

# Self-Assessed Financial Literacy in Housing Markets\*

Boaz Abramson

Department of Economics  
Stanford University  
boaza@stanford.edu

Andres Yany

Department of Economics  
Stanford University  
ayany@stanford.edu

July 2021

## Abstract

This paper introduces a novel dimension of household heterogeneity that plays an important role in housing markets. Households who self-assess themselves to be more financially literate are 1) more likely to own a house and 2) take on higher leverage on their home. We solve a heterogeneous agent portfolio choice model to infer the role of mortgage terms and of expectations on future house prices for the empirical patterns. We find that households with higher levels of self-assessed financial literacy are in fact better at the parts of the transaction that are relevant to them, namely access to more accommodating mortgage terms when they are young and better risk-return trade-offs when they are old. Moreover, by ignoring heterogeneity in financial literacy, standard models introduce quantitatively substantial biases in evaluating housing market policies. Housing demand elasticity with respect to wealth is downsized by approximately 40% when taking financial literacy into account.

---

\*For valuable guidance we thank Monika Piazzesi and Martin Schneider. We have also benefited from helpful suggestions by Adrien Auclert, Rodrigo Carril, Liran Einav, Pete Klenow, Sean Myers, Alessandra Peter and Chris Tonetti, as well as participants of the Stanford Macro Seminar.

# 1 Introduction

This paper asks why households differ in their housing and mortgage choices. A large literature on portfolio choice with housing relates such choices to standard household characteristics such as age, income and wealth (e.g. [Campbell and Cocco 2003](#), [Cocco 2004](#)). These models assume households have similar home ownership preferences, have the same expectations on future house prices and face the same credit constraints in the housing markets. In these models, the only reason some households become renters rather than owners is due to insufficient wealth or because they are too young and unwilling to save enough to buy a house. However, the data tells a different story. In fact, age, income and wealth can account for only a small share of the cross sectional variation in housing behavior. Hence, the standard model is limited in its ability to capture the way decisions are made in the housing markets. The evaluation of housing market policies based on such models might therefore be misleading.

We study the role of financial literacy in housing markets. First, we show that financial literacy is a dimension of household heterogeneity that matters for housing outcomes. Exploiting a novel feature of the recent Survey of Consumers Finances (SCF), we document a robust relationship between self-assessed financial literacy and behavior in housing markets. The basic stylized fact is that households reporting higher levels of financial literacy are 1) more likely to own a house rather than rent one, and 2) tend to take on more levered positions to finance their house. The relationship is economically meaningful even after controlling for the potential confounding factors - e.g. education, income, wealth, age, gender, degree of risk aversion, usage of financial intermediaries, etc.

Following the empirical findings, we then ask what are the channels through which financial literacy matters for housing tenure and leverage. We consider two main candidates. First, households reporting higher levels of financial literacy might search for better deals in the credit markets, thereby paying lower interest rates on their mortgages and facing laxer down-payment requirements. Second, expectations on future house prices might be important. More literate households might expect different risk-return trade-offs in the housing market due to access to better property deals.

To examine the role of these channels for the observed empirical patterns, we solve a standard heterogeneous life-cycle model of portfolio choice with housing. As we cannot observe households beliefs on future house prices nor the mortgage contracts they engage in, a structural model is in place. The key new feature in the model is that we allow for heterogeneity along the dimension of self-assessed financial literacy. First, households may face different mortgage terms, appealing to the possibility that financially literate

households search and negotiate for cheaper and larger credit. We model heterogeneity in mortgage terms by allowing both the mortgage spread and the down-payment requirement to differ across households. Second, households with higher self-assessed financial literacy may have access to better investment opportunities in housing markets, due to ,e.g., sophisticated search skills. We model this form of heterogeneity by allowing both the mean return and volatility of the distribution from which future house values are drawn to differ by literacy.

While there might be additional potential channels that can rationalize our empirical findings, we focus on those which are more intuitive and are arguably the most important for housing outcomes. Introducing heterogeneity along the mortgage terms households face and the expectations they have on future house values allows the model to overwhelmingly account for the stylized facts. As in the data, model-implied homeownership rates and loan-to-value ratios are increasing with self-assessed financial knowledge, after controlling for income, wealth and age. By fitting the data, our model suggests that heterogeneity in mortgage terms and in expectations on future house prices can indeed account for the link between self-assessed financial literacy and housing outcomes that is evident in the data.

To implement the model quantitatively, we use SCF micro data on balance sheets, income, and demographic characteristics of a representative sample of U.S. households. Following the meaningful variation apparent in the data, we categorize households into three types by their self-assessed financial knowledge: low, intermediate and high. The estimation needs to identify four groups of parameters: 1) mean expected return on the housing asset, 2) expected volatility of this return, 3) down-payment constraints, and 4) mortgage spreads. To accommodate for heterogeneity in financial literacy, each group of parameters consists of a tuple of three. We estimate the parameters by applying a Simulated Method of Moments (SMM) design. The data moments we use are house ownership and loan-to-value (LTV) ratios across the three groups of literacy and across three age groups of households: young, middle-aged and old. We also estimate the discount factor in order to match the aggregate wealth-to-income ratio in the economy.

We find that households who self-assess themselves as more literate face laxer constraints in the credit markets - they pay a lower spread when borrowing against the value of their house, and are required to a lower down-payment on their mortgage. More literate households also face better risk-return trade-offs in the housing market - the Sharpe ratio of the housing asset is increasing with financial literacy, suggesting that more literate households engage in better investment deals. On average, our estimates suggest that differences in credit markets are quantitatively more important for explaining why house-

holds with higher levels of self-assessed financial literacy are more likely to own a house and tend to take on higher leverage. The empirical analysis points out that differences in ownership rates are more stark across literacy groups, compared to differences in loan-to-value ratios. The estimation therefore uses the literacy types to mostly accommodate for the differences in home-ownership. For this, credit constraints are more important than expectations on future prices. Laxer credit constraints matter more for ownership decisions, while expectations matter more for leverage.

The quantitative exercise illustrates how the role of financial literacy varies across the life cycle. By alternatively allowing for only one of the two channels of heterogeneity, we show that households who self-assess themselves as more financially sophisticated tend to be better at those parts of the transaction that are more relevant to them. At the beginning of life, when mortgage terms matter more for housing decisions, those with higher levels of self-assessed literacy manage to search for lower mortgage spreads and down-payment requirements. Later on, when returns on housing are more relevant, they make deals with better risk-return trade-offs.

A main result of our analysis is that incorporating financial literacy heterogeneity in an otherwise standard housing portfolio choice model matters for evaluating housing market policies. Housing demand elasticity with respect to wealth is downsized by approximately 40% when taking financial literacy into consideration. To see this, we compare our model to a benchmark model in which heterogeneity in financial literacy plays no role. That is, we estimate the model while restricting housing market parameters to be equal across the reported levels of literacy. This allows us to quantitatively evaluate the cost of abstracting from the role financial literacy plays in housing markets.

The benchmark estimation targets the average home-ownership rate and loan-to-value ratio across the three age-groups. While this benchmark model can fit the targeted aggregate life cycle dynamics of both housing tenure and leverage, ignoring the role of financial literacy introduces substantial biases in the evaluation of housing market policies. We compute the aggregate housing demand elasticity to a wealth shock with and without financial literacy heterogeneity. This echoes debated policies that are designed to incentivize house ownership. A 10% increase in the wealth of young households leads to a 10% increase in housing demand in the benchmark model, but only to a 6.4% increase in housing demand in the heterogeneous agent model. Similarly, a 10% wealth shock targeted at the young and poor households leads to a 20% increase in housing demand in the benchmark model, but only to a 11% increase in housing demand in the heterogeneous agent model. Evaluating the effect of housing market policies to encourage ownership therefore crucially depends on whether or not heterogeneity in financial literacy is taken

into account.

The underlying reason for this discrepancy is the following. While the benchmark model is able to match the life cycle dynamics of housing tenure and leverage, this comes at an overlooked cost. By ignoring the observed heterogeneity in financial literacy, the model over-estimates the correlations between housing outcomes and wealth and age. The benchmark model matches life cycle housing moments using the wrong mechanism - it underestimates the large variability in ownership and leverage within wealth and age bins as it appears in the data. Indeed, when a wealth shock hits, households with the same wealth and age adjust their demand for housing in the same fashion. By incorporating heterogeneity in the observed literacy dimension, the heterogeneous model is able to significantly reduce the correlation between ownership and wealth and age. Since households of different literacy types face different mortgage terms and have different expectations on future house prices, they now respond differently to a wealth shock. For example, the elasticity of housing demand with respect to wealth is lower for low literacy types since the credit constraints they face are more strict. The effect of a wealth shock in the heterogeneous agent model is therefore mitigated.

Indeed, adding a new source of heterogeneity in any dimension will mechanically reduce the excessive correlations between wealth and housing outcomes that is generated by the benchmark model. To what degree does heterogeneity in a certain dimension matter for housing markets is therefore a quantitative question. Our policy experiment suggests that self-assessed financial literacy plays an important role in the housing markets and should hence be incorporated in housing choice models.

This paper is the first to study the role of financial sophistication in the housing markets. Optimal portfolio choice models have increasingly highlighted the role of financial literacy in the stock markets as a source of heterogeneity that can help solve traditional household finance puzzles. [Lusardi, Michaud and Mitchell \(2017\)](#) shows that incorporating financial literacy in an otherwise classic stochastic consumption-savings life cycle model enhances its ability to generate the substantial wealth inequality observed in the data. Other studies relate financial literacy to stock market participation ([Van Rooij, Lusardi and Alessie \(2011\)](#)), diversification ([Guiso and Jappelli \(2008\)](#)) and performance in financial markets ([Campbell, Ramadorai and Ranish \(2013\)](#), [Hackethal, Haliassos and Jappelli \(2012\)](#), [Calvet, Campbell and Sodini \(2009\)](#)). We provide first evidence that financial literacy is also important in the housing markets. We then show that this relationship is quantitatively important for evaluating housing market policies.

More broadly, we point to an important limitation of the standard life-cycle models of portfolio choice with housing. These models were originally developed by [Campbell and](#)

Cocco (2003), Cocco (2004) and Yao and Zhang (2004) and have since been used to evaluate housing market policies. For example, Li, Liu and Yao (2009) analyze how changes in lending conditions affect households decisions, and Bajari et al. (2013) predict how severe a housing bust is expected to be. Our paper shows that in order to precisely measure the way households respond to changes in the housing market, we should incorporate heterogeneity in financial literacy to the standard model.

Our paper relates to the a large literature on the role of credit conditions in housing markets. Mian and Sufi (2009), Mian and Sufi (2011) and Di Maggio and Kermani (2017) provide empirical evidence that laxer credit conditions inflate house prices via increased demand for housing. Other studies (e.g. Landoigt, Piazzesi and Schneider (2015), Kiyotaki, Michaelides and Nikolov (2011)) calibrate life cycle models in order to assess the significance of cheap credit in a general equilibrium framework. Our paper suggests that credit conditions are not homogeneous across households. Low financial literacy types face tighter constraints and pay a larger spread on their mortgage. This in turn translates to a lower elasticity of housing demand with respect to wealth.

With respect to the role of expectations in housing markets, several papers have used surveys to measure expectations on future house prices. Notably, Case, Quigley and Shiller (2003) shows that the expectations of home buyers in the U.S. are largely affected by recent experience. Following a house-boom, households expect large price appreciations and low levels of risk. Piazzesi and Schneider (2009) show that even at the height of the boom there are households who believe it is a good time to buy a house because house prices will rise further. Several papers have studied the role of expectations in shaping behavior in housing markets<sup>1</sup>. Our paper contributes to this literature by showing that households who are more literate than others might hold different expectations on future house prices, e.g. due to sophisticated search skills. Our quantitative analysis suggests that expectations can be elicited based on observable literacy types.

A contribution of this paper is to highlight the relevance of *self-assessed* financial literacy in the housing markets over traditional measures of literacy. Most of the literature relating financial sophistication to economic behavior has solely focused on testable measures of literacy, most prominently surveys measuring the ability of respondents to answer simple questions with respect to inflation, interest rates, and diversification. While these methods arguably capture a dimension of financial literacy, it is unclear whether people make choices based on their success in such tests<sup>2</sup>. In fact, our data shows that

---

<sup>1</sup>Examples include Burnside, Eichenbaum and Rebelo (2011), Glaeser, Gottlieb and Gyourko (2012) and Glaeser and Nathanson (2017).

<sup>2</sup>In fact, respondents typically do not receive feedback on their answers. Therefore they do not necessarily know how literate they actually are.

this is not the case. The relationship between such measures and house tenure decisions dissipates once we control for observed demographic characteristics. On the contrary, our novel measure of self-assessed literacy is a remarkably powerful and robust predictor.

There are two possible ways to reconcile why self assessed literacy matters for housing outcomes. First, respondents might be better at evaluating their own literacy level. If this is the case, rather than presenting respondents with somewhat arbitrary questions on inflation and interest rates, we should simply ask them to self-assess their literacy. Second, self-assessed financial literacy might measure the degree of self-confidence in financial markets, which in turn can be important for how households behave in the housing markets. For example, households who self-assess themselves as highly literate might have over-optimistic beliefs on the risk-return trade-off they face in the housing markets and therefore tend to own and lever more. Our quantitative results suggest that this mechanism alone cannot underlie the empirical patterns. While the estimation strategy cannot distinguish between heterogeneous beliefs on future housing returns and heterogeneity in the fundamental distribution of returns, we show that heterogeneity in fundamental mortgage terms is crucial to match the stylized facts. That is, self-assessed financial literacy is in fact a proxy for literacy itself. This is important for future work studying the relationship between financial literacy and economic behavior.

The paper proceeds as follows. Section 2 presents the stylized facts relating self-assessed financial knowledge to housing market outcomes. Section 3 introduces a heterogeneous agent life-cycle model of optimal portfolio choice with housing. Section 4 discusses the Simulated Method of Moments estimation approach. Section 5 discusses the estimation results, while Section 6 evaluates the costs of abstracting from heterogeneity in financial literacy. Section 7 concludes.

## 2 Empirical Results on Self-Assessed Financial Literacy

We begin by documenting some patterns in survey data. The 2016 SCF wave offers a novel approach to measuring financial literacy and relating it to economic behavior. To the best of our knowledge, this is the first population representative sample to measure self-assessed financial literacy. The 2016 wave asks respondents the following:

*“On a scale from zero to ten, where zero is not at all knowledgeable about personal finance and ten is very knowledgeable about personal finance, what number would you (and your partner) be on the scale?”*

Figure 1 plots the SCF data on self-assessed financial literacy and housing market outcomes - the proportion of households who own a house (left panel) and the ratio of col-

lateralized debt to house value for home owners (right panel). The basic stylized fact is that households who self-assess themselves as more financially literate are 1) more likely to own a house and 2) tend to take a more levered position on their house.

