/(1-\gamma)$ of housing equity. In both cases, the household's problem reduces to the model exposed in section 4.1 where the real asset is housing equity:

$$\begin{aligned} A_t &= p_t^H H_t - B_t, \\ R &= \frac{R_{t+1}^H - R^m}{1-\gamma} + R^m, \\ \underline{A}_t &= (1-\gamma)p_t^H \underline{H}. \end{aligned}$$

When the real asset is housing equity, the minimum investment requirement increases with the price of housing p_t^H and decreases with the maximum loan-to-value ratio (*i.e.* increases with the intensity of the borrowing constraint on the mortgage market).

4.3 Discussion of the set-up

We think the simple set-up presented in section 4.1, although highly stylized, captures some key features of the Chinese economy. First, Chinese households have limited access to credit, as in our model. Second, they have a limited choice of saving instruments, which mainly consist of deposits. The interest rate on deposits is notoriously low, in part due to the policy of setting a ceiling for deposit rates (Laurens and Maino 2007, Green 2005). With strict capital controls, households do not have access to foreign financial markets where they might benefit from higher returns.[4] On the contrary, the rate of return on housing has been large, due to substantial increases in real estate prices over the last years.

In the model, we take both rates of return, on deposits (r) and housing equity R, as exogenous. The rate on deposits is regulated by the government and is therefore set exogenously. As for the return on housing equity, it depends on equilibrium on this market and is exogenous to a single household, our unit of analysis.

In section 4.2, we interpret the real asset as an equity share in housing. Several authors have argued that the need to buy a house, together with borrowing constraints, might generate strong savings for young households in order to finance the required down-payment. The decision to buy a home is generally presented as exogenous (e.g. Wang and Wen, 2012). A shortcoming of this argument is that households have the possibility to rent a home instead of buying it. In our model, the decision to buy a home is endogenous. As we will show, this dramatically changes the results as it destroys the link between housing and saving. To recover this link, it is necessary to assume a wedge between the rate of return of housing and that of financial instrument such as deposits.

To simplify the exposition, we have assumed that households can only buy the real asset at $t = 2$. Assuming they can buy it at $t = 1$ would not change the results, but would generate more cases. The fact that households can only buy a house in their second period of life is also consistent with our data: Figures 1 and 2 show the distribution of home purchase age in the 2002 and 2007 surveys, with a mode at 42 in 2002 and 33 in 2007.

4.4 Understanding high saving from young households

We now solve for the optimal plan in the simple setting of section 4.1.

Consider first the case when the rates of return on deposits and real assets are equal: $r = R$. Then, both assets are perfect substitutes. In that case, the minimum investment requirement $\underline{a}$ (and therefore the price of housing) does not have any impact on saving behavior: if a household cannot meet the minimum requirement at $t = 2$, it will simply buy deposits instead, without any loss on utility.

The following Proposition derives saving in both periods in the case $r = R$.

Proposition 1. *If $r = R$, the demand for deposits and real assets are given by $d_1 = s_1^*$ and*

[4] Bacchetta, Benhima, and Kalantzis (2012) show that it can be optimal for the government to implement capital controls and let the interest rate on saving deviate from the world interest rate. However, they argue that in the case of China, the government should choose a domestic interest rate on deposits higher than the world rate, not lower.

$d_2 + \zeta_2 a_2 = s_2^*$, where

$$s_1^* = \frac{\beta(1+\beta)(1+rd_0) - \frac{g}{r}}{1+\beta+\beta^2}, \quad s_2^* = \frac{\beta^2[r(1+rd_0)+g]}{1+\beta+\beta^2}.$$

If $s_2^* \leq \underline{a}$, asset holdings s_2^* only consist of deposits, otherwise the distribution of assets between deposits and real assets is undetermined.

Proof. See Appendix 7.2. □

When $r = R$, the saving behavior is entirely driven by the life-cycle motive. A steeper income profile between the first and second period of life (a higher g), leads to lower saving when young and larger saving when middle-aged, i.e. to a more pronounced inverted-U shape.

Corollary 1. *When $r = R$ and the real asset is housing equity as in section 4.2, saving from young households is independent of housing prices p^H.*

The absence of link between housing prices and saving when $r = R$ comes from the fact that households are free to choose between owning and renting a home. When housing prices are high enough, the required down payment to buy a home exceeds saving in the second period. Then, households are not willing to deviate from the saving profile corresponding to the life-cycle motive. They prefer to rent a home instead.

Consider now the case of a low return on saving instruments $r < R$. Then, the real asset strictly dominates deposits. Suppose the household is rich enough to meet the minimum investment requirement (7). Then, the household buys deposits $d_1 = s_1^*$ at $t = 1$ and invests all its saving in the real asset at $t = 2$, with $a_2 = s_2^*$. Call this a plan of type (a).

Note that the amount of assets purchased in the second period, s_2^*, increases with the intertemporal wealth $w = (1 + rd_0) + g/r$. Suppose now that the household has a lower intertemporal wealth, such that $s_2^* < \underline{a}$. Then, it cannot invest in the real assets without distorting its saving profile. There are two possibilities. The household can choose to distort its saving profile to buy the minimum amount of real assets $\underline{a}$ at $t = 2$. This requires reducing consumption both at $t = 2$, but also at $t = 1$ since the household tries to smooth consumption. We refer to this as a type-(b) plan. Or the household can just invest in deposits at $t = 1$ and $t = 2$, without buying the real asset, which we refer to as a type-(c) plan. The terms of the trade-off are the following: by choosing a type-(b) plan, the households reaps the benefits of the higher return of the real asset between periods 2 and 3, but suffers from the disutility of distorting its consumption-saving profile.

The following Proposition shows that households with a large enough intertemporal wealth choose a type (a) plan, while households with an intermediary wealth choose a type (b) plan, and households with a low enough wealth choose a type (c) plan.

Proposition 2. *There is a function φ strictly decreasing on $[1, +\infty)$, with $\varphi(1) = \frac{1+\beta+\beta^2}{\beta^2}$, $\lim_{x\to+\infty} \varphi(x) = 1$, and such that:*

- *For an intertemporal wealth* $w \geq \frac{\underline{a}}{r}\varphi(1)$*, the household chooses the type-(a) plan with* $d_1^a = s_1^*$*,* $d_2^a = 0$*, and* $a_2^a = s_2^* > \underline{a}$*.*
- *For* $\frac{\underline{a}}{r}\varphi\left(\frac{R}{r}\right) \leq w < \frac{\underline{a}}{r}\varphi(1)$*, the household chooses the type-(b) plan with*

$$d_1^b = \frac{\beta(1 + rd_0) - \frac{g - \underline{a}}{r}}{1 + \beta} > s_1^*,$$

$d_2^b = 0$*, and* $a_2^b = \underline{a}$*.*

- *For* $w < \frac{\underline{a}}{r}\varphi\left(\frac{R}{r}\right)$*, the household chooses the type-(c) plan with* $d_1^c = s_1^*$*,* $d_2^c = s_2^*$*, and* $a_2^c = 0$*.*

In addition, the difference between deposits d_1 *in plans of type (b) and (a),* $d_1^b - d_1^a$ *increases with* $\underline{a}$ *and decreases with* w*.*

Proof. See Appendix 7.2. □

A household with an intermediate level of intertemporal wealth finds it optimal to save a larger fraction of its income when young. This way, the household accumulates enough wealth to meet the minimum investment requirement in its middle age, and benefits from the larger return of the real asset when old.

Figure 3 plots the saving of the young household D_1/Y as a function of the minimum investment requirement $\underline{A}/Y$. For households that choose plans of type (b) (middle segment of Figure 3), the proposition predicts that a higher value of $\underline{a} = \underline{A}/Y$ leads to more saving in their young age (d_1^b increases with $\underline{a}$). For "rich" or "poor" households, saving does not depend of $\underline{A}/Y$.

In the second period, saving of households choosing type-(b) plans is given by $\underline{a} - d_1^b$ which can be shown to be also increasing in $\underline{a}$ when $r(1 + \beta) > 1$, a plausible assumption. Finally, in their third period, households dissave. Their saving is equal to $-\underline{A}$, with a saving rate equal to $1/(R-1)$ independent of $\underline{a}$.

If we interpret the real asset as an equity share in housing, as in section 4.2, the minimum investment requirement $\underline{A}$ is an increasing function of the price of housing. Therefore, the model shows how a higher price of housing can lead to more saving by young households with an intermediate level of intertemporal wealth. On the contrary, for rich households with type-(a) plans and poor households with type-(c) plans, the price of housing should not have any effect on saving.

This leads to testable empirical implications. Consider a household with a type-(b) plan, *i.e.* a household with an intermediate level of intertemporal wealth. According to the model, we expect the saving rate of this household to be negatively correlated with the income-price ratio when young (increasing with the housing price and decreasing with the income). On the contrary, for rich households, the saving rate in the young and middle age should not be correlated with the income-price ratio.

5 Empirical Results

This section presents empirical evidence on the relation between household savings and housing prices, using as a guide insights from the theoretical model outlined in Section 4.

Our identification strategy consists in exploiting both the geographical and household dimensions of the data. We compare households with different shares of young members facing different real estate prices in provincial real estate markets. Consider households with similar characteristics except regarding their share of young members and the price of housing in their province. The mechanism described in the model implies that the difference in the saving rate of households with a larger share of young members and households with a lower share of young members, should be higher in provincial real estate markets with a higher housing price.

