

How Does Financial Literacy Affect Mortgage Default?*

Sumit Agarwal

National University of Singapore

Souphala Chomsisengphet

Office of the Comptroller of the Currency

Yunqi Zhang

National University of Singapore

* The views expressed in this paper are those of the authors and may not reflect those of the Office of the Comptroller of the Currency.

How Does Financial Literacy Affect Mortgage Delinquency?

Abstract

This paper uses a dataset from one of the leading subprime lenders in America, containing detailed information on borrower and loan characteristics, finds that borrowers from the financial industry, who have higher financial literacy, are less likely to become delinquent. This effect cannot be explained by borrower characteristics such as income and education, loan terms, property characteristics, or geographic effects. We also find there are variations in this effect of financial literacy for different types of borrowers or different kinds of loans. Our results indicate that financial literacy plays an important part in repayment behavior.

Key words: financial literacy, mortgages, financial crisis, household finance

JEL classification: D10, R20, G01, G21

1. Introduction

Households have been increasingly involved in financial markets and have access to increasing numbers of new financial products and services. Before the outbreak of the subprime crisis, there had been an increasing number of people holding consumer credit and mortgages. Alternative financial services have also grown fast in recent years.¹ As a consequence, people need to make more financial decisions than before. Many financial products are complex and have high requirements on investors' financial sophistication. People's financial literacy, however, are normally poor.² Researchers cast doubt on the assumption that people can make financial decisions rationally, which stimulates studies on financial literacy. There are some comments that the spike of mortgage delinquency rates in the subprime crisis are, to a large extent, attributed to individuals' poor financial literacy (Akerlof and Shiller, 2010; Boeri and Guiso, 2008). However, empirical evidence on the role of financial literacy in mortgage delinquency is scarce.

In this paper we use unique proprietary data containing detailed information on borrower characteristics and mortgage origination and performance, and comprehensively analyze how financial literacy affects mortgage borrowers' repayment behavior. Financial literacy means the "ability to process economic information and make informed decisions about financial planning, wealth accumulation, pensions, and debt" (Lusardi and Mitchell, 2014).³ We hypothesize that people with better financial literacy can make better financial planning such that their income

¹ See FINRA Investor Education Foundation (2009), Lusardi (2011), and Lusardi and Mitchell (2014).

² See the review by Agarwal et al. (2011), and Lusardi and Mitchell (2014).

³ Definition of financial literacy varies in the literature. Particularly, some studies distinguish financial expertise from financial literacy (see McDaniel, Martin, and Maines (2002), depending on how first-hand information and second-hand information help to build this ability. We do not make such a distinction.

and wealth is sufficient to pay their periodical debt services, so they are less likely to become delinquent.

To measure financial literacy, we construct a dummy variable based on whether the borrower works in the financial industry. We control for socioeconomic characteristics including education level and family income to address potential correlation between financial literacy and socioeconomic factors. We also consider mortgage contract terms and disentangle the effects of financial literacy on financial budgeting from the effects of financial literacy on loan choices. Additionally, we take into account the geographical effects and the property-type fixed effects. We find that borrowers from the financial industry are 14% less likely to become delinquent than borrowers from the non-financial industry. This lower likelihood of delinquency cannot be explained by socioeconomic factors such as education and income, nor can it be explained by mortgage characteristics, location and origination year.

We analyze the subsample comprising full-documentation loans to eliminate the concern that borrowers may exaggerate income when applying non-full-documentation loans. We also examine the period from 2005 to 2007 during which the income trends of the financial industry are similar to those of the non-financial industry. The results are similar.

We compare the mark-to-market loan-to-value ratio (LTV) of high-literacy borrowers' delinquent mortgages and non-high-literacy borrowers' delinquent mortgages. We find no difference in the mark-to-market LTV and eliminate the possibility that the difference in the delinquency rates of high and low financial literacy borrowers comes from their different incentives to strategically default.

We construct two other dependent variables for robustness tests. We compute the proportion of the time when a mortgage is in delinquency, and the ratio of the number of a

borrower's missed payments to the number of the borrower's total payments. We get similar results using the different delinquency measures.

This paper contributes to the literature in the following aspects. First, many studies on mortgage delinquency have emerged since the subprime crisis of 2007. These studies focus on the role of loan characteristics, trigger events like unemployment and divorce, equity position, debt-to-income ratio, credit history, local housing markets, and macroeconomic conditions in mortgage delinquency.⁴ However, there is little research on how financial literacy affects mortgage delinquency. To our knowledge, there is only one related paper: Gerardi, Goette, and Meier (2013) survey information on about three hundred subprime borrowers and find that numerical ability has statistically significant correlation with the borrowers' mortgage default. We contribute to the literature by focusing on the comprehensive ability to handle financial issues. Our dataset also increases the geographical coverage from three states in Gerardi, Goette, and Meier (2013) to the entire U.S..

Second, we add to the existing literature on financial literacy and financial education programs. There is worldwide evidence that households' financial literacy is poor.⁵ That challenges the rational individual assumption in conventional economic models and spurs innovation in both theoretical and empirical research. In theoretical research, some recent works incorporate financial literacy in the model of consumers' financial decision (Delavande,

⁴ See the review by Jones and Sirmans (2014).

⁵ See Agarwal et al. (2015), Annamaria Lusardi and Mitchell (2009), Agnew et al. (2011), Lusardi and Mitchell (2011), Lusardi, Mitchell, and Curto (2010), Bateman et al. (2012), Banks and Oldfield (2007), Behrman et al. (2012), Cole, Sampson, and Zia (2011), Hastings and Mitchell (2011), Lusardi and Tufano (2009), and McArdle, Smith, and Wills (2011).

Rohwedder, and Willis, 2008; Jappelli and Padula, 2013; Hsu, 2011; Lusardi, Michaud, and Mitchell, 2013). On other hand, many studies provide empirical evidence on the association between financial literacy and household behavior. Several studies document the relation between financial literacy and individual's behavior on assets management.⁶ At the same time, the effects of financial literacy on liabilities have also been discovered. Agarwal et al. (2009) find that people at different ages have different performance regarding interest rates of a wide range of financial products. Moore (2003), Lusardi and Tufano (2009), Lusardi and Scheresberg (2013), and Mottola (2013) argue that less financially literate borrowers are more likely to have high-cost debts; Agarwal and Mazumder (2013) find people with higher cognitive ability make less mistakes in using credit card and loan choice; Stango and Zinman (2009) conclude that less financially sophisticated borrowers have larger-size loans and less wealth. Bucks and Pence (2008) find that people are not aware of the specific contents of their loan contracts, particularly the interest rate. These studies indicate that literacy may affect loan choice. Agarwal et al. (2011) review the literature on financial counseling and education, and find that though the literature shows some evidence that financial education programs are correlated with improved financial decisions, further studies are needed to give strong evidence that the improved outcomes are from better financial literacy obtained from the education programs, due to the limitations of studies on education program: decline in delinquency is because of the threat of oversight and the transaction cost of counseling instead of the information acquired during counseling (Agarwal et al., 2014). Our research avoids such limitation and improves our understanding of

⁶ See Banks and Oldfield (2007), Christelis, Jappelli, and Padula (2010), Lusardi and Mitchell (2007), McArdle, Smith, and Wills (2011), and van Rooij, Lusardi, and Alessie (2011).

this issue.