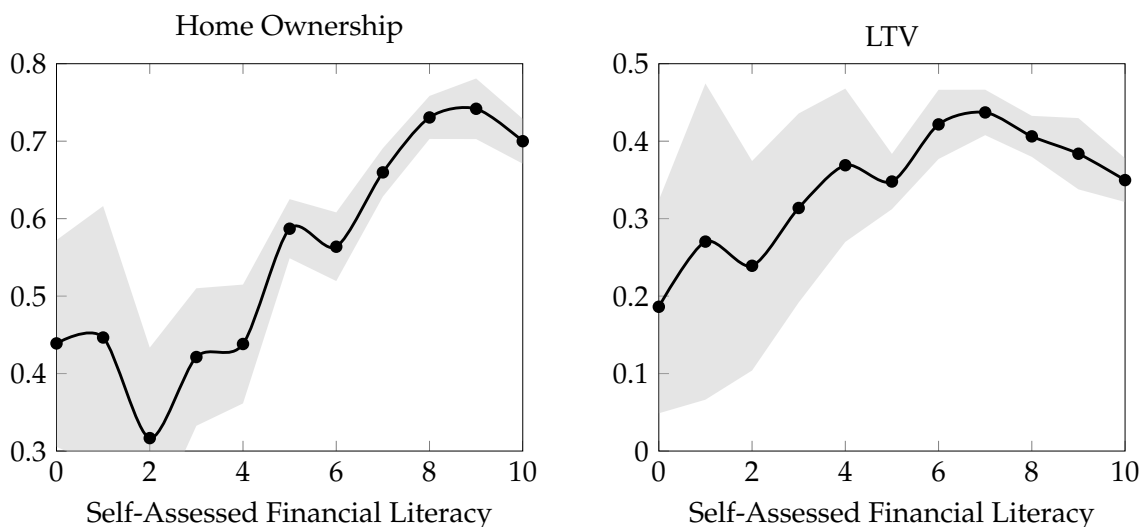


Figure 1: Self-Assessed Financial Literacy in the Housing Markets

Notes: SCF data. Each dot is a weighted average of households that self-assess their financial knowledge on a scale of 0-10. House ownership measures whether or not the household owns a ranch/farm/mobile home/house/condo. The Loan-To-Value ratio is computed for home owners as the ratio of housing collateralized debt to house value. Lines are kernel-weighted local polynomial regressions. 95% confidence intervals are plotted. Standard errors are computed using the “scfcombo” Stata package in order to account for the SCF complex sample specification as well as the multiple imputation process.

How should we interpret this new measure of financial literacy? To address this question, Table 1 provides descriptive statistics for three groups of households: those who assess their financial literacy to be low (0-4 on scale), intermediate (5-7) and high (8-10)<sup>3</sup>.

First, households who report higher levels of financial literacy might indeed be more literate. As this table illustrates, they are more educated, they score higher grades in finance related questions<sup>4</sup>, and are more likely to consult with financial advisers. Second, self-assessed financial literacy might simply measure the degree of self-confidence in financial markets. Indeed, households with higher levels of self-assessed financial literacy

<sup>3</sup>While our results are robust to the exact pooling of households into groups, the data suggests a significant intra-group variation in outcomes, larger than the inter-group variability

<sup>4</sup>The grade is computed as the number of correct answers to the following questions: 1) Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than today, exactly the same as today, or less than today with the money in this account? 2) Suppose you had \$100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than \$102, exactly \$102, or less than \$102? 3) Do you think that the following statement is true or false: buying a single company’s stock usually provides a safer return than a stock mutual fund?



are willing to take on more risk and tend to participate more in the stock markets.

With respect to demographic characteristics, households who self-assess their literacy to be higher than others are more likely to be males and tend to be wealthier.

One concern with this literacy measure is that it is simply a linear combination of other, already observed, household characteristics. To alleviate these concerns, we treat self-assessed financial literacy as the dependent variable in an OLS regression model where all the outcomes from Table 1 serve as predictors. The  $R^2$  from this regression is only 0.1 – suggesting that financial literacy is a meaningful measure over and above its relation with other observables.

We now argue that self-assessed financial literacy is an important and robust predictor of house ownership and loan-to-value ratios. That is, the relationship illustrated in Figure 1 is not due to potential confounders. To establish this argument, we examine the linear cross-sectional relationships between self-assessed financial literacy and the housing market outcomes. We specify the following linear model:

$$Y_i = \beta_{low}FK_{low,i} + \beta_{high}FK_{high,i} + \Gamma X_i + \epsilon_i \quad (1)$$

where  $Y_i$  is the outcome of interest,  $FK_{low,i}$  is an indicator equal to one in case the household reports its financial literacy to be low (0-4 on the 0-10 scale), and  $FK_{high,i}$  is the equivalent for household that assess their financial literacy to be high (8-10 on the scale). The omitted group consists of the intermediate literacy types. The vector of covariates  $X_i$  consists of an age polynomial, education attainment levels, a gender dummy, total wealth and income, self-assessed risk preference, the traditional measures of financial literacy and dummies for usage of financial advisers when investing and borrowing.  $\epsilon_i$  is the normally distributed error term<sup>5</sup>.

Table 2 reports the main results. Consistent with Figure 1, the first (second) column depicts the unconditional positive correlation between financial literacy and home ownership (LTV). As illustrated in the figure, differences in ownership rates are more stark than differences in loan-to-value ratios. Households who report high levels of financial literacy are more likely to own a house relative to those who place themselves in the intermediate range (the change in odds ratio is 1.64), which are in turn more probable to be owners relative to those belonging to the low category (by an estimated change in odds ratio of 2.24). In terms of LTV, there doesn't seem to be much difference between the high

---

<sup>5</sup>In order to account for both the multiple imputation process and the dual-frame complex sample which are features of the SCF data, standard errors are computed using the “scfcombo” Stata package.

Table 1: Descriptive Statistics

Dependent Variable	Low FL	Intermediate FL	High FL
<b>A. Demographics</b>			
Age	51.80 (16.8)	50.77 (16.2)	54.47 (16.4)
Gender	0.62 (0.48)	0.71 (0.45)	0.74 (0.44)
Income	48,867 (132,113)	68,203 (64,066)	78,564 (79,444)
Wealth (log)	10.00 (2.64)	11.15 (2.11)	11.72 (2.00)
Education Level	2.26 (1.07)	2.79 (1.03)	2.87 (1.02)
<b>B. Financial Indicators</b>			
Financial Literacy Score	1.86 (0.88)	2.12 (0.87)	2.24 (0.84)
Self-Assessed Financial Risk	2.79 (2.66)	4.09 (2.53)	4.37 (2.87)
Use of Adviseries: Borrowing	0.40 (0.49)	0.52 (0.50)	0.59 (0.49)
Use of Adviseries: Investing	0.44 (0.50)	0.58 (0.49)	0.63 (0.48)
Stock Market Participation	0.28 (0.49)	0.52 (0.50)	0.55 (0.49)
Equity Share of Financial Assets	0.44 (0.31)	0.43 (0.29)	0.43 (0.28)
Number of Stocks Held	0.48 (1.90)	0.78 (3.43)	1.58 (6.22)
<b>Number of Observations</b>	<b>2,168</b>	<b>10,083</b>	<b>12,136</b>

Notes: Households are divided into three groups according to their self reported financial knowledge: Low (0-4 on scale), intermediate (5-7) and high (8-10). Income is the sum of wage income, income from retirement and social security funds, from self managed businesses and transfers. Gender is defined as proportion of males. Total wealth is defined by the SCF as the balance between total assets and total debt. The education level is a categorical variable that ranges from 1 (no high-school) to 4 (academic degree). Financial literacy is measured as the number of correct answers to the three questions specified in footnote 4. Self assessed financial risk is reported by households on a 0-10 scale, where 0 is “not at all willing to take financial risk”. Use of financial advisories is a dummy equal one if the households reports using advisers when borrowing/investing. Stock market participation is an indicator equal to one if the household has equity in directly held stocks or mutual funds. Equity share is the ratio of equity to total financial assets. Capital gains are the nominal dollar gains on directly held stocks and mutual funds. Number of stocks measures the number of different directly held stocks in a household’s portfolio.

and intermediary literacy types. Conditional on owning a house, the loan-to-value ratio of low types is 7.9% lower than that of the benchmark intermediate group.

Columns 3-4 then add the demographic controls, as well as education attainment levels<sup>6</sup>. Indeed, the magnitude of the relationship between self-assessed financial literacy and housing outcomes is weakened. However, the coefficients  $\beta_{low}$  and  $\beta_{high}$  are still economically and statistically significant. To interpret the sizable coefficients, the intermediate literacy households are 48% more likely to own a house with respect to low literacy types and are 24% less likely be home owners relative to the high types. This suggests that financial literacy is a source of heterogeneity that is economically meaningful for ownership and leverage over and above its relation to age, wealth, income, education and gender.

In columns 5-6 we also control for the usage of financial intermediaries and for the ability of households to answer finance related questions on inflation, interest rates and diversification. If self-assessed literacy and the traditional test-based measures of literacy are equivalent measures of sophistication, we should expect  $\beta_{low}$  and  $\beta_{high}$  to converge to zero. Not only is this not the case, but rather the test-based measures are not as powerful in predicting home ownership as the self-assessed measure. The way people evaluate their own knowledge matters for housing market outcomes. This is important for future work studying the relationship between financial literacy and economic behavior. Finally, in columns 7-8 we also control for households willingness to take risk. If self-assessed financial literacy simply proxies for over-confidence, this should be captured by the degree to which households are happy to take risk. The results show that this is not case.<sup>7</sup>

While we are not claiming to identify a causal relationship between self-assessed financial literacy and housing outcomes, Table 2 does suggest that self-assessed financial literacy is an economically important and robust predictor of house ownership and loan-to-value ratios<sup>8</sup>. Households that see themselves as more financially literate are more likely to own a house, and take on larger mortgages.

---

<sup>6</sup>We also control for wealth quartiles and an age polynomial.

<sup>7</sup>To further alleviate such concerns, our results are practically unchanged when we also control for participation in the stock market and for equity shares.

<sup>8</sup>For robustness, we consider additional variations of the regression Equation 1, including incorporating the continuous version of our literacy measures and interacting this measure with other control variables to allow for differential relationships with respect to housing market outcomes. Results are quantitatively similar. These results are also robust to the way we divide households into literacy categories.

Table 2: Prediction Regressions: Self-Assessed Financial Literacy in the Housing Markets

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(9)
	H.Own.	LTV	H.Own.	LTV	H.Own.	LTV	H.Own.	LTV
Self-Ass. Fin. Lit.								
Low	-0.805*** (0.085)	-0.079** (0.316)	-0.392*** (0.142)	-0.051*** (0.022)	-0.402*** (0.143)	-0.047*** (0.22)	-0.446*** (0.143)	-0.048*** (0.21)
High	0.494*** (0.051)	-0.025** (0.011)	0.217** (0.088)	0.011 (0.009)	0.208** (0.090)	0.008 (0.009)	0.220** (0.090)	0.005 (0.009)
Educ. Level								
High-School			-0.049 (0.130)	0.037** (0.016)	-0.074 (0.132)	0.032** (0.017)	-0.084 (0.132)	0.024** (0.014)
Some College			-0.215 (0.131)	0.064*** (0.017)	-0.229* (0.131)	0.056*** (0.017)	-0.220* (0.131)	0.037** (0.015)
Bachelors+			-0.619*** (0.158)	0.103*** (0.019)	-0.655*** (0.164)	0.090*** (0.019)	-0.623*** (0.165)	0.052* (0.016)
Age			0.043*** (0.011)	-0.001 (0.001)	0.042*** (0.011)	-0.002 (0.002)	0.043*** (0.011)	-0.01*** (0.002)
Male			0.146 (0.101)	0.000 (0.011)	0.130 (0.102)	-0.004 (0.011)	0.155 (0.104)	-0.012 (0.01)
ln(wealth)			1.209*** (0.117)	-0.105*** (0.015)	1.210*** (0.118)	-0.109*** (0.015)	1.220*** (0.118)	-0.0104 (0.013)
ln(income)			-0.119* (0.069)	0.160*** (0.006)	-0.122* (0.071)	0.157*** (0.006)	-0.121* (0.071)	0.119 (0.006)
Fin. Lit. Questions								
Inflation					0.218 (0.182)	-0.042* (0.023)	0.207 (0.184)	-0.045** (0.022)
Interest Rate					0.124 (0.173)	-0.073*** (0.024)	0.121 (0.175)	-0.066*** (0.023)
Diversification					0.116 (0.198)	-0.062** (0.026)	0.123 (0.202)	-0.073*** (0.026)
Ad. Borrowing					0.318*** (0.079)	0.027*** (0.010)	0.328*** (0.079)	0.008 (0.008)
Ad. Investing					-0.222*** (0.073)	-0.002 (0.008)	-0.205*** (0.073)	0.007 (0.008)
Self. Ass. Fin. Risk							-0.051*** (0.013)	0.007*** (0.002)
Observations	24,112	15,007	24,112	15,007	24,112	15,007	24,112	15,007
R <sup>2</sup>	0.055	0.048	0.414	0.384	0.417	0.389	0.418	0.489

Notes: House ownership measures whether or not the household owns a ranch/farm/mobile home/house/condo. The Loan-To-Value ratio is computed for home owners as the ratio of housing collateralized debt to house value. \*\*\* is significant at 1%; \*\* is significant at 5%; \* is significant at the 10% level. Standard errors are computed using the "scfcombo" Stata package in order to account for the SCF complex sample specification as well as the multiple imputation process. The explanatory variables are self-assessed financial literacy (low, high), education level (high school, some college, bachelors), age, gender, ln(wealth), ln(income), financial literacy questions (inflation, interest rate, diversification), use of advisories (borrowing, investing) and self-assessed financial risk.

### 3 Model

Motivated by the empirical patterns, we now ask what are the channels through which financial literacy matters for housing tenure and leverage. We consider two intuitive candidates. First, households reporting higher levels of financial literacy might search for better deals in the credit markets, thereby paying lower interest rates on their mortgages and facing laxer down-payment requirements. Second, expectations on future house values might be important. More literate households might expect different risk-return trade-offs in the housing markets due to access to more beneficial property deals. While there might be additional potential channels that can rationalize our empirical findings, we focus on those which are arguably the most important for housing outcomes. To examine the role of mortgage terms and expectations on future returns, we solve a standard heterogeneous life-cycle model of portfolio choice with housing.

#### 3.1 Household Problem

Households live for a finite number of discrete periods  $t = 0, \dots, T$ , with a probability of survival from period  $t - 1$  to period  $t$  of  $\lambda_t$ , and  $\lambda_{T+1} = 0$ . Household  $i$  enters the model with an innate level of financial literacy  $f_i$  which is the key new feature of this study. Our models abstracts from the possibility that financial literacy might evolve throughout life as a result of different experiences in financial markets. While this is an interesting avenue, it also requires making assumptions on the way in which such literacy is accumulated<sup>9</sup>.