More specifically, our baseline specification is the following:

$$\begin{aligned} s_{i,t} = a_{i,t}\frac{y_{i,t}}{p^H_{j,t}} + b_{i,t}\log(y_{i,t}) + c_{i,t}\text{share}^{20-30}_{i,t} + d_{i,t}\text{share}^{30-40}_{i,t} \\ + e_{i,t}\frac{y_{i,t}}{p^H_{j,t}} \times \text{share}^{20-30}_{i,t} + f_{i,t}\frac{y_{i,t}}{p^H_{j,t}} \times \text{share}^{30-40}_{i,t} + g_{i,t}\log(y_{i,t}) \times \text{share}^{20-30}_{i,t} + h_{i,t}\log(y_{i,t}) \times \text{share}^{30-40}_{i,t} \\ + \mathbf{K}_{ij,t}\text{controls}_{ij,t} + \epsilon_{i,t} \quad (11) \end{aligned}$$

where i denotes a household, j a province, and t a date. The variable $s_{i,t}$ denotes the saving rate of household i in year t, $y_{i,t}$ its net income, $\text{share}^{20-30}_{i,t}$ ($\text{share}^{30-40}_{i,t}$) its share of members aged between 20 and 30 (30 and 40), and $p^H_{j,t}$ denotes the average price of housing in province j and date t. As explained in Section 4, the relevant variable to capture the effect of housing on saving is the income-price ratio $\frac{y_{i,t}}{p^H_{j,t}}$. Our identification strategy focuses on the interactions $\frac{y_{i,t}}{p^H_{j,t}} \times \text{share}^{20-30}_{i,t}$ and $\frac{y_{i,t}}{p^H_{j,t}} \times \text{share}^{30-40}_{i,t}$. If the mechanism of the model is present in the data, the corresponding coefficients should be negative. This identification strategy should address concerns that provincial housing prices might be correlated with other variables likely to affect the saving rate. It assumes that the reaction of saving rates to those omitted variables is the same across households with different shares of young members.

The data consists of repeated cross-sections for the years 1995, 2002, and 2007. We compute the average provincial housing price using the value of homes per square meter for households owning their home.[5] There are between 11 and 16 provinces in the sample, depending on years (see Table 9 in the Appendix). To avoid extreme values of the saving rate, we exclude the top and bottom 0.5% values for this variable. For the years 2002 and 2007, we use population-based weights to correct for the fact that the sample sizes are not proportional to regional or provincial

[5]In the CHIP survey households were asked to estimate the market value of their housing. These self-reported housing values and self-reported housing areas are used to calculate each household's housing price per square meter, and then the provincial average housing price per square meter.

populations.[6] Importantly, we control for the household income, both directly and interacted with the share of young members. This should capture the cross-sectional variation coming from the numerator of the income-price ratio and make sure that our results are driven by housing prices. A potential issue with our methodology is that we cannot control for provincial fixed effects. To address this issue, we include among controls several variables aggregated at the provincial level. The baseline regression includes average provincial income and consumption as control variables. We experiment with other variables in robustness checks later. Descriptive statistics of the main variables used in the regressions are presented in Table 10 of the Appendix.

5.1 Baseline specification

Table 1 presents a first set of results for this baseline specification. The dependent variable is the saving rate of each of the households in the dataset. Columns (1), (4), and (7) present a benchmark regression which only includes control variables. Columns (2), (5), and (8) add the variables of interest, i.e. the income-price ratio y/p^H and the share of young households $share^{20-30}$ and $share^{30-40}$. Finally, columns (3), (6), and (9) add the interaction terms described above. For each coefficient, p-values are reported in brackets.

Control variables include age dummies, the size of the household (in log), the number of generations in the household, and the household income and gross wealth (in log)[7]. Unfortunately, the 2007 survey does not report total wealth any more. For that year, we used the value of housing as a proxy for gross wealth for households owning their home and set the variable to zero for renters. As explained above, we also control for the average income and consumption of the province. Among variables with statistically significant coefficients at conventional levels, age dummies capture the age-profile described in Section 3, the household size (number of generations) is negatively (positively) correlated with saving rates, and the effect of household income is positive: richer households save comparatively more. As for provincial average income and consumption, households in richer cities tend to have a higher saving rate but the average city consumption has a negative sign. A likely explanation of the latter is that prices are higher in cities with a high average (nominal) consumption, and therefore households save less in these areas, conditional on their nominal income.

In columns (2), (5), and (8), the income-price ratio has a negative sign, as predicted in the theoretical framework of Section 4: higher house prices, for a given income level (and therefore a lower income-price ratio), are associated with higher household saving. The coefficient is statistically significant at 1% for 1995 and 2002, but not for 2007.

Finally, columns (3), (6), and (9) present the full regression. The coefficient of the interaction

[6] Weights were not available for 1995.

[7] Gross wealth is defined as the sum of total financial assets, durable goods value, house value, fixed assets and other assets

Table 1: Baseline regressions

	1995			2002			2007		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(Intercept)	2.45	3.15	2.94	1.3	1.86	1.82	1.51	1.6	1.54
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Age 30-34	-0.04	0.00	0.01	0.00	-0.03	-0.03	-0.03	-0.06	-0.06
	[0.1]	[0.96]	[0.83]	[0.88]	[0.43]	[0.4]	[0.07]	[0.00]	[0.01]
Age 35-39	-0.04	0.01	0.02	-0.03	-0.06	-0.06	-0.04	-0.07	-0.07
	[0.09]	[0.7]	[0.54]	[0.24]	[0.07]	[0.05]	[0.02]	[0.00]	[0.00]
Age 40-44	-0.09	-0.04	-0.03	-0.06	-0.06	-0.06	-0.1	-0.09	-0.09
	[0.00]	[0.24]	[0.38]	[0.01]	[0.03]	[0.02]	[0.00]	[0.00]	[0.00]
Age 45-49	-0.11	-0.06	-0.05	-0.1	-0.09	-0.1	-0.12	-0.1	-0.1
	[0.00]	[0.04]	[0.09]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Age 50-54	-0.08	-0.04	-0.03	-0.06	-0.05	-0.05	-0.06	-0.04	-0.04
	[0.00]	[0.2]	[0.37]	[0.02]	[0.04]	[0.03]	[0.00]	[0.02]	[0.02]
Age 55-59	-0.03	0.01	0.02	0.02	0.02	0.02	-0.04	-0.03	-0.02
	[0.31]	[0.66]	[0.45]	[0.49]	[0.43]	[0.46]	[0.04]	[0.17]	[0.23]
Age 60-64	0.01	0.06	0.06	0.03	0.03	0.03	-0.02	-0.01	0
	[0.69]	[0.07]	[0.05]	[0.4]	[0.37]	[0.37]	[0.46]	[0.8]	[0.96]
Age 65-69	0.01	0.06	0.06	0.11	0.11	0.11	-0.06	-0.05	-0.04
	[0.79]	[0.1]	[0.07]	[0.00]	[0.00]	[0.00]	[0.01]	[0.06]	[0.1]
Log(size)	0.01	-0.01	0.00	-0.07	-0.07	-0.06	-0.06	-0.06	-0.06
	[0.81]	[0.65]	[0.93]	[0.04]	[0.05]	[0.1]	[0.01]	[0.01]	[0.01]
# of generations	-0.03	-0.03	-0.03	0.03	0.03	0.03	0.02	0.02	0.02
	[0.07]	[0.13]	[0.08]	[0.07]	[0.07]	[0.12]	[0.04]	[0.06]	[0.09]
Log(income)	0.22	0.29	0.23	0.12	0.16	0.19	0.15	0.16	0.16
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Log(gross wealth)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	[0.44]	[0.43]	[0.42]	[0.94]	[0.86]	[0.75]	[0.93]	[0.87]	[0.78]
Log(avg prov. income)	0.83	0.48	0.57	0.54	0.51	0.51	0.61	0.59	0.59
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Log(avg prov. consumption)	-1.08	-0.81	-0.87	-0.66	-0.68	-0.68	-0.74	-0.73	-0.73
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$		-47.13	-34.83		-9.18	-8		-2.33	-1.37
		[0.00]	[0.03]		[0.00]	[0.01]		[0.24]	[0.5]
share^{20-30}		0.1	0.11		0.01	0.01		0.05	0.06
		[0.00]	[0.00]		[0.72]	[0.81]		[0.03]	[0.02]
share^{30-40}		0	0.01		0.08	0.08		0.1	0.1
		[0.91]	[0.84]		[0.06]	[0.07]		[0.00]	[0.00]
Log(income) $\times$ share^{20-30}			0.11			-0.09			0.06
			[0.06]			[0.19]			[0.2]
Log(income) $\times$ share^{30-40}			0.12			-0.11			-0.07
			[0.01]			[0.03]			[0.01]
$\frac{y}{p^H} \times \text{share}^{20-30}$			-141.93			-23.52			-20.04
			[0.00]			[0.12]			[0.03]
$\frac{y}{p^H} \times \text{share}^{30-40}$			-20.04			-16.22			-0.11
			[0.53]			[0.15]			[0.99]
R^2	0.07	0.08	0.08	0.06	0.06	0.06	0.08	0.09	0.09
$adj.R^2$	0.07	0.07	0.08	0.06	0.06	0.06	0.08	0.08	0.09
N	6583	6583	6583	6431	6431	6431	9155	9155	9155

p-values in brackets

term $\frac{y_{i,t}}{p^H_{j,t}} \times share^{20-30}_{i,t}$ is negative for all three years, consistent with the identification strategy. It is statistically significant at the 5% level for 1995 and 2002. The estimation is somewhat less accurate for 2002 with a p-value of 12%. The coefficients of the $\frac{y_{i,t}}{p^H_{j,t}} \times share^{30-40}_{i,t}$ are negative but have very larger standard errors and cannot be distinguished from 0 at conventional levels of statistical significance. Overall, our baseline regressions lend support to the existence of a link between housing prices and saving, in line with the mechanism highlighted in the theoretical section of this paper.

