Financial Literacy and Mortgage Stress

Mingzhi Hu
School of Management
Chinese Academy of Housing and Real Estate
Zhejiang University of Technology, Hangzhou, China
hu_mingzhi@outlook.com

Zhenguo Lin
Hollo School of Real Estate
College of Business
Florida International University, Florida, USA
zlin@fiu.edu

Yingchun Liu
Department of Finance, Insurance, Real Estate and Law
G. Brint Ryan College of Business
University of North Texas, Texas, USA
yingchun.liu@unt.edu

November 14, 2023

Abstract

This paper examines the effect of financial literacy on mortgage stress. Using data from the Panel Study of Income Dynamics (PSID), we find that borrowers with high levels of financial literacy are 60.3 percent less likely to suffer from mortgage stress than borrowers with low levels of financial literacy after controlling for observables. Our estimated results are robust to potential sample selection bias and functional misspecification. In addition, we also find that the effect of financial literacy varies across borrowers of different ages. Further analysis reveals strong cross effects of financial literacy and quantitative reasoning on mortgage stress.

Keywords: Financial literacy; Mortgage stress; Financial industry; Household finance
JEL: D10; R20; G10; G21

Financial Literacy and Mortgage Stress

November 14, 2023

Abstract

This paper examines the effect of financial literacy on mortgage stress. Using data from the Panel Study of Income Dynamics (PSID), we find that borrowers with high levels of financial literacy are 60.3 percent less likely to suffer from mortgage stress than borrowers with low levels of financial literacy after controlling for observables. Our estimated results are robust to potential sample selection bias and functional misspecification. In addition, we also find that the effect of financial literacy varies across borrowers of different ages. Further analysis reveals strong cross effects of financial literacy and quantitative reasoning on mortgage stress.

Keywords: Financial literacy; Mortgage stress; Financial industry; Household finance

JEL: D10; R20; G10; G21

1 Introduction

Housing is the single biggest asset on the balance sheets of most households [Bian et al., 2018, Chen et al., 2020, Hardin et al., 2023], and mortgages are the most significant debts for most families in nearly all market economies [Agarwal et al., 2017]. The expansion in mortgage credit in the early-to-mid 2000s driven by looser lending constraints in the U.S. mortgage market has a profound impact on real estate and financial markets [Justiniano et al., 2019]. Easier access to credit led to a record homeownership rate in 2004 when it reached the peak of 69.2% and resulted in an unprecedented rise in home prices [Acolin et al., 2016, Di Maggio and Kermani, 2017, Favara and Imbs, 2015]. The expansion of credit and increase in homeownership rate co-occurred despite sharply declining relative (and in some areas absolute) income growth over the same time [Mian and Sufi, 2009], resulting in a huge increase in mortgage debts [Justiniano et al., 2019]. Evidence suggests that the average homeowner extracts 25 to 30 cents for every dollar increase in home equity [Mian and Sufi, 2011]. With the dramatic decline in house prices starting in 2007, late payments on mortgages and defaults prevailed [Foote et al., 2008, Mayer et al., 2009]. Data from CoreLogic showed that more than half of U.S. subprime mortgages originated in 2006 and 2007 became defaulted after 5 years, and many more homeowners were behind on their mortgage payments [Gerardi et al., 2013].

Most previous studies focus on the reasons for the expansion of mortgages to high-risk borrowers, such as the rapid increase in credit supply, the relaxation of underwriting standards of mortgage credit, and inattentive actors in the financial market [Coval et al., 2009, Gennaioli et al., 2012, Jiang et al., 2014, Mian and Sufi, 2009], while too little attention has been paid to the question of why many borrowers take out mortgages that they could not repay [Gerardi et al., 2013]. In particular, there has been little discussion about the role of borrower's financial literacy in making decisions involved in the mortgage process [Bajo and Barbi, 2018, Duca and Kumar, 2014, Gathergood and Weber, 2017a], although several studies on the relationship between financial literacy and participation in financial

markets exist [Hsiao and Tsai, 2018, Klapper et al., 2013, Van Rooij et al., 2011]. This study examines how financial literacy affects mortgage stress.

Closely related to this study, two papers attempt to investigate whether there is a difference in mortgage delinquency or default between financial and non-financial professionals [Agarwal et al., 2017], and across borrowers with different levels of ability to perform basic mathematical calculations [Gerardi et al., 2013]. In particular, Agarwal et al. [2017] find that borrowers working in the financial industry are less likely to become delinquent than others by using a dataset from a leading American subprime lender that consists of residential mortgage loans originated from 1997 to 2007. Gerardi et al. [2013] present empirical evidence that numerical ability is negatively associated with the propensity to mortgage default by using a survey of subprime mortgage borrowers who took out mortgages in 2006 or 2007. This study differs from them in two important ways. First, Agarwal et al. [2017] argue that one of the reasons why financial professionals are associated with lower mortgage delinquency might be that "financial professionals [likely] have higher financial literacy". However, they don't have information about borrowers' financial literacy, so they are unable to test whether their statement is true. By utilizing the information of both financial professionals and financial literacy in the PSID data, we show that it is the financial literacy that helps to reduce mortgage stress. This is a significantly different result between our paper and Agarwal et al. [2017], i.e., we find it is the actual knowledge that reduces mortgage stress, not just working in the financial industry. Given the information in our data and our analysis, we can estimate the effect of improving financial literacy on mortgage stress. Indeed, we find that borrowers with high levels of financial literacy are about 60.3 percent less likely to suffer from mortgage stress than borrowers with low levels of financial literacy, after controlling for observables. Our estimated results are robust to potential sample selection bias and functional misspecification. The policy implications of these analyses and estimates are important. By improving financial literacy, we can greatly reduce mortgage stress that is detrimental to homeowners. In addition, we also find variations in the effect of

financial literacy across borrowers of different ages. Further analysis shows that quantitative reasoning skills can further reduce mortgage stress for borrowers with high levels of financial literacy, while quantitative reasoning ability alone has no significant effect on mortgage stress. Overall, this study highlights the important role of financial literacy in household's mortgage stress.

Second, we use the data from a nationally representative survey (i.e., PSID) including all types of mortgages rather than subprime mortgages in Agarwal et al. [2017] and Gerardi et al. [2013]. [1] After the 2007-2010 subprime crisis, lenders adopt much strict lending guidelines, and subprime mortgages no longer exist. Our analysis has a much broader policy implication and complements the previous studies.

The rest of this paper is structured as follows. Section 2 describes our data and presents some summary statistics. Section 3 analyzes the effect of financial literacy on mortgage stress. Section 4 conducts several robustness checks on the results. Section 5 conducts the heterogeneous effects of financial literacy, and Section 6 concludes with remarks.

2 Data

The data used in the study are from the Panel Study of Income Dynamics (PSID). The PSID is conducted by the University of Michigan Survey Center, and it is a longitudinal household survey started in 1968 with a sample of over 18,000 individuals living in about 5,000 families in the United States. Individuals in each household were followed annually from 1968 to 1997, and biannually after 1997. Since 2009, the PSID adds a question related to the household's mortgage stress, in which the PSID questionnaire asks each respondent whether he/she or anyone in his/her family living there is currently behind his/her mortgage payments. We use data from four waves of the survey after 2009 for our analysis.

[1] According to the data from Mortgage Bankers Association, subprime mortgages consisted of 6 to 14 percent of the total residential mortgage market during the period of 2004 to 2007, and they essentially disappeared after the 2007-2010 subprime crisis due to lenders' much strict lending guidelines. Subprime borrowers are often associated with higher debt-to-income ratio (DTI), higher loan-to-value ratio (LTV), higher interest rates, and lower credit scores, compared with their prime counterparts.

The PSID dataset is appropriate for the current study in several respects. First, the dataset is a nationally representative sample of the U.S. population, and it contains detailed household demographic information (i.e., gender, age, marital status, number of children, race, and religious belief), socioeconomic characteristics (i.e., education, self-employment, household wealth, and household income), and mortgage loan features (i.e., interest rate and interest rate type (adjustable-rate mortgage or fixed-rate mortgage)). With the information on housing value, we are able to obtain current loan-to-value (LTV) ratios as well as debt-to-income (DTI) ratios. These details allow us to control for the common variables that are found to be critical in explaining mortgage decisions [Agarwal et al., 2017, Gerardi et al., 2013, Lin et al., 2016]. Also, since 2009 the respondents in the PSID are asked whether they have had difficulties recently making their mortgage payments. [2] Mortgage stress is defined as an indicator variable that takes the value of 1 for respondents who report that they or anyone in their family living there are currently behind their mortgage payments, and zero otherwise.

Second, the PSID also asks each respondent about the industry they are currently working in. The literature suggests that mortgage decisions are different among individuals working in different industries due to the differences in the access to advice networks, the degree of concern about reputations for loan repayment, and the level of financial literacy [Agarwal et al., 2017]. In this paper, we also examine the effect of working in the financial industry on the mortgage stress of borrowers. [3] In addition, the PSID conducted a supplementary study

[2] In the PSID, the respondents are asked the following question: "Some people have had difficulties recently making their mortgage or loan payments. Are you, or anyone in your family living there currently behind on your (mortgage/loan) payments?".