#### Income Process

As in standard models, household face idiosyncratic income shocks. In each period until retirement at age  $t = Ret$  households are endowed with labor income  $Y_t$  that follows an exogenous stochastic process. Following retirement, the household receives a constant fraction  $\theta_{Ret}$  of its last labor income  $Y_{Ret}$ . Following [Cocco, Gomes and Maenhout, 2005](#), the income process is given by:

$$\log Y_t = \begin{cases} f(t) + \log \bar{Y}_t + \log \hat{Y}_t^i + u_t & t \leq Ret \\ \log(\theta_{Ret}) + f(Ret) + \log \bar{Y}_{Ret} + \log \hat{Y}_{Ret}^i & t > Ret \end{cases} \quad (2)$$

---

<sup>9</sup>We assume financial literacy constant throughout life since, in our simulation exercise, every household is simulated only 1 period. This is explained in detail in section 4.3.1.

where  $f(t)$  is a deterministic life cycle profile and  $u_t$  is an idiosyncratic temporary shock distributed as  $N(0, \sigma_u^2)$ .  $\bar{Y}_t$  and  $\hat{Y}_t^i$  are the aggregate and idiosyncratic components of income, both following a random walk in logs:

$$\begin{aligned} \log \bar{Y}_t &= \log \bar{Y}_{t-1} + \bar{\epsilon}_t \\ \log \hat{Y}_t^i &= \log \hat{Y}_{t-1}^i + \hat{\epsilon}_t^i \end{aligned}$$

where  $\bar{\epsilon}_t$  is distributed  $N(0, \sigma_{\bar{\epsilon}}^2)$  and  $\hat{\epsilon}_t^i$  is distributed  $N(0, \sigma_{\hat{\epsilon}}^2)$ . The shocks  $\bar{\epsilon}_t, \hat{\epsilon}_t^i, u_t$  are uncorrelated. These assumptions allow us to denote the permanent shock to income as:

$$\epsilon_t^{\bar{Y}} = \bar{\epsilon}_t + \hat{\epsilon}_t^i \sim N(0, \sigma_{\bar{Y}}^2).$$

## Preference

Households choose the amount of housing services  $S_t$  and the numeraire  $C_t$  to be consumed each period in order to maximize lifetime utility given by:

$$E_0 \left\{ \sum_{t=0}^T \beta^t \left[ \prod_{j=0}^t \lambda_j \lambda_{t+1} u(C_t, S_t) + \prod_{j=0}^t \lambda_j (1 - \lambda_{t+1}) D_t \right] \right\}$$

where  $D_t$  is the bequest utility in case of death. The per period utility  $u(C_t, S_t)$  is assumed to be strictly increasing and concave in both  $C_t$  and  $S_t$ . Households can consume housing services in two ways: either by renting or by owning a house. Denote by  $\tau_t \in \{0, 1\}$  the tenure choice at period  $t$ , with  $\tau_t = 1$  indicating ownership. A house of quality  $H_t$  provides housing services according to the technology<sup>10</sup>:

$$S_t = H_t$$

The functional form of the per-period utility function is the standard Cobb-Douglas:

$$u(C_t, S_t) = \frac{[C_t^\rho S_t^{1-\rho}]^\gamma}{1-\gamma}$$

where  $\gamma$  is the relative risk aversion parameter and  $\rho$  measures the intra-temporal substitution between housing and other consumption goods. The bequest utility  $D_t$  is a function of total wealth at the conclusion of period  $t$ ,  $W_t$ , as well as house prices:

$$D_t(W_t, P_t) = \frac{\bar{D}(W_t^i/P_t^\rho)^{1-\gamma}}{1-\gamma}$$

where  $\bar{D}$  mediates the importance of bequest motives relative to other consumption. The functional form of the bequest function is chosen to ensure value function homogeneity.

<sup>10</sup>Many models of portfolio choice with housing incorporate an age-dependent preference for tenure which is driven by exogenous forces such as uncertainty regarding changes in workplace and household size. Following Landvoigt (2017) we also solve a specification of the model where  $S_t = \phi(\tau_t, t)H_t$  and  $\phi(\tau_t, t) = 1 + (1 - \tau_t)e^{-\kappa t}$ .  $\kappa$  then regulates the age-dependent preference to own. Since our baseline model fits the housing market data patterns, we proceed without incorporating an age-dependent preference.

## House Quality and Prices

Houses serve not only as a consumption good but also as an asset that the households can save in. Each quality unit of the housing asset sells for a price of  $P_t$ , and can be rented for a price of  $P_t^r$  in the rental market. For simplicity, we assume the rental price is pegged to the selling price, that is  $P_t^r = \alpha P_t$ . Aggregate shocks to the economy are captured by innovations to the per-unit price  $P_t$  which follows a random walk in logs:

$$G_t^P = \frac{P_t}{P_{t-1}} = \exp\{\epsilon_t^P\}$$

where  $\epsilon_t^P \sim N(d_P, \sigma_P^2)$  and  $d_P$  is the deterministic drift in log house price growth. We assume that the vector of innovations to permanent income and house prices  $(\epsilon_t^{\hat{Y}}, \epsilon_t^P)$  is independent across time with a variance matrix of:

$$\text{Var}(\epsilon_t^{\hat{Y}}, \epsilon_t^P) = \begin{bmatrix} \sigma_{\hat{Y}}^2 & \sigma_{\hat{Y}P} \\ \sigma_{\hat{Y}P} & \sigma_P^2 \end{bmatrix}.$$

Aggregate shocks to house prices might hence be contemporaneously correlated with permanent shocks to income. The quality of an owner-occupied home  $H_t$  is itself stochastic and evolves according to the idiosyncratic process:

$$H_{t+1} = Q_i(H_t) = (1 + g_{i,t+1})H_t$$

where  $g_{i,t} \sim N(\mu(f_i), \sigma^2(f_i))$  is i.i.d across time. The house quality evolution depends on the financial sophistication of the household that owns it. Households with higher self-assessed financial literacy may have access to better investment opportunities in the housing markets, due to ,e.g., sophisticated search skills. We model this form of heterogeneity by allowing both the mean return and volatility of the distribution from which future house values are drawn to differ by literacy.

## Collateral Constraints and Default

Households are allowed to save  $B_t$  in a risk free asset which generates  $R$  units of return at  $t + 1$  for each unit of the numeraire saved in  $t$ . When borrowing households pay a financial-literacy dependent interest rate spread of  $\varrho(f_i) > 0$  appealing to the possibility that financially literate households search and negotiate for cheaper credit. Households also face an idiosyncratic collateral constraint:

$$B_t \geq \begin{cases} 0 & \tau_t = 0 \\ -[1 - \delta(f_i)] P_t H_t & \tau_t = 1 \end{cases}.$$

Thus only house owners can borrow against the value of their house, and can finance their purchase up to a ratio of  $(1 - \delta(f_i))$  of the value of their house. The down-payment requirement can vary across different levels of financial literacy, alluding to the possibility that financially literate households might have access to larger credit. Households do not have an option of defaulting upon their mortgage. The choice of modeling mortgages as non-contingent debt relies on the fact that defaults are very rare in that data. Strictly speaking, default seems an important feature for models studying the Great Recession dynamics, but not for the long run steady state which we focus on.

### Budget Constraints

When specifying the budget constraint we distinguish between two cases: households who have previously rented a house at period  $t - 1$ , and those who previously owned a house.

#### Case 1: Previous Renters

The budget constraint for a household who was renting in period  $t - 1$  is given by:

$$C_t + B_t + P_t H_t \left\{ (1 - \tau_t)\alpha + \tau_t (1 + \psi) \right\} = R B_{t-1} + Y_t \quad (3)$$

where  $\psi$  accounts for the proportional maintenance cost  $\psi P_t H_t$  that an owner must pay every period to offset depreciation. This household enters the period with wealth equal to accrued savings  $R B_{t-1}$  and contemporary income  $Y_t$ . It chooses how much to consume, how much to save in bonds, whether or not to purchase a house (in which case it can also borrow), and which quality of housing to consume.

#### Case 2: Previous Owners

Previous owners gets to choose whether or not to sell their house, in which case they can buy or rent a different quality of housing. If it chooses to sell, it pays a proportional transaction cost  $\nu P_t H_t$ . We denote the decision of whether to sell or not by  $\xi_t = \{0, 1\}$  where  $\xi_t = 1$  indicates selling. A previous owner faces the following budget constraint:

$$C_t + B_t + \xi_t P_t H_t \left\{ (1 - \tau_t)\alpha + \tau_t (1 + \psi) \right\} + P_t Q_i(H_{t-1}) [(1 - \xi_t)\psi + \xi_t \nu] = \left[ R_t + 1_{\{B_{t-1} < 0\}} \varrho(f_i) \right] B_{t-1} + Y_t + P_t Q_i(H_{t-1}) \quad (4)$$



Finally, previous owners might be forced to sell their house. The model implementation consists of an age-dependent exogenously specified moving shock  $M_t$ , where  $M_t = 1$  indicates that the household is forced to moved. We incorporate this feature to capture life-cycle shocks to the household which are not captured by the model and induce selling the house and adjusting quality, whether it is through the owner-occupied market or the rental market.

## Bellman Equations

The recursive nature of the problem allows us to state it in terms Bellman equations. Denote  $X_t = \{f_i, W_t, P_t, \tau_{t-1}, Q_t, \hat{Y}_t, M_t\}$  the tuple of state variables. In addition, denote  $Z_t = \{\tau_t, H_t, C_t, B_t, \xi_t\}$  the choice variable tuple. The following problem specifies the household value function at  $t < Ret - 1$ <sup>11</sup>:

$$V_t(f_i, W_t, P_t, \tau_{t-1}, Q_t, \hat{Y}_t, M_t) = \lambda_t \left\{ \max_{Z_t} u(C_t, S_t) + \beta \mathbb{E}_t^i [V_{t+1}(f_i, W_{t+1}, P_{t+1}, \tau_t, Q_{t+1}, \hat{Y}_{t+1}, M_{t+1})] \right\} + (1 - \lambda_t) D(W_t, P_t) \quad (5)$$

where  $\hat{Y}_t = \bar{Y}_t \hat{Y}_t^i$  is the permanent income component. The problem is subject to the collateral constraint:

$$B_t \geq \begin{cases} 0 & \tau_t = 0 \\ - [1 - \delta(f_i)] P_t H_t & \tau_t = 1 \end{cases}. \quad (6)$$

The budget constraint is:

$$\begin{aligned} W_t &= C_t + B_t + P_t H_t [(1 - \tau_t) \alpha + \tau_t (1 + \psi)] & \tau_{t-1} &= 0 \\ W_t &= C_t + B_t + (1 - \xi_t) \psi P_t Q_i (H_{t-1}) \dots & \tau_{t-1} &= 1. \end{aligned} \quad (7)$$

$$\dots + (1 - \xi_t) P_t Q_t (1 - \nu) + \xi_t [P_t H_t \{(1 - \tau_t) \alpha + \tau_t (1 + \psi)\}]$$

Finally, the evolution of the endogenous state variables, total wealth and house quality, is given by:

$$\begin{aligned} W_{t+1} &= (R + 1_{\{L_t < 0\}} \rho(f_i)) B_t + (1 - \nu) P_{t+1} Q_{t+1} + Y_{t+1} \\ Q_{t+1} &= \tau_t Q_i(H_t) = \tau_t (1 + g_{i,t+1}) H_t \end{aligned}.$$

<sup>11</sup>The Bellman equations for  $t \geq Ret - 1$  are given in Appendix B

The above can be solved by employing standard dynamic programming methods. In order to reduce the state space dimensionality and efficiently compute the policy functions, Appendix B presents a transformed and equivalent problem. The solution relies on the homothetic nature of the problem.

## 3.2 Discussion

A solution to the model yields optimal housing choices for households with a given income, wealth, age, and self-assessed financial literacy. Notably, we do not model the supply side of the economy or solve for equilibrium prices in the housing markets. Instead, our goal is to infer the role of expectations and credit conditions in shaping house tenure and leverage choices. In any competitive equilibrium, households take current house prices as given. Thus, inferring expectations over future house values as well as current credit markets fundamentals from observed choices in data is a well-defined exercise.

### Optimal Choices

Which households choose to own in the model? Under complete markets (i.e. absent of collateral constraints) the choice of tenure is independent of the house quality choice. That is, the household problem can be solved in two steps. First, households decide whether to live in an owner occupied or rental house. Independently, they also decide on the quality of their home. The reason is simply that each household, regardless of its wealth, can both purchase and rent any house it likes. With no borrowing constraints, purchasing a high-quality house simply means taking on larger debt. Facing this non-segmented housing markets, the decision to own versus rent is based solely on a comparison between the per-unit user cost – the cost of buying one quality unit of housing and selling it after one period – and the per-unit rental cost. In terms of our housing market parameters, this is more likely when the risk-return trade off is higher (that is, when  $\mu$  is larger and  $\sigma$  lower) and when the mortgage spread  $\varrho$  is smaller.

In the presence of collateral constraints, this is no longer the case. Poor and middle-income households cannot afford high quality homes. They therefore face a segmented housing market - they can rent high value houses but can own only lower quality ones. Since home-owners are required to supply at least the down-payment on the house value, these are households who also like to save. In the model, these are households that face a downward sloping income profile or are sufficiently wealthy - typically the older households. Once a house is purchased, transaction costs make it harder to adjust the size of the house later in life, leading to inertia in housing choices.

In the absence of collateral constraints, young households would optimally take a short position in the safe asset. For these households, it is optimal to offset their riskless future income by borrowing in the risk-less bond. However, with collateral constraints this is possible only to the extent that households own a house. The minimum down-payment requirement  $\delta$  and mortgage spread  $\varrho$  therefore play an important role in the extensive margin decision of ownership. For those deciding to own, the expected risk-adjusted return  $\frac{\mu}{\sigma}$  plays an important role in determining the amount of leverage to take on the house. The higher the Sharpe ratio is, the more levered home-owners tend to be. To sum, for ownership credit constraints are more important than expectations on future returns. Expectations matter more for leverage.

Finally, in order to match moving rates as observed in data, we add age-dependent moving shocks,  $M_t$ , appealing to the fact that home owners might sell their house due to non-financial forces such as job re-allocations.

## 4 Calibration and Estimation

### 4.1 Data

In order to estimate the channels through which self-assessed financial knowledge translates into behavior in the housing markets, we use the 2016 cross-section of the SCF. The survey data include information on balance sheets, income, and demographic characteristics of a representative sample of U.S households. As discussed in Section 2, the measurement of self-assessed financial knowledge was first introduced to the 2016 questionnaire, limiting us to the use of this particular wave. We use the summary extract public data of the SCF and focus on families for which the head of household is aged between 25 and 80, the time period considered in our model. Total wealth is defined by the SCF as the balance between total assets (financial and non-financial) and total debt, coded as “networth”. We omit households with total net-worth larger than 7 million dollars. Our model is not suitable for describing the life-cycle wealth dynamics of rich households who rely much more on stock market capital gains and non-traditional retirement income sources. Applying the SCF sampling weights, these households consist of about 11% of effective observations<sup>12</sup>. We define total labor income as the sum of wage income, income from retirement and social security funds, income from self managed businesses and transfers from other sources. We use the variable “houses” as the value of the house (for owners) which is defined by the SCF as the value of the primary residence. Mortgage debt for house owners

---

<sup>12</sup>Abstracting from the stock market in the model thus seems reasonable for the lower 90% of households.

is coded by the SCF as “mrthel” and includes all forms of debt which are collateralized against the value of the house. Self assessed financial knowledge is measured by “knowl” and is categorized into three groups, as discussed in section 2.

## 4.2 Calibration

Table 3 shows the calibrated model parameters that we do not estimate. The preference parameters are taken from the literature. Risk aversion  $\gamma$  is set to 3 and the Cobb-Douglas weight on housing services,  $\rho$ , is set to 13% based on [Piazzesi, Schneider and Tuzel \(2007\)](#). The parameter  $\bar{D}$  that governs the strength of the bequest motive is 10 to match the average home-ownership rate at the age of 80.