5.2 Non-linearity of the link between housing prices and saving

To test the prediction that the main effect of the model should disappear for rich households (who do not face the same financial constraint), we run regressions on subsets defined by wealth quartiles (Tables 2 to 4). For 1995, the coefficient of the interaction term $\frac{y_{i,t}}{p^H_{j,t}} \times share^{20-30}_{i,t}$ gets its large value in absolute value term in the second quartile. For both 2002 and 2007, the interaction term is only statistically significant in the first quartile (at the 10% for 2002 and 5% level for 2007). These results confirm the non-linearity predicted by the model. The result for 2007 should however be taken with caution: as explained above, gross wealth for the year is proxied by housing wealth.

To get a sense of the magnitude of the effect, consider the first quartile of 2002. The estimated coefficient is -88, for an average income-price ratio equal to 3.3×10^{-3}. Other things equal, a housing prices that are 20% higher should reduce the ratio by 0.610^{-3}, which would mean the saving rate of a young individual aged 20 to 30 would be 5 percentage points higher. Similarly, housing prices that are 50% higher would be associated with a saving rate for this age of individual that is 10 percentage points higher.

5.3 Robustness checks

Finally, tables 5 to 7 present a set of alternative specifications to check the robustness and validity of our main results. For convenience, the baseline specification is reproduced in column (1). Column (2) replaces the age dummy variables by a continuous measure of age and its square. In column (3), we control for the competitive saving motive of Wei and Zhang (2011). We add dummy variables indicating whether there is a son in the household and whether the household consists of a single young man, the provincial sex ratio and its interactions with the former dummy variables. Column (4) controls for the education of the top income member, and column (5) for its marital status. Overall, the coefficients of interest are remarkably stable throughout. In column (6), we use an alternative definition of the saving rate where we exclude expenditures on durables from the definition of consumption. Finally, column (7) uses the number of young members in the household instead of the share.

Table 2: Regressions on subsets defined by wealth quartiles for 1995

	(1) All sample	(2) Quartile 1	(3) Quartile 2	(4) Quartile 3	(5) Quartile 4
(Intercept)	2.94	2.28	2.77	3.72	3.6
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Age 30-34	0.01	0.08	0.02	-0.06	-0.02
	[0.83]	[0.19]	[0.76]	[0.35]	[0.74]
Age 35-39	0.02	0.09	0.01	-0.06	0.04
	[0.54]	[0.21]	[0.87]	[0.35]	[0.63]
Age 40-44	-0.03	0.00	-0.01	-0.06	-0.04
	[0.38]	[0.96]	[0.89]	[0.28]	[0.52]
Age 45-49	-0.05	0.00	-0.08	-0.09	-0.04
	[0.09]	[0.95]	[0.21]	[0.13]	[0.54]
Age 50-54	-0.03	-0.03	-0.03	-0.07	0.03
	[0.37]	[0.57]	[0.61]	[0.21]	[0.68]
Age 55-59	0.02	0.03	-0.02	0.01	0.06
	[0.45]	[0.64]	[0.71]	[0.85]	[0.36]
Age 60-64	0.06	0.12	0.04	0.03	0.04
	[0.05]	[0.05]	[0.56]	[0.59]	[0.52]
Age 65-69	0.06	0.05	0.07	0.05	0.09
	[0.07]	[0.41]	[0.32]	[0.47]	[0.24]
Log(size)	0.00	-0.08	-0.02	-0.01	0.1
	[0.93]	[0.21]	[0.81]	[0.92]	[0.17]
# of generations	-0.03	-0.02	-0.02	-0.02	-0.06
	[0.08]	[0.64]	[0.59]	[0.46]	[0.08]
Log(income)	0.23	0.3	0.15	0.29	0.24
	[0.00]	[0.00]	[0.03]	[0.00]	[0.00]
Log(gross wealth)	0.00	0.00	-0.01	-0.04	-0.01
	[0.42]	[0.48]	[0.9]	[0.4]	[0.6]
Log(avg prov. income)	0.57	0.2	1.43	0.46	0.11
	[0.00]	[0.52]	[0.00]	[0.18]	[0.79]
Log(avg prov. consumption)	-0.87	-0.44	-1.7	-0.8	-0.48
	[0.00]	[0.1]	[0.00]	[0.01]	[0.19]
$\frac{y}{p^H}$	-34.83	-45.49	-8.98	-32.38	-72.18
	[0.03]	[0.20]	[0.82]	[0.32]	[0.03]
share^{20-30}	0.11	0.17	0.11	0.06	0.08
	[0.00]	[0.02]	[0.10]	[0.32]	[0.28]
share^{30-40}	0.01	-0.08	-0.01	0.11	-0.04
	[0.84]	[0.24]	[0.93]	[0.08]	[0.55]
Log(income) $\times$ share^{20-30}	0.11	0.02	0.3	-0.04	0.23
	[0.06]	[0.88]	[0.05]	[0.74]	[0.07]
Log(income) $\times$ share^{30-40}	0.12	-0.09	0.19	-0.02	0.34
	[0.01]	[0.4]	[0.07]	[0.82]	[0.00]
$\frac{y}{p^H} \times \text{share}^{20-30}$	-141.93	-45.94	-224.05	-102.7	-183.19
	[0.00]	[0.68]	[0.02]	[0.15]	[0.01]
$\frac{y}{p^H} \times \text{share}^{30-40}$	-20.04	82.5	-21.17	-48.65	-44.36
	[0.53]	[0.28]	[0.76]	[0.41]	[0.47]
R^2	0.08	0.08	0.09	0.08	0.09
$adj.R^2$	0.08	0.07	0.08	0.07	0.08
N	6583	1620	1672	1640	1643

p-values in brackets

Table 3: Regressions on subsets defined by wealth quartiles for 2002

	(1) All sample	(2) Quartile 1	(3) Quartile 2	(4) Quartile 3	(5) Quartile 4
(Intercept)	1.82	1.97	1.38	2.61	1.83
	[0.00]	[0.02]	[0.16]	[0.01]	[0.01]
Age 30-34	-0.03	0.01	-0.06	0.00	-0.06
	[0.4]	[0.91]	[0.38]	[0.98]	[0.36]
Age 35-39	-0.06	-0.04	-0.07	-0.04	-0.1
	[0.05]	[0.52]	[0.28]	[0.57]	[0.15]
Age 40-44	-0.06	-0.06	-0.09	-0.04	-0.07
	[0.02]	[0.34]	[0.12]	[0.44]	[0.18]
Age 45-49	-0.1	-0.1	-0.11	-0.07	-0.12
	[0.00]	[0.07]	[0.03]	[0.2]	[0.01]
Age 50-54	-0.05	-0.03	-0.04	0.00	-0.11
	[0.03]	[0.59]	[0.39]	[0.92]	[0.02]
Age 55-59	0.02	0.00	-0.02	0.08	-0.01
	[0.46]	[0.96]	[0.71]	[0.14]	[0.9]
Age 60-64	0.03	0.08	0.07	0.09	-0.12
	[0.37]	[0.26]	[0.34]	[0.23]	[0.1]
Age 65-69	0.11	0.1	0.14	0.13	0.08
	[0.00]	[0.21]	[0.09]	[0.09]	[0.32]
Log(size)	-0.06	-0.21	0.09	0.05	-0.11
	[0.1]	[0.00]	[0.21]	[0.45]	[0.17]
# of generations	0.03	0.12	-0.01	-0.06	0.04
	[0.12]	[0.00]	[0.7]	[0.1]	[0.33]
Log(income)	0.19	0.17	0.16	0.31	0.21
	[0.00]	[0.00]	[0.01]	[0.00]	[0.00]
Log(gross wealth)	0.00	0.00	0.02	-0.02	0.01
	[0.75]	[0.8]	[0.73]	[0.71]	[0.55]
Log(avg prov. income)	0.51	0.56	0.52	0.26	0.59
	[0.00]	[0.00]	[0.00]	[0.04]	[0.00]
Log(avg prov. consumption)	-0.68	-0.75	-0.68	-0.49	-0.78
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$	-8	-14.49	-2.67	-15.56	-10.14
	[0.01]	[0.18]	[0.78]	[0.03]	[0.03]
share^{20-30}	0.01	-0.08	-0.09	0.08	-0.08
	[0.81]	[0.40]	[0.27]	[0.34]	[0.45]
share^{30-40}	0.08	0.09	0.07	0.08	0.1
	[0.07]	[0.32]	[0.40]	[0.35]	[0.31]
Log(income) $\times$ share^{20-30}	-0.09	-0.11	-0.32	-0.37	0.16
	[0.19]	[0.56]	[0.16]	[0.05]	[0.29]
Log(income) $\times$ share^{30-40}	-0.11	0.02	-0.07	-0.34	-0.08
	[0.03]	[0.89]	[0.61]	[0.01]	[0.52]
$\frac{y}{p^H} \times \text{share}^{20-30}$	-23.52	-87.79	49.03	-2.87	-31.48
	[0.12]	[0.08]	[0.3]	[0.93]	[0.14]
$\frac{y}{p^H} \times \text{share}^{30-40}$	-16.22	-23.4	-35.29	22.19	-28.81
	[0.15]	[0.55]	[0.26]	[0.29]	[0.1]
R^2	0.06	0.07	0.05	0.06	0.08
$adj.R^2$	0.06	0.06	0.04	0.04	0.07
N	6431	1602	1629	1606	1594