[3] In line with the U.S. Department of Commerce and the Bureau of the Census, the PSID classifies 19 types of industries: 1. Agriculture, Forestry, Fishing, and Hunting; 2. Mining; 3. Utilities; 4. Construction; 5. Manufacturing; 6. Wholesale Trade; 7. Retail Trade; 8. Transportation and Warehousing; 9. Information; 10. Finance and Insurance; 11. Real Estate and Rental and Leasing; 12. Professional, Scientific, and Technical Services; 13. Management, Administrative and Support, and Waste Management Services; 14. Educational Services; 15. Health Care and Social Assistance; 16. Arts, Entertainment, and Recreation; 17. Accommodations and Food Services; 18. Other Services (Except Public Administration); 19. Public Administration and Active Duty Military. Similar to Agarwal et al. [2017], we construct an indicator variable of financial professionals, which equals 1 if the borrower works in the finance and insurance industry and 0 otherwise.

in 2016, which is the PSID's wellbeing and daily life supplement (PSID-WB). The PSID-WB questionnaire includes 11 sections and is designed as a relatively brief self-administered instrument to collect information on wellbeing, personality traits, and everyday skills. [4] In a section of PSID-WB, the financial literacy of respondents is directly measured by six questions that test the ability to understand and effectively use financial skills. [5] These questions have been previously used in the Health and Retirement Study, the English Longitudinal Study of Ageing, and the Survey of Health, Ageing and Retirement in Europe [Banks et al., 2010]. We consider a respondent to have high levels of financial literacy if he/she correctly answers all these 6 questions. We use the unique identification numbers assigned by the PSID to link the supplement data (PSID-WB) to our main data. We drop observations with missing values, and our final four waves panel dataset contains over 5,000 observations.

Table 1 provides summary statistics for the full sample, subsamples working in financial vs. non-financial industries, and subsamples with high vs. low levels of financial literacy. In total, there are 5,099 households in our sample, of which 6.37% work in the financial industry, and 46.38% have high levels of financial literacy. In the dataset, 3.02% of respondents have mortgage stress. For the interest rate type (FRM vs. ARM), 91.96% of borrowers choose fixed-rate mortgages, and the average interest rate is 4.70%. The LTV and DTI ratios, two measures that reflect the willingness and ability of borrowers to repay the mortgage debt, are 65.32% and 15.55%, respectively [Agarwal et al., 2017, Bajo and Barbi, 2018, Gathergood and Weber, 2017b,a]. The average age of the respondents in our sample is about 48 years

[4] There are 11 sections (sections A-K) in the PSID-WB questionnaire, Section A: life satisfaction and flourishing, Section B: positive and negative emotions, Section C: activities and experienced wellbeing, Sections D and E: personality and self-efficacy, Section F: trust-hostility, Section G: achievement orientation, Section H: verbal reasoning, Section I: health literacy, Section J: quantitative reasoning, and Section K: financial literacy.

[5] Six questions are designed to measure the financial literacy of respondents, which are: 1) If you buy a drink for 85 cents and pay with one dollar, how much change should you get back?, 2) A shop is having a half-price (50% off) sale. Before the sale, a sofa costs $300. How much will it cost on sale?, 3) If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease?, 4) A used car dealer is selling a car for $6,000. This is two thirds of what it cost new. How much did the car cost new?, 5) If 5 people all have the winning numbers in the lottery and the prize is $2 million, how much will each of them get?, and 6) Suppose you have $200 in a savings account. The account earns 10 percent interest each year. How much would you have in the account at the end of two years?

old, 83.11% of respondents are white, 74.21% have a college or above education, and 10.51% are self-employed. For household composition, 76.6% of respondents are married, and the number of children under 18 years of age living at home is 0.84. The average annual household income and wealth are $117,289 and $257,585, respectively.

Preliminary results from Table 1 reveal three interesting observations that motivate us to examine the issues further. First, households working in the financial industry have a much lower rate of mortgage stress (1.54%) than those working in the non-financial industry (3.12%). This result is in fact consistent with the finding by Agarwal et al. [2017] that financial professionals are much less likely to have mortgage delinquency than non-financial professionals. Second, mortgage stress differs greatly between borrowers with high levels of financial literacy and borrowers with low levels of financial literacy. Households with low levels of financial literacy have a much higher rate of mortgage stress (4.5%) than those with high levels of financial literacy (1.31%), and the difference is 3.19 percentage points and statistically significant at the 1% level. Third, there are clear differences in household and loan characteristics between financial professionals and non-professionals as well as between households with high levels of financial literacy and those with low levels of financial literacy. For example, households with higher income, wealth, and education tend to work in the financial industry and have a high level of financial literacy. The mortgage interest rate tends to be lower for financial professionals and those with high levels of financial literacy. These systematic differences in observables highlight the importance of controls in the following analysis we conduct.

Table 1: Summary statistics

Variable	Full sample	Working industry				Financial literacy			
		Financial industry	Other industries	Difference	p-value	Higher levels	Lower levels	Difference	p-value
Mortgage stress	0.0302	0.0154	0.0312	0.0158	0.1068	0.0131	0.0450	0.0319	0.0000
Financial industry	0.0637					0.0820	0.0479	0.0341	0.0000
Financial literacy	0.4638	0.5969	0.4548	0.1421	0.0000				
FRM	0.9196	0.9046	0.9206	0.0160	0.3049	0.9032	0.9338	0.0306	0.0001
Interest rate	4.6958	4.3262	4.7210	0.3948	0.0001	4.3505	4.9945	0.6440	0.0000
LTV	0.6532	0.7107	0.6493	0.0614	0.0028	0.6322	0.6713	0.0391	0.0001
DTI	0.1555	0.1554	0.1555	0.0001	0.9765	0.1493	0.1609	0.0116	0.0000
Male	0.8613	0.8154	0.8645	0.0491	0.0132	0.8989	0.8288	0.0701	0.0000
Age	48.032	46.000	48.171	2.1707	0.0029	48.039	48.027	0.0114	0.9745
Married	0.7660	0.7538	0.7669	0.0131	0.5918	0.8106	0.7275	0.0831	0.0000
Number of children	0.8353	0.9292	0.8289	0.1003	0.1233	0.8427	0.8288	0.0139	0.6634
White	0.8311	0.8677	0.8287	0.0390	0.0691	0.9488	0.7293	0.2195	0.0000
Religion	0.1243	0.0923	0.1265	0.0342	0.0705	0.1285	0.1207	0.0078	0.3977
College	0.7421	0.9108	0.7306	0.1802	0.0000	0.8816	0.6214	0.2602	0.0000
Self-employed	0.1051	0.1231	0.1039	0.0192	0.2754	0.1061	0.1042	0.0019	0.8265
Household income	117,289	170,356	113,676	56,680	0.0000	138,954	98,548	40,406	0.0000
Household wealth	257,585	365,122	250,265	114,858	0.0058	361,981	167,280	194,701	0.0000
Observations	5,099	325	4,774			2,365	2,734		

Note: This table reports the means of loan and borrower characteristics for the full sample, subsample of individuals working in the financial industry, subsample of individuals working in non-financial industries, subsample of individuals with higher levels of financial literacy, and subsample of individuals with lower levels of financial literacy. Mortgage stress is an indicator variable that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in the financial industry and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. FRM is an indicator of fixed-rate mortgage that equals 1 if the interest rate on the loan is fixed and 0 otherwise. Interest rate is the current interest rate on the loan; LTV is the loan-to-value ratio, i.e., the ratio between the principal of the loan and the value of the real property; DTI is the ratio between the total monthly mortgage payments and the total monthly income of the household. Male is an indicator variable of male borrowers. Age is the age of borrower. Married is an indicator variable of being married. Number of children refers to the number of children under 18 years of age in the household. White is an indicator variable of the borrower being white. Religion is an indicator variable of the borrower with religious beliefs. College is an indicator variable of the borrower having a college degree or higher. Self-employed is an indicator variable of the borrower's main job being self-employed only. Household income refers to the total household income. Household wealth refers to the total value of household assets excluding home equity. Data source comes from four waves of the PSID.

3 Main Empirical Results

We use the following logit regression to examine the relationship between financial literacy and mortgage stress. We also include an indicator variable of working in the financial industry in the regression suggested by Agarwal et al. [2017].

$$\begin{aligned} Mortgage\,stress_{ijt} = \beta_0 + \beta_1 Financial\,industry_{ijt} + \beta_2 Financial\,literacy_{ijt} \\ + \alpha\, X_{ijt} + \theta_j + \delta_t + \epsilon_{ijt} \end{aligned} \tag{1}$$

where the dependent variable $Mortgage\,stress_{ijt}$ is an indicator variable of mortgage stress that equals 1 if borrower i in state j and year t has had difficulties recently making mortgage payments and 0 otherwise. $Financial\,industry_{ijt}$ is an indicator variable of financial profession that equals 1 if the borrower works in a financial industry and 0 otherwise. $Financial\,literacy_{ijt}$ is an indicator variable of financial literacy that equals 1 if the borrower has a high level of financial literacy and 0 otherwise. A respondent is considered to have high levels of financial literacy if he/she correctly answers all six questions designed to evaluate financial skills in daily life. X_{ijt} is a vector of loan characteristics and household socioeconomic and demographic controls, as described in Table 1. We control for four mortgage variables in our analysis as suggested by previous studies [Agarwal et al., 2017, Bajo and Barbi, 2018, Duca and Kumar, 2014, Gathergood and Weber, 2017a, Gerardi et al., 2010]: mortgage interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. On top of the common household characteristics that could impact mortgage decisions, we also consider several important household demographic and socioeconomic variables. For example, we add an indicator of self-employment to account for those who are already exposed to high risk in the labor market and may therefore be more likely to have mortgage stress [Agarwal et al., 2017, Heaton and Lucas, 2000, Van Rooij et al., 2011]. We also control for religious beliefs since religion, as a set of social norms, not only cultivates an individual's ethical behavior, but also promotes risk aversion in the mortgage market [Con-

klin et al., 2022]. Year dummies (δ_t) and state dummies (θ_j) are also included to control for the time trend and time-invariant characteristics at the state level. Finally, ϵ_{ijt} is the error term.