The rest of this article is organized as follows. Section 2 describes our data and the methodology. Section 3 studies loan choices. Section 4 analyzes the effect of financial literacy on delinquency and Section 5 investigates the heterogeneity of the effect. Section 6 presents the robustness tests. Section 7 concludes the paper.

2. Data and Methodology

2.1. Data

The data used in this paper are obtained from one of the leading subprime lenders in the United States of America, which originated subprime loans all over the United States. The data consist of residential mortgage loans originated from 1996 to 2007 across 8,900 cities and 50 states (and the District of Columbia) in the United States. The data contain extensive information on borrower, loan, and property characteristics. First, it has detailed information on each borrower's demographic characteristics, financial status, and credit rating. We know the borrower's age, gender, marital status, years of residence, and employment status. We also have information on the co-borrower's education, gender, combined monthly income of the borrower and the co-borrower, etc. Meanwhile, the data contain the debt-to-income ratio, which equals the total monthly housing expenses for the borrower and co-borrower combined divided by the total monthly income for the borrower and co-borrower combined, and times one hundred. We have FICO score, a proxy we normally use to measure credit quality. Most importantly, we know the borrower's occupation and we are going to use this information to proxy financial literacy. The details are explained in Section 2.2.

Second, we have detailed information on both loan origination and performance. For

origination, we have information on the interest rate, the loan amount, the loan term, the combined LTV, the interest rate type (ARM or FRM), the lien status, the documentation type, the servicer's name, whether the loan has prepayment penalty, whether the loan is a subprime mortgage and the purpose of the loan (to purchase house or refinance existing mortgage). The data also keep a record of the performance history. We know, for each month, whether the borrower has payments overdue, as well as how many payments have been missed. Last, the data provide information on property characteristics. The data contain the occupancy status, the property type, and the postal code of the property.

2.2. Methodology

The main research question of the paper is how financial literacy affects delinquency. We use the following logistic regression and least squares regression to test the relation between financial literacy and mortgage delinquency:

$$Pr(D_i = 1) = 1 / (1 + \exp(\alpha + \beta * L_i + \sum \lambda_{ij} * x_{ij})). \quad (1)$$

$$Y_i = \alpha + \beta * L_i + \sum \lambda_{ij} * x_{ij} + \varepsilon_i. \quad (2)$$

In the logistic regression (1), the dependent variable D is a dummy measuring mortgage delinquency. We define delinquency as when a loan has overdue payments. For each month, we can identify whether a borrower becomes delinquent. In the baseline specification, we consider whether a borrower becomes delinquent during the course of the mortgage tenure, regardless of how many times the borrower becomes delinquent. Therefore, in our baseline model, the dependent variable D equals 1 if the borrower becomes delinquent at least once and 0 otherwise. In a robustness test we also construct continuous dependent variables based on the frequency of delinquency and the extent of delinquency, which are explicitly introduced in Section 6.4. The corresponding least squares regression is equation (2).

The most important independent variable L is the measure of financial literacy. With the unique dataset we know whether or not a borrower works in the financial industry: the dataset includes a variable briefly describing the borrower's job; we can infer whether the borrower works in an industry related to finance. We define the financial industry as jobs related to the knowledge of interest rate computation, risk management, macroeconomics, financial planning or other aspects important to liabilities management. We take into account banking, insurance, fixed-income security, stock and derivatives investment, risk analysis and management, accounting, asset management, financial regulation, economists, professor of economics-related disciplines, and so on. People of these professions are more likely to have received good training on financial knowledge and skills, and hence, have higher financial literacy.⁷ We construct a dummy variable for financial literacy, which equals 1 if the borrower works in the financial industry and 0 otherwise.⁸ Note that we merely use this dummy to measure the borrower's financial literacy and

⁷ One concern is that when an individual enters financial industry, this individual may not gain financial literacy immediately, so the employment sector cannot be a proxy for financial literacy. However, we have information on professional years, namely, how long had this individual been in this industry. 97.4% of financial industry borrowers had been working in financial industry for more than one year when they applied the loans. Therefore, we expected most of them have higher financial literacy than non-financial industry borrowers. In an unreported test, we exclude financial industry borrowers with working experience less than one year and get similar results. Another issue is if individual exits financial industry, this individual may not loss financial literacy immediately. So in our sample, some non-financial industry borrowers may have as much financial literacy as financial industry borrowers have. However, that only make our estimates downward biased and do not affect our qualitative inference.

⁸ Borrowers from financial industry may get subprime loans if their perceived risk is high. For instance, they may be risk loving and therefore have bad credit history. That does not affect our qualitative inference: in our paper, we

we do not require the causality relation between the borrower's profession and financial literacy level. It is possible that some people are talented in finance and choose to join the financial industry. That possibility does not affect our results.

More specifically, the occupation variable briefly describes the borrower's job.⁹ We can infer the employment sector of the borrower from the description. The inference is realized in three steps. First, we sort observations by the job description variable in alphabetical order, then we scan the value of the description variable for all observations and pick out strings of letters related to finance. We pick about two hundred strings related to banking, insurance, fixed-income security, stock and derivatives investment, risk analysis and management, accounting, asset management, financial regulation, economists, and financial research.¹⁰ Second, we let the financial literacy dummy equals one if the job description variable includes one of these strings of letters. Third, we keep all the observations whose financial literacy dummy equals 1, and manually check these observations one by one to find out borrowers who have been misclassified as high financial literacy borrower.¹¹ We reset the value of financial literacy dummy to zero for these observations.

show that financial industry borrowers are *less* likely to become delinquent because of financial literacy; therefore, any factor that makes financial industry borrowers riskier can only make the estimates the lower bound of the real effects.

⁹ Some examples are "DRIVER", "HOUSE KEEPING", "DEAL/CASINO", and "LOAN CONSULTANT/LENDING BANK".

¹⁰ Some examples are "STOCK", "MORTG" (related to mortgage), "BOND", "FINANC" (related to finance), "LAON" (typo of "LOAN"), "INSURANCE", "PENSION", and "MONEY".

¹¹ For example, there may be "BODYGUARD/BANK" who works in a bank but the job does not require financial literacy, or "CIGARETTE STOCKER" who does not work in financial industry but the job title includes "STOCK".

We include a set of controls (x_i) from borrower characteristics to property characteristics. Financial literacy may lead to lower delinquency rates through two mechanisms. First, borrowers with better financial literacy are more likely to choose suitable mortgages. This is one way finance literacy decreases the probability of delinquency (Moore, 2003; Lusardi and Tufano, 2009; Lusardi and Scheresberg, 2013; Mottola, 2013). Another mechanism is conditional on certain mortgage contract borrowers with better financial literacy being able to make better decisions on consumption and investment such that they are able to pay off the debts. We include mortgage term variables including the loan amount, the annual percentage interest rate, the adjustable rate mortgage (ARM) dummy, the first lien dummy, the interest-only loan dummy, the documentation status, the subprime dummy, the combined LTV, and the debt-to-income ratio in the regression to disentangle the second mechanism from the first one.

Second, we pick out the effects of education on repayment behavior. The borrower's education level is absent in our dataset. Nevertheless, our dataset includes the co-borrower's years of schooling and we use it as the proxy for the borrower's education, as Mare (1991) points out that people with similar education level tend to marry each other. In our sample, 93.2% of the borrowers were married and the difference in the ages of the borrower and the co-borrower is less than ten years. We consider them couples. For the other 6.8%, the co-borrowers are most likely to be the borrowers' relatives or close friends who have close relationship with the borrowers. They tend to have similar education levels. In an unreported robustness test, we restrict our sample to those 93.2% for which we think the borrower and co-borrower are couples and get similar results.