p-values in brackets

Table 4: Regressions on subsets defined by wealth quartiles for 2007

	(1) All sample	(2) Quartile 1	(3) Quartile 2	(4) Quartile 3	(5) Quartile 4
(Intercept)	1.54	0.59	1.37	2.02	3.11
	[0.00]	[0.14]	[0.01]	[0.00]	[0.00]
Age 30-34	-0.06	-0.06	0.01	-0.15	0.00
	[0.01]	[0.15]	[0.86]	[0.00]	[0.96]
Age 35-39	-0.07	-0.08	-0.02	-0.15	0.02
	[0.00]	[0.07]	[0.68]	[0.00]	[0.67]
Age 40-44	-0.09	-0.1	-0.05	-0.13	-0.02
	[0.00]	[0.01]	[0.21]	[0.00]	[0.69]
Age 45-49	-0.1	-0.17	-0.01	-0.13	-0.06
	[0.00]	[0.00]	[0.88]	[0.00]	[0.18]
Age 50-54	-0.04	-0.05	0.01	-0.09	0.02
	[0.02]	[0.18]	[0.74]	[0.01]	[0.68]
Age 55-59	-0.02	-0.04	0.07	-0.1	0.02
	[0.23]	[0.32]	[0.08]	[0.01]	[0.73]
Age 60-64	0.00	0.01	0.02	-0.06	0.1
	[0.96]	[0.9]	[0.58]	[0.14]	[0.03]
Age 65-69	-0.04	-0.05	0.01	-0.1	0.05
	[0.1]	[0.29]	[0.88]	[0.04]	[0.3]
Log(size)	-0.06	-0.09	-0.15	-0.02	-0.12
	[0.01]	[0.02]	[0.00]	[0.73]	[0.03]
# of generations	0.02	0.02	0.02	0.01	0.03
	[0.09]	[0.37]	[0.43]	[0.6]	[0.23]
Log(income)	0.16	0.12	0.21	0.19	0.21
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Log(gross wealth)	0.00	0	-0.1	-0.02	-0.09
	[0.78]	[0.00]	[0.00]	[0.48]	[0.00]
Log(avg prov. income)	0.59	0.54	0.7	0.7	0.48
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Log(avg prov. consumption)	-0.73	-0.58	-0.72	-0.86	-0.66
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$	-1.37	7.39	-3.69	-1.27	-11.64
	[0.5]	[0.16]	[0.44]	[0.74]	[0.00]
share^{20-30}	0.06	0.08	0.09	0.05	-0.04
	[0.02]	[0.13]	[0.12]	[0.31]	[0.60]
share^{30-40}	0.1	0.12	0.09	0.14	0.16
	[0.00]	[0.02]	[0.07]	[0.00]	[0.01]
Log(income) $\times$ share^{20-30}	0.06	0.18	0.00	0.02	0.19
	[0.2]	[0.07]	[0.98]	[0.87]	[0.06]
Log(income) $\times$ share^{30-40}	-0.07	0.04	0.02	-0.2	-0.22
	[0.01]	[0.55]	[0.85]	[0.00]	[0.00]
$\frac{y}{p^H} \times \text{share}^{20-30}$	-20.04	-53.62	-31.79	-1.31	-20.69
	[0.03]	[0.02]	[0.2]	[0.95]	[0.19]
$\frac{y}{p^H} \times \text{share}^{30-40}$	-0.11	-38.33	3.6	11.46	24.12
	[0.99]	[0.03]	[0.84]	[0.33]	[0.02]
R^2	0.09	0.13	0.11	0.09	0.08
$adj.R^2$	0.09	0.12	0.11	0.08	0.07
N	9155	2292	2309	2349	2067

p-values in brackets

To summarize our results, we find support in the data for a housing price effect on saving behavior: households with young members aged 20 to 30 tend to save more when they face a larger price of housing relative to their income. The estimated coefficient is statistically significant at conventional levels for 1995 and 2007, but the estimation is less accurate for 2002, where statistical significance just falls short of the 10% level. This effect seems to display some non-linearity, as predicted by the model: for all three years, we find a statistically significant coefficient for lower levels of wealth. We do not find a similar effect for young members aged 30 to 40.

Table 5: Robustness checks for 1995

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(income)	0.23	0.23	0.23	0.24	0.25	0.27	0.24
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$	-34.83	-36.15	-36.53	-35.21	-39.07	-30.47	-31.42
	[0.03]	[0.02]	[0.03]	[0.03]	[0.02]	[0.03]	[0.06]
share^{20-30}	0.11	0.11	0.12	0.11	0.12	0.08	
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.01]	
share^{30-40}	0.01	0.06	0.01	0.01	0.01	0.00	
	[0.84]	[0.01]	[0.85]	[0.74]	[0.67]	[0.90]	
Log(income) $\times$ share^{20-30}	0.11	0.11	0.12	0.11	0.1	0.08	
	[0.06]	[0.07]	[0.05]	[0.06]	[0.11]	[0.11]	
Log(income) $\times$ share^{30-40}	0.12	0.13	0.12	0.11	0.11	0.09	
	[0.01]	[0.00]	[0.01]	[0.01]	[0.02]	[0.02]	
$\frac{y}{p^H} \times \text{share}^{20-30}$	-141.93	-144.38	-142.89	-142.58	-137.09	-119.57	
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	
$\frac{y}{p^H} \times \text{share}^{30-40}$	-20.04	-18.79	-22.84	-20.54	-16.07	-15.5	
	[0.53]	[0.55]	[0.48]	[0.52]	[0.62]	[0.57]	
Age dummies	yes	no	yes	yes	yes	yes	yes
Age		-0.01					
		[0.00]					
Age^2		0.00					
		[0.00]					
Household with son			-0.15				
			[0.64]				
Prov. sex ratio			-0.35				
			[0.11]				
Single young man			-0.18				
			[0.68]				
Household with son $\times$ sex ratio			0.17				
			[0.6]				
Single young man $\times$ sex ratio			0.18				
			[0.68]				
Secondary education				0.03			
				[0.00]			
Tertiary education				-0.03			
				[0.13]			
Married					-0.12		
					[0.06]		
Single					-0.04		
					[0.56]		
Widow					0.12		
					[0.62]		
Number 20-30							0.04
							[0.00]
Number 30-40							0.01
							[0.36]
Log(income) x number 20-30							0.02
							[0.29]
Log(income) $\times$ number 30-40							0.03
							[0.05]
$\frac{y}{p^H}$ x number 20-30							-33.11
							[0.00]
$\frac{y}{p^H}$ x number 30-40							-5.92
							[0.56]
R^2	0.08	0.08	0.08	0.08	0.08	0.12	0.08
$adj.R^2$	0.08	0.08	0.08	0.08	0.08	0.11	0.08
N	6583	6861	6583	6583	6464	6583	6583

p-values in brackets

Note: The regressions also control for the size of the household, the number of generations, wealth, and average provincial income and consumption. In column (6), an alternative definition is used for the saving rate, where education expenditures are excluded from consumption.

Table 6: Robustness checks for 2002

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(income)	0.19	0.19	0.21	0.21	0.19	0.2	0.19
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$	-8	-8.61	-10.71	-8.08	-8.13	-4.11	-7.59
	[0.01]	[0.01]	[0.00]	[0.01]	[0.01]	[0.10]	[0.02]
share^{20-30}	0.01	0.02	0.00	0.01	0.02	0.02	
	[0.81]	[0.50]	[0.94]	[0.76]	[0.59]	[0.42]	
share^{30-40}	0.08	0.09	0.08	0.08	0.1	0.08	
	[0.07]	[0.00]	[0.05]	[0.06]	[0.02]	[0.02]	
Log(income) $\times$ share^{20-30}	-0.09	-0.05	-0.11	-0.09	-0.11	-0.13	
	[0.19]	[0.44]	[0.14]	[0.19]	[0.12]	[0.01]	
Log(income) $\times$ share^{30-40}	-0.11	-0.09	-0.13	-0.11	-0.12	-0.11	
	[0.03]	[0.08]	[0.01]	[0.03]	[0.02]	[0.01]	
$\frac{y}{p^H} \times \text{share}^{20-30}$	-23.52	-19.94	-19.81	-23.36	-21.47	11.36	
	[0.12]	[0.17]	[0.19]	[0.12]	[0.16]	[0.32]	
$\frac{y}{p^H} \times \text{share}^{30-40}$	-16.22	-16.13	-13.52	-15.42	-14.68	-3.18	
	[0.15]	[0.15]	[0.23]	[0.17]	[0.19]	[0.71]	
Age dummies	yes	no	yes	yes	yes	yes	yes
Age		-0.02					
		[0.00]					
Age^2		0.00					
		[0.00]					
Household with son			1.05				
			[0.02]				
Prov. sex ratio			-0.39				
			[0.22]				
Single young man			-0.63				
			[0.24]				
Household with son $\times$ sex ratio			-1.09				
			[0.02]				
Single young man $\times$ sex ratio			0.67				
			[0.23]				
Secondary education				-0.02			
				[0.52]			
Tertiary education				-0.06			
				[0.03]			
Married					-0.05		
					[0.22]		
Single					0.04		
					[0.42]		
Widow					-0.02		
					[0.78]		
Number 20-30							0.00
							[0.96]
Number 30-40							0.04
							[0.00]
Log(income) x number 20-30							-0.02
							[0.25]
Log(income) $\times$ number 30-40							-0.03
							[0.02]
$\frac{y}{p^H}$ x number 20-30							-5.97
							[0.16]
$\frac{y}{p^H}$ x number 30-40							-2.6
							[0.42]
R^2	0.06	0.06	0.07	0.07	0.07	0.09	0.06
$adj.R^2$	0.06	0.06	0.06	0.06	0.06	0.09	0.06
N	6431	6765	6431	6406	6406	6431	6431

p-values in brackets

Note: The regressions also control for the size of the household, the number of generations, wealth, and average provincial income and consumption. In column (6), an alternative definition is used for the saving rate, where education expenditures are excluded from consumption.