We estimate a series of different specifications by gradually increasing the number of controlled variables in X_{ijt} and lastly adding financial literacy to study their effects on mortgage stress. The estimated coefficients, standard errors, significance levels, and marginal effects are reported in Table 2. We begin with the simplest specification by controlling for financial industry, year and state fixed effects only, and we report the results in column (1) of Table 2. The estimated results indicate that, without controlling for household and loan characteristics, borrowers working in the financial industry are 2.65 percentage points less likely to suffer from mortgage stress, and the difference is statistically significant at the 1% level. Given that the average mortgage stress is 3.02 percent in the data (see Table 1), the 2.65 percentage points decrease represents an 87.75% decrease in the likelihood of households experiencing mortgage stress if they are working in the financial industry. [6] As a first step toward measuring the effect of working in the financial industry on mortgage stress, in Specification 2 we further control for loan and household demographic & socioeconomic characteristics. The results are reported in column (2) of Table 2. After controlling for household and loan characteristics, the marginal effect of working in the financial industry decreases to 2.52% and remains significant at the 10% level. A similar finding has been found by Agarwal et al. [2017] who use a dataset from a leading American subprime lender and show a lower probability of being delinquent for financial professionals compared with non-financial professionals.

Agarwal et al. [2017] suggest that the financial occupation may lead to lower delinquency rates through various mechanisms, and one of which is that financial professionals may have high financial literacy. To test this potential mechanism, we further add the financial literacy variable in the regression model by utilizing this information from the PSID data. The results

[6] 2.65 / 3.02 = 87.75%

are reported in column (3) of Table 2. As expected, we find that financial literacy is negatively associated with mortgage stress. In particular, borrowers with high levels of financial literacy are 1.82 percentage points less likely to have mortgage stress holding everything else constant, and the difference is significant at the 1% level. This result indicates that homeowners with a high level of financial literacy are 60.3% (1.82 / 3.02 = 60.3%) less likely to experience mortgage stress. Interestingly, when the financial literacy of borrowers is controlled for, the impact of working in the financial industry becomes insignificant. This provides supportive evidence on the potential mechanism suggested by Agarwal et al. [2017] that the lower delinquency rates of financial professionals relative to non-financial professionals can be explained by the difference in financial literacy. Put differently, our results suggest that it is the actual knowledge of financial literacy to help reduce mortgage stress. [7]

The estimated coefficients of other control variables are also expected. For example, higher interest rates increase the cost of borrowing and imply a higher probability of becoming delinquent and suffering from mortgage stress [Agarwal et al., 2017]. Engaging in self-employment leads to a higher likelihood of having mortgage stress since self-employers have higher income uncertainty [Agarwal et al., 2017, Heaton and Lucas, 2000, Van Rooij et al., 2011]. Several studies have documented that religious beliefs impact an individual's economic preferences such as risk aversion and consequently economic decision-making and behaviors [Benjamin et al., 2016]. We in fact find evidence that households with a religion are less likely to have mortgage stress [Conklin et al., 2022, Hilary and Hui, 2009].

[7] We also use the dominance analysis to determine the relative importance of control variables [Grömping, 2007]. The dominance statistics are derived as a weighted average marginal contribution to the overall fit statistic. If control variable A has a larger dominance statistic than control variable B, control variable A has a larger marginal contribution to the explained variance than control variable B. The results in Table A1 in the appendix suggest that 5.62% of the explained variance for mortgage stress can be attributed to financial literacy, and only 0.48% can be attributed to financial industry.

Table 2: Impact of financial profession and financial literacy on mortgage stress

	(1)		(2)		(3)	
	Coef.	Marginal effect	Coef.	Marginal effect	Coef.	Marginal effect
Financial industry	-0.8232**	-0.0265	-0.8921*	-0.0252	-0.7986	-0.0224
	(0.4176)		(0.5362)		(0.5055)	
Financial literacy					-0.6481***	-0.0182
					(0.2290)	
FRM			-0.3571	-0.0101	-0.4050	-0.0114
			(0.2927)		(0.2954)	
Interest rate			0.2162***	0.0061	0.2116***	0.0059
			(0.0435)		(0.0441)	
LTV			0.8161	0.0230	0.7709	0.0217
			(0.8092)		(0.8383)	
DTI			5.2734***	0.1489	5.5196***	0.1552
			(1.4411)		(1.5228)	
Male			-0.9646***	-0.0272	-0.9758***	-0.0274
			(0.3512)		(0.3580)	
Age			-0.0046	-0.0001	-0.0049	-0.0001
			(0.0117)		(0.0120)	
Married			0.4243	0.0120	0.4618	0.0130
			(0.3267)		(0.3322)	
Number of children			0.1833**	0.0052	0.1787**	0.005
			(0.0886)		(0.0894)	
White			-1.3187***	-0.0372	-1.1491***	-0.0323
			(0.2447)		(0.2504)	
Religion			-0.6209**	-0.0175	-0.6375**	-0.0179
			(0.2622)		(0.2601)	
College			-0.2521	-0.0071	-0.1377	-0.0039
			(0.2151)		(0.2209)	

Table 2: Continued

	(1)		(2)		(3)	
	Coef.	Marginal effect	Coef.	Marginal effect	Coef.	Marginal effect
Self-employed			1.0568***	0.0298	1.0262***	0.0288
			(0.2682)		(0.2728)	
Ln(Household income)			-0.2039	-0.0058	-0.1483	-0.0042
			(0.2690)		(0.2821)	
Ln(Household wealth)			-0.0758***	-0.0021	-0.0774***	-0.0022
			(0.0190)		(0.0190)	
Other controls						
State fixed effects		Yes		Yes		Yes
Year fixed effects		Yes		Yes		Yes
Observations		4,542		4,542		4,542
R-squared		0.0534		0.2432		0.2492

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the full sample. Results are estimated with logistic regressions. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. FRM is an indicator of fixed-rate mortgage that equals 1 if the interest rate on the loan is fixed and 0 otherwise. Interest rate is the current interest rate on the loan; LTV is the loan-to-value ratio, i.e., the ratio between the principal of the loan and the constructed value of the real property; DTI is the ratio between the total monthly mortgage payments and the total monthly income of the household. Male is an indicator variable of male borrowers. Age is the age of the borrower. Married is an indicator variable of being married. Number of children refers to the number of children under 18 years of age in the household. White is an indicator variable of the borrower being white. Religion is an indicator variable of the borrower with religious beliefs. College is an indicator variable of the borrower having a college degree or higher. Self-employed is an indicator variable of the borrower's main job being self-employed only. Household income refers to the total household income. Household wealth refers to the total value of household assets excluding home equity. The logit model predicts the outcome on the basis of independent variables and drops whatever observations that perfectly predict one or the other outcome, and 557 observations are dropped during estimation, see pages 1328-1329 for a good discussion on this issue in Stata Base Reference Manual Release 17. We test the significance of the difference in the coefficient of Financial Industry in specifications (2) and (3) by using seemingly unrelated regression [Wooldridge, 2010]. The results show that the change in the coefficient of Financial Industry from significance (-0.8921, $p<0.10$) to non-significance (-0.7986) is indeed statistically significant (χ^2=3.30, p-value=0.0694). Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

4 Robustness Check

There are two potential issues, which may bias our estimated results. They include (1) sample selection bias, and (2) functional misspecification. After addressing these two issues, we also conduct a falsification test, in which we randomly and artificially assign financial literacy to households. Our findings are robust to all these issues.

4.1 Sample selection bias

We observe the mortgage stress only for mortgage-indebted households. If the subsample of mortgage-indebted households is not a random sample of the entire households, then our previous estimators may suffer from sample selection bias. To correct this sample selection bias, we implement the Heckman copula model with the Joe copula [Smith, 2003, Hasebe, 2013], which consists of two equations: a selection equation and an outcome equation. The selection equation estimates the probability of households being selected into homeowners with mortgages by using the full sample including renters, and the control variables include household social and economic variables such as gender, age, marital status, number of children, race, education, household income, and household wealth, which are found to be critical in explaining homeownership [Arimah, 1997, Gathergood and Weber, 2017a, Gyourko et al., 1999]. In the outcome equation, we correct for the potential sample selection bias. Table 3 reports the maximum likelihood estimation of the Heckman copula model. As shown in Table 3, our main results are robust to the potential sample selection bias: the marginal effect of financial literacy is -0.0197 (compared to -0.0182 in column 3 of Table 2). Table 3 also reports a coefficient Kendall's τ, which ranges between -1 and 1; A value of Kendall's τ closer to 1 (-1) means a stronger (negative) dependence between the error terms in the selection and the outcome equations, and therefore stronger evidence of sample selection. Table 3 shows that the estimated Kendall's τ is -0.6995, which implies a strong and negative dependence. Furthermore, the LR test of independent equations shows evidence of rejection

to the null of the independent errors.