Third, the borrower's profession is correlated with certain socioeconomic characteristics. For instance, borrowers from the financial industry may earn higher incomes which leads to lower delinquency rates, and one may argue the effects of financial literacy on delinquency are

effectively the effects of income on repayment behavior. To address this issue, in the regression we include combined monthly income, including the income of both the borrower and co-borrower to measure the borrower's ability to pay off the debt more accurately.

Fourth, we also control for the geographic fixed effects and origination year. The financial industry may cluster in certain areas (such as New York City) and market conditions in these regions may be different from those in others. Additionally, credit supply may be different in different periods of time. In some years, lenders relax loan screening standard and mortgages originated in these years are more likely to become delinquent (Mian and Sufi, 2009).

We also control for other variables associated with the borrower's repayment behavior: FICO score, the purpose of borrowing the mortgage, occupation status, property type, and servicer fixed effects. We further control for other demographic characteristics including gender, borrower age, marital status, borrower residence years, self-employed or not, and citizenship to further wipe out potential omitted variables problems.

2.3.Descriptive Statistics

Table 1 shows the summary statistics of the demographic characteristics, financial literacy dummy, loan origination, and performance information for the full sample.¹² 3.6% of the borrowers in our sample work in the financial industry.¹³ 98.9% of the loans are subprime loans,

¹² We restrict the value of the variables within reasonable boundaries as follows: 0<borrower's age < 80; 0< borrower's residence years < 100, 1<debt-to-income ratio <100, 300< FICO score < 850, 0<co-borrower's years of schooling < 40, combined LTV >1, and combined monthly income < 30000. The original sample size is 152,829.

¹³ According to Bureau of Labor Statistics, the employment sector of Financial Activities accounts for 5.3% of the total labor force. See http://www.bls.gov/emp/ep_table_201.htm.

a source of the financial crisis in 2007. 71.3% of the borrowers are males, in part reflecting the gender disparity in socioeconomic status.

[Insert Table 1 About Here]

Then we divide the full sample into two subsamples based on the financial literacy dummy: one comprises borrowers working in the financial industry; the other includes borrowers working in the non-financial industry. Table 2 presents the summary statistics of these two subsamples. The delinquency rate for the financial-industry borrowers is 9.7%, 15.5% lower than the delinquency rate for the non-financial-industry borrowers. The difference in the delinquency rates implies financial literacy may have an impact on delinquency. The average FICO score is higher for the financial-industry borrowers. That is consistent with the correlation of financial literacy and FICO score in the literature: “households that can maintain a high FICO score show that they have the discipline and knowledge to plan their financial matters effectively” (Amromin et al., 2011). The correlation indicates that borrowers’ professions are correlated with financial literacy, supporting the effectiveness of our literacy dummy. The co-borrower’s years of schooling, a proxy for the borrower’s education, is similar across the two subsamples, suggesting that the difference in the delinquency rates is not likely to come from difference in education. The two subsamples have different average combined income: borrowers from the financial industry earn 27.7% more than their counterparts from the non-financial industry. Almost all the loan characteristics are similar across these two subsamples, except that the financial-industry borrowers have bigger loan size, probably because of their higher income. We control for all these variables in our regression to eliminate the disturbances from these factors.

[Insert Table 2 About Here]

3. Loan Choices

Before the analysis of mortgage borrowers' repayment behavior, we first explore how financial literacy affects their loan choices. Borrowers with higher financial literacy are expected to be smarter in loan choices, and they are more likely to have loans with lower cost. We consider prepayment penalty. If a loan has prepayment penalty, the borrower would be charged a punitive fee if the loan is prepaid. Since the right to prepay is a call option and has value for the borrower, prepayment penalty decreases the value of the option and increases the loan cost. If borrowers from the financial industry perform better in loan choices, they are less likely to have prepayment penalty.^{14 15}

We use equation (3) to test our hypothesis. We use logistic regression and control for *all* variables in Table 4 except prepayment penalty years. The results are in Table 3. Estimates for control variables are omitted. Borrowers from the financial industry are 21.1% less likely to have prepayment penalty. The results indicate that borrowers with higher financial literacy tend to have low-cost loans and are consistent with our hypothesis.

[Insert Table 3 About Here]

¹⁴ Borrowers may forgo prepayment option to obtain lower interest rates. In the regression we control for the interest rate to address this concern.

¹⁵ In an unreported analysis, we find that financial literacy is also correlated with lower interest rates. However, we are unable to establish causality between financial literacy and interest rate. Financial industry borrowers are less risky and hence are of better credit quality, and the lower interest rates may result from their better credit quality rather than their better decision.

4. Loan Performance: The Baseline Result

Then we rigorously analyze how financial literacy affects borrowers' repayment behavior using logistic regression (1).

First of all, we regress the delinquency dummy on the financial literacy measure. We include FICO score, purpose of borrowing the loan, occupancy-status, and property type, servicer and origination year fixed effect. Column (1) of Table 4 presents the results. We report the odds ratio rather than the coefficient to measure the marginal effects of financial literacy. The effect of financial literacy dummy is negative and statistically significant at 1% level, and consistent with our hypothesis and what Table 2 suggests. Additionally, the magnitude of the effect is important: the odds ratio is about 0.87, meaning that if the borrowers are from the financial industry, the odds of delinquency are about 13% lower, holding everything else constant.¹⁶

Next we take into account factors that are correlated with financial literacy and may affect borrowers' repayment behavior, to eliminate alternative explanations.

4.1. Mortgage Characteristics

Literature shows that individuals with better financial literacy choose loans with lower cost (Moore, 2003; Lusardi and Tufano, 2009; Lusardi and Scheresberg, 2013; Mottola, 2013). In Column (2) of Table 4, we include loan characteristics: the interest rate and a set of dummies for

¹⁶ The effect can be considered the lower bound for the effects of financial literacy. One important reason is there may be noise in our literacy measure: some financial industry borrowers identified in our sample may be only slightly better in financial literacy than non-financial industry borrowers.

ARM, first-lien loans, prepayment penalty, interest-only loan, full documentation and subprime loans. Borrowers with lower financial literacy may not be good at controlling their debts lower than the affordable level, and they may choose larger loans (Stango and Zinman, 2009). Therefore in Column 3 of Table 4 we put in variables including the loan amount, LTV and the debt-to-income ratio.

[Insert Table 4 About Here]

The results indicate that the loan characteristics have statistically significant impacts on delinquency. For instance, the borrowers who choose ARM are 7.4% more likely to become delinquent, consistent with the findings by Cunningham and Capone (2014). Borrowers who provide full documentation are about 17.2% less likely to become delinquent, as Jiang, Nelson, and Vytlačil (2014), and LaCour-Little and Yang (2013) find. Nevertheless, both the significance and the magnitude of the odds ratio of financial literacy do not change when we control for these loan characteristics, suggesting the effect of financial literacy on repayment behavior captured by the estimated odds ratio of the literacy dummy does not function through loan choice. The results indicate that it is insufficient to prevent borrowers from high-cost loans, such as the enactment of the anti-predatory lending law.