Table 7: Robustness checks for 2007

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log(income)	0.16	0.16	0.16	0.17	0.16	0.17	0.16
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
$\frac{y}{p^H}$	-1.37	-1.57	-1.38	-1.47	-0.92	-3	-0.92
	[0.50]	[0.42]	[0.49]	[0.47]	[0.65]	[0.09]	[0.66]
share^{20-30}	0.06	0.1	0.06	0.05	0.07	0.04	
	[0.02]	[0.00]	[0.02]	[0.03]	[0.01]	[0.06]	
share^{30-40}	0.1	0.09	0.1	0.1	0.1	0.08	
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	
Log(income) $\times$ share^{20-30}	0.06	0.07	0.06	0.05	0.06	0.05	
	[0.2]	[0.11]	[0.17]	[0.22]	[0.19]	[0.18]	
Log(income) $\times$ share^{30-40}	-0.07	-0.08	-0.07	-0.08	-0.07	-0.09	
	[0.01]	[0.01]	[0.01]	[0.01]	[0.02]	[0.00]	
$\frac{y}{p^H} \times \text{share}^{20-30}$	-20.04	-18.31	-20.76	-20.36	-19.72	-19.81	
	[0.03]	[0.04]	[0.03]	[0.03]	[0.03]	[0.02]	
$\frac{y}{p^H} \times \text{share}^{30-40}$	-0.11	1.24	-0.17	-0.01	-0.19	0.83	
	[0.99]	[0.84]	[0.98]	[1]	[0.98]	[0.88]	
Age dummies	yes	no	yes	yes	yes	yes	yes
Age		-0.01					
		[0.00]					
Age^2		0.00					
		[0.00]					
Household with son			-0.22				
			[0.41]				
Prov. sex ratio			-0.08				
			[0.63]				
Single young man			0.05				
			[0.88]				
Household with son $\times$ sex ratio			0.21				
			[0.44]				
Single young man $\times$ sex ratio			-0.06				
			[0.85]				
Secondary education				-0.05			
				[0.00]			
Tertiary education				-0.08			
				[0.00]			
Married					-0.07		
					[0.00]		
Single					-0.16		
					[0.00]		
Widow					-0.11		
					[0.01]		
Number 20-30							0.03
							[0.00]
Number 30-40							0.03
							[0.00]
Log(income) x number 20-30							0.01
							[0.35]
Log(income) $\times$ number 30-40							-0.03
							[0.00]
$\frac{y}{p^H}$ x number 20-30							-5.82
							[0.01]
$\frac{y}{p^H}$ x number 30-40							1.19
							[0.47]
R^2	0.09	0.08	0.09	0.09	0.09	0.09	0.09
$adj.R^2$	0.09	0.08	0.09	0.09	0.09	0.09	0.09
N	9155	9900	9155	9155	9154	9155	9155

p-values in brackets

Note: The regressions also control for the size of the household, the number of generations, wealth, and average provincial income and consumption. In column (6), an alternative definition is used for the saving rate, where education expenditures are excluded from consumption.

6 Conclusion

This paper has revisited the puzzle of the high savings rates of Chinese households, focusing on the role of housing prices and borrowing constraints. Insights from a theoretical model suggest that for financially constrained households (i.e. for households that face severe borrowing contraints and need to make substantial downpayments for the purchase of their houses), higher house prices will be associated with higher saving rates when the return on financial instruments is low compared to that on real estate. We present key stylized facts and empirical results on Chinese household savings rates, based on detailed micro data from the CHIP database. Regression results lend substantial support to the model. We find in particular that higher house prices are indeed associated with higher savings, ceteris paribus. This is especially the case in the age groups and the wealth levels predicted by the model. The results were subjected to a battery of robustness tests, which confirmed the main findings.

These results shed a new light on household savings in China, which have attracted much attention in the policy and academic debate. They may be especially relevant now that the housing sector in China is under intense scrutiny.

Our analysis, however, makes clear that the housing sector per se is not the cause of elevated saving rates. Rather, we find that high saving rates result from the combined effect of high housing prices, borrowing constraints, and the absence of alternative savings instruments that yield reasonable returns.

References

Ang, J., 2009. Household Saving Behaviour in an Extended Lifecycle Model: A Comparative Study of China and India. Journal of Development Studies, 45(8), 1344-1359.

Bacchetta, P., Benhima, K., Kalantzis, Y., 2012. Capital controls with international reserve accumulation: Can this be optimal? CEPR Discussion Paper 8753, Center for Economic Policy Research.

Baldacci, E., G. Callegari, D. Coady, D. Ding, M. Kumar, P. Tommasino, and J. Woo, 2010. Public Expenditures on Social Programs and Consumption in China. IMF Working Paper WP/10/69.

Banerjee, A., X. Meng and N. Qian, 2010. The Lifecycle Model and Household Savings: Micro Evidence from Urban China. Unpublished Manuscript, Peking University, http://afd.pku.edu.cn/files/09.pdf.

Barnett, S. and R. Brooks, 2010. China: Does Government Education and Health Spending Boost Consumption? IMF Working Paper WP/10/16.

Bernanke, B., 2005. The Global Saving Glut and the US Current Account Deficit. Remarks at the Homer Jones Lecture, St Louis, 14 April 2005.

Bernanke, B., 2005. The Chinese Economy: Progress and Challenges. Remarks at the Chinese Academy of Social Sciences, Beijing, China, December 15, 2006.

Bostic, R., S. Gabriel and G. Painter, 2009. Housing Wealth, Financial Wealth, and Consumption: New Evidence from Micro Data. Regional Science and Urban Economics, 39, 79-89.

Bracke, T., M. Bussiere, M. Fidora and R. Straub, 2009. A Framework for Assessing Global Imbalances. The World Economy, Wiley Blackwell, 33(9), 1140-1174.

Cagetti, M., 2003. Wealth Accumulation over the Life Cycle and Precautionary Savings. Journal of Business and Economic Statistics, 21(3), 339-353.

Campbell, J., and J. Cocco, 2005. How do House Prices Affect Consumption? Evidence from Micro Data. Harvard Business School Working Paper, 1, 1-57.

Case, K. E., J. M. Quigley and R. J. Shiller, 2005. Comparing Wealth Effects: The Stock Market versus the Housing Market. Advances in Macroeconomics, 5(1), article 1.

Chamon, M., and E. Prasad, 2010. Why Are Saving Rates of Urban Households in China Rising ? American Economic Journal : Macroeconomics, 2:1, 93-130.

Chamon, M., K. Liu and E. Prasad, 2010. Income Uncertainty and Household Savings in China. IMF Working Paper WP/10/289.

Chen, C.-L., C.-M. Kuan, and C.-C. Lin, 2007. Saving and Housing of Taiwanese Households: New Evidence from Quantile Regression Analyses. Journal of Housing Economics, 16, 102-126.

Chen, Y., M. Funke and A. Mehrota, 2011. What Drives Urban Consumption in Mainland China? The Role of Property Price Dynamics. BOFIT Discussion Papers 13/2011, Bank of Finland, Institute for Economies in Transition.

Coeurdacier, N., K. Jin and S. Guibaud, 2012. Credit Constraints and Growth in a Global Economy. Mimeo.

Deaton, A., 1989. Saving and Liquidity Constraints. NBER Working Papers, 1, 1-29.

Deaton, A. 2005. “Measuring Poverty in a Growing World (or Measuring Growth in a Poor World).” Review of Economics and Statistics, 87(1): 1-19.

Eichen, M. and M. Zhang, 1993. Annex: The 1988 Household Sample Survey—Data Description and Availability. In K. Griffin and R. Zhao, eds., The Distribution of Income in China, New York: St. Martin’s Press.

Engelhardt, G. V., 1996. House Prices and Home Owner Savings Behavior. Regional Science and Urban Economics, 26(3), 313-336.

Feng, J. He, L., and H. Sato, 2011. Public Pension and Housing Saving: Evidence from China. Journal of Comparative Economics, 39(4), 470-485.

Gervais, M., 2002. Housing taxation and capital accumulation. Journal of Monetary Economics 49 (7), 1461–1489.

Green, S., 2005. Making monetary policy work in china: a report from the money market front line. Stanford Center for International Development, Working Paper 245.

Handelshøyskole, N., 2011. Why Is the Chinese Savings Profile Upside Down? A LCH-Study of Chinese Household Savings. Master's thesis, Norwegian School of Economics.

Horioka, C., and J. Wan, 2007. The Determinants of Household Saving in China: a Dynamic Panel Analysis of Provincial Data. Journal of Money, Credit and Banking, 39(8), 2077-2096.

Jin, Y., 2011. Housing Wealth and Consumption in the Chinese Urban Economy: Evidence from Micro-Data. Undergraduate Honors Thesis under the direction of D. Romer, U.C. Berkeley, http://econ.berkeley.edu/sites/default/files/jin_yizhou.pdf.

Ju, J. and S.-J. Wei, 2010. Domestic Institutions and the Bypass Effect of Financial Globalization. American Economic Journal: Economic Policy, American Economic Association, 2(4), 173-204, November.

Khan, A.R., K. Griffin, C. Riskin, and R. Zhao, 1992. Household Income and its Distribution in China. China Quarterly, 132, 1029-1061. Kraay, A. 2000. Household Saving in China. World Bank Economic Review, 14(3), 545-570.

Laurens, B. J., Maino, R., 2007. China: Strengthening monetary policy implementation. IMF Working Paper WP/07/14, International Monetary Fund.