Table 3: Controlling for selection bias
(Heckman Copula model)

	(1)
Financial industry	-0.0063
	(0.0045)
Financial literacy	-0.0197**
	(0.0091)
Kendall's τ	-0.6995
LR test of independent equations	
χ^2	1789.6
p value	0.0000
Other controls	
Household characteristics	Yes
Loan characteristics	Yes
State fixed effects	Yes
Year fixed effects	Yes
Observations	8,255

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the full sample. Results are estimated with the Heckman copula model that controls for selection bias. The coefficient of Kendall's measures the dependence between the error terms in the selection and the outcome equations. Our main results are robust to the use of other copula functions. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has high levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

4.2 Functional misspecification

Our previous estimations assume linear impacts of control variables on mortgage stress. If this assumption is invalid, our previous estimations may be biased due to functional misspecification. To deal with this potential issue, we apply the commonly-used propensity

score matching (PSM) method, the identification of which compares treatment and control units with similar values on the propensity score (i.e., the conditional probability of being treated given a set of control variables).

The PSM method does not depend on the assumption of linear impacts of control variables on mortgage stress. There are three steps to examine the impact of working in the financial industry or financial literacy by using the PSM approach. First, we estimate a propensity score (i.e., the probability of being treated) for all observations based on household and loan characteristics obtained from logit regression. Second, we match each observation in the treatment group (financial professionals or those with high levels of financial literacy) to the observations in the control group (nonfinancial professionals or those with low levels of financial literacy) by using various matching algorithms. There are many matching algorithms that differ in how the matched households from the control group are selected. In this paper, we use three matching methods that are commonly used in the literature: radius matching, kernel matching, and one-to-four nearest-neighbor matching. [8] Third, we re-estimate the baseline regression by using the sample of matched pairs. The estimation results of the PSM method are reported in Table 4. Consistent with the previous results, we find that borrowers with high levels of financial literacy are 1.52 to 2.83 percentage points less likely to suffer from mortgage stress than those with low levels of financial literacy, and the differences remain statistically significant at the 1% level. However, financial professionals do not have a significantly lower probability of having mortgage stress than non-financial professionals, holding other things equal. In other words, our previous results are robust to functional misspecification.

As suggested by Rosenbaum and Rubin [1983], the matching procedure needs to balance the distribution of the control variables across treatment and control groups in the matched sample. We now use two methods to check the matching quality of our matched sample.

[8] Typically, we should choose a small number of matches. Simulations in Abadie et al. [2004] suggest that using one-to-four matches performed well in terms of mean-squared error. Our results are also robust by using one-to-one, one-to-two, and one-to-three matches.

Table 4: Controlling for model misspecification
(Propensity score matching (PSM) method)

	Neighborhood matching	Radius matching	Kernel matching
Financial industry	-0.0092	-0.0132	-0.0114
	(0.0089)	(0.0313)	(0.0074)
Financial literacy	-0.0176***	-0.0283***	-0.0152***
	(0.0060)	(0.0049)	(0.0056)
Other controls			
Household characteristics	Yes	Yes	Yes
Loan characteristics	Yes	Yes	Yes

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the matched sample with different matching algorithms. Results are estimated with the propensity score matching method that controls for model misspecification. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

First, the PSM estimator is only defined in the region of common support [Dehejia and Wahba, 2002, Heckman et al., 1997]. The common support condition states that any combination of observations in the treatment group can also be observed in the control group. Matching an incomparable control group to the treatment group can cause estimation bias. Results of the common support condition test are shown in Table 5. We can see that the majority of observations in the treatment and control groups fall into the region of common support, implying that our PSM estimation is reliable.

Second, we summarize household and loan characteristics (control variables) for the treatment and control groups before and after matching (see Table 6). It is apparent that the two groups are more comparable along almost all control variables. For example, before matching, the difference in white between the treated and control groups is 0.0390 (Panel A: treatment group includes financial professions) and 0.2195 (Panel B: treatment group in-

cludes households with high levels of financial literacy), and the difference shrinks to 0.0031 and 0.0016 after matching. This finding suggests that after matching, the control variables are balanced across treatment and control groups, indicating a good matching quality of the matched sample.

Table 5: Common support condition test

	Off support	On support
Untreated: Other industries	467	4307
Treated: Financial industry	0	325
Untreated: Low levels of financial literacy	81	2635
Treated: High levels of financial literacy	12	2353

The result from the mean difference test in Table 6 is straightforward, but one potential concern is that the mean differences of control variables may be susceptible to variations in the unit of measurement. To deal with this issue, we calculate the standardized bias of each control variable before and after matching:

$$Standardized\,bias_i = \frac{\left|\bar{X}_{i,treat} - \bar{X}_{i,control}\right|}{\sqrt{(S^2_{i,treat} + S^2_{i,control})/2}} \tag{2}$$

where $Standardized\,bias_i$ refers to the standardized bias of the covariate X_i. $\bar{X}_{i,treat}$ refers to the mean value of X_i in the treated group. $\bar{X}_{i,control}$ refers to the mean value of X_i in the control group. $\left|\bar{X}_{i,treat} - \bar{X}_{i,control}\right|$ refers to the absolute value of the difference between $\bar{X}_{i,treat}$ and $\bar{X}_{i,control}$. $S^2_{i,treat}$ refers to the variance of X_i in the treated group. $S^2_{i,control}$ refers to the variance of X_i in the control group. Obviously, the standardized bias of each control variable is independent of its unit of measurement.

Table 6: Mean difference test: Balancing quality of propensity score matching

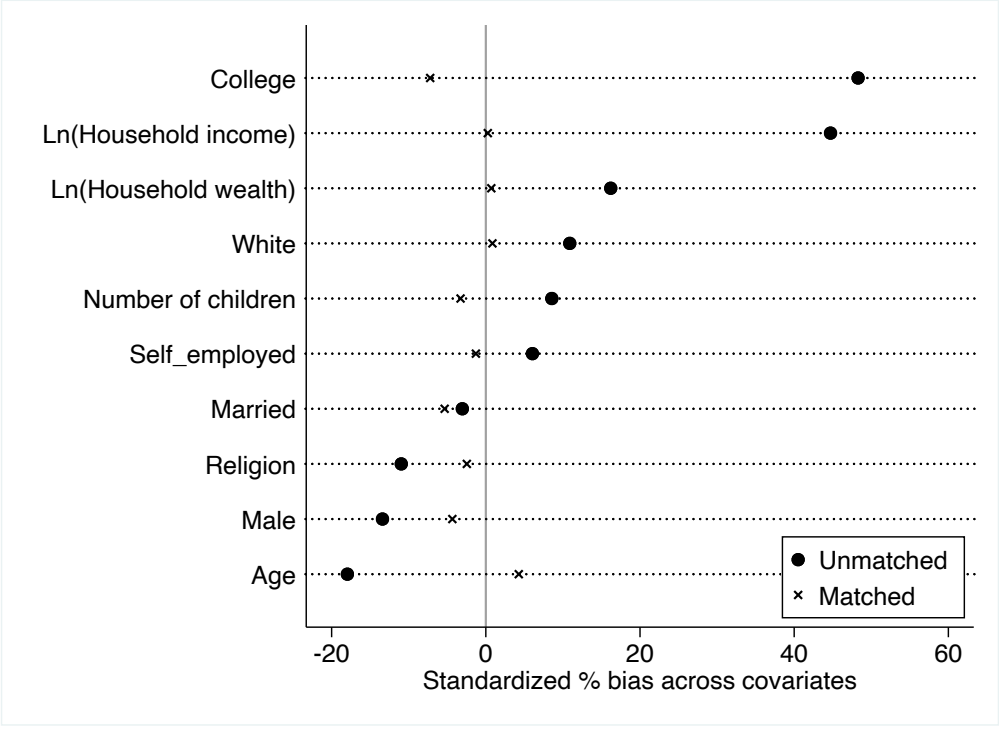
Panel A: Financial industry	Group	Mean		Differences	p-value
		Treated	Control		
Male	Unmatched	0.8154	0.8645	0.0491	0.0130
	Matched	0.8154	0.8313	0.0159	0.5960
Age	Unmatched	46.000	48.171	2.1710	0.0030
	Matched	46.000	45.483	0.5170	0.5720
Married	Unmatched	0.7539	0.7669	0.0130	0.5920
	Matched	0.7539	0.7767	0.0228	0.4930
Children	Unmatched	0.9292	0.8289	0.1003	0.1230
	Matched	0.9292	0.9677	0.0385	0.6810
White	Unmatched	0.8677	0.8287	0.0390	0.0690
	Matched	0.8677	0.8646	0.0031	0.9080
Religion	Unmatched	0.0923	0.1265	0.0342	0.0710
	Matched	0.0923	0.1000	0.0077	0.7400
College	Unmatched	0.9108	0.7306	0.1802	0.0000
	Matched	0.9108	0.9377	0.0269	0.1950
Self-employed	Unmatched	0.1231	0.1039	0.0192	0.2750
	Matched	0.1231	0.1272	0.0041	0.8750
Ln(Household income)	Unmatched	11.734	11.433	0.3010	0.0000
	Matched	11.734	11.732	0.0020	0.9750
Ln(Household wealth)	Unmatched	9.9493	9.2357	0.7136	0.0060
	Matched	9.9493	9.9180	0.0313	0.9270
Panel B: financial literacy	Group	Mean		Differences	p-value
		Treated	Control		
Male	Unmatched	0.8989	0.8288	0.0701	0.0000
	Matched	0.8984	0.8900	0.0084	0.3490
Age	Unmatched	48.038	48.027	0.0110	0.9740
	Matched	47.984	47.873	0.1110	0.7630
Married	Unmatched	0.8106	0.7275	0.0831	0.0000
	Matched	0.8096	0.8022	0.0074	0.5210
Children	Unmatched	0.8427	0.8288	0.0139	0.6630
	Matched	0.8449	0.8261	0.0188	0.5700
White	Unmatched	0.9488	0.7293	0.2195	0.0000
	Matched	0.9486	0.9502	0.0016	0.8010
Religion	Unmatched	0.1285	0.1207	0.0078	0.3980
	Matched	0.1292	0.1392	0.0100	0.3150
College	Unmatched	0.8816	0.6214	0.2602	0.0000
	Matched	0.8810	0.8764	0.0046	0.6300
Self-employed	Unmatched	0.1061	0.1042	0.0019	0.8270
	Matched	0.1058	0.1043	0.0015	0.8660
Ln(Household income)	Unmatched	11.624	11.304	0.3200	0.0000
	Matched	11.616	11.610	0.0060	0.7580
Ln(Household wealth)	Unmatched	9.8428	8.7953	1.0475	0.0000
	Matched	9.8173	9.6627	0.1546	0.2360

The standardized bias of the control variables before and after matching is shown in Figure 1. Again, the standardized bias of most control variables decreases significantly after matching, indicating that our PSM method has mitigated the significant distribution differences that existed before matching.