The income process specified in (2) is calibrated based on the results of [Cocco, Gomes and Maenhout \(2005\)](#). The deterministic part  $f(t)$  follows a three-degree polynomial in age. We use the coefficients characterizing the life-cycle profile of high-school graduates estimated by [Cocco, Gomes and Maenhout \(2005\)](#) from PSID data, and adapt them to fit our income specification. The cyclical component has the usual hump shape. Based on these results, the annual standard deviation of the permanent shock is set to about 10%. The correlation between innovations to house price and permanent income,  $\sigma_{\hat{Y}P}$ , is set to zero based on [Flavin and Yamashita \(2002\)](#).  $\theta_R$  is set to be 0.7 in accordance with [Cocco, Gomes and Maenhout \(2005\)](#).

Moving to prices, the risk-free interest rate  $R$  is calculated as the average real yield of a 1-year treasury bond between 1972-2017. The volatility of house price growth is set to 15% This number reflects both idiosyncratic risk, which [Landoigt, Piazzesi and Schneider \(2015\)](#) and [Case and Shiller \(1990\)](#) estimate to be between 9% and 15%, and aggregate housing risk which [Flavin and Yamashita \(2002\)](#) estimate to be between 5% and 9%. We estimate the drift in house price growth to be equal to that implied by the estimated income process, which is 0.5%. This ensures that the ratio between house prices growth and income growth is stationary. To compute the rent-to-price ratio  $\alpha$  we use the FHFA aggregate price index and deflate it by the CPI of house rental prices. The long run value of this series is consistent with [Davis, Lehnert and Martin \(2008\)](#) and [Sommer, Sullivan and Verbrugge \(2013\)](#). The maintenance cost accrued by house owners in order to offset depreciation is set as a 1% share of the house value, and is line with other values in the housing literature.

We estimate the age-dependent probability of moving,  $M_t$ , from the 2010 Census data. To identify moving for reasons that are exogenous to our model (e.g. marriage, divorce, disaster loss) we use the 2015 American Housing Survey which asks respondents for mov-

ing circumstances. The life-cycle mobility shock is estimated to be downward sloping with age. Finally, survival rates  $\lambda_t$  are calculated from The National Center of Health Statistics mortality rates.

Table 3: Calibrated Parameters

Parameter	Notation	Value
Relative risk aversion	$\gamma$	3
Housing services weight in utility	$1 - \rho$	0.13
Relative income at retirement	$\theta_{Ret}$	0.7
Bequest motives	$\bar{D}$	10
Permanent shock volatility	$\sigma_{\hat{Y}}^2$	0.0106
Transitory shock volatility	$\sigma_u^2$	0.0738
Riskless rate	$R$	1.01
Maintenance cost	$\psi$	0.01
House price shock volatility	$\sigma_P^2$	0.0144
Rent to price ratio	$\alpha$	0.05
Transaction cost	$\nu$	0.08

## 4.3 Estimation Procedure

### 4.3.1 Simulated Method of Moments Approach

We estimate the remaining model parameters by applying a Simulated Method of Moments (SMM) to the cross-sectional 2016 SCF data. Denote the set of these parameters, which we specify below in more detail, by the vector  $\eta$ . For every model period  $t$  we simulate a large number of  $I_t$  households from the SCF distribution of this age cohort. Denote this sampled data by  $\Omega_t = \{f_{i,t}, \tau_{i,t-1}, W_{it}, P_t Q_{it}, \hat{Y}_{it}\}_{i=1}^{I_t}$  where  $f_{it}$  is the self-assessed financial knowledge category household  $i$  belongs to,  $\tau_{i,t-1}$  denotes the house ownership status at the beginning of the period,  $W_{it}$  is the total wealth and  $P_t Q_{it}$  is the house value for owners as defined in Section 4.1. Since the data does not distinguish between the permanent income component  $\hat{Y}_{it}$  and the temporary shock  $u_{it}$ , we decompose the observed labor income  $Y_{it}$  by simulating  $u_{it}$  from its specified distribution. Similarly, we simulate a moving shock  $M_{it}$  for each household based on the calibrated probabilities.

Given the simulated state variables and values for both the calibrated parameters from Table 3 and  $\eta$ , we can solve for the period  $t$  optimal policies for each household  $i$ , denoted by  $Z_t(\Omega_t, \eta) = \{\tau_{i,t}, H_{i,t}, C_{i,t}, B_{i,t}, \bar{\zeta}_{i,t}\}_{i=1}^{I_t}$ . By simulating the period  $t + 1$  shocks to idiosyncratic income ( $\epsilon_{i,t+1}^{\hat{Y}}$ ), aggregate housing prices ( $\epsilon_{t+1}^P$ ) and owner specific house quality growth ( $g_{i,t+1}$ ) we can map these policies into year  $t + 1$  state variables. Together, this

yields the simulated sample of next period’s state variables,  $\hat{\Omega}_{t+1}(\Omega_t, \eta)$ . We then estimate  $\eta$  by minimizing (in an SMM fashion) the distance between the simulated samples  $\hat{\Omega}_{t+1}$  and the data samples  $\Omega_{t+1}$ .

The estimation relies on two assumptions. Since our data is not of panel structure, the observed households at  $t$  are not followed into period  $t + 1$ . Comparing the simulated samples  $\hat{\Omega}_{t+1}$  and the data samples  $\Omega_{t+1}$  thus assumes a birth cohort invariant distribution of  $\Omega_t$ . While in general this distribution may differ across birth cohorts (e.g. if the joint distribution of tenure and age was different in the past compared to what it is today), our SMM approach only requires invariance across two consecutive birth cohorts. The second assumption follows from the enduring nature of self-assessed financial knowledge in the model. While the data suggests that older households report higher levels of financial knowledge, our model assumes it to be innate. However, this would be a concern to our SMM estimation only to the extent that financial knowledge evolves in between two consecutive periods. Appendix C specifies the estimation procedure, including the estimation of standard errors, in more detail. In what follows we discuss the data moments we target and the parameters we thereby estimate.

### 4.3.2 Parameters and Moments

We want to study the role of expectations on future house prices and of mortgage terms in generating the relationship between financial literacy and housing outcomes apparent in the data. For this, the model entertains heterogeneity across different levels of self-assessed financial literacy. This heterogeneity is captured by the 4 groups of parameters we estimate: 1) the expected returns on the housing asset,  $\mu(f_i)$ ; 2) the expected volatility of these returns  $\sigma(f_i)$ ; 3) the mortgage interest rate spread  $\rho(f_i)$ ; and 4) the minimum down-payment requirement  $\delta(f_i)$ .

As data moments in the SMM objective function we use 6 groups of moments: house ownership ratio and loan-to-value ratio, where we divide households into three age groups to capture life cycle dynamics. We define the young as households between the ages of 25 and 40, the middle aged as those between the ages 41 and 60, and the old as those older than 60. Each of these moments is computed across the three types of financial literacy. We also estimate the discount factor  $\beta$  in order to match aggregate wealth in the data<sup>13</sup>. This gives us 13 moments and 19 parameters. The choice of these moments relies mainly

---

<sup>13</sup>In an additional exercise, we also allow for heterogeneity in the discount factor across self-assessed financial literacy groups. Data moments in this case are augmented with wealth by literacy groups. Such heterogeneity doesn’t seem to play an important role, as estimates are basically equal across levels of self assessed literacy.

on their interaction with model policies, as discussed in Section 3.2.

## 5 Results

Table 4 lists the estimated parameters across literacy groups. Households who self-assess themselves as more literate face laxer constraints in the credit markets. They pay a lower spread when borrowing against the value of their house, and are required to a lower down-payment on their mortgage. More literate households also face better risk-return trade-offs in the housing markets. While both the expected return and volatility of this return are estimated to be hump-shaped with respect to literacy, the overall Sharpe ratio is slightly increasing – from 1.08 for the low group, 1.2 for the intermediate group, and 1.7 for the higher literacy households. On average, our estimates suggest that differences in credit markets are quantitatively more important for explaining the stylized facts. As the empirical analysis in Section 2 points out, differences in ownership rates are more stark across literacy groups, compared to differences in loan-to-value ratios. The estimation therefore uses the literacy types to mostly accommodate for the differences in home-ownership. For this, credit constraints are more important than expectations on future prices. Laxer credit constraints matter for ownership decisions, but expectations matter more for leverage. In Section 5.2 we offer a more nuanced view, according to which households who self-assess themselves as financially sophisticated tend to be better at those parts of the transaction that are more relevant to them.

Table 4: Estimated Parameters: Full Model

<u>Estimates</u>	<u>Low Lit.</u>	<u>Int. Lit.</u>	<u>High Lit.</u>
Expected return $\hat{\mu}(f)$	0.04 (0.0004)	0.07 (0.0005)	0.05 (0.0001)
Volatility $\hat{\sigma}(f)$	0.037 (0.0005)	0.058 (0.005)	0.031 (0.002)
Sharpe Ratio $\frac{\hat{\mu}(f)}{\hat{\sigma}(f)}$	1.07 (0.02)	1.2 (0.1)	1.7 (0.1)
Mortgage spread $\hat{q}(f)$	0.04 (0.0002)	0.028 (0.002)	0.009 (0.0006)
Min. down-payment $\hat{\delta}(f)$	0.2 (0.005)	0.154 (0.0005)	0.143 (0.01)

### 5.1 Model Fit

Figure 2 shows the fit of our estimation to the data. The heterogeneous agent model is able to closely mimic the stylized facts it was designed to target. As in the data, model-

implied home-ownership rates and loan-to-value ratios are increasing with self-assessed financial knowledge. Both in the data and in the model, home-ownership rates exhibit a strong slope in financial literacy for all age groups. In fact, literacy seems to be as important for tenure as age. The difference across levels of financial literacy is less stark when considering loan-to-value ratios, and shows up only for middle-aged and old households. Relating back this observation to the estimated parameters in Table 4, we are able to explain identification. The estimation uses the literacy types to mostly accommodate for the differences in home-ownership. For this, credit constraints are more important than expectations on future prices. Laxer credit constraints matter for ownership decisions, but expectations matter more for leverage

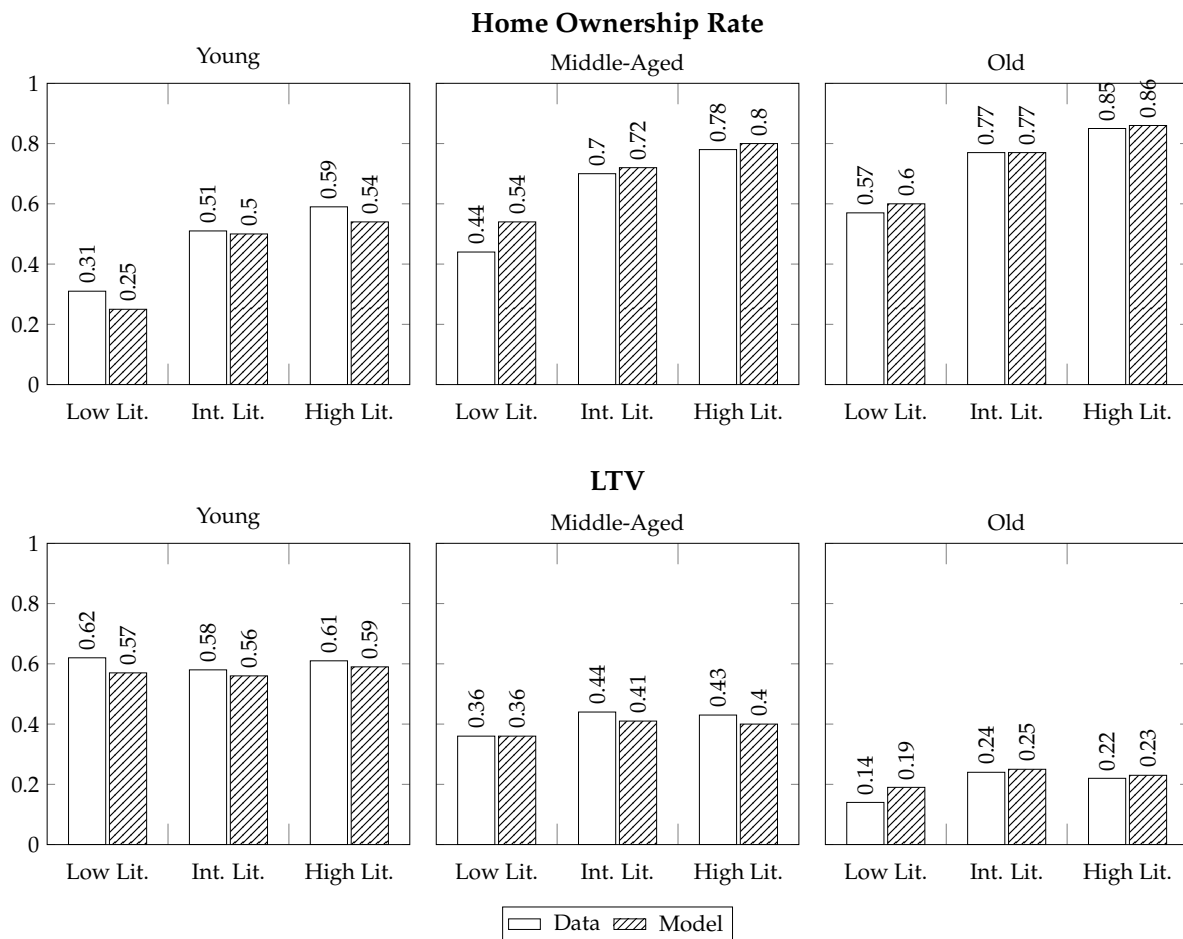


Figure 2: Full Model: Target and Model Generated Moments

Note: The figure compares model generated moments (red bars) to SCF data moments (blue bars). The Loan-to-Value ratio is averaged across all home-owners and is computed in the SCF as the ratio of all debt which is collateralized against the house, divided by the value of the house. Young households are those in which the household head is 40 years old or younger, middle-aged are those between the ages of 41 and 60, and the old are those older than 61. Low literacy households are those self-assessing their knowledge to be between 1 and 4 on the 1 – 10 scale, intermediate literacy households are those self-assessing their knowledge to be between 5 and 7 and high literacy households are those self-assessing their knowledge to be between 8 and 10.



By fitting the data, our model suggests that heterogeneity in mortgage terms and in expectations on future house prices can account for the link between self-assessed financial literacy and housing outcomes that is evident in the data. A model in which households that report higher financial literacy levels also shop for better deals in the mortgage markets and have better risk-return trade-offs on the housing asset can rationalize our empirical findings.

## 5.2 Interpreting Self-Assessed Financial Literacy

How important are each of these two channels in generating the documented stylized facts? In what follows we show that if we want to better infer how financially sophisticated individuals differ, we need to take a life cycle perspective. We argue that households who self-assess themselves as financially sophisticated tend to be better at those parts of the transaction that are more relevant to them. That is, financial literacy matters differently throughout life. At the beginning of life, when mortgage terms matter more for housing decisions, those with higher self-assessed literacy manage to shop for lower mortgage spreads and down-payment requirements. Later on, when returns on housing is more relevant, they make deals with better risk-return trade-offs.