Li S. and C. Luo, 2010. Re-estimating the Income Gap between Urban and Rural Households in China. In M. K. Whyte, ed., One Country, Two Societies: Rural-Urban Inequality in Contemporary China, 105-121, Cambridge, MA: Harvard University Press.

Li, S., C. Luo, Z. Wei and X. Yue, 2008. Appendix: The 1995 and 2002 Household Surveys: Sampling Methods and Data Descriptions. In B. Gustafsson, S. Li and T. Sicular, eds., Inequality and Public Policy in China, New York: Cambridge University Press.

Li, S.-M., 2010. Mortgage Loan as a Means of Home Finance in Urban China: A Comparative Study of Guangzhou and Shanghai. Housing Studies, 25(6), 857-876

Lucas, R., 1990. Why Doesn't Capital Flow from Rich to Poor Countries? American Economic Review, 80(2), 92-6.

Luo, C., S. Li, T. Sicular, Q. Deng and X. Yue, 2011. Appendix I: The 2007 Surveys: Sampling Methods and Data Description. Western University, CIBC Working Paper, http://economics.uwo.ca/centres/cibc/wp2011/RisingInequality/AppendixI.pdf.

Luo, C., X. Yue, and S. Li (2011). A Critique of Wang Xiaolu's Estimation of the Gray income (dui Wang Xiaolu huisi shouru gusuande zhiyi). Bijiao, 52, 146-158.

Ma, G., and W. Yi, 2010. China's High Savings Rate: Myth and Reality. Bank for International Settlements Working Paper No. 312.

Meng, X., 2003. Unemployment, Consumption Smoothing, and Precautionary Saving in Urban China. Journal of Comparative Economics, 31(3), 465-485.

Modigliani, F., and S. L. Cao, 2004. The Chinese Saving Puzzle and the Lifecycle Hypothesis. Journal of Economic Literature, 42(1), 145-170.

Movshuk, O., 2009. Household Saving over the Life Cycle: International Evidence from Micro Data. Working Paper, University of Toyoma, 1, 1-30.

Nabar, M., 2011. Targets, Interest Rates and Household Saving in Urban China. IMF Working Paper, 223, 1-29.

National Bureau of Statistics of China, 2010a. Sales Price Indices of Buildings in 70 Medium-Large Sized Cities (2010.02). www.stats.gov.cn, February.

National Bureau of Statistics of China, 2010b. Sales Price Indices of Buildings in 70 Medium-Large Sized Cities (90 sq.meters and below) (2010.02). www.stats.gov.cn, February.

Sato, H., T. Sicular and X. Yue, 2011. Housing Ownership, Incomes and Inequality in China, 2002-2007. Western University, CIBC Working Paper No. 2011-12. http://economics.uwo.ca/centres/cibc/wp2011/Sato_Sicular_Yue12.pdf.

Song, J., T. Sicular and X. Yue, 2011. Appendix II: The 2002 and 2007 CHIP Surveys: Sampling, Weights, and Combining the Urban, Rural, and Migrant Samples. Western University, CIBC Working Paper, http://economics.uwo.ca/centres/cibc/wp2011/RisingInequality/AppendixII.pdf.

Wang, X. and W. T. Woo, 2011. The Size and Distribution of Hidden Household Income in China. Asian Economic Papers, 10(1), 1-26.

Wang, X. and Y. Wen, 2012. Housing Prices and the High Chinese Savings Rate Puzzle. China Economic Review, 23(2), 265-283.

Wei, S.-J., X. Zhang and Y. Liu, 2012. Status Competition and Housing Prices. NBER Working Papers 18000, National Bureau of Economic Research.

Wei, S.-J., X. Zhang, 2011. The Competitive Saving Motive: Evidence from Rising Sex Ratios and Savings Rates in China. Journal of Political Economy, 119(3), 511 - 564.

Yang, D. T., J. Zhang and S. Zhou, 2011. Why Are Savings Rates So High in China? IZA Discussion Paper No. 5465.

Yellen, J. L., 2011. Aggregate Demand and the Global Economic Recovery. Remarks at the Federal Reserve Bank of San Francisco, San Francisco, California, November 29, 2011.

Zhou, X., 2011. In: Banque de France Financial Stability Review on "global imbalances and financial stability", 2011, 165-169.

7 Appendix

7.1 Data

Our empirical analysis uses urban household survey data from the 1995, 2002 and 2007 rounds of the China Household Income Project (CHIP). Detailed explanation of the data can be found in Eichen and Zhang (1993), Li et al. (2008), Luo et al. (2011), and Song et al. (2011). Here we mention some features of the data relevant to our analysis.

The CHIP survey samples are subsamples of the NBS annual urban household survey sample. The 1995 and 2002 CHIP urban samples cover 12 provinces; in 2007 the sample was expanded by four additional provinces (Table 8).[8] In all years the provinces were selected so as to cover China's major regions (East, Center, West, and municipalities with provincial status, which, due to their distinctiveness, are treated as a separate regional category). The sample sizes are not proportional to regional or provincial populations, and Song et al. (2011) therefore suggests use of population-based weights. Song et al. (2011) use population information from China's 2000 census and the 2005 population sample survey to develop weights for the 2002 and 2007 CHIP samples. We use these weights in our analysis of the 2002 and 2007 data. For 1995 we do not use weights.

The CHIP datasets contain detailed information about household income and expenditures, as well as other relevant individual and household variables. Some of this information is supplied to CHIP by the NBS from its annual survey, and some of it is additional information collected using a supplementary questionnaire designed by the CHIP research group. Information on household income and expenditures, which we use to calculate household savings, are supplied by the NBS, which collects them using household income and expenditure diaries.

[8] The provinces in the 1995 and 2002 CHIP surveys were Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Hubei, Guangdong, Chongqing (included in Sichuan in 1995; separate province in 2002), Sichuan, Yunnan and Gansu. In 2007 Shanghai, Zhejiang, Fujian, and Henan were added.

Table 8: CHIP Urban Household Survey Samples

	Provinces	Households	Individuals
1995	11	6934	21696
2002	12	6835	20632
2007	16	10000	29262

The NBS's measure of urban disposable income (*ke zhipei shouru*), includes employee income (wages, salaries and other compensation), net self-employment income, net property income, and transfer income from public and private sources net of taxes and fees. Consumption expenditures include household spending on food; clothing and footwear; household appliances; medical care and health; transport and communications; recreational, educational, and cultural services; housing; and other miscellaneous items.

Certain biases in the NBS urban survey sample are well known, and these are shared by the CHIP sample. Informal rural-urban migrant households are not well represented. The sample is largely composed of households that are long-term, formal urban residents with urban residence permits (*hukou*). Rich households are also thought to be under-represented, and the degree of under-representation has probably increased in recent years. With the growth of private businesses and privately owned assets in China, and with the emergence of an ultra-rich segment of society, rich households have become less willing to participate in the surveys, and households that participate are more likely to underreport certain types of income. The resulting "grey" income question has been the subject of some discussion recently (Wang and Woo 2011, Luo, Yue and Li 2011).

If the propensity to save of under-represented groups and out of grey income differs from that of other groups and out of reported income, then the estimates of savings based on the NBS and CHIP data will be biased (Kraay 2010). Chamon and Prasad (2011) examine the savings rates of richer households in the NBS urban household survey and find that they do show higher propensity to save. It is likely, therefore, that average savings rates calculated using the NBS urban survey data are understated. We mention this issue, as it should inform the analysis and interpretation of results.

Some researchers have noted problems with the NBS measure of disposable income. As discussed in Khan et al. (1992), the NBS income measure does not include imputed rents on owner-occupied housing. Excluding imputed rents will cause some overstatement of household savings rates (Ma and Yi 2010). NBS income also does not include the value of consumption subsidies, although these have been substantially reduced following the elimination of planned allocation of consumer goods in the early 1990s and with the privatization of urban housing, which was basically completed by the early 2000s. Subsidized rental housing continues for a small segment of urban households. In our analysis we have used estimates of imputed rents developed by Sato et al. (2011) to compute a measure of income that includes imputed rents, and we recalculate savings rates accordingly (to recalculate savings rates, we include imputed rents in both income and consumption). Due to lack of consistent information on consumption subsidies across years, our income variable does not include such subsidies.

NBS income includes income received through cash-transfer programs, such as unemployment insurance, pensions, and the minimum living guarantee program, but it does not include employer contributions to pension and social insurance programs for employees. Also, it does not include the value of subsidies implicit in social programs and public services such as education. Some researchers (e.g., Li and Luo 2010) have argued that the market value of such programs should be estimated and included as an income component. Due to difficulties in collecting the relevant information and differing views about what should be included and how it should be computed, as well as the fact that most international studies do not include such types of income, here we do not include them. The exclusion of these income components may cause some underestimation of the level of income and of savings.

7.2 Model: proofs of the Propositions

Consider the case $r < R$. We start by showing that optimal plans fall into the three categories described in section 4.4: type (a), (b), or (c).

Denote λ_t the Lagrange multiplier associated to the budget constraint in period t (see equations (2)- (4)), γ_2 the Lagrange multiplier associated to the minimum investment requirement (7), and δ_2 the Lagrange multiplier associated to the borrowing constraint at $t = 2$.[9]

The first-order conditions with respect to A_2 and d_2 are:

$$\zeta_2 \lambda_2 = \zeta_2 \beta R \lambda_3 + \gamma_2,$$
$$\lambda_2 = \beta r \lambda_3 + \delta_2.$$

Suppose $D_2 = 0$. The multiplier γ_2 can be zero or strictly positive. If $\gamma_2 = 0$, then, from (7), we have $A_2 > \underline{A}$. This is a plan of type (a). If $\gamma_2 > 0$, then the household invests exactly $A_2 = \underline{A}$ at $t = 1$. This is a plan of type (b). Suppose now $D_2 > 0$, then $\delta_2 = 0$ and $\gamma_2 = \zeta_2 \beta (r - R) \lambda_3$. Since $r < R$, this implies $\gamma_2 = 0$ and therefore $\zeta_2 = 0$. Then, the household does not invest in the real asset. This is a plan of type (c).