4.2. Education

The potential relation between education and repayment behavior may affect our inference on the relationship between financial literacy and delinquency. It is possible that people with better education are more likely to enter the financial industry, and higher education level makes them less likely to become delinquent. So the lower delinquency rates of borrowers from the

financial industry do not result from their better financial literacy but their better education. As discussed in Section 2.2, we proxy the borrower education by the co-borrower's years of schooling to address this issue.

The results are in Column (4) of Table 4. Education has a statistically significant impact on the borrowers' delinquency, but the magnitude is very small, only 0.6%. Moreover, after we control for education, the effect of financial literacy on repayment behavior is still statistically significant and the magnitude of the odds ratio does not have any material change. Therefore, the relation between financial literacy and delinquency we have identified is not the impact of education on repayment behavior.

4.3. Income

In Table 2 we find borrowers from the financial industry have higher average income. Higher income represents higher ability to repay debts. Thus, the financial-industry borrowers are less likely to become delinquent simply because they have higher income. We incorporate the natural logarithm of the combined monthly income to eliminate the effects of income. Column (5) of Table 4 presents the results. Income has no statistically significant effects on repayment behavior, suggesting that after considering debt-to-income ratio, the income *per se* is not an important predictor of delinquency rate. More importantly, the estimates of financial literacy remain virtually unchanged. Borrowers from the financial industry are still about 14.1% less likely to become delinquent and the estimates are significant at 1% level. Thus, we can eliminate the possibility that the effect of financial literacy on delinquency is from the correlation of financial literacy and income.

4.4. Local Real Estate Markets

Another concern is that in some areas, the finance industry may account for a larger portion of the local employment than in other areas because the finance industry has high spill-over effects and may cluster in some places. Real estate markets in different locations, which affect borrowers' repayment behavior, are partially separated and they are affected by local economic conditions and other geographic attributes. It is possible that the financial industry clusters in some areas where the real estate market condition is better, and borrowers from these areas have lower delinquency rates. This negative correlation of the financial industry and delinquency rates due to different real estate markets may contaminate our estimation.

In previous regression from Column (1) to Column (5) of Table 4 we have taken into account time-varying factors by allowing the standard errors to cluster at the MSA level.¹⁷ To further address this issue, we control for MSA fixed effects in Column (6) of Table 4 to account for time-constant effects.¹⁸ The impact of financial literacy on delinquency is not affected by the incorporation of MSA fixed effects: the effect is still significant at 1% level and the magnitude of odds ratio is similar. Thus the impact of financial literacy on delinquency is not related to the local real estate markets.

4.5. Propensity Score Matching

We also construct a more balanced sample by propensity score matching. We obtain the propensity score by regressing the literacy dummy on loan characteristics that are used as control

¹⁷ Throughout the paper, the standard errors are clustered at the MSA level unless otherwise stated.

¹⁸ The number of observations of Column (6) is lower than those of others because some MSAs do not have delinquent loans.

variables in Table 4 with logistic regression. Then we use nearest neighborhood matching to construct the sample and repeat regression (1). Table 5 shows the results. Financial literacy dummy is still correlated with lower delinquency likelihood, and the magnitude of the estimates is close to those in Table 4.^{19 20}

5. Heterogeneity of the Impact

We have displayed that better financial literacy reduces borrowers' delinquency. In this section, we first focus on financial industry borrowers who have professional certificates, senior positions, or loan-related jobs. They tend to have even higher financial literacy than the average level of financial industry borrowers, therefore the effect of financial literacy on delinquency rate should be bigger.

Then we test how the effect differs depending on borrower age. We divide the sample into different subsamples based on borrower age, and compare the estimates across different subsamples.²¹

¹⁹ We exploit k -nearest neighbor matching. Table 5 shows the case when k equals 5. When k gradually decreases, the odds ratio for the literacy dummy remains in the range between 0.85 to 0.89.

²⁰ We also examine the relationship between mathematical ability and loan performance, which is documented by Gerardi, Goette, and Meier (2013). We consider people specializing in statistics, engineering, programming and mechanics borrowers with high mathematical ability, and we find that the relationship between mathematical ability and loan performance is insignificant. The reason could be that our measure of mathematical ability is not a direct measure and may not be accurate enough to capture the effects.

²¹ Another way is to construct an interaction of literacy dummy and each of the demographic and loan characteristics variables. The features of our sample, however, do not allow us to use this method. For about 96%

5.1. Senior and More Professional Staff

First, instead of considering all people from the financial industry, we focus on the following people in the financial industry: senior staff (such as team leader, vice president or CEO), people with professional certificates (such as CFA or CPA), borrowers whose jobs require high financial knowledge level (such as analysts or financial planners), and people with loan-related jobs (such as mortgage brokers or loan officers). They tend to have higher financial literacy than the average level of the financial industry. Therefore, we should observe a bigger effect of the financial literacy on delinquency. We define these people as high-financial literacy borrowers, drop financial industry borrowers who are not high-financial literacy borrowers, and repeat regression (1).^{22 23} The results are in Table 6. The odds ratio is 0.774, indicating that those senior and more professional staff from financial-industry are about 22.6% less likely to become delinquent as compared with the non-financial-industry borrowers. The magnitude of the effect is much bigger than the one in Section 4 (from 12.5% to 14.4%) and consistent with our hypothesis.

[Insert Table 6 About Here]

of the observations, the literacy dummy equals 0. Thus for the interaction of this literacy dummy to any other variable, at least 96% of the observations would be 0, and the interaction would be highly correlated with the literacy dummy, which leads multi-collinearity problem and make coefficient estimates insignificant. Therefore we do not use the interaction method.

²² Only in Section 5.1, the financial literacy dummy is define based on senior staff, staff with professional certificates, and loan-related staff. In all other sections, we use the financial literacy dummy defined in Section 2.2.

²³ Note we control for income which is highly correlated with salary and wealth.

5.2. Age

The relation between age and financial literacy has been widely studied. Theoretical research predicts that financial literacy follows a hump-shaped pattern during the life-cycle: Jappelli and Padula (2013) endogenize financial literacy in a multi-period life-cycle model, and find that households invest in financial literacy until retirement and after that financial literacy is falling. This finding is supported by empirical studies (Agarwal et al., 2007; Chen and Volpe, 1998; Lusardi, Mitchell, and Curto, 2010; Lusardi and Mitchell, 2011). Our study also confirms this viewpoint. We divide our sample into three parts: borrowers younger than 35, between 35 and 50, and older than 50. Then we estimate regression (1) for the three subsamples respectively. Table 7 presents the results. For borrowers younger than 35 (Column (1)), the effect is significant and the magnitude of the effect is 19.7%, bigger than 14.4% for the full sample. The effect is even larger for borrowers older than 50 (Column (3)), which is 23.5%. For borrowers between 35 and 50, however, the estimate becomes insignificant. This result is consistent with the findings in previous studies. One interpretation is young and old borrowers on average have lower financial literacy. The difference in financial literacy between borrowers from the financial industry and borrowers from the non-financial industry is bigger for young and old borrowers than for middle-age borrowers. Therefore, the effect of financial literacy on delinquency is greater for young and old people.²⁴

²⁴ One might doubt the significance level (5% and 10%) is not high. The substantial difference in the point estimates can alleviate this concern.