To further show that financial literacy incrementally matters over working in the financial industry, we conduct two additional analyses. [9] First, we regress mortgage stress on financial literacy and other controls based on the matched sample obtained by using financial industry, loan and borrower characteristics as the matching variables with the one-to-one matching method. The results, as reported in column (1) of Table 7, confirm the positive role of financial literacy in reducing mortgage stress. Second, we repeat the PSM exercise within the subsamples of "Financial industry = 1" and "Financial industry = 0", and the results are reported in columns (2) and (3) of Table 7, respectively. Again, we find that the coefficient of financial literacy remains negative and significant. These findings reaffirm our previous results.

[9] We thank an anonymous referee to suggest these two analyses.

(a) Financial industry



(b) Financial literacy

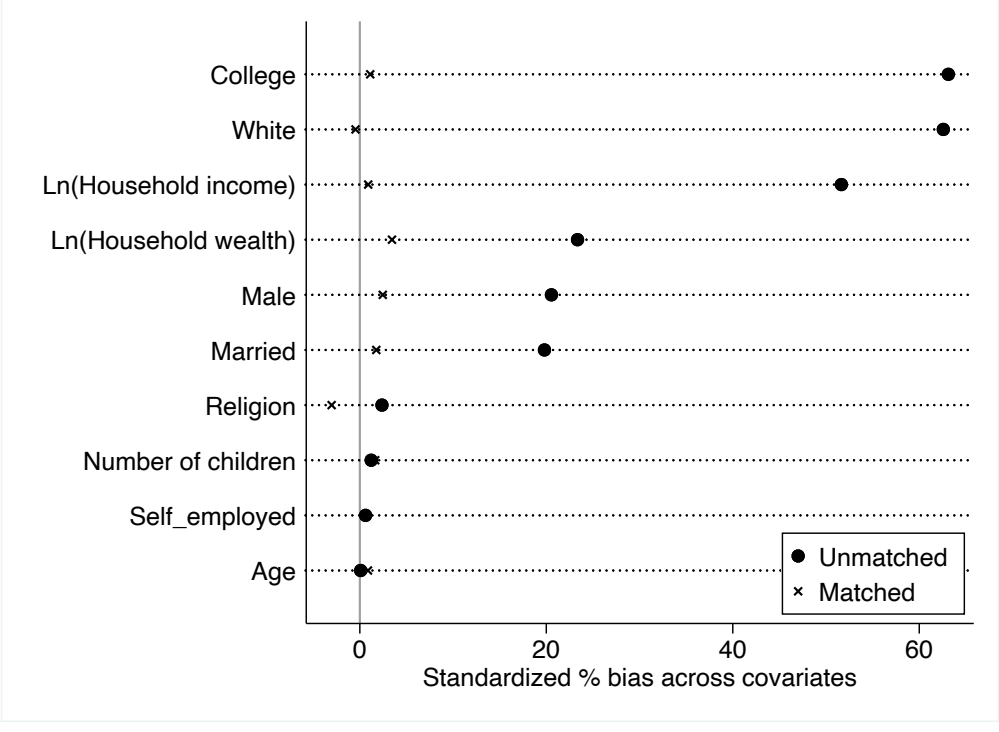


Figure 1: Unconfoundedness condition test: standardized bias

Table 7: Controlling for model misspecification

(Propensity score matching (PSM) method)

	(1) Full sample		(2) Financial industry = 1		(3) Financial industry = 0	
	Coef.	Marginal effect	Coef.	Marginal effect	Coef.	Marginal effect
Financial literacy	-0.6636**	-0.0171	-2.6966*	-0.0325	-0.8413***	-0.0154
	(0.2974)		(1.4853)		(0.2620)	
Financial industry	-1.3947	-0.0358				
	(0.8659)					
Other controls						
Household characteristics	Yes		Yes		Yes	
Loan characteristics	Yes		Yes		Yes	

Note: This table presents the regression results of mortgage stress on financial industry (only in column (1)), financial literacy, loan and borrower characteristics using the matched sample with the one-to-one matching algorithm. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

4.3 Falsification test

Following Agarwal et al. [2017], we now conduct a falsification test. The falsification test artificially assigns financial literacy to the households and then estimates its effect on mortgage stress, which is similar to a placebo test. A placebo test measures the treatment effect of false treatment. For example, to test the therapeutic effect of a drug, patients can be divided into two groups: one group takes the drug, as the treatment group, and the other group takes a placebo (such as a sugar pill), as the control group. During the experiment, patients do not know whether they are taking the drug or a placebo. Since 2,365 households in our data are reported to have high levels of financial literacy, we randomly and artificially choose 2,365 households in the sample and "treat" them to have high levels of

financial literacy (i.e., randomly assigned pseudo treatment group). We then rerun the baseline regression model by using the randomly assigned pseudo treatment group to replace the original households with high levels of financial literacy. To obtain consistent results, we repeat this random assignment and regression estimation thirty times. We report our results in Table 8, and three points are worth noting. First, there are only two random assignments, i.e., assignments (12) and (20), with significant coefficients of financial literacy for the pseudo treatment group (i.e., the households who are randomly assigned to have high levels of financial literacy) as opposed to the pseudo control group (i.e., the households who are randomly assigned to have low levels of financial literacy). Second, with respect to the two random assignments having significant coefficients of financial literacy, the signs of the coefficients are opposite and unexpectedly positive. Third, it is interesting to see that the coefficients of financial industry are negative and significant in almost all cases. These results in fact provide supportive evidence to the prior analysis that working in the financial industry would have an effect on mortgage stress if financial literacy is not controlled for.

Table 8: Falsification test

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Financial industry	-0.8865*	-0.8937*	-0.8956*	-0.8760*	-0.8894*	-0.8724	-0.9037*	-0.8942*	-0.8869*	-0.8851*
	(0.5328)	(0.5332)	(0.5381)	(0.5295)	(0.5356)	(0.5337)	(0.5396)	(0.5379)	(0.5387)	(0.5373)
Financial literacy	-0.1846	0.0189	-0.1224	0.1881	-0.2622	-0.1300	-0.0894	0.0121	0.1628	0.0480
	(0.1853)	(0.1863)	(0.1842)	(0.1856)	(0.1824)	(0.1862)	(0.1861)	(0.1855)	(0.1849)	(0.1850)
Other controls										
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542
	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Financial industry	-0.8894*	-0.9125*	-0.8892*	-0.9010*	-0.9075*	-0.8936*	-0.8920*	-0.9074*	-0.8887*	-0.9193*
	(0.5306)	(0.5383)	(0.5351)	(0.5342)	(0.5342)	(0.5374)	(0.5363)	(0.5442)	(0.5367)	(0.5388)
Financial literacy	0.1439	0.3814**	0.0745	0.0842	0.2383	-0.0285	-0.0001	0.2691	-0.1015	0.3165*
	(0.1838)	(0.1824)	(0.1834)	(0.1813)	(0.1871)	(0.1825)	(0.1785)	(0.1828)	(0.1875)	(0.1855)
Other controls										
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542

Table 8: Continued

	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
Financial industry	-0.9394*	-0.8679	-0.8944*	-0.8950*	-0.8892*	-0.8777	-0.8941*	-0.9038*	-0.8880*	-0.8923*
	(0.5345)	(0.5375)	(0.5385)	(0.5388)	(0.5351)	(0.5348)	(0.5363)	(0.5412)	(0.5355)	(0.5403)
Financial literacy	0.2571	0.3025	0.0546	-0.0374	-0.0215	-0.1824	-0.0768	-0.2880	0.0814	0.1059
	(0.1797)	(0.1919)	(0.1816)	(0.1836)	(0.1848)	(0.1869)	(0.1862)	(0.1832)	(0.1798)	(0.1834)
Other controls										
Household characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Loan characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542	4,542

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the full sample. 2,365 borrowers have higher levels of financial literacy in the full sample and we randomly assign 2,365 borrowers to have higher levels of financial literacy, which consist of the pseudo treatment group. We run the baseline regression model with the random assignments thirty times. Results are estimated with logistic regressions. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% levels, respectively.

Our results are also robust to two other issues. [10] First, in order to mitigate the bias from household unobservables, we take advantage of our panel data by using panel regression (i.e., the random effect model) to re-estimate our baseline results, which are reported in Table 9. Consistent with our earlier results, the effect of financial literacy is 1.66 percentage points less likely to have mortgage stress (compared to 1.82 percentage points in column (3) of Table 2), and it is significant at the 5% level. In addition, the impact of financial occupation is still not significant. Therefore, the results after controlling for unobserved heterogeneity of households remain consistent with our previous estimations.