To see this, consider the following exercise. We begin by simulating the heterogeneous agent model while shutting off heterogeneity in expectations on future house values. That is, we simulate a model where we maintain the estimates of mortgage spreads  $\rho(f_i)$  and minimum down-payment requirement  $\delta(f_i)$  from Table 4, but set the expected returns on the housing asset,  $\mu$ , and the expected volatility of these returns  $\sigma$  to be equal across literacy types. Specifically, we use the weighted average  $\mu$  and  $\sigma$  from Table 4. The results of this exercise are given in Figure 4 in the appendix. The fit of the model without risk-return heterogeneity (yellow bars) with respect to the data (blue bars) becomes worse for the old, but actually better for the young. This can be seen in both housing market outcomes, and across the three literacy types. This suggests that heterogeneity in expectations of future house prices matters more for the old than it does for the young. Older households do not tend to borrow as much, and therefore utilize their literacy toward access to better risk-return trade-offs in the housing markets.

Next, we consider the analog case in which we allow heterogeneity in expectations on future house values but dismiss of heterogeneity in mortgage terms by setting  $\rho$  and  $\delta$  to their weighted average values. The results are illustrated in Figure 5 in the appendix. The fit of the model without heterogeneity in mortgage terms (yellow bars) with respect to the data (blue bars) deteriorates more for middle-aged households, but less so for older

households. For the middle-aged, this model does worse in terms of matching both home-ownership and loan-to-value for the low-literacy households as well as the loan-to-value ratio for the high literacy group. At the same time, for the old households there doesn't seem to be much of a difference between the two models. . Since middle-aged households tend to borrow more against the value of their house, the more literate among these households tend to be better at shopping in the credit markets - they pay lower mortgage spreads and are required to pay less collateral. These results suggest that heterogeneity in mortgage terms matters more for the younger households. Overall, we conclude that sophisticated households are better at the parts of the housing markets that is more relevant to their stage in life. This is consistent with a rational attention allocation where agents choose to learn about what is most important for them.

### **5.3 Self-Assessed Financial Literacy and Self-Confidence**

Self-assessed financial literacy might measure the degree of self-confidence in financial markets, which in turn can be important for how households behave in the housing markets. For example, households who self-assess themselves as highly literate might have over-optimistic beliefs on the risk-return trade-off they face in the housing markets. If this is the case, they would be more likely to own house and take on more levered positions.

Our quantitative results suggest that this mechanism alone cannot underlie the empirical patterns. A limitation of the estimation strategy is that we cannot distinguish between heterogeneous beliefs on future housing returns and heterogeneity in the fundamental distribution of returns. The reason is that both home-ownership and LTV are household policies and do not depend on actual realizations. However, we show that heterogeneity in fundamental mortgage terms is important for matching the stylized facts. As seen in Figure 5, a model without heterogeneity in mortgage terms fundamentals produces a diminishing relationship between literacy types and the loan-to-value ratios. Regardless of whether we interpret heterogeneity in expectations on future house prices as subjective beliefs or fundamental differences in returns, this channel alone cannot generate the relationship we observe in the cross section. That is, self-assessed financial literacy is, at least partly, a proxy for literacy itself.

## **6 How Important is Financial Literacy in Housing Markets?**

We have thus far shown that households who self-assess to be more financially literate face cheaper and laxer mortgage terms and have access to better risk-return trade-offs in

housing deals. But how important is accommodating for this heterogeneity in an otherwise standard model of portfolio choice with housing? We argue that absent financial literacy heterogeneity, the evaluation of housing market policies can be largely misleading. The main result is that the estimated housing demand elasticity with respect to wealth is downsized by approximately 40% when taking financial literacy into account.

## 6.1 Benchmark Estimation

We begin by estimating a benchmark model in which heterogeneity in self-assessed literacy plays no role. That is, we estimate the model while restricting parameters to be equal across levels of literacy. The data moments are the life cycle ownership rates and loan-to-value ratios, as well as aggregate wealth. We are therefore left with 5 parameters and 7 moments. The estimated parameters for the benchmark specification are shown in Table 5.

Table 5: Estimated Parameters: Benchmark Model

Estimates	Values
Expected return $\hat{\mu}$	0.07 (0.0003)
Volatility $\hat{\sigma}$	0.064 (0.0005)
Sharpe Ratio $\left(\frac{\hat{\mu}}{\hat{\sigma}}\right)$	1.17 (0.005)
Mortgage spread $\hat{q}$	0.028 (0.0002)
Min. down-payment $\hat{\delta}$	0.17 (0.0001)
Discount factor $\hat{\beta}$	0.95 (0.0009)

Notes: SMM Estimates from the benchmark model.

Expected house quality gains and the volatility of these gains are estimated to be 7% per year and 6.4% per year, respectively. Mortgage spread is estimated at 2.8% and the minimum down-payment requirement is 17%. The discount factor is estimated to be 0.95. By estimating these parameters, the benchmark model closely matches the life cycle dynamics of housing tenure and leverage as well as the aggregate wealth-to-income ratio in the economy. Figure 3 shows this by plotting model generated moments against the data moments.

If the standard housing model can account for the life cycle dynamics we see in the data, why should we consider augmenting it with heterogeneity in financial literacy? First, the benchmark model cannot account for the stylized facts presented in Section 2.

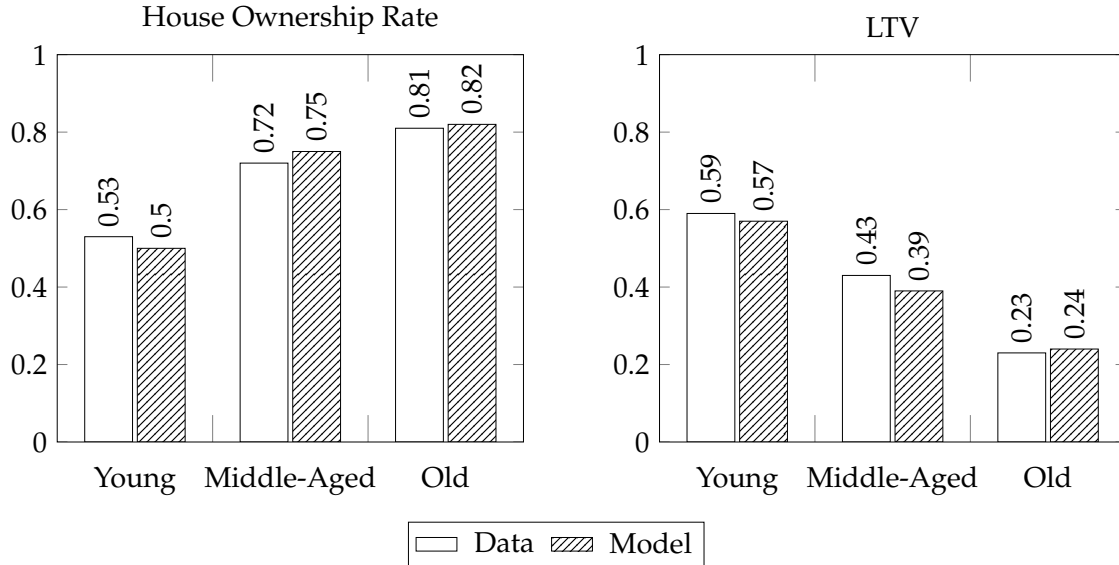


Figure 3: Benchmark Model: Target and Model Generated Moments

Note: The figure compares model generated moments (red bars) to SCF data moments (blue bars). The Loan-to-Value ratio is averaged across all home-owners and is computed in the SCF as the ratio of all debt which is collateralized against the house, divided by the value of the house. Young households are those in which the household head is 40 years old or younger, middle-aged are those between the ages of 41 and 60, and the old are those older than 61.

Absent of heterogeneity in housing market parameters, the benchmark model can differentiate between households who report different levels of literacy only to the extent that in the data this heterogeneity correlates with wealth, age, previous housing tenure and previous house value - the remaining initial conditions. If literacy matters only through its correlation with these confounding factors, the benchmark model should capture the observed heterogeneity in housing outcomes. By feeding the observed joint distribution of self-assessed literacy with wealth, age, and previous house market outcomes, the benchmark estimation shows that financial literacy matters over and above its relation to these variables. More importantly, ignoring the role of financial literacy in the housing markets has quantitatively significant policy implications.

## 6.2 Housing Demand Elasticity

We compute the aggregate housing demand elasticity to a wealth shock with and without financial literacy heterogeneity. This echoes debated policies that are designed to encourage house ownership. A 10% increase in the wealth of young households leads to a 10% increase in housing demand in the benchmark model, but only to a 6.4% increase in housing demand in the heterogeneous agent model. Similarly, a 10% wealth shock targeted at

the young and poor households leads to a 20% increase in housing demand in the benchmark model, but only to a 11% increase in housing demand in the heterogeneous agent model. The corresponding housing demand elasticities are shown in Table 6. Evaluating the effect of housing market policies to encourage ownership therefore crucially depends on whether or not heterogeneity in financial literacy is taken into account.

Table 6: Housing Demand Elasticity - With and Without Financial Literacy

Population	Benchmark Model	Heterogeneous Model
Young	1	0.64
Young and Poor	2	1.1

*Notes:* Housing demand elasticity is computed as the percentage increase in home-ownership in response to a percentage increase in wealth. The first row shows this elasticity for young households (aged 25-40) whereas the second row focuses on young and poor household (from the bottom quantile of the wealth distribution).

The underlying reason for this discrepancy is the following. While the benchmark model is able to match the life cycle dynamics of housing tenure and leverage, this comes at an overlooked cost. By ignoring the observed heterogeneity in financial literacy, the model over-estimates the correlations between housing outcomes and wealth and age. The benchmark model matches life cycle housing moments using the wrong mechanism - it underestimates the large variability in ownership and leverage within wealth and age bins as it appears in the data.

To see this, we regress the model-generated home-ownership and loan-to-value policy functions on wealth and age variables. Focusing on control variables  $X$  that can both be computed as part of the model and appear in our empirical analysis in Section 2 — age, age squared, wealth and wealth quartiles — the following specification is chosen to closely mimic Equation 1:

$$Y_i = \beta_{low}FK_{low,i} + \beta_{high}FK_{high,i} + \Gamma X_i + \epsilon_i \quad (8)$$

We then compare the estimates obtained from this exercise to the estimates obtained from regressing Equation 8 on the SCF data. The results are shown in Table 7.<sup>14</sup>

<sup>14</sup>Comparing the model-generated estimates to those from the SCF data requires estimates and standard errors be computed in a similar fashion. As discussed in Section 2, in order to accommodate for the complex sampling design of the SCF, estimates and standard errors are computed by applying a bootstrapping routine. We therefore follow this routine for computing the model-implied estimates. We draw 1,000 bootstrap samples from the the SCF distribution of the model state variables. We then apply the policy functions on each sample to simulate model-implied regression estimates.

Table 7: Data and Model Regressions

	Home Ownership			LTV		
	Data	Bench.	Full	Data	Bench.	Full
Low Fin. Lit.	-0.644*** (0.119)	-0.206*** (0.071)	-0.711*** (0.035)	-0.078** (0.027)	0.00 (0.002)	-0.060*** (0.002)
High Fin. Lit.	0.215** (0.087)	0.427*** (0.035)	0.327*** (0.025)	0.21** (0.009)	-0.009*** (0.00)	0.012*** (0.00)
Age	0.031** (0.014)	0.407*** (0.01)	0.235*** (0.005)	0.004 (0.003)	-0.007*** (0.000)	-0.006*** (0.000)
Age <sup>2</sup>	0.000** (0.000)	-0.004** (0.000)	-0.002*** (0.000)	0.000*** (0.000)	0.000** (0.000)	0.000*** (0.000)
$\ln(\frac{wealth}{income})$	0.928*** (0.088)	5.87*** (0.088)	2.55*** (0.084)	-0.09*** (0.019)	-0.184*** (0.001)	-0.175*** (0.001)
$\frac{wealth}{income}$ Q2	0.613*** (0.071)	0.076 (0.075)	1.55*** (0.039)	-0.038 (0.029)	0.088*** (0.004)	0.033*** (0.002)
$\frac{wealth}{income}$ Q3	0.957*** (0.184)	-1.293*** (0.164)	1.281*** (0.010)	-0.134*** (0.049)	-0.039*** (0.005)	-0.009*** (0.004)
$\frac{wealth}{income}$ Q4	0.122 (0.315)	— (—)	1.464*** (0.170)	-0.139** (0.067)	-0.117*** (0.006)	-0.193*** (0.005)
R <sup>2</sup>	0.361	0.794	0.541	0.336	0.804	0.730

Notes: Households are divided into three groups according to their self reported financial knowledge: Low (0-4 on scale), intermediate (5-7) and high (8-10). Total wealth is defined by the SCF as the balance between total assets and total debt, and income is the sum of incomes and transfers from all sources. Households are assigned to wealth-to-income quartiles. \*\*\* is significant at 1%; \*\* is significant at 5%; \* is significant at the 10% level. Standard errors in the data are computed using the "scfcombo" Stata package in order to account for the SCF complex sample specification as well as the multiple imputation process. Standard errors in the model are computed by simulating 1,000 bootstrap samples from the SCF data. The wealth-to-income fourth quartile is omitted from the home-ownership logit regression since all simulated households who belong to this quartile end up owning a house.

The benchmark model (column 2) generates an excessive co-movement of home-ownership with wealth and age. For a one-unit increase in  $\log(w)$ , the expected change in the home-ownership log odds is 0.928 in the data (column 1) but 5.87 in the benchmark model. Similarly, for a one-year increase in age, the expected change in the home-ownership log odds is 0.031 in the data but 0.4 in the benchmark model. A one unit increase in  $\log(w)$  is associated with a 18.4% reduction in loan-to-value in the benchmark model (column 5), compared to only a 9% decline in the actual data (column 4). In order to match the life cycle variation observed in the data, the model uses the variation in observed state variables. Since heterogeneity in financial literacy plays no role, the only source of such variation comes from wealth, age, and previous tenure and house value. Indeed, the housing demand elasticity of young households with respect to wealth is estimated to be 1. A wealth shock induces a large reaction of housing demand.

By incorporating heterogeneity in the observed literacy dimension, the heterogeneous-agent model significantly reduces the bias in the correlation between ownership and wealth and age. To see this, we repeat the exercise described by Equation 8 and compare the estimates obtained from the full heterogeneous agent model, to those in the data. As Table 7 shows, the model is able to closely match the relationship between literacy and housing market outcomes throughout the wealth and age distribution. The estimated home-ownership log odds ratio between the low (high) financial literacy group and the omitted intermediate group is  $-0.644$  (0.215) in the data (column 1) and  $-0.711$  (0.327) in the full heterogeneous-agent model (column 3). Similarly, the estimated increase (decrease) in the loan-to-value ratio between the omitted group and the low (high) literacy group is 0.078 (0.021) in the data (column 4) and 0.06 (0.012) in the financial literacy model (column 6).

Since households of different literacy types face different mortgage terms and have different expectations on future house prices, they now respond differently to a wealth shock. To provide some intuition, the elasticity of housing demand with respect to wealth is lower for low literacy types since the credit constraints they face are more strict. Indeed, the housing demand elasticity of young households with respect to wealth is estimated to be 0.64. The effect of a wealth shock in the heterogeneous-agent model is therefore mitigated.