[Insert Table 7 About Here]

6. Robustness Tests

6.1. *Income Verification and Income Trends*

There are two issues regarding our income measure. First, in order to obtain larger loan amount, borrowers may exaggerate their income when applying non-full-documentation loans, so we may have measurement error of income. To address this concern, we analyze the subsample comprising full-documentation loans: the borrower needs to provide documents, such as W-2 form, pay stub, tax returns, and income statement, to verify their income when applying the loans. Using the subsample we estimate equation (1) and the results are in Column (1) of Table 8.

[Insert Table 8 About Here]

Another issue is that the income trends of financial industry may be different from those of the other sectors, which make our income measure inaccurate. We restrict our sample to 2005-2007: we focus on loans originated after 1 January 2005, and analyze their performance information before 2008 (we do not have performance information after 2007). We compute the ratio of median personal total income of non-financial industry and median personal total income of financial industry for 2005, 2006 and 2007 respectively, based on the American Community Survey (ACS) data. The ratio is 0.526 for 2005, 0.505 for 2006, and 0.525 for 2007, so there is little difference in the income trends of financial industry and non-financial industry during this period. We estimate equation (1) on this subsample and the results are in Column (2) of Table 8.

The estimates for both subsamples indicate that the income exaggeration and income

trends concerns do not affect our results. The effects of financial literacy on delinquency are -15.1% and -13.7% for the two subsamples respectively, similar to the results in Section 4 (from -14.4% to -12.5%).

6.2. Strategic Default

A possible argument undermining our results is that the strategic default behavior of the high-literacy borrowers and the low-literacy borrowers is different, which leads to the different delinquency rates of the two types of borrowers. Strategic default occurs when the property value is substantially lower than the mortgage balance: the borrower can choose to foreclose the house rather than pay off the mortgage debt.²⁵ To address this issue, we analyze the mark-to-market LTV of all delinquent mortgages. Mark-to-market LTV is the ratio of the loan balance to the market price of the house and reflects the borrowers' tendency to strategically default.

We extract all mortgages that have ever been delinquent, match their loan characteristics information to their performance information of the months when they are delinquent, and get a panel dataset. First, we compare the mark-to-market LTV of high-literacy borrowers' loans and low-literacy borrowers' loans when they are delinquent. Table 9 shows the summary statistics of the mark-to-market LTV for the two groups of borrowers. The means for the two groups are similar and are far less than 100%, suggesting that the put option is out of the money. The pattern is confirmed in Figure 1 which plots the kernel density of the mark-to-market LTV of the two

²⁵ If the borrower chooses foreclosure, the amount forgiven by the lender becomes taxable income to the borrower. Therefore the borrower's "payoff" from the foreclosure equals the difference between the outstanding balance and the property value minus the additional tax incurred. The borrower chooses foreclosure when this payoff is positive.

groups: for both groups, the mark-to-market LTV of almost all the loans are below 100%. Besides, the distributions of the mark-to-market LTV for the two groups are similar. Therefore there is little difference in the tendency to strategically default across the two groups.

[Insert Table 9 About Here]

To rigorously analyze the difference in the mark-to-market LTV, we apply equation (3) to the panel data, with the mark-to-market LTV as the dependent variable and using generalized least squares regression, to examine whether the mark-to-market LTV when a loan is delinquent is correlated to the financial literacy of the borrower. The results are in Table 10. The coefficient of the financial literacy dummy is insignificant, indicating that there is no statistically significant difference in the mark-to-market LTV between the two types of borrowers. Therefore, there is no difference in the tendency to strategically default.

[Insert Table 10 About Here]

6.3. *Falsification*

The key strategy of the paper is to assign higher financial literacy to people from the financial industry. In this section we conduct two falsifications: we change the way we assign higher financial literacy.

First, there are 4849 borrowers from the financial industry in our sample. Now we randomly choose 4849 observations in the sample and consider them high-literacy borrowers. Since the high financial literacy is randomly assigned, the high-literacy borrowers should not

perform better than others. We call them pseudo-high-literacy borrowers. Then we repeat the regression (2) to see whether the pseudo-high-literacy borrowers have lower delinquency rates. Since the assignment is random, every time it repeats, the assigned observations change. To get a consistent result, we repeat the random assignment and regression thirty times. The results are in Table 11. For simplicity we only report the odds ratio of the financial literacy dummy. Only one of the thirty random assignments the pseudo-high-literacy borrowers have lower delinquency rates. Therefore, the pseudo-high-literacy borrowers do not consistently perform better than the other borrowers.

[Insert Table 11 About Here]

Then we use an alternative strategy. The high financial literacy is assigned not to people from the financial industry but to other professions. We assign high financial literacy to engineers, and servicemen and police officers. Then we conduct regression (2). The results are in Table 12. The estimates of the pseudo financial literacy are not statistically significant for both the engineers, and the servicemen and police officers. The pseudo financial literacy has no impact on delinquency.

[Insert Table 12 About Here]

6.4. Alternative Delinquency Measures

The dependent variable we have used so far is a dummy variable indicating whether the borrower has become delinquent or not. We also construct another two measures of mortgage

delinquency as dependent variables. First, we define *delinq_ratio* as the fraction of time when the mortgage is in delinquency, i.e., the fraction of time when there are payments over due. Second, we define *missed_ratio* as the number of missed mortgage payments divided by the number of total payments. The difference between these two measures is that the first one only considers whether the mortgage is in delinquency in certain month, while the second one also considers the extent of delinquency (Gerardi, Goette, and Meier 2013).

Table 13 presents the summary statistics for the two delinquency measures for borrowers from the financial industry and the non-financial industry respectively. The means of both *delinq_ratio* and *missed_ratio* are lower for borrowers from the financial industry, with similar standard deviations, suggesting that better financial literacy may reduce the frequency of delinquency.

[Insert Table 13 About Here]

Then we use the two delinquency measures as the dependent variable in equation (2) respectively to conduct generalized least-squares regression. Table 14 presents the results. Column (1) is the result for *delinq_ratio* and the coefficient of financial literacy is negative, suggesting that better financial literacy makes borrowers become delinquent less frequently by 9.0% (0.0026 divided by 0.029). The result is similar for *missed_ratio*. Therefore, the improvement of borrower repayment behavior due to higher financial literacy is robust when we use different delinquency measures.

[Insert Table 14 About Here]

7. Conclusion

This paper uses a unique dataset including borrower, loan, and property characteristics, analyze how subprime mortgage borrowers' financial literacy affects the delinquency rates, and contribute to financial literacy literature and mortgage delinquency literature. We use borrowers' profession as the proxy of financial literacy. We find that borrowers working in the financial industry are 14.9% less likely to become delinquent than borrowers from the non-financial industry. The results support our hypothesis that better financial literacy reduces borrowers' delinquency rates. We control for loan characteristics and find that the effect of financial literacy on delinquency rates does not result from better loan choices; we also find the effect cannot be explained by education, a factor correlated with financial literacy, or geographic effects which may be correlated with the distribution of the financial industry. We exclude the influence of strategic default. We construct two alternative delinquency measures as dependent variables and find the effect is robust.

One may argue that the social stigma of delinquency could be felt more keenly for those who work in the finance industry and the lower delinquency rates result from social stigma. For example, it may be more disgraceful for financial-industry borrowers than for non-financial industry borrowers to become delinquent. Unfortunately we do not have any direct measure on social stigma. However, there is a literature on the relationship between employment sector and moral, which alleviates this concern. For instance, Cohn, Fehr, and Maréchal (2014) study the business culture in banking industry and find that people working in banking industry makes people more dishonest. In other words, financial industry borrowers are less "moral". Therefore, if there is any difference in the social stigma of delinquency, we tend to expect that the social stigma is *less* important for the delinquency of financial industry borrowers.