Table 9: Impact of financial profession and financial literacy on mortgage stress

(The random effect model)

	Coef.	Marginal effect
Financial industry	-1.5927	-0.0278
	(1.1293)	
Financial literacy	-0.9510**	-0.0166
	(0.3815)	
Year of 2011	0.3931	0.0069
	(0.3042)	
Year of 2013	0.9084***	0.0159
	(0.3169)	
Year of 2015	0.1016	0.0018
	(0.3890)	
Other controls		
Household characteristics		Yes
Mortgage characteristics		Yes
Observations		4,542

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the full sample. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

[10] We would like to thank the two anonymous reviewers for these comments.

Second, we redefine an indicator variable of financial literacy by using the last question only (i.e., related to interest rates and compounding) as a proxy of financial literacy. The results are reported in Table 10. The estimated coefficient of *Financial literacy* is negative and statistically significant at the 1% level, and the estimated coefficient of *Financial industry* is still insignificant. These findings are in line with the baseline results reported in column (3) of Table 2.

Table 10: Alternative definition of financial literacy

	Coef.	Marginal effect
Financial industry	-0.7877	-0.0221
	(0.4993)	
Financial literacy	-0.7565***	-0.0212
	(0.2324)	
Other controls		
Household characteristics		Yes
Mortgage characteristics		Yes
State fixed effects		Yes
Year fixed effects		Yes
Observations		4,542
R-squared		0.2522

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using the full sample. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower correctly answers the last question related to interest rates and compounding and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

5 Further analysis

We have found the effect of financial literacy on mortgage stress. In this section, we examine the heterogeneity of the effect of financial literacy on mortgage stress across borrowers of

different ages. In addition, we also examine the role of quantitative reasoning in mortgage stress. Gerardi et al. [2013] present empirical evidence showing that borrowers' ability to perform basic mathematical calculations is negatively associated with the propensity to default on their mortgages.

We first examine the heterogeneity of the effect of financial literacy on mortgage stress between young and old households. We group borrowers into two categories according to their ages: old households with age 45 and above, and the rest of the young households. We focus on the age difference since financial literacy has differential impacts on homeownership, the choice of mortgage types, and leverage position between young and old households [Gathergood and Weber, 2017b]. The estimated results by using the sub-samples of young and old households are shown in Panel A of Table 11. Again, the impact of working in the financial industry on mortgage stress is not significant for both young and old households, providing supportive evidence for our previous results. Consistent with our findings, results in Table 11 also show a negative relationship between financial literacy and mortgage stress, while the effect of financial literacy on mortgage stress is stronger (statistically and economically more significant) for old households than that for young households. The results by using alternative cut-off age of 40 to define young and old households, as shown in Panel B of Table 11, remain consistent.

Next, we examine the relationship between quantitative reasoning and mortgage stress. The PSID-WB asks respondents some numerical questions to fill in to test their quantitative reasoning. [11] We define an indicator variable of quantitative reasoning that equals 1 if a household correctly answers all the questions and 0 otherwise. 38.8% of respondents correctly answer all the questions, and they are considered to have high levels of quantitative reasoning. The results by controlling for quantitative reasoning, household and loan characteristics, and state and year fixed effects are reported in column (1) of Table 12. The estimated coefficient

[11] In a section of PSID-WB (In Section J: Quantitative Reasoning), several questions drawn from the 2012 Health and Retirement Study are designed to measure mathematical (or "quantitative") reasoning of respondents.

Table 11: Heterogeneity effects between young and old household groups

Panel A: Cut-off age of 45	(1) Young household		(2) Old household	
	Coef.	Marginal effect	Coef.	Marginal effect
Financial industry	-0.9956	-0.0298	-0.6485	-0.0205
	(1.1669)		(0.5375)	
Financial literacy	-0.6570*	-0.0196	-0.7855**	-0.0249
	(0.3532)		(0.3425)	
Other controls				
Household characteristics	Yes		Yes	
Loan characteristics	Yes		Yes	
State fixed effects	Yes		Yes	
Year fixed effects	Yes		Yes	
Observations	1,842		2,089	
R-squared	0.2711		0.2970	
Panel B: Cut-off age of 40	(1) Young household		(2) Old household	
	Coef.	Marginal effect	Coef.	Marginal effect
Financial industry	-1.8655	-0.0549	-0.7049	-0.0225
	(2.3359)		(0.6263)	
Financial literacy	-0.3669	-0.0108	-0.9135***	-0.0291
	(0.4077)		(0.2956)	
Other controls				
Household characteristics	Yes		Yes	
Loan characteristics	Yes		Yes	
State fixed effects	Yes		Yes	
Year fixed effects	Yes		Yes	
Observations	1,217		2,667	
R-squared	0.3162		0.2749	

Note: This table presents the regression results of mortgage stress on financial industry, financial literacy, loan and borrower characteristics using two subsamples: young household (age 44 and below) and old stage (age 45 and above). Results are estimated with logistic regressions. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

of quantitative reasoning is negative and statistically significant at the 5% level, indicating that good mathematical reasoning abilities are negatively associated with the likelihood of having mortgage stress. This result is in fact in line with Gerardi et al. [2013]. However, after controlling for financial industry and financial literacy, the magnitude of the effect substantially decreases and becomes statistically insignificant, as shown in column (2) of Table 12. Meanwhile, the effect of financial literacy on mortgage stress is negative and statistically significant at the 5% level. These findings indicate that the negative relation between quantitative reasoning and mortgage stress observed in column (1) of Table 12 is largely due to a lack of control for financial literacy in the specification.

We next examine the cross effect of financial literacy and quantitative reasoning on mortgage stress. The regression results by replacing the indicator variable of financial literacy with cross-effects of low/high financial literacy and low/high quantitative reasoning are reported in column (3) of Table 12. The findings in Table 12 suggest that quantitative reasoning skills can further reduce the probability of having mortgage stress for borrowers with high levels of financial literacy, while quantitative reasoning ability alone has no significant effect on mortgage stress.

As a robustness check, we use alternative definitions of mortgage stress, financial literacy, and quantitative reasoning. In Table 13, columns (1) and (2) report the results with continuous variables of mortgage stress (i.e., the number of months behind on mortgage payments), financial literacy (i.e., the correct number of questions designed to measure the financial literacy of respondents), and quantitative reasoning (i.e., the correct number of questions designed to measure the mathematical reasoning of respondents). In order to see how the effect of financial literacy on mortgage stress varies over age, we add an interaction term between financial literacy and age (continuous), and columns (3) of Table 13 report the results. The findings indicate that the effect increases with the borrower's age. Overall, Table 13 reaffirms our previous findings.

Table 12: The cross effect of financial literacy and quantitative reasoning

	(1)		(2)		(3)	
	Coef.	Marginal effect	Coef.	Marginal effect	Coef.	Marginal effect
Quantitative reasoning	-0.4806**	-0.0136	-0.2966	-0.0083		
	(0.2192)		(0.2201)			
Financial industry			-0.7434	-0.0209	-0.7455	-0.0209
			(0.5031)		(0.5110)	
Financial literacy			-0.5645**	-0.0159		
			(0.2320)			
Cross effect of financial literacy and quantitative reasoning						
(Ref.: Low levels of financial literacy & low levels of quantitative reasoning)						
High levels of financial literacy & high levels of quantitative reasoning					-0.7981***	-0.0224
					(0.3024)	
High levels of financial literacy & low levels of quantitative reasoning					-0.6691**	-0.0188
					(0.3041)	
Low levels of financial literacy & high levels of quantitative reasoning					-0.3904	-0.0110
					(0.2833)	
Other controls						
Household characteristics		Yes		Yes		Yes
Loan characteristics		Yes		Yes		Yes
State fixed effects		Yes		Yes		Yes
Year fixed effects		Yes		Yes		Yes
Observations		4,542		4,542		4,542
R-squared		0.2441		0.2505		0.2507

Note: This table presents the regression results of mortgage stress on quantitative reasoning, financial industry, financial literacy, interaction dummies between financial industry and financial literacy, loan and borrower characteristics using the full sample. Results are estimated with logistic regressions. The dependent variable is an indicator variable of mortgage stress that equals 1 if the borrower has had difficulties recently making mortgage or loan payments and 0 otherwise. Quantitative reasoning is an indicator variable that equals 1 if the borrower has higher levels of quantitative reasoning and 0 otherwise. Financial industry is an indicator variable that equals 1 if the borrower works in a financial profession and 0 otherwise. Financial literacy is an indicator variable that equals 1 if the borrower has higher levels of financial literacy and 0 otherwise. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. ∗, ∗∗, and ∗∗∗ indicate significance at the 10%, 5%, and 1% levels, respectively.

Table 13: Impact of financial profession and financial literacy on mortgage stress

(Alternative measurement of variables)

	(1)	(2)	(3)
Financial industry	-6.1842	-4.2691	-5.8594
	(4.1113)	(4.6588)	(4.0915)
Financial literacy	-1.5452**	-2.0417*	-0.7243
	(0.7834)	(1.1604)	(0.9223)
Quantitative reasoning		-1.2853	
		(1.2378)	
Financial literacy × Age			-0.0162***
			(0.0046)
Other controls			
Household characteristics	Yes	Yes	Yes
Mortgage characteristics	Yes	Yes	Yes
State fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	4,542	4,542	4,542
R-squared	0.2592	0.2609	0.2677

Note: Based on the baseline equation, column (1) reports the results by changing the dependent variable with a continuous variable of mortgage stress (i.e., the number of months behind on mortgage payments) and the independent variable with a continuous variable of financial literacy (i.e., the correct number of questions designed to measure the financial literacy of respondents). Based on the specification in column (1), column (2) further adds a continuous variable of quantitative reasoning (i.e., the correct number of questions designed to measure the mathematical reasoning of respondents) and its interaction term with the continuous variable of financial literacy. Column (3) reports the results from the baseline equation by adding the interaction term between financial literacy and age. Household characteristics include gender, age, marital status, number of children, race, religion, education, self-employment, household income, and household wealth. Loan characteristics include the amount of interest rate, interest rate type (adjustable-rate mortgage or fixed-rate mortgage), LTV, and DTI. Robust standard errors are reported in parentheses. $*$, $**$, and $***$ indicate significance at the 10%, 5%, and 1% levels, respectively.