We should point out that adding a new source of heterogeneity in any dimension will mechanically reduce the excessive correlations between wealth and housing outcomes that is generated by the benchmark model and therefore reduce housing demand elasticities. To what degree does heterogeneity in a certain dimension matter for housing markets is therefore a quantitative question. Our policy experiment suggests that self-

assessed financial literacy plays an important role in the housing markets and should hence be incorporated in housing choice models.

Finally, it is important to note that the heterogeneous-agent model was estimated to match the relationships between literacy, housing tenure and leverage across three specific age groups. Remarkably, Table 7 illustrates that it manages to closely capture the average relationship between literacy and housing outcomes across all wealth and age bins. We believe this is a quantitative success of our model.

## 7 Summary and Concluding Remarks

This paper introduces a novel dimension of household heterogeneity that plays an important role in housing markets. First, we document new facts linking self-assessed financial literacy to housing outcomes. We show that households that see themselves as more literate are 1) more likely to own a house and 2) take on more debt against the value of the house. The relationship is economically meaningful and robust after controlling for the potential confounding factors.

Motivated by these empirical patterns, we then ask what are the channels through which financial literacy matters for housing tenure and leverage? We focus on the role of mortgage terms and expectations on future house prices, arguably the most relevant drivers of housing market behavior. We incorporate heterogeneity in financial literacy along these channels in an otherwise standard portfolio choice with housing. We find that households who self-assess themselves as more literate face laxer constraints in the credit markets - they pay a lower spread when borrowing against the value of their house, and are required to a lower down-payment on their mortgage. More literate households also face better risk-return trade-offs in housing markets, suggesting that more literate households engage in better investment deals. Our estimates suggest that differences in credit markets are quantitatively more important for explaining the stylized facts. Moreover, financial literacy matters differently throughout life. At the beginning of life, when mortgage terms matter more for housing decisions, those with higher self-assessed literacy manage to shop for lower mortgage spreads and down-payment requirements. Later on, when returns on housing are more relevant, they make deals with better risk-return trade-offs.

The main takeaway is that accommodating for heterogeneity in financial literacy is quantitatively important for assessing housing market policies. Absent financial literacy heterogeneity, the evaluation of housing market policies can be largely misleading. The estimated housing demand elasticity with respect to wealth is downsized by ap-



proximately 40% when taking financial literacy into account. By ignoring the observed heterogeneity in financial literacy, models without financial literacy over-estimate the correlations between housing outcomes and wealth and age. By uncovering this important dimension of household heterogeneity, our model is able to significantly reduce this bias.

We conclude that financial literacy matters for housing outcomes, and that heterogeneity in mortgage terms and expectations on future house prices can be the underlying mechanism for why literacy matters. Future work evaluating housing market policies should therefore take heterogeneity in financial literacy into consideration.

## References

- Bajari, Patrick, Phoebe Chan, Dirk Krueger, and Daniel Miller.** 2013. "A dynamic model of housing demand: Estimation and policy implications." *International Economic Review*, 54(2): 409–442.
- Burnside, Craig, Martin Eichenbaum, and Sergio Rebelo.** 2011. "Understanding Booms and Busts in Housing Prices." mimeo Northwestern.
- Calvet, Laurent E, John Y Campbell, and Paolo Sodini.** 2009. "Measuring the financial sophistication of households." *American Economic Review*, 99(2): 393–98.
- Campbell, John Y, and Joao F Cocco.** 2003. "Household risk management and optimal mortgage choice." *The Quarterly Journal of Economics*, 118(4): 1449–1494.
- Campbell, John Y, Tarun Ramadorai, and Benjamin Ranish.** 2013. "Getting better: Learning to invest in an emerging stock market." Available at SSRN, 2176222.
- Case, Karl E, and Robert J Shiller.** 1990. "Forecasting prices and excess returns in the housing market." *Real Estate Economics*, 18(3): 253–273.
- Case, Karl E, John M Quigley, and Robert J Shiller.** 2003. "Home-buyers, Housing and the Macroeconomy."
- Cocco, Joao F.** 2004. "Portfolio choice in the presence of housing." *The Review of Financial Studies*, 18(2): 535–567.
- Cocco, Joao 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.
- Davis, Morris A, Andreas Lehnert, and Robert F Martin.** 2008. "The Rent-price ratio for the aggregate stock of owner-occupied housing." *Review of Income and Wealth*, 54(2): 279–284.
- Di Maggio, Marco, and Amir Kermani.** 2017. "Credit-induced boom and bust." *The Review of Financial Studies*, 30(11): 3711–3758.

- Flavin, Marjorie, and Takashi Yamashita.** 2002. "Owner-occupied housing and the composition of the household portfolio." *American Economic Review*, 92(1): 345–362.
- Glaeser, Edward L, and Charles G Nathanson.** 2017. "An extrapolative model of house price dynamics." *Journal of Financial Economics*, 126(1): 147–170.
- Glaeser, Edward L, Joshua D Gottlieb, and Joseph Gyourko.** 2012. "Can cheap credit explain the housing boom?" In *Housing and the financial crisis*. 301–359. University of Chicago Press.
- Guiso, Luigi, and Tullio Jappelli.** 2008. "Financial literacy and portfolio diversification."
- Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli.** 2012. "Financial advisors: A case of babysitters?" *Journal of Banking & Finance*, 36(2): 509–524.
- Kiyotaki, Nobuhiro, Alexander Michaelides, and Kalin Nikolov.** 2011. "Winners and losers in housing markets." *Journal of Money, Credit and Banking*, 43(2-3): 255–296.
- Landvoigt, Tim.** 2017. "Housing demand during the boom: The role of expectations and credit constraints." *The Review of Financial Studies*, 30(6): 1865–1902.
- Landvoigt, Tim, Monika Piazzesi, and Martin Schneider.** 2015. "The housing market (s) of San Diego." *American Economic Review*, 105(4): 1371–1407.
- Li, Wenli, Haiyong Liu, and R Yao.** 2009. "Housing over Time and over the Life Cycle." *Federal Reserve*.
- Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S Mitchell.** 2017. "Optimal financial knowledge and wealth inequality." *Journal of Political Economy*, 125(2): 431–477.
- Mian, Atif, and Amir Sufi.** 2009. "The consequences of mortgage credit expansion: Evidence from the US mortgage default crisis." *The Quarterly Journal of Economics*, 124(4): 1449–1496.
- Mian, Atif, and Amir Sufi.** 2011. "House prices, home equity-based borrowing, and the US household leverage crisis." *American Economic Review*, 101(5): 2132–56.
- Pakes, Ariel, and David Pollard.** 1989. "Simulation and the asymptotics of optimization estimators." *Econometrica: Journal of the Econometric Society*, 1027–1057.
- Piazzesi, Monika, and Martin Schneider.** 2009. "Momentum traders in the housing market: survey evidence and a search model." *American Economic Review*, 99(2): 406–11.
- Piazzesi, Monika, Martin Schneider, and Selale Tuzel.** 2007. "Housing, consumption and asset pricing." *Journal of Financial Economics*, 83(3): 531–569.
- Sommer, Kamila, Paul Sullivan, and Randal Verbrugge.** 2013. "The equilibrium effect of fundamentals on house prices and rents." *Journal of Monetary Economics*, 60(7): 854–870.
- Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie.** 2011. "Financial literacy and stock market participation." *Journal of Financial Economics*, 101(2): 449–472.
- Yao, Rui, and Harold H Zhang.** 2004. "Optimal consumption and portfolio choices with risky housing and borrowing constraints." *The Review of Financial Studies*, 18(1): 197–239.

# Appendix

## Contents

<b>A</b>	<b>Figures and Tables</b>	<b>36</b>
<b>B</b>	<b>Dynamic Programming Solution</b>	<b>38</b>
B.1	Bellman Equation for $t \geq Ret$ . . . . .	38
B.2	Bellman Equation for $t = Ret - 1$ . . . . .	39
B.3	Transformed Model . . . . .	40
B.3.1	Transformed Model for $t \geq Ret$ . . . . .	40
B.3.2	Transformed Model for $t < Ret - 1$ . . . . .	42
B.3.3	Transformed Model for $t = Ret - 1$ . . . . .	43
<b>C</b>	<b>Estimation Procedure</b>	<b>44</b>
C.1	Standard Errors . . . . .	46

# A Figures and Tables

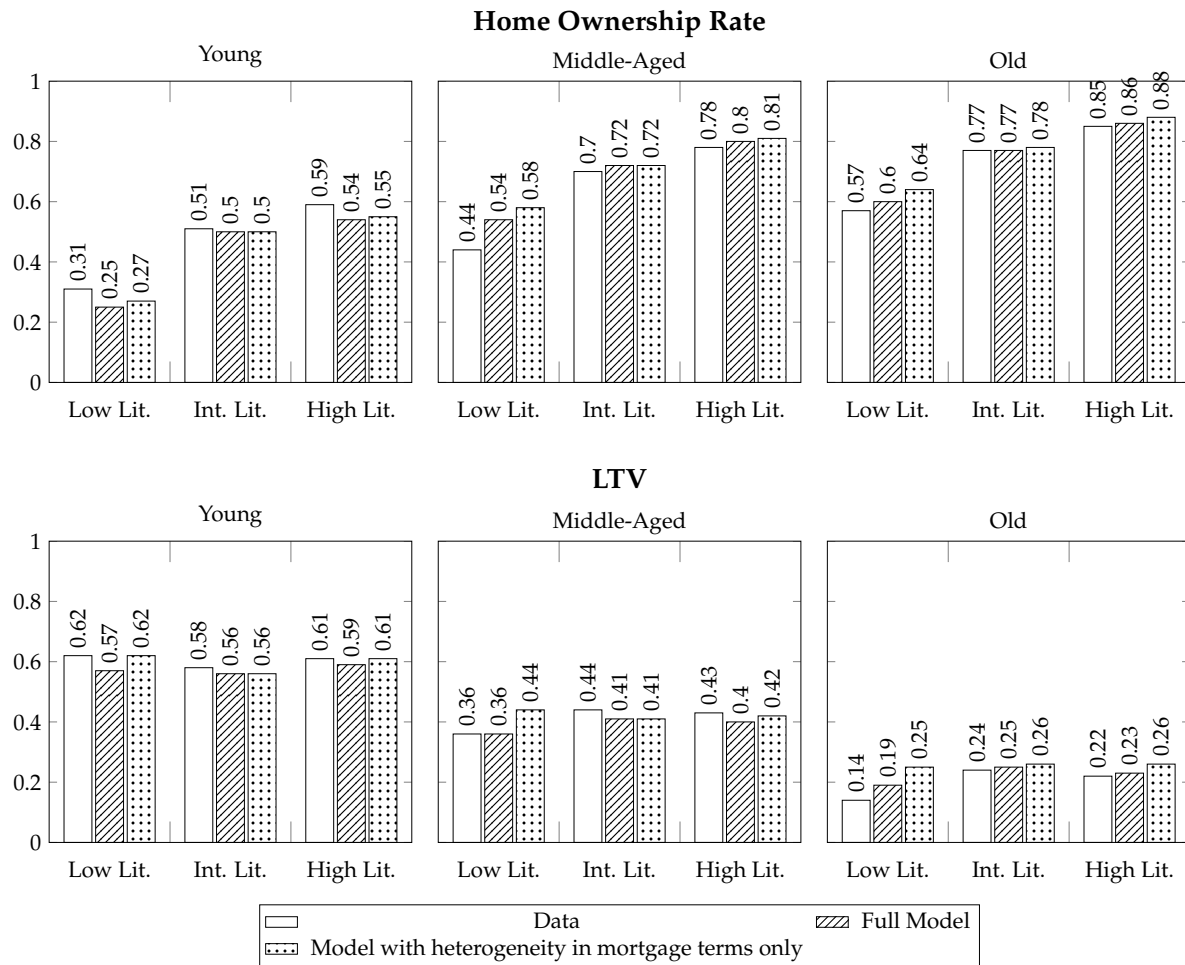


Figure 4: Shutting off Heterogeneity in Expectations on Future Prices

Note: The figure compares between 1) SCF data moments; 2) The full heterogeneous model generated moments ; and 3) Moments generated by a model in which mean expected return ( $\mu(f_i)$ ) and volatility ( $\sigma(f_i)$ ) are set to their benchmark model estimates (Table 5) for all literacy groups  $f_i = \{Low, Int, High\}$  whereas estimates of mortgage spread ( $q(f_i)$ ) and down-payment requirements ( $\delta(f_i)$ ) are taken from the full heterogeneous-agent model.

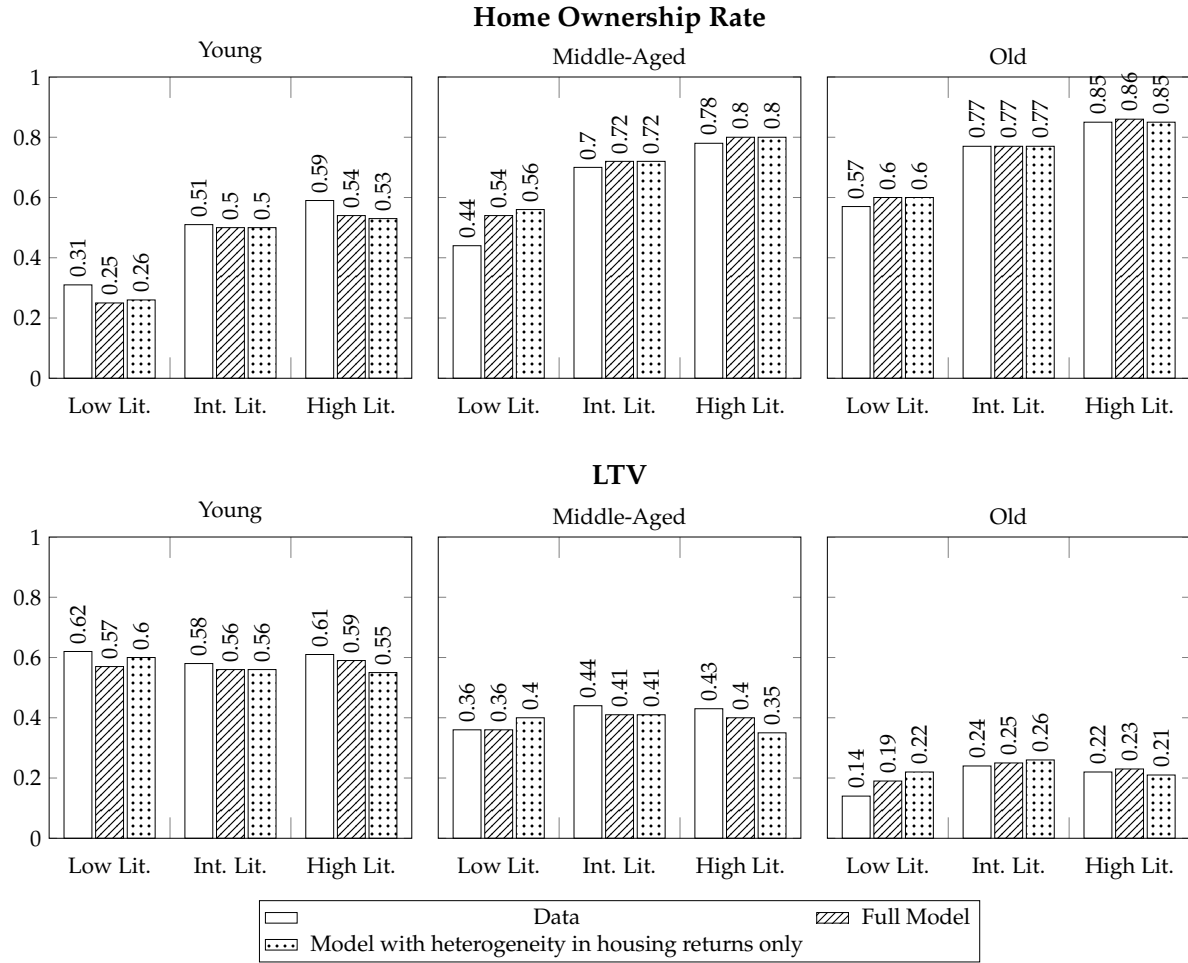


Figure 5: Shutting off Heterogeneity in Mortgage Terms

Note: The figure compares between 1) SCF data moments; 2) The full heterogeneous model generated moments; and 3) Moments generated by a model in which mortgage spread ( $\varrho(f_i)$ ) and down-payment requirements ( $\delta(f_i)$ ) are set to their benchmark model estimates (Table 5) for all literacy groups  $f_i = \{Low, Int, High\}$  whereas estimates of mean expected return ( $\mu(f_i)$ ) and volatility ( $\sigma(f_i)$ ) are taken from the full heterogeneous agent model.