Denote U^a, U^b, U^c, and d_1^a, d_1^b, d_1^c, the utilities and first-period deposits associated to those three types.

With logarithmic utility, it is straightforward to show that optimal plans of type (a) and (c)

[9] Assumption 1 makes sure that the optimal level of deposits d_1 is non-negative. For notational convenience, we discard the borrowing constraint at $t = 1$ (5).

are characterized by the following demand for deposits:

$$d_1^a = d_1^c = s_1^* = \frac{\beta(1+\beta)(1+rd_0) - \frac{g}{r}}{1+\beta+\beta^2},$$
$$A_2^a = d_2^c = s_2^* = \frac{\beta^2[r(1+rd_0)+g]}{1+\beta+\beta^2}.$$

s_1^* is positive by assumption 1.

To be feasible, an optimal plan of type (a) requires that $A_2^c > \underline{A}$. This condition can be rewritten $w > \frac{1+\beta+\beta^2}{\beta^2}\frac{\underline{a}}{r}$.

An optimal plan of type (b) is characterized by the following demand for deposits:

$$d_1^b = \frac{\beta(1+rd_0) - \frac{g-\underline{a}}{r}}{1+\beta}.$$

We have $d_1^b > d_1^a, d_1^c$ when $w < \frac{1+\beta+\beta^2}{\beta^2}\frac{\underline{a}}{r}$.

For $w < \frac{1+\beta+\beta^2}{\beta^2}\frac{\underline{a}}{r}$, a plan of type (a) is not feasible, so the optimal plan has to be of type (a) or (b). When the intertemporal wealth w approaches $\underline{a}$ from above, it can be shown that consumption C_1^b in a type-(b) plan goes to 0, Then, utility U^b goes to $-\infty$ and we have $U^b < U^c$. On the contrary, when $w = \frac{1+\beta+\beta^2}{\beta^2}\frac{\underline{a}}{r}$, the type (b)-plan coincides with the type (a)-plan which dominates the type (c)-plan because of the larger return on saving at $t=2$. So, $U^b > U^a$. Therefore, there is a threshold $w^* \in \left(\frac{\underline{a}}{r}, \left[\frac{1+\beta+\beta^2}{\beta^2}\right]\frac{\underline{a}}{r}\right)$ such that $U^a = U^b$ when $w = w^*$, $U^a > U^b$ when $w < w^*$ and $U^a < U^b$ when $w > w^*$. Replacing C_t^b and C_t^c by their value in $U^a = U^b$, we can show that the threshold w^* satisfies:

$$(1+\beta+\beta^2)\log w^* - (1+\beta+\beta^2)log(1+\beta+\beta^2) = \\ (1+\beta)\log\big(w^* - \frac{\underline{a}}{r}\big) + \beta^2\log\big(\frac{\underline{a}}{r}\frac{R}{r}\big) - (1+\beta)\log(1+\beta) - \beta^2\log(\beta^2). \quad (12)$$

This can be further simplified to

$$\left(\frac{w^*}{1+\beta+\beta^2}\right)^{1+\beta+\beta^2} = \left(\frac{w^* - \frac{\underline{a}}{r}}{1+\beta}\right)^{1+\beta}\left(\frac{\frac{\underline{a}}{r}\frac{R}{r}}{\beta^2}\right)^{\beta^2}. \quad (13)$$

Note that w^* is homogeneous of degree 1 with respect to $\frac{\underline{a}}{r}$. We can therefore look for a solution $w^* = \omega\frac{\underline{a}}{r}$. Equation (13) then simplifies to $\psi_{\frac{1+\beta+\beta^2}{\beta^2}}(\omega) = \frac{R}{r}$ where ψ_u is defined by

$$\psi_u(x) = \frac{x(1+(x-1)^{-1})^{u-1}}{u(1+(u-1)^{-1})^{u-1}}.$$

The function ψ_u is continuous, differentiable and strictly decreasing on $(1,u]$, with $\psi_u(u) = 1$ and

$\lim_{x \to 1} \psi_u(x) = +\infty$. Therefore, it has an inverse on that interval. Denote φ the inverse of $\psi_{\frac{1+\beta+\beta^2}{\beta^2}}$ defined on $[1, +\infty)$. The mapping φ has the properties stated in Proposition 2. The threshold is then given by $w^* = \frac{\underline{a}}{r}\varphi\left(\frac{R}{r}\right)$.

Finally, the difference $d_1^b - d_1^a$ is given by:

$$d_1^b - d_1^a = \frac{1}{1+\beta}\left[\frac{\underline{a}}{r} - \frac{\beta^2}{1+\beta+\beta^2}w\right],$$

which increases with $\underline{a}$ and decreases with w.

This proves Proposition 2.

The proof of Proposition 1 is straightforward: when $r = R$, the set of intertemporal wealth w such that $\frac{\underline{a}}{r}\varphi\left(\frac{R}{r}\right) \leq w < \frac{\underline{a}}{r}\varphi(1)$ is empty and type (b) plans are impossible. In addition, deposits and the real asset become perfect substitutes. Then, $d_1 = s_1^*$ and $d_2 + \zeta_2 a_2 = s_2^*$.

7.3 Descriptive statistics

	1995	2002	2007
Number of individuals	21726	20636	30340
Number of households	6931	6835	10000
Number of provinces	11	13	16

Table 9: General descriptive statistics

Variable	#	First Quartile	Mean	Median	Std Deviation	Third Quartile
Saving rate 1995	6931	-0.296	-0.189	-0.090	0.605	0.079
Saving rate 2002	6835	-0.090	0.010	0.097	0.588	0.276
Saving rate 2007	10000	0.013	0.146	0.205	0.404	0.380
Saving rate (durables excl. from cons.) 1995	6931	-0.211	-0.107	-0.035	0.539	0.126
Saving rate (durables excl. from cons.) 2002	6835	-0.003	0.118	0.152	0.328	0.313
Saving rate (durables excl. from cons.) 2007	10000	0.066	0.202	0.242	0.339	0.405
Age of top income member 1995	6931	37.000	45.242	43.000	11.839	54.000
Age of top income member 2002	6835	37.000	44.785	44.000	11.202	51.000
Age of top income member 2007	10000	39.000	47.583	46.000	12.511	55.000
Share of age 20-30 1995	6931	0.000	0.128	0.000	0.213	0.250
Share of age 20-30 2002	6835	0.000	0.101	0.000	0.171	0.250
Share of age 20-30 2007	10000	0.000	0.110	0.000	0.181	0.333
Share of age 30-40 1995	6931	0.000	0.216	0.000	0.283	0.500
Share of age 30-40 2002	6835	0.000	0.169	0.000	0.258	0.333
Share of age 30-40 2007	10000	0.000	0.176	0.000	0.262	0.333
Income-price ratio 1995	6931	5.679	10.204	8.625	6.613	12.835
Income-price ratio 2002	6835	26.238	48.010	41.083	30.436	61.828
Income-price ratio 2007	10000	10.906	22.632	17.643	18.190	28.515
Province sex ratio 1995	6931	0.970	0.983	0.990	0.031	0.996
Province sex ratio 2002	6835	0.963	0.975	0.978	0.022	0.990
Province sex ratio 2007	10000	0.955	0.979	0.976	0.030	0.990
Income 1995	6931	9060.000	14072.281	12408.000	8094.421	16801.000
Income 2002	6835	14121.000	24058.299	20734.000	15108.310	29601.830
Income 2007	10000	26600.000	47874.254	39452.295	34298.772	58796.073
Gross wealth 1995	6931	10700.000	38040.598	23000.000	57808.805	45075.000
Gross wealth 2002	6835	30975.000	84204.178	58000.000	110939.838	102226.664
Gross wealth 2007 (proxy)	10000	25000.000	78711.389	50000.000	111214.095	100000.000
Province avg income 1995	6931	11781.082	14398.718	12996.085	4405.832	15849.212
Province avg income 2002	6835	19992.933	24622.886	21443.821	7444.149	26421.314
Province avg income 2007	10000	36851.245	49999.609	42057.876	16039.973	65177.959
Province avg consumption 1995	6931	13428.837	16256.617	14846.711	4707.770	17806.564
Province avg consumption 2002	6835	19279.240	23341.262	21936.925	5688.893	23239.282
Province avg consumption 2007	10000	31246.502	40125.269	35575.853	12438.105	49102.835