6 Conclusion

This paper examines the effect of financial literacy on mortgage stress. Using data from Panel Study of Income Dynamics (PSID), we find that borrowers with high levels of financial literacy are 60.3 percent less likely to suffer from mortgage stress than borrowers with low levels of financial literacy, after controlling for household characteristics, loan features,

working in the financial industry, and geographic and time fixed effects. Our estimated results are robust to potential sample selection bias and functional misspecification. In addition, we also find that the effect of financial literacy on mortgage stress is stronger for old households than for young households. Further analysis reveals strong cross effects of financial literacy and quantitative reasoning on mortgage stress. Our findings suggest that quantitative reasoning skills can further reduce the probability of having mortgage stress for borrowers with high levels of financial literacy, while quantitative reasoning ability alone has no significant effect on mortgage stress.

This paper makes two contributions to the current literature. First, the paper contributes to the body of literature on the determinants of mortgage stress. The massive expansion in the supply of mortgage debts since the late 1990s leads to a rapidly growing number of risky and subprime borrowers in the 2000s [Favara and Imbs, 2015, Justiniano et al., 2019]. A sharp rise in mortgage stress and delinquency rates was followed by the fall in mortgage supply [Mayer et al., 2009, Mejia, 1999, Mocetti and Viviano, 2017]. Different from previous studies examining the reasons for mortgage delinquency from the supply side, we focus on the effect of borrowers' financial literacy on mortgage stress from the demand side. Our results highlight the important role of financial literacy in household's mortgage stress, which are helpful for mortgage pricing and prediction of mortgage delinquency. Second, the paper also enriches the literature on the consequences of financial literacy by linking it to the decisions involved in the mortgage process, as previous studies most focus on its impacts on individual's investment decisions in the financial market [Hsiao and Tsai, 2018, Jappelli and Padula, 2013, Klapper et al., 2013, Van Rooij et al., 2011, 2012].

Mortgage defaults and foreclosures are detrimental to household financial health, physical health, and mental health, and they are also very costly to financial institutions. Our findings raise an important policy question about the best way to reduce mortgage defaults and foreclosures. Previously, it has been argued that risky mortgage products such as subprime mortgages are an important culprit in causing mortgage defaults, and financial literacy

can help reduce defaults in subprime mortgages [Agarwal et al., 2017, Gerardi et al., 2013]. We find that borrowers with high levels of financial literacy are 60.3 percent less likely to experience mortgage stress in the general mortgage market. This finding suggests the potential for significant benefits from increased financial education for all mortgage borrowers. A previous study by Collins and Schmeiser [2013] found that foreclosure counseling could reduce incidences of foreclosure. The strong quantitative results we find in the effect of financial literacy on mortgage stress suggest that implementing policies to enhance financial literacy can effectively reduce mortgage stress in the broader market. This reduction in mortgage stress is particularly crucial, given its role as a precursor to mortgage defaults and foreclosures.

References

Alberto Abadie, David Drukker, Jane Leber Herr, and Guido W. Imbens. Implementing Matching Estimators for Average Treatment Effects in Stata. *Stata Journal*, 4(3):290–311, 2004. ISSN 1536867X. doi: TheStataJournal. URL http://www.stata-journal.com/article.html?article=st0072%5Cnhttp://www.stata-journal.com/sjpdf.html?article=st0072.

Arthur Acolin, Jesse Bricker, Paul Calem, and Susan Wachter. Borrowing Constraints and Homeownership. *American Economic Review*, 106(5):625–629, may 2016. ISSN 0002-8282. doi: 10.1257/aer.p20161084. URL https://pubs.aeaweb.org/doi/10.1257/aer.p20161084.

Sumit Agarwal, Souphala Chomsisengphet, and Yunqi Zhang. How Does Working in a Finance Profession Affect Mortgage Delinquency? *Journal of Banking and Finance*, 78:1–13, may 2017. ISSN 03784266. doi: 10.1016/j.jbankfin.2017.01.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0378426617300298.

Ben Arimah. The Determinants of Housing Tenure Choice in Ibadan, Nigeria. *Urban Studies*, 34(1):105–124, 1997. ISSN 0042-0980. doi: 10.1080/0042098976294. URL http://journals.sagepub.com/doi/10.1080/0042098976294.

Emanuele Bajo and Massimiliano Barbi. Financial Illiteracy and Mortgage Refinancing Decisions. *Journal of Banking and Finance*, 94:279–296, 2018. ISSN 03784266. doi: 10.1016/j.jbankfin.2018.08.001. URL https://doi.org/10.1016/j.jbankfin.2018.08.001.

James Banks, Cormac O 'Dea, and Zoë Oldfield. Cognitive Function, Numeracy and Retirement Saving Trajectories. *The Economic Journal*, 120(548):F381–F410, nov 2010. ISSN 0013-0133. doi: 10.1111/j.1468-0297.2010.02395.x. URL https://academic.oup.com/ej/article/120/548/F381-F410/5089502.

Daniel J. Benjamin, James J. Choi, and Geoffrey Fisher. Religious Identity and Economic Behavior. *Review of Economics and Statistics*, 98(4):617–637, oct 2016. ISSN 0034-6535. doi: 10.1162/REST_a_00586. URL https://direct.mit.edu/rest/article/98/4/617-637/58335.

Xun Bian, Zhenguo Lin, and Yingchun Liu. House Price, Loan-to-Value Ratio and Credit Risk. *Journal of Banking and Finance*, 92:1–12, 2018. ISSN 03784266. doi: 10.1016/j.jbankfin.2018.04.006. URL https://doi.org/10.1016/j.jbankfin.2018.04.006.

Yan Chen, Ming Jiang, and Onur Kesten. An Empirical Evaluation of Chinese College Admissions Reforms through a Natural Experiment. *Proceedings of the National Academy of Sciences*, 117(50):31696–31705, dec 2020. ISSN 0027-8424. doi: 10.1073/pnas.2009282117. URL http://www.pnas.org/lookup/doi/10.1073/pnas.2009282117.

J. Michael Collins and Maximilian D. Schmeiser. The Effects of Foreclosure Counseling for Distressed Homeowners. *Journal of Policy Analysis and Management*, 32(1):83–106, jan 2013. ISSN 0276-8739. doi: 10.1002/pam.21670. URL https://onlinelibrary.wiley.com/doi/10.1002/pam.21670.

James Conklin, Moussa Diop, and Mingming Qiu. Religion and Mortgage Misrepresentation. *Journal of Business Ethics. Forthcoming*, may 2022. ISSN 0167-4544. doi: 10.1007/s10551-021-04831-2. URL https://link.springer.com/10.1007/s10551-021-04831-2.

Joshua Coval, Jakub Jurek, and Erik Stafford. The Economics of Structured Finance. *Journal of Economic Perspectives*, 23(1):3–25, jan 2009. ISSN 0895-3309. doi: 10.1257/jep.23.1.3. URL https://pubs.aeaweb.org/doi/10.1257/jep.23.1.3.

Rajeev Dehejia and Sadek Wahba. Propensity Score-Matching Methods for Nonexperimental Causal Studies. *Review of Economics and Statistics*, 84(1):151–161, 2002. ISSN 0034-6535. doi: 10.1162/003465302317331982. URL http://www.mitpressjournals.org/doi/10.1162/003465302317331982.

Marco Di Maggio and Amir Kermani. Credit-Induced Boom and Bust. *The Review of Financial Studies*, 30(11):3711–3758, nov 2017. ISSN 0893-9454. doi: 10.1093/rfs/hhx056. URL https://academic.oup.com/rfs/article/30/11/3711/3903098.

John V. Duca and Anil Kumar. Financial Literacy and Mortgage Equity Withdrawals. *Journal of Urban Economics*, 80:62–75, mar 2014. ISSN 00941190. doi: 10.1016/j.jue.2013.08.003. URL https://linkinghub.elsevier.com/retrieve/pii/S0094119013000600.

Giovanni Favara and Jean Imbs. Credit Supply and the Price of Housing. *American Economic Review*, 105(3):958–992, mar 2015. ISSN 0002-8282. doi: 10.1257/aer.20121416. URL https://pubs.aeaweb.org/doi/10.1257/aer.20121416.

Christopher L. Foote, Kristopher Gerardi, Lorenz Goette, and Paul S. Willen. Just the Facts: An Initial Analysis of Subprime's Role in the Housing Crisis. *Journal of Housing Economics*, 17(4):291–305, 2008. ISSN 10511377. doi: 10.1016/j.jhe.2008.09.005. URL http://dx.doi.org/10.1016/j.jhe.2008.09.005.

John Gathergood and Jörg Weber. Financial Literacy, Present Bias and Alternative Mortgage Products. *Journal of Banking and Finance*, 78:58–83, 2017a. ISSN 03784266. doi: 10.1016/j.jbankfin.2017.01.022. URL http://dx.doi.org/10.1016/j.jbankfin.2017.01.022.