## B Dynamic Programming Solution

Equation 5 specifies the problem faced by household  $i$  at period  $t < Ret - 1$ . For completeness, we will first specify the equivalent problem for  $t \geq Ret - 1$ . Next, we will present a transformation to the model that serves two purposes. The first is improving on efficiency of computation by reducing the state space. As seen below, we are able to dispose of both the permanent income  $\hat{Y}_t$  and the house price index  $P_t$ , thereby allowing for enhanced speed in the estimation procedure. Second, the transformed problem is the basis for comparing the model output to the Survey of Consumers Finance survey data.

### B.1 Bellman Equation for $t \geq Ret$

Define the state variable tuple  $X_t^{Ret} = \{f_i, W_t, P_t, \tau_{t-1}, Q_t, Y_{Ret}, M_t\}$ . The following Bellman equation specifies the household value function after retirement, i.e. for  $t \geq Ret$  :

$$\begin{aligned} \tilde{V}_t(f_i, W_t, P_t, \tau_{t-1}, Q_t, Y_{Ret}, M_t) = & \lambda_t \left\{ \max_{Z_t} u(C_t, S_t) \dots \right. \\ & \left. + \beta \mathbb{E}_t^i [\tilde{V}_{t+1}(f_i, W_{t+1}, P_{t+1}, \tau_t, Q_{t+1}, Y_{Ret}, M_{t+1})] \right\} \dots \quad (9) \\ & + (1 - \lambda_t) D(W_t, P_t) \end{aligned}$$

where  $Z_t$  is the vector of policy variables defined as  $Z_t = \{C_t, H_t, B_t, \tau_t, \zeta_t\}$ . The problem is subject to the same collateral constraint as before:

$$B_t \geq \begin{cases} 0 & \tau_t = 0 \\ -(1 - \delta(f_i)) P_t H_t & \tau_t = 1 \end{cases}$$

subject to the budget constraint:

$$\begin{aligned} W_t &= C_t + B_t + P_t H_t [(1 - \tau_t)\alpha + \tau_t(1 + \psi)] & \tau_{t-1} &= 0 \\ W_t &= C_t + B_t + (1 - \zeta_t)\psi P_t Q_i(H_{t-1})\dots & \tau_{t-1} &= 1 \\ & \dots + (1 - \zeta_t)P_t Q_t (1 - \nu) + \zeta_t [P_t H_t \{(1 - \tau_t)\alpha + \tau_t(1 + \psi)\}] \end{aligned}$$

Note that, compared to the Bellman equation for  $t < Ret - 1$ , the state variable for income is now  $Y_{Ret}$ . The reason is that, for  $t \geq Ret$ , the household gets a fixed portion  $\theta_{Ret}$  of his last income  $Y_{Ret}$ . and therefore, in case the household decides to default, she pays a cost proportional to  $Y_{Ret}$ . Similarly as in the problem before retirement, the evolution of

the endogenous state variables, total wealth and house quality is given by:

$$\begin{aligned} W_{t+1} &= (R + 1_{\{B_t < 0\}}\varrho(f_i))B_t + (1 - \nu(f_i))P_{t+1}Q_{t+1} + \theta_{Ret}Y_{Ret} \\ Q_{t+1} &= \tau_t Q_i(H_t) = \tau_t(1 + g_{i,t+1})H_t \end{aligned}$$

## B.2 Bellman Equation for $t = Ret - 1$

Next, consider the problem faced by household  $i$  one period before retirement, i.e. at period  $t = Ret - 1$ . This period is particularly relevant since the continuation value function is given by  $\tilde{V}_R(\cdot)$ , whereas the current value function is given by  $V_{R-1}(\cdot)$ . More specifically, using the notation of  $X_t$  and  $X_t^{Ret}$  previously defined, the household value function is:

$$\begin{aligned} V_t(X_t) &= \lambda_t \left\{ \max_{Z_t} u(C_t, S_t) + \beta \mathbb{E}_t^i[\tilde{V}_{t+1}(X_{t+1}^{Ret})] \right\} \dots \\ &\quad + (1 - \lambda_t)D(W_t, P_t) \end{aligned} \quad (10)$$

Note that the state variable in the current value function is the permanent income component at  $t = Ret - 1$ , i.e.  $\hat{Y}_{Ret-1}$ , whereas in the continuation value function it is the total income at  $t = Ret$ , i.e.  $Y_{Ret}$ . This problem is subject to the same collateral constraint and budget constraint as in  $t < Ret - 1$  (i.e. equations 6, 7 and ??). Similarly, the evolution of the endogenous state variables is given by:

$$\begin{aligned} W_{t+1} &= \left[ R + 1_{\{B_t < 0\}}\varrho(f_i) \right] B_t + (1 - \nu(f_i))P_{t+1}Q_{t+1} + Y_{t+1} \\ Q_{t+1} &= \tau_t(1 + g_{i,t+1})H_t \end{aligned}$$

where  $Y_{Ret} = \hat{Y}_{Ret-1} \exp(f(Ret) + \hat{\epsilon}_{Ret}^i + \bar{\epsilon}_{Ret} + u_{Ret})$ .<sup>15</sup>

---

<sup>15</sup>This expression follows from the definition of the income process:

$$\log(Y_{Ret}) = f(Ret) + \log(\bar{Y}_{Ret}) + \log(\hat{Y}_{Ret}^i) + u_{Ret}$$

Hence:

$$\begin{aligned} Y_{Ret} &= \exp(f(Ret) + \log(\bar{Y}_{Ret}) + \log(\hat{Y}_{Ret}^i) + u_{Ret}) \\ &= \exp(f(Ret) + \log(\bar{Y}_{Ret-1}) + \log(\hat{Y}_{Ret-1}^i) + \hat{\epsilon}_{Ret}^i + \bar{\epsilon}_{Ret} + u_{Ret}) \\ &= \exp(f(Ret) + \log(\hat{Y}_{Ret-1}) + \hat{\epsilon}_{Ret}^i + \bar{\epsilon}_{Ret} + u_{Ret}) \\ &= \hat{Y}_{Ret-1} \exp(f(Ret) + \hat{\epsilon}_{Ret}^i + \bar{\epsilon}_{Ret} + u_{Ret}) \end{aligned}$$

## B.3 Transformed Model

### B.3.1 Transformed Model for $t \geq Ret$

By backward induction, consider first the problem defined by equation 9 for households who are retired or are about to retire at the end of the period, i.e. for  $t \geq Ret$ . We normalize all the quantities of the model by total income  $Y_{Ret}$  and use the notation  $\tilde{x}$  to denote the normalized variables:

$$\tilde{w}_t = \frac{W_t}{Y_{Ret}}, \quad \tilde{q}_t = \frac{Q_t P_t}{Y_{Ret}}, \quad \tilde{c}_t = \frac{C_t}{Y_{Ret}}$$

$$\tilde{b}_t = \frac{B_t}{Y_{Ret}}, \quad \tilde{h}_t = \frac{P_t H_t}{Y_{Ret}}$$

Denote  $\tilde{v}_t(f_i, \tilde{w}_t, \tau_{t-1}, \tilde{q}_t, M_t) = \frac{V_t(f_i, W_t, P_t, \tau_{t-1}, Q_t, Y_{Ret}, M_t)}{(Y_{Ret} P_t^{-\rho})^{1-\gamma}}$  the normalized value function. In addition denote the normalized policy variables  $\tilde{z}_t = \{\tau_t, \tilde{b}_t, \tilde{h}_t, \tilde{c}_t, \tilde{\zeta}_t\}$  and the new endogenous state variables  $\tilde{x}_t = \{f_i, \tilde{w}_t, \tau_{t-1}, \tilde{q}_t, M_t\}$ . Finally, the normalized bequest function is  $d(\tilde{w}_t) = \bar{D} \frac{\tilde{w}_t^{1-\gamma}}{1-\gamma}$ .

Then the household problem in 9 can be re-written as follows:

$$\begin{aligned} \tilde{v}_t(\tilde{x}_t) = & \lambda_t \left[ \max_{\tilde{z}_t} u(\tilde{c}_t, \tilde{s}_t) + \beta \mathbb{E}_t \left[ v_{t+1}(\tilde{x}_{t+1}) \left( G_{t+1}^P \right)^{-\rho(1-\gamma)} \right] \right] \dots \\ & + (1 - \lambda_t) d(\tilde{w}_t) \end{aligned}$$

where  $\tilde{s}_t = \phi(\tau_t, t) \tilde{h}_t$  and  $G_t^P = \frac{P_t}{P_{t-1}} = \exp\{\epsilon_t^P\}$ , with  $\epsilon_t^P \sim N(0, \sigma_p^2)$ . This problem is subject to the normalized collateral constraint:

$$\tilde{b}_t \geq \begin{cases} 0 & \tau_t = 0 \\ -[1 - \delta(f_i)] \tilde{h}_t & \tau_t = 1 \end{cases} \quad (11)$$

subject to the normalized budget constraint:

$$\tilde{w}_t = \begin{cases} \tilde{c}_t + \tilde{b}_t + \tilde{h}_t [(1 - \tau_t)\alpha + \tau_t(1 + \psi(f_i))] & \tau_{t-1} = 0 \\ \tilde{c}_t + \tilde{b}_t + (1 - \tilde{\zeta}_t) \tilde{q}_t (1 - v(f_i)) \dots & \tau_{t-1} = 1 \\ \dots + (1 - \tilde{\zeta}_t) \psi \tilde{q}_t + \tilde{\zeta}_t \tilde{h}_t [(1 - \tau_t)\alpha + \tau_t(1 + \psi(f_i))] & \end{cases} \quad (12)$$

Finally, the evolution of the endogenous state variables is:



$$\begin{aligned}\tilde{w}_{t+1} &= \left[ R + 1_{\{\tilde{b}_t < 0\}} \varrho(f_i) \right] \tilde{b}_t + (1 - \nu(f_i)) \tilde{q}_{t+1} + \theta_{Ret} \\ \tilde{q}_{t+1} &= \tau_t (1 + g_{i,t+1}) \tilde{h}_t G_{t+1}^P\end{aligned}\quad (13)$$

Note 2 state variables have been taken away: the price  $P_t$  and the income at retirement  $Y_{Ret}$ , thus improving the efficiency of computation.

To solve this problem, it can be divided and simplified into two smaller sub-problems:

- The "moving" problem ( $M$ ). Define the value function of a household who was a previous renter ( $\tau_{t-1} = 0$ ) or that has sold its house ( $\tau_{t-1} = 1, \zeta_t = 1$ ), conditional on surviving to period  $t + 1$ :

$$\begin{aligned}\tilde{v}_t^M(f_i, \tilde{w}_t) &= \max_{\tilde{c}_t, \tilde{b}_t, \tau_t, \tilde{h}_t} u(\tilde{c}_t, \tilde{h}_t) + \beta \mathbb{E}_t \left[ \tilde{v}_{t+1}(f_i, \tilde{w}_{t+1}, \tau_t, \tilde{q}_{t+1}, M_{t+1}) \left( G_{t+1}^P \right)^{-\rho(1-\gamma)} \right] \\ \text{s.t. } \tilde{w}_t &= \tilde{c}_t + \tilde{b}_t + (1 - \tau_t) \alpha \tilde{h}_t + \tau_t \tilde{h}_t (1 + \psi(f_i)) \\ \tilde{b}_t &\geq -(1 - \delta(f_i)) \tau_t \tilde{h}_t\end{aligned}$$

- The "forced-to-stay" problem ( $S$ ). Define the value function of a household who was a home owner and is forced to stay in its house ( $\tau_{t-1} = 1, \zeta_t = 0, \tau_t = 1$ ), conditional on surviving to period  $t + 1$ :

$$\begin{aligned}\tilde{v}_t^S(f_i, \tilde{w}_t, \tilde{q}_t) &= \max_{\tilde{c}_t, \tilde{b}_t} u(\tilde{c}_t, \tilde{q}_t) + \beta \mathbb{E}_t \left[ \tilde{v}_{t+1}(f_i, \tilde{w}_{t+1}, \tau_t, \tilde{q}_{t+1}, M_{t+1}) \left( G_{t+1}^P \right)^{-\rho(1-\gamma)} \right] \\ \text{s.t. } \tilde{w}_t + \nu \tilde{q}_t &= \tilde{c}_t + \tilde{b}_t + (1 + \psi(f_i)) \tilde{q}_t \\ \tilde{b}_t &\geq -(1 - \delta(f_i)) \tilde{q}_t\end{aligned}$$

Note that in these 2 sub-problems the evolution of the endogenous state variables is defined by 13. With this, the household problem can then be rewritten as follows:

$$\begin{aligned}\tilde{v}_t(f_i, \tilde{w}_t, \tau_{t-1}, \tilde{q}_t, M_t) &= \lambda_t \left[ \tau_{t-1} \left[ (1 - M_t) \max \left\{ \tilde{v}_t^M(f_i, \tilde{w}_t), \tilde{v}_t^S(f_i, \tilde{w}_t, \tilde{q}_t) \right\} \dots \right. \right. \\ &\quad \left. \left. + M_t \tilde{v}_t^M(f_i, \tilde{w}_t) \right] + (1 - \tau_{t-1}) \tilde{v}_t^M(f_i, \tilde{w}_t) \right] \dots \\ &\quad + (1 - \lambda_t) d(\tilde{w}_t)\end{aligned}$$

If the household is a previous owner ( $\tau_{t-1} = 1$ ) then the household can either be forced to move by an unexpected shock ( $M_t = 1$ ) or not ( $M_t = 0$ ). If his not forced to move ( $M_t = 0$ ), then she can choose between two alternatives:

- Selling the house ( $\zeta_t = 1$ ) and choose problem  $M$  with value function  $\tilde{v}_t^M(f_i, \tilde{w}_t)$ .

b) Stay in the current house ( $\zeta_t = 0$ ) and choose problem  $S$  with value function  $v_t^S(f_i, \tilde{w}_t, \tilde{q}_t)$ .