One limitation of our paper is that we focus on subprime loans. Subprime loans are important as they are the culprits of the subprime crisis. However, it would be interesting to study the prime borrowers and Alt-A borrowers. Further investigation in this would generate helpful conclusions and we leave this issue for future research.

References

- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2014. "The Effectiveness of Mandatory Mortgage Counseling: Can One Dissuade Borrowers from Choosing Risky Mortgages?." *National Bureau of Economic Research Working Paper*, no. No. 19920.
- Agarwal, Sumit, Gene Amromin, Itzhak Ben-david, Souphala Chomsisengphet, and Douglas D. Evanoff. 2011. "Financial Counseling , Financial Literacy , and Household Decision Making." In *Financial Literacy: Implications for Retirement Security and the Financial Marketplace*, 181. doi:10.2139/ssrn.1628975.
- Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas D. Evanoff. 2015. "Financial Literacy and Financial Planning: Evidence from India." *Journal of Housing Economics* 27. Elsevier Inc.: 4–21. doi:10.2139/ssrn.1728831.
- Agarwal, Sumit, John C Driscoll, Xavier Gabaix, and David Laibson. 2009. "The Age of Reason : Financial Decisions Over the Lifecycle." *Brookings Papers on Economic Activity*, 51–101.
- Agarwal, Sumit, and Bhashkar Mazumder. 2013. "Cognitive Abilities and Household Financial Decision Making." *American Economic Journal: Applied Economics* 5 (1): 193–207. doi:10.1257/app.5.1.193.
- Agnew, Julie R., Lisa R. Szykman, Stephen P. Utkus, and Jean a. Young. 2011. "Trust, Plan Knowledge and 401(k) Savings Behavior." *Journal of Pension Economics and Finance* 11 (01): 1–20. doi:10.1017/S1474747211000230.
- Akerlof, George A., and Robert J. Shiller. 2010. "Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism." *Princeton University Press*.
- Annamaria Lusardi, and Olivia S Mitchell. 2009. "How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness." *National Bureau of Economic Research Working Paper* No.15350.
- Authority, Financial Industry Regulatory. 2010. "Financial Capability in the United States: National Survey, Executive Summary."
- Banks, James, and Zoë Oldfield. 2007. "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving." *Fiscal Studies* 28 (2): 143–70. doi:10.1111/j.1475-5890.2007.00052.x.

- Bateman, Hazel, Christine Eckert, John Geweke, Jordan Louviere, Susan Thorp, and Stephen Satchell. 2012. "Financial Competence and Expectations Formation: Evidence from Australia*." *Economic Record* 88 (280): 39–63. doi:10.1111/j.1475-4932.2011.00766.x.
- Behrman, Jere R, Olivia S Mitchell, Cindy K Soo, and David Bravo. 2012. "How Financial Literacy Affects Household Wealth Accumulation." *American Economic Review* 102 (3): 300–304. doi:10.1257/aer.102.3.300.
- Boeri, Tito, and Luigi Guiso. 2008. "'The Subprime Crisis: Greenspan's Legacy.'" *The First Global Financial Crisis of the 21st Century* 73 73: 37.
- Bucks, Brian, and Karen Pence. 2008. "Do Borrowers Know Their Mortgage Terms?" *Journal of Urban Economics* 64 (2): 218–33. doi:10.1016/j.jue.2008.07.005.
- Chen, Haiyang, and Ronald P. Volpe. 1998. "An Analysis of Personal Financial Literacy among College Students." *Financial Services Review* 7 (2): 107–28.
- Christelis, Dimitris, Tullio Jappelli, and Mario Padula. 2010. "Cognitive Abilities and Portfolio Choice." *European Economic Review* 54 (1). Elsevier: 18–38. doi:10.1016/j.euroecorev.2009.04.001.
- Cohn, Alain, Ernst Fehr, and Michel André Maréchal. 2014. "Business Culture and Dishonesty in the Banking Industry." *Nature*, no. 516: 86–89. doi:10.1038/nature13977.
- Cole, Shawn, Thomas Sampson, and Bilal Zia. 2011. "Prices or Knowledge ? What Drives Demand for Financial Services in Emerging Markets ?" *The Journal of Finance* 66 (6): 1933–67.
- Cunningham, Donald F., and Charles A. Capone. 2014. "The Relative Termination Experience of Adjustable to Fixed-Rate Mortgages." *The Journal of Finance* 45 (5): 1687–1703.
- Delavande, Adeline, Susann Rohwedder, and Robert J. Willis. 2008. "Preparation for Retirement, Financial Literacy and Cognitive Resources." *Michigan Retirement Research Center Research Paper* 2008-190.
- Gerardi, Kristopher, Lorenz Goette, and Stephan Meier. 2013. "Numerical Ability Predicts Mortgage Default." *Proceedings of the National Academy of Sciences of the United States of America* 110 (28): 11267–71. doi:10.1073/pnas.1220568110.
- Hastings, Justine S, and Olivia S Mitchell. 2011. "How Financial Literacy and Impatience Shape Retirement Wealth and Investment Behaviors." *National Bureau of Economic Research Working Paper* No.16740.
- Hsu, Joanne W. 2011. "Aging and Strategic Learning: The Impact of Spousal Incentives on Financial Literacy." *Division of Research & Statistics and Monetary Affairs, Federal Reserve Board*. Oxford University Press.

- Jappelli, Tullio, and Mario Padula. 2013. "Investment in Financial Literacy and Saving Decisions." *Journal of Banking & Finance* 37 (8). Elsevier B.V.: 2779–92. doi:10.1016/j.jbankfin.2013.03.019.
- Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlačil. 2014. "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency." *The Review of Economics and Statistics* 96 (1): 1–18.
- Jones, Tim, and G. Stacy Sirmans. 2014. "Mortgage Default and Foreclosure : An Analysis of the Underlying Determinants , Associated Losses , and Valuation." *American Real Estate Society Working Paper*.
- LaCour-Little, Michael, and Jing Yang. 2013. "Taking the Lie out of Liar Loans: The Effect of Reduced Documentation on the Performance and Pricing of Alt-A and Subprime Mortgages." *Journal of Real Estate Research* 35 (4): 507–53.
- Lusardi, Annamaria. 2011. "Americans' Financial Capability." *National Bureau of Economic Research Working Paper* No.17103.
- Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S. Mitchell. 2013. "Optimal Financial Knowledge and Wealth Inequality." *National Bureau of Economic Research Working Paper* No. 18669.
- Lusardi, Annamaria, and Olivia S Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52: 5–44.
- Lusardi, Annamaria, and Olivia S. Mitchell. 2007. "Baby Boomer Retirement Security: The Roles of Planning, Financial Literacy, and Housing Wealth." *Journal of Monetary Economics* 54 (1): 205–24. doi:10.1016/j.jmoneco.2006.12.001.
- . 2011. "Financial Literacy around the World: An Overview." *Journal of Pension Economics and Finance* 10 (04): 497–508. doi:10.1017/S1474747211000448.
- Lusardi, Annamaria, Olivia S. Mitchell, and Vilsa Curto. 2010. "Financial Literacy among the Young." *Journal of Consumer Affairs* 44 (2): 358–80. doi:10.1111/j.1745-6606.2010.01173.x.
- Lusardi, Annamaria, and Carlo de Bassa Scheresberg. 2013. "Financial Literacy and High-Cost Borrowing in the United States." *National Bureau of Economic Research Working Paper* No.18969.
- Lusardi, Annamaria, and Peter Tufano. 2009. "Debt Literacy, Financial Experiences, and Overindebtedness." *National Bureau of Economic Research Working Paper* No. 14808.
- Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56 (1): 15–32.