John Gathergood and Jörg Weber. Financial literacy: A Barrier to Home Ownership for the Young? *Journal of Urban Economics*, 99:62–78, 2017b. ISSN 00941190. doi: 10.1016/j.jue.2017.02.001. URL http://dx.doi.org/10.1016/j.jue.2017.02.001.

Nicola Gennaioli, Andrei Shleifer, and Robert Vishny. Neglected Risks, Financial Innovation, and Financial Fragility. *Journal of Financial Economics*, 104(3):452–468, 2012. ISSN 0304405X. doi: 10.1016/j.jfineco.2011.05.005. URL http://dx.doi.org/10.1016/j.jfineco.2011.05.005.

Kristopher Gerardi, Lorenz Goette, and Stephan Meier. Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data - Federal Reserve Bank of Atlanta. 2010. URL https://www.frbatlanta.org/research/publications/wp/2010/10.aspx.

Kristopher Gerardi, Lorenz Goette, and Stephan Meier. Numerical Ability Predicts Mortgage Default. *Proceedings of the National Academy of Sciences*, 110(28):11267–11271, jul 2013. ISSN 0027-8424. doi: 10.1073/pnas.1220568110. URL http://www.pnas.org/cgi/doi/10.1073/pnas.1220568110.

Ulrike Grömping. Estimators of Relative Importance in Linear Regression Based on Variance Decomposition. *The American Statistician*, 61(2):139–147, may 2007. ISSN 0003-1305. doi: 10.1198/000313007X188252. URL http://www.tandfonline.com/doi/abs/10.1198/000313007X188252.

Joseph Gyourko, Peter Linneman, and Susan Wachter. Analyzing the Relationships among Race, Wealth, and Home Ownership in America. *Journal of Housing Economics*, 8(2): 63–89, jun 1999. ISSN 10511377. doi: 10.1006/jhec.1999.0242. URL https://linkinghub.elsevier.com/retrieve/pii/S1051137799902425.

William G. Hardin, Mingzhi Hu, and Zhenguo Lin. Culture, Wealth, Gender, Real Estate, and Consumption in China: It's a Boy. *Journal of Real Estate Research*, 45(1):55–82, jan 2023. ISSN 0896-5803. doi: 10.1080/08965803.2022.2045053. URL https://www.tandfonline.com/doi/full/10.1080/08965803.2022.2045053.

Takuya Hasebe. Copula-Based Maximum-Likelihood Estimation of Sample-Selection Models. *The Stata Journal: Promoting communications on statistics and Stata*, 13(3):547–573, sep 2013. ISSN 1536-867X. doi: 10.1177/1536867X1301300307. URL http://journals.sagepub.com/doi/10.1177/1536867X1301300307.

John Heaton and Deborah Lucas. Portfolio Choice and Asset Prices: The Importance of

Entrepreneurial Risk. *The Journal of Finance*, 55(3):1163–1198, jun 2000. ISSN 00221082. doi: 10.1111/0022-1082.00244. URL http://doi.wiley.com/10.1111/0022-1082.00244.

James Heckman, Hidehiko Ichimura, and Petra Todd. Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme. *The Review of Economic Studies*, 64(4):605–654, 1997. doi: 10.2307/2971733. URL http://www.jstor.org/stable/2971733.

Gilles Hilary and Kai Wai Hui. Does Religion Matter in Corporate Decision Making in America? *Journal of Financial Economics*, 93(3):455–473, sep 2009. ISSN 0304405X. doi: 10.1016/j.jfineco.2008.10.001. URL https://linkinghub.elsevier.com/retrieve/pii/S0304405X0900083X.

Yu Jen Hsiao and Wei Che Tsai. Financial Literacy and Participation in the Derivatives Markets. *Journal of Banking and Finance*, 88:15–29, 2018. ISSN 03784266. doi: 10.1016/j.jbankfin.2017.11.006. URL https://doi.org/10.1016/j.jbankfin.2017.11.006.

Tullio Jappelli and Mario Padula. Investment in Financial Literacy and Saving Decisions. *Journal of Banking and Finance*, 37(8):2779–2792, aug 2013. ISSN 03784266. doi: 10.1016/j.jbankfin.2013.03.019. URL https://linkinghub.elsevier.com/retrieve/pii/S0378426613001623.

Wei Jiang, Ashlyn Aiko Nelson, and Edward Vytlacil. Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency. *The Review of Economics and Statistics*, 96(1):1–18, mar 2014. ISSN 0034-6535. doi: 10.1162/REST_a_00387. URL https://direct.mit.edu/rest/article/96/1/1/58122/Liar-s-Loan-Effects-of-Origination-Channel-and.

Alejandro Justiniano, Giorgio E. Primiceri, and Andrea Tambalotti. Credit Supply and the Housing Boom. *Journal of Political Economy*, 127(3):1317–1350, jun 2019. ISSN 0022-

3808. doi: 10.1086/701440. URL https://www.journals.uchicago.edu/doi/10.1086/701440.

Leora Klapper, Annamaria Lusardi, and Georgios A. Panos. Financial Literacy and Its Consequences: Evidence from Russia during the Financial Crisis. *Journal of Banking and Finance*, 37(10):3904–3923, 2013. ISSN 03784266. doi: 10.1016/j.jbankfin.2013.07.014. URL http://dx.doi.org/10.1016/j.jbankfin.2013.07.014.

Zhenguo Lin, Yingchun Liu, and Jia Xie. Immigrants and Mortgage Delinquency. *Real Estate Economics*, 44(1):198–235, feb 2016. ISSN 10808620. doi: 10.1111/1540-6229.12092. URL https://onlinelibrary.wiley.com/doi/10.1111/1540-6229.12092.

Christopher Mayer, Karen Pence, and Shane M. Sherlund. The Rise in Mortgage Defaults. *Journal of Economic Perspectives*, 23(1):27–50, jan 2009. ISSN 0895-3309. doi: 10.1257/jep.23.1.27. URL https://pubs.aeaweb.org/doi/10.1257/jep.23.1.27.

Luis Mejia. Availability of Credit and Loan Default: A Look at the Commercial Mortgage Supply Cycle. *Journal of Real Estate Research*, 18(1):175–196, jan 1999. ISSN 0896-5803. doi: 10.1080/10835547.1999.12090989. URL https://www.tandfonline.com/doi/full/10.1080/10835547.1999.12090989.

Atif Mian and Amir Sufi. The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *Quarterly Journal of Economics*, 124(4):1449–1496, nov 2009. ISSN 0033-5533. doi: 10.1162/qjec.2009.124.4.1449. URL https://academic.oup.com/qje/article-lookup/doi/10.1162/qjec.2009.124.4.1449http://qje.oxfordjournals.org/lookup/doi/10.1162/qjec.2009.124.4.1449.

Atif Mian and Amir Sufi. House Prices, Home Equity–Based Borrowing, and the US Household Leverage Crisis. *American Economic Review*, 101(5):2132–2156, aug 2011. ISSN 0002-8282. doi: 10.1257/aer.101.5.2132. URL https://pubs.aeaweb.org/doi/10.1257/aer.101.5.2132.

Sauro Mocetti and Eliana Viviano. Looking behind Mortgage Delinquencies. *Journal of Banking and Finance*, 75:53–63, 2017. ISSN 03784266. doi: 10.1016/j.jbankfin.2016.11.002. URL http://dx.doi.org/10.1016/j.jbankfin.2016.11.002.

Paul Rosenbaum and Donald Rubin. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika*, 70(1):41–55, 1983. ISSN 00063444. doi: 10.1093/biomet/70.1.41. URL http://www.jstor.org/stable/2335942.

Murray D. Smith. Modelling Sample Selection Using Archimedean Copulas. *The Econometrics Journal*, 6(1):99–123, jun 2003. ISSN 1368-4221. doi: 10.1111/1368-423X.00101. URL https://academic.oup.com/ectj/article/6/1/99-123/5074166.

Maarten Van Rooij, Annamaria Lusardi, and Rob Alessie. Financial Literacy and Stock Market Participation. *Journal of Financial Economics*, 101(2):449–472, 2011. ISSN 0304405X. doi: 10.1016/j.jfineco.2011.03.006. URL http://dx.doi.org/10.1016/j.jfineco.2011.03.006.

Maarten C.J. Van Rooij, Annamaria Lusardi, and Rob J.M. Alessie. Financial Literacy, Retirement Planning and Household Wealth. *The Economic Journal*, 122(560):449–478, may 2012. ISSN 0013-0133. doi: 10.1111/j.1468-0297.2012.02501.x. URL https://academic.oup.com/ej/article/122/560/449-478/5079981.

Jeffrey M. Wooldridge. *Econometric Analysis of Cross Section and Panel Data*. The MIT Press, Cambridge, MA, second edition, jan 2010.

Appendix

Table A1: Dominance statistics of variables

	Dominance statistics	Marginal contribution
Financial industry	0.0003	0.48%
Financial literacy	0.0035	5.62%
Mortgage characteristics		
FRM	0.0010	1.61%
Interest rate	0.0100	16.05%
LTV	0.0091	14.61%
Debt-to-income ratio	0.0103	16.53%
Household characteristics		
Male	0.0030	4.82%
Age	0.0006	0.96%
Married	0.0011	1.77%
Number of children	0.0019	3.05%
White	0.0069	11.08%
Religion	0.0019	3.05%
College	0.0005	0.80%
Self-employed	0.0019	3.05%
Household income	0.0040	6.42%
Household wealth	0.0063	10.11%