If the household is a previous owner and she is forced to move then she faces problem  $M$ . On the other hand, if a household is a previous renter ( $\tau_{t-1} = 0$ ) then she can only face the previous renter problem  $v_t^M(f_i, \tilde{w}_t)$ . Finally, the dynamic problem can be solved recursively starting from period  $T$ , where  $\tilde{v}_T(f_i, \tilde{w}_T) = d(\tilde{w}_T)$ .

### B.3.2 Transformed Model for $t < Ret - 1$

Next, consider the case of a household of age  $t < Ret - 1$ . In this case we normalize quantities by the permanent income  $\hat{Y}_t$ :

$$w_t = \frac{W_t}{\hat{Y}_t}, \quad q_t = \frac{Q_t P_t}{\hat{Y}_t}, \quad c_t = \frac{C_t}{\hat{Y}_t}$$

$$b_t = \frac{B_t}{\hat{Y}_t}, \quad h_t = \frac{P_t H_t}{\hat{Y}_t}$$

Denote by  $x_t = \{w_t, \tau_{t-1}, q_t\}$  the vector of state variables and

$$v_t(f_i, w_t, \tau_{t-1}, q_t, M_t) = \frac{V_t(f_i, W_t, P_t, \tau_{t-1}, Q_t, \hat{Y}_t, M_t)}{(\hat{Y}_t P_t^{-\rho})^{1-\gamma}}$$

the normalized value function. In addition let  $z_t = \{\tau_t, \zeta_t, h_t, c_t, b_t\}$  the vector of policy variables in the normalized problem. It will be useful to note that  $w_t = \frac{\hat{Y}_{t-1}}{\hat{Y}_t} ((R + 1_{\{b_{t-1} < 0\}} \varrho(f_i)) b_{t-1} + (1 - \nu) q_t + \exp\{f(t) + u_t\})$ . Finally, the normalized bequest function is  $d(w_t) = \bar{D} \frac{w_t^{1-\gamma}}{1-\gamma}$ .

Then, the household problem in 5 can be re-written as follows:

$$v_t(f_i, w_t, \tau_{t-1}, q_t, M_t) = \lambda_t \left\{ \max_{z_t} u(c_t, s_t) \dots \right.$$

$$+ \beta \mathbf{E}_t [v_{t+1}(f_i, w_{t+1}, \tau_t, q_{t+1}, M_{t+1}) [G_t^Y (G_t^P)^{-\rho}]^{1-\gamma}] \dots$$

$$\left. + (1 - \lambda_t) d(w_t) \right\}$$

where  $G_t^Y = \frac{\hat{Y}_t}{\hat{Y}_{t-1}} = \exp(\bar{\epsilon}_t + \hat{\epsilon}_t^i)$  and  $G_t^P$  defined as above. It is useful to define  $\epsilon_t^{\hat{Y}} = \bar{\epsilon}_t + \hat{\epsilon}_t^i$  so that  $\epsilon_t^{\hat{Y}} \sim N(0, \sigma_{\hat{Y}}^2)$  and  $G_t^Y = \exp(\epsilon_t^{\hat{Y}})$ . Note that the innovations  $(\bar{\epsilon}_t, \epsilon_t^i, \epsilon_t^{\hat{P}}, u_t)$

are independent across time. . Similarly as in the problem after retirement, we have thus eliminated two state variables: the price  $P_t$  and permanent income  $\hat{Y}_t$ . This problem is subject to the normalized collateral constraint defined above (equation 11). The budget constraint is given by equation 12.

To solve the problem, define the same two sub-problems as in the case  $t \geq Ret$ , but for the period before retirement:

- Problem  $M$ :

$$\begin{aligned} v_t^M(f_i, w_t) &= \max_{c_t, b_t, \tau_t, h_t} u(c_t, s_t) + \beta \mathbb{E}_t[(v_{t+1}(x_{t+1})[G_{t+1}^Y (G_{t+1}^P)^{-\rho}]^{1-\gamma}] \\ \text{s.t. } w_t &= c_t + b_t + (1 - \tau_t)\alpha h_t + \tau_t h_t(1 + \psi) \\ b_t &\geq -(1 - \delta)\tau_t h_t \end{aligned}$$

- Problem  $S$ :

$$\begin{aligned} v_t^S(f_i, w_t, q_t) &= \max_{c_t, b_t} u(c_t, q_t) + \beta \mathbb{E}_t[(v_{t+1}(x_{t+1})[G_{t+1}^Y (G_{t+1}^P)^{-\rho}]^{1-\gamma}] \\ \text{s.t. } w_t + v q_t &= c_t + b_t + (1 + \psi)q_t \\ \text{s.t. } b_t &\geq -(1 - \delta)q_t \end{aligned}$$

The general household problem can then be written as follows:

$$\begin{aligned} v_t(x_t) &= \lambda_t \left[ \tau_{t-1} \left[ (1 - M_t) \max \left\{ v_t^M(f_i, w_t), v_t^S(f_i, w_t, q_t) \right\} \dots \right. \right. \\ &\quad \left. \left. + M_t v_t^M(f_i, w_t) \right] + (1 - \tau_{t-1}) v_t^M(f_i, w_t) \right] \\ &\quad + (1 - \lambda_t) d(w_t) \end{aligned} \tag{14}$$

Basically, in 14 the household faces the same choices as in the after retirement problem. Finally, the evolution of the endogenous state variables is given by:

$$\begin{aligned} w_{t+1} &= [(R + 1_{\{b_t < 0\}} \varrho(f_i)) b_t + (1 + g_{i,t+1}) G_{t+1}^P \tau_t h_t (1 - v(f_i))] \frac{1}{G_{t+1}^Y} \\ &\quad + \exp\{f(t+1) + u_{t+1}\} \\ q_{t+1} &= \tau_t h_t \frac{G_{t+1}^P}{G_{t+1}^Y} (1 + g_{i,t+1}) \end{aligned}$$

### B.3.3 Transformed Model for $t = Ret - 1$

Finally, we complete the transformation by obtaining the relationship between the transformed problems for  $t \geq Ret$  and  $t < Ret - 1$ . Specifically, at  $t = Ret - 1$ , the household

problem is defined by equation 10. Using the notation defined above and some algebra, the normalized household problem in this case can be written as:

$$v_t(x_t) = \lambda_t \left\{ \max_{z_{Ret-1}} u(c_t, s_t) + \beta \mathbb{E}_t[(\tilde{v}_{t+1}(\tilde{x}_{t+1})[G_{t+1}^Y (G_{t+1}^P)^{-\rho} \exp\{f(t+1) + u_{t+1}\}]^{1-\gamma})] \right\} \\ + (1 - \lambda_t)d(w_t)$$

Note that the current value function is  $v_t(\cdot)$  while the continuation value function is  $\tilde{v}_t(\cdot)$ . Also note that in the continuation value, the expression  $\exp\{f(t+1) + u_{t+1}\}$  is the factor that allows to convert normalized variables by  $\hat{Y}_{Ret}$  in variables normalized by  $Y_{Ret}$ , i.e. for instance  $w_{Ret} = \tilde{w}_{Ret} \exp\{f(Ret) + u_{Ret}\}$ . This problem is subject to the same collateral constraint and budget constraint as in the problem before retirement. Next, define the evolution of the endogenous state variables:

$$\hat{w}_{t+1} = \exp\{-f(t+1) - u_{t+1}\} w_{t+1} \\ = \exp\{-f(t+1) - u_{t+1}\} \left[ [(R + 1_{\{l_t < 0\}}) \varrho(f_i)] l_t \dots \right. \\ \left. + \tau_t h_t G_{t+1}^P (1 + g_{i,t+1}) (1 - \nu(f_i)) \right] \frac{1}{G_{t+1}^Y} + \exp\{f(t+1) + u_{t+1}\}$$

and :

$$\tilde{q}_{t+1} = \exp\{-f(t+1) - u_{t+1}\} q_{t+1} \\ = \exp\{-f(t+1) - u_{t+1}\} \tau_t h_t \frac{G_{t+1}^P}{G_{t+1}^Y} (1 + g_{i,t+1})$$

Finally, to solve this problem, we can define the similar sub-problems  $M$ , and  $S$  as in the problem before retirement, but noting that the continuation value is computed with the value function  $\tilde{v}_t(\cdot)$  and the expectation term includes the factor  $\exp\{f(t+1) + u_{t+1}\}$ .

## C Estimation Procedure

We estimate the housing market parameters by applying a Simulated Method of Moments (SMM) to the cross-sectional 2016 SCF data. Denote these parameters by the vector  $\eta = \{\mu_g, \sigma_g^2, \delta, \varrho\}$ .

The estimation proceeds as follows. we begin by simulating a large number of  $I$  households. For each household  $i = 1 \dots I$  and for each age cohort  $t = 25 \dots 79$ :

- a) Compute the value of the normalized state variables at age  $t$  for household  $i$ ,  $\Omega_t^{i,n} = \{f_i, w_{it}, \tau_{it-1}, q_{it}, M_{it}, u_{it}\}$
- i. Using the cross-sectional distribution of the 2016 SCF data, draw the vector of variables  $\Omega_t^i = \{f_{i,t}, \tau_{i,t-1}, W_{it}, P_t Q_{it}, Y_{it}\}$ . That is, the non normalized state variables.
  - ii. Using Markov approximations, draw the idiosyncratic temporary shock  $u_{it}$ , distributed as  $N(0, \sigma_u)$ .
  - iii. Calculate the normalized vector of state variables  $\Omega_t^{i,n}$  using  $q_{i,t} = \frac{P_t Q_{it}}{\hat{Y}_{it}}$  and  $w_{it} = \frac{W_{it}}{\hat{Y}_{it}}$  where the variables are normalized by the permanent component of income  $\hat{Y}_{it} = Y_{it} \exp(f(t) + u_{it})$ . Post retirement, variables are normalized using total income at retirement.
  - iv. Using the age-dependent distribution of moving, draw the moving shock  $M_{it}$ .
- b) Based on the state variables and the policy functions, compute the optimal policy for each household  $i$  at age  $t$ , i.e.  $z_{i,t+1}(\Omega_t^{i,n}) = \{\tau_{i,t}, b_{i,t+1}, h_{i,t+1}, c_{i,t+1}, \xi_{i,t+1}\}$ .
- c) Using Markov approximations, simulate the period  $t + 1$  aggregate house price shock  $\epsilon_{i,t+1}^P$  and owner specific house quality growth  $g_{i,t+1}$ . Moreover, if  $t < Ret - 1$ , simulate the period  $t + 1$  permanent shock to income  $\epsilon_{i,t+1}^{\hat{Y}}$  and idiosyncratic temporary shock  $u_{it+1}$ .
- d) Finally, compute the evolution of wealth  $w_{i,t+1}$  and housing quality  $q_{i,t+1}$  using the evolution equations from Section B.

This yields realizations of state variables  $\hat{\Omega}_{t+1}^{i,n}(\Omega_t^{i,n}, \eta)$  and policy variables  $z_{i,t+1}(\Omega_t^{i,n}, \eta)$  for each household  $i = 1 \dots I$  and each age cohort  $t = 26 \dots 80$ . Note that these realizations depend on the value of the parameters  $\eta$ . Next, we compare model simulated moments to moments in the data.

For the benchmark model, we define 3 age groups (young  $y$  for age  $t = 26 \dots 40$ , middle-aged  $m$  for  $t = 41 \dots 60$  and old  $o$  for  $t = 61 \dots 80$ ) and compute the weighted average of ownership  $\bar{\tau}_a$ , and loan-to-value ratio  $L\bar{T}V_a$  (computed only among current owners in that period) for each age group  $a \in \{y, m, o\}$ :

$$\bar{\Theta}_a(\eta) = \frac{1}{|a|} \sum_{t \in a} \sum_{i=1}^I \omega_i \hat{\Theta}_{i,t}(\eta) \quad , \hat{\Theta} = \{\tau, LTV\}.$$

where the weights  $\omega_i$  are extracted from the 2016 SCF data and  $|a|$  is the number of periods in each of the age groups. This yields 6 moments which we compare to the em-

pirical moments  $\bar{\Theta}_a$  computed using the 2016 SCF data. Additionally, we add aggregate wealth as a moment. The set of 5 parameters  $\eta = \{\mu, \sigma, \rho, \delta, \beta\}$  is estimated to minimize the mean of the square error of the 7 simulated moments with respect to their empirical counterpart :

$$\hat{\eta}_{SMM} = \underset{\eta}{\operatorname{argmin}} \frac{1}{7} \left( \sum_{\Theta \in \{\tau, LTV\}} \sum_{a \in \{y, m, o\}} \left( \bar{\Theta}_a - \hat{\Theta}_a(\eta) \right)^2 + (\bar{w} - \hat{w}(\eta))^2 \right).$$

where  $\bar{w}$  is the weighted average wealth-to-income in the data and  $\hat{w}(\eta)$  is its model counterpart. Regarding the model with heterogeneity in financial literacy, the simulated moments are computed for each age group defined previously but now also for each level of financial literacy  $f = \{low, medium, high\}$ . Thus:

$$\bar{\Theta}_{a,f}(\eta) = \frac{1}{|a|} \sum_{t \in a} \sum_{i \in f} \omega_i \hat{\Theta}_{i,t}(\eta) \quad , \hat{\Theta} = \{\tau, LTV\}.$$

This yields 19 simulated moments. Similarly, the set of 13 parameters  $\eta = \{\mu(f_i), \sigma(f_i), \rho(f_i), \delta(f_i), \beta\}$  is estimated to minimize the mean of the square error of the 19 simulated moments  $\hat{\Theta}_{a,f}$  with respect to their empirical counterpart  $\bar{\Theta}_{a,f}$ :

$$\hat{\eta}_{SMM} = \underset{\eta}{\operatorname{argmin}} \frac{1}{19} \left( \sum_{\Theta \in \{\tau, LTV\}} \sum_{a \in \{y, m, o\}} \sum_{f \in \{low, med, high\}} \left( \bar{\Theta}_{a,f} - \hat{\Theta}_{a,f}(\eta) \right)^2 + (\bar{w} - \hat{w}(\eta))^2 \right)$$

## C.1 Standard Errors

We estimate the standard errors of the estimated SMM parameters based on the seminal work of [Pakes and Pollard \(1989\)](#). Given a data sample of size  $n$ , denote the set of targeted moments as  $M_n$  and the model simulated moments as  $S(\eta)$ . Hence, the problem can be rewritten as:

$$\hat{\eta}_{SMM} = \underset{\eta}{\operatorname{argmin}} (M_n - S(\eta))' (M_n - S(\eta))$$

Denote by  $J$  the Jacobian matrix of the function  $\eta \rightarrow S(\eta)$  and  $n.s$  the number of observations simulated with the model. Moreover, denote as  $\Omega$  the asymptotic variance of  $M_n$ . Then, it can be shown that:

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{d} N \left( 0, \left(1 + \frac{1}{s}\right) (J'J)^{-1} J' \Omega J (J'J)^{-1} \right)$$

We compute  $J$  numerically by calculating small changes of the function  $\eta \rightarrow S_n(\eta)$  at  $\eta = \hat{\eta}$ . Furthermore, we estimate  $\Omega$  by bootstrap. With this, the asymptotic variance of  $\hat{\eta}$  is:

$$\text{Var}(\hat{\eta}) = \frac{1}{n} \left(1 + \frac{1}{s}\right) (J'J)^{-1} J' \hat{\Omega} J (J'J)^{-1}$$