- McArdle, John J., James P. Smith Smith, and Robert Wills. 2011. "Cognition and Economic Outcomes in the Health and Retirement Survey." *Explorations in the Economics of Aging*, University of Chicago Press, 209–33.
- McDaniel, Linda, Roger D. Martin, and Laureen a. Maines. 2002. "Evaluating Financial Reporting Quality: The Effects of Financial Expertise vs. Financial Literacy." *The Accounting Review* 77 (s-1): 139–67. doi:10.2308/accr.2002.77.s-1.139.
- 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): 1449–96.
- Moore, Danna L. 2003. "Survey of Financial Literacy in Washington State: Knowledge, Behavior, Attitudes, and Experiences." *Washington State Department of Financial Institutions*.
- Mottola, Gary R. 2013. "In Our Best Interest: Women, Financial Literacy, and Credit Card Behavior." *Numeracy* 6 (2): 4.
- Stango, Victor, and Jonathan Zinman. 2009. "Exponential Growth Bias and Household Finance." *The Journal of Finance* 64 (6): 2807–49. <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2009.01518.x/full>.
- Van Rooij, Maarten, Annamaria Lusardi, and Rob Alessie. 2011. "Financial Literacy and Stock Market Participation." *Journal of Financial Economics* 101 (2). Elsevier: 449–72. doi:10.1016/j.jfineco.2011.03.006.

Figure 1 Mark-to-Market LTV

This figure shows the kernel density of the mark-to-market LTV of financial industry borrowers and non-financial industry borrowers when loans are delinquent. The solid line is for the financial industry borrowers and the dashed line is for the non-financial industry borrowers.

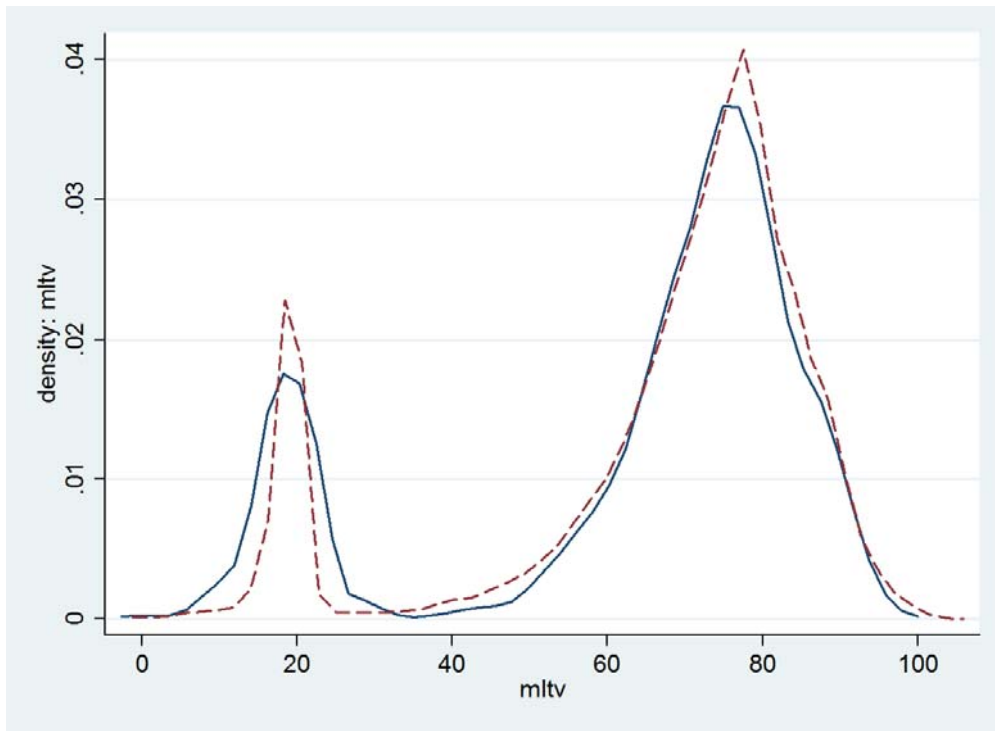


Table 1 Summary Statistics

This table presents the summary statistics of the full sample. We extract the performance information of loans and match this information to the origination data, creating a cross-sectional dataset. We restrict the value of the variables within reasonable boundaries as introduced in Section 2.3.

Variable	Obs	Mean	Std. Dev.	Min	Max
Delinquency dummy	136619	0.111	0.314	0	1
Literacy dummy	136619	0.036	0.186	0	1
Interest rate	136619	9.658	1.596	3.000	23.472
Term	136619	352.993	32.143	120	480
Loan amount	136619	192250.5	140080.7	6350	1400000
Combined LTV	136619	87.995	13.900	4.546	121.818
Debt-to-income ratio	136619	28.027	9.617	1.001	86.670
ARM	136619	0.600	0.490	0	1
First lien	136619	0.803	0.398	0	1
Prepayment penalty	136619	0.708	0.455	0	1
Prepayment penalty years	136619	1.606	1.147	0	5
Interest only	136619	1.155	0.362	1	2
Full documentation	136619	0.706	0.456	0	1
Brokered Loan	136619	0.879	0.326	0	1
Subprime	136619	0.989	0.105	0	1
FICO score	136619	621.786	58.498	400	824
Purpose: refinance=1	136619	0.587	0.492	0	1
Loan age	136619	7.343	8.525	1	109
Owner occupied	136619	0.954	0.210	0	1
Co-borrower school year	136619	12.870	3.403	0	39
Combined income	136619	7607.403	4057.861	537	29999
Male	136619	0.713	0.453	0	1
Minority	136619	0.277	0.448	0	1
Borrower age	136619	41.543	10.548	18	79
Married	136619	0.954	0.210	0	1
Borrower residence years	136619	5.956	6.264	0	99
Self employed	136619	0.189	0.392	0	1
Citizen	136619	0.960	0.195	0	1

Table 2 Comparison between Borrowers from the Financial and the Non-financial Industry

This table compares the mean of loan characteristics for financial industry and non-financial industry borrowers.