Table 10: Descriptive statistics for the variables used in the regressions

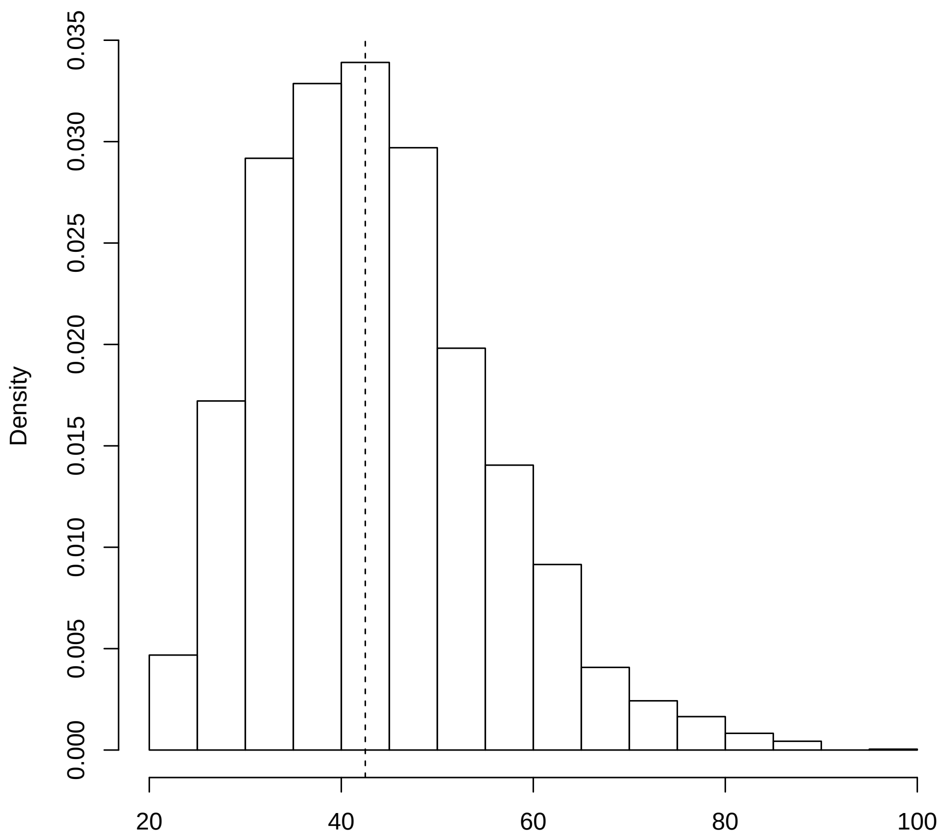

Figure 1: Distribution of home purchase age in 2002

Note: The dashed line indicates the mode of the distribution.

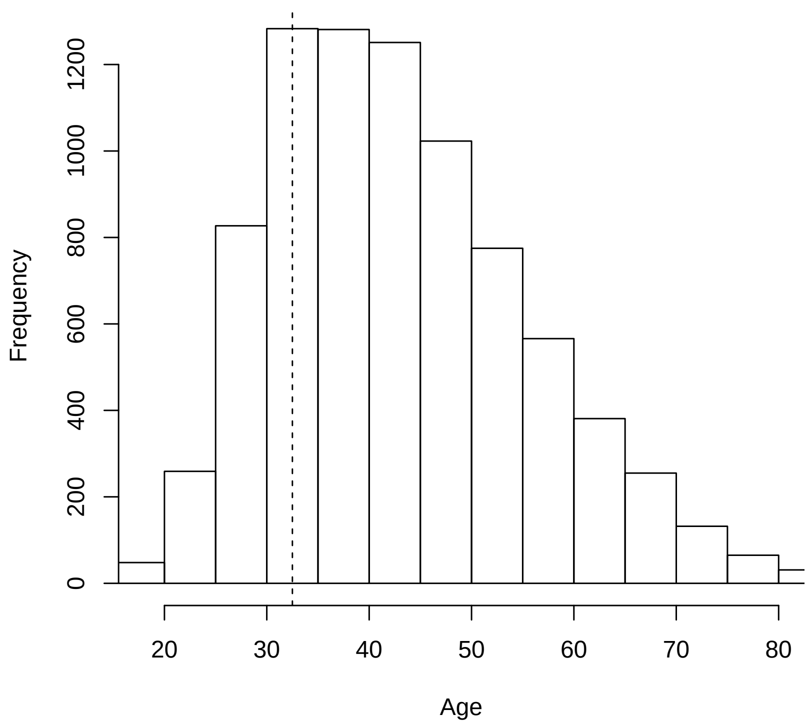

Figure 2: Distribution of home purchase age in 2007

Note: The dashed line indicates the mode of the distribution.

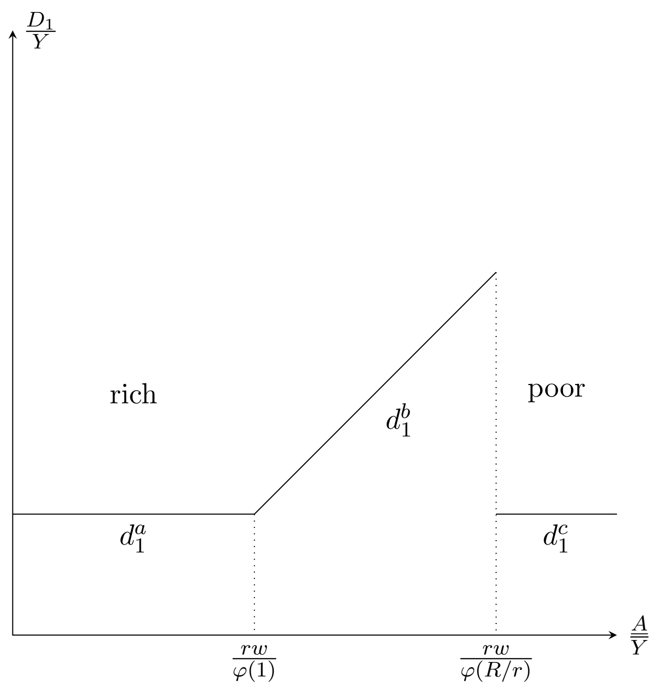

Figure 3: Saving of young household as a function of minimum investment requirement

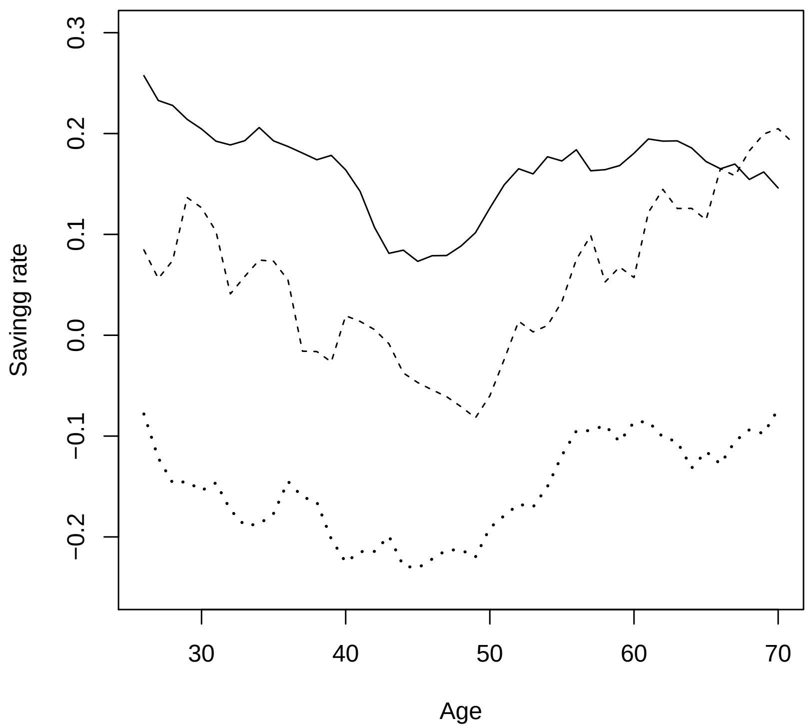

Figure 4: Average saving rate by age of top income member

Note: Solid line: 2007, dashed line: 2002, dotted line: 1995. Saving rates take into account imputed rents, consumption includes durables and education expenses. The age relates to the household member with the highest income. Averages are computed as a three-year moving average.

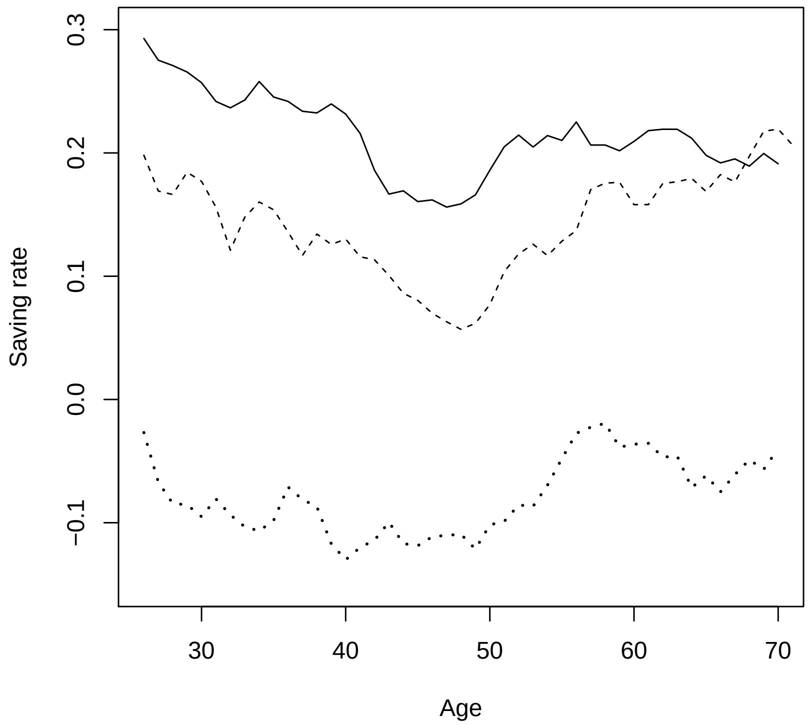

Figure 5: Average saving rate by age of top income member (excluding durables from consumption) Note: Solid line: 2007, dashed line: 2002, dotted line: 1995. Saving rates take into account imputed rents, consumption excludes durables but includes education expenses. The age relates to the household member with the highest income. Averages are computed as a three-year moving average.