Variable	Non-financial industry	Financial industry
Delinquency dummy	0.112	0.097
Interest rate	9.661	9.582
Term	352.998	352.870
Loan amount	191443.8	214022.1
Combined LTV	87.941	89.452
Debt-to-income ratio	28.068	26.939
ARM	0.599	0.610
First lien	0.804	0.781
Prepayment penalty	0.709	0.684
Prepayment penalty years	1.609	1.526
Interest only	1.154	1.181
Full documentation	0.708	0.654
Brokered Loan	0.879	0.877
Subprime	0.989	0.991
FICO score	621.480	630.029
Purpose: refinance=1	0.588	0.544
Loan age	7.329	7.728
Owner occupied	0.955	0.920
Co-borrower school year	12.854	13.308
Combined income	7542.680	9354.283
Male	0.719	0.535
Minority	0.277	0.286
Borrower age	41.565	40.942
Married	0.954	0.941
Borrower residence years	5.971	5.546
Self employed	0.187	0.240
Citizen	0.960	0.974

Table 3 Loan Choice

This table presents the regression of prepayment penalty dummy on financial literacy and other information using the full sample. The dependent variable is an indicator variable that equals 1 if the loan has prepayment penalty and 0 otherwise, and we use logistic regression. We report odds ratio to measure the marginal effects. The controls variables comprise all controls in Table 4 except prepayment penalty years. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent Variable	prepayment penalty dummy
Literacy Dummy (Coefficient/odds ratio)	0.789*** (-3.42)
Control Variables	Yes
MSA fixed effects	Yes
Constant	Yes
Property type fixed effects	Yes
Origination year fixed effects	Yes
Servicer fixed effects	Yes
R-squared/Pseudo R-squared	0.422
Observations	133135

Table 4 Baseline Results

This table presents the logistic regression of the delinquency dummy on financial literacy and other information using the full sample. The dependent variable is an indicator variable that equals 1 if the borrower becomes delinquent at least once and 0 otherwise. We report odds ratio of each variable to measure the marginal effects. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent Variable:						
Delinquency dummy	(1)	(2)	(3)	(4)	(5)	(6)
Literacy dummy	0.870** (-2.27)	0.870** (-2.22)	0.874** (-2.17)	0.875** (-2.15)	0.859** (-2.52)	0.856** (-2.55)
FICO score	0.987*** (-56.23)	0.987*** (-52.32)	0.986*** (-57.79)	0.986*** (-57.88)	0.986*** (-56.91)	0.986*** (-56.02)
Purpose: refinance=1	0.660*** (-16.73)	0.727*** (-11.75)	0.871*** (-4.48)	0.870*** (-4.49)	0.901*** (-3.45)	0.908*** (-3.12)
Loan age	1.150*** (41.79)	1.165*** (36.48)	1.166*** (36.23)	1.166*** (36.27)	1.167*** (36.18)	1.168*** (36.02)
Owner occupied	0.828*** (-3.57)	0.933 (-1.39)	0.770*** (-5.24)	0.769*** (-5.26)	0.779*** (-4.75)	0.776*** (-4.83)
Interest rate		1.220*** (15.56)	1.203*** (12.86)	1.204*** (12.91)	1.205*** (12.56)	1.209*** (12.12)
Term		1.003*** (5.29)	1.002*** (4.21)	1.002*** (4.19)	1.002*** (3.96)	1.002*** (3.76)
ARM		1.074* (1.65)	1.004 (0.08)	1.003 (0.07)	0.986 (-0.30)	1.001 (0.03)
First lien		0.984 (-0.30)	1.264*** (4.13)	1.255*** (4.02)	1.319*** (4.38)	1.262*** (3.01)
Prepayment penalty		0.987 (-0.38)	0.97 (-0.83)	0.969 (-0.86)	0.971 (-0.84)	0.952 (-1.19)
Interest only		1.039 (1.09)	0.988 (-0.35)	0.988 (-0.37)	0.984 (-0.47)	0.988 (-0.39)
Full documentation		0.728*** (-10.90)	0.672*** (-14.27)	0.674*** (-14.20)	0.663*** (-13.86)	0.665*** (-13.32)
Brokered loans		0.836*** (-4.39)	0.825*** (-4.49)	0.826*** (-4.47)	0.818*** (-4.74)	0.828*** (-4.01)
Subprime		1.691 (1.57)	1.685 (1.56)	1.699 (1.59)	1.732 (1.64)	0.966 (-0.06)
Log loan amount			0.949* (-1.86)	0.954* (-1.69)	0.933 (-1.47)	0.956 (-0.82)
Combined LTV			1.018*** (18.56)	1.018*** (18.51)	1.016*** (15.76)	1.016*** (14.46)

Table 4 (continued)

	(1)	(2)	(3)	(4)	(5)	(6)
Debt-to-income ratio			1.014*** (8.98)	1.014*** (8.94)	1.014*** (6.38)	1.015*** (6.73)
Co-borrower school year				0.994** (-1.96)	0.993** (-2.42)	0.992*** (-2.60)
Log combined income					1.035 (0.63)	1.033 (0.59)
Male					0.957** (-1.97)	0.958* (-1.84)
Minority					1.176*** (4.58)	1.138*** (3.82)
Borrower age					0.992*** (-7.10)	0.992*** (-6.81)
Married					1.103* (1.85)	1.104* (1.83)
Borrower residence year					0.991*** (-4.31)	0.991*** (-4.16)
Self employed					0.974 (-0.94)	0.979 (-0.73)
Citizen					1.322*** (4.37)	1.325*** (4.43)
MSA fixed effects	No	No	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Property type fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Origination year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Servicer fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Pseudo R-squared	0.213	0.225	0.229	0.229	0.231	0.236
Observations	136619	136619	136619	136619	136619	135440

Table 5 Propensity Score Matching

This table presents the logistic regression of the delinquency dummy on financial literacy and other information using the sample constructed by propensity score matching. The dependent variable is an indicator variable that equals 1 if the borrower becomes delinquent at least once and 0 otherwise. We report odds ratio of each variable to measure the marginal effects. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable: Delinquency dummy	
Literacy dummy	0.850*** (-2.63)
Control Variables	Yes
MSA fixed effects	Yes
Constant	Yes
Property type fixed effects	Yes
Origination year fixed effects	Yes
Servicer fixed effects	Yes
Pseudo R-squared	0.244
Observations	28587

Table 6 Senior Staff, Loan-related Staff and Staff with Certificates

This table presents the logistic regression of the delinquency dummy on high financial literacy dummy and other information. The high literacy dummy equals 1 if the borrower is from financial industry and is senior staff, staff with professional certificates, or staff with jobs requiring high financial knowledge level, and 0 otherwise. We exclude *financial industry* borrowers whose high literacy dummy is 0. We report odds ratio of each variable to measure the marginal effects. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable: Delinquency dummy	
High literacy dummy	0.774** (-2.26)
Control Variables	Yes
MSA fixed effects	Yes
Constant	Yes
Property type fixed effects	Yes
Origination year fixed effects	Yes
Servicer fixed effects	Yes
Pseudo R-squared	0.237
Observations	131133

Table 7 Borrower Age

This table presents the logistic regression of the delinquency dummy on financial literacy and other information. We divide our sample into three subsamples and run the logistic regression: borrowers younger than 35 (Column (1)), borrowers between 35 and 50 (Column (2)), and borrowers older than 50 (Column (3)). We report odds ratio of each variable to measure the marginal effects. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable: Delinquency dummy			
	(1) age <35	(2) 35<=age<=50	(3) age>50
Literacy dummy	0.803** (-1.96)	0.932 (-0.83)	0.765* (-1.76)
Control Variables	Yes	Yes	Yes
MSA fixed effects	Yes	Yes	Yes
Constant	Yes	Yes	Yes
Property type fixed effects	Yes	Yes	Yes
Origination year fixed effects	Yes	Yes	Yes
Servicer fixed effects	Yes	Yes	Yes
Pseudo R-squared	0.257	0.24	0.232
Observations	38272	68575	27271

Table 8 Income Verification and Income Trends

This table presents the logistic regression of the delinquency dummy on financial literacy and other information. We focus on subsample comprising full-documentation loan (Column (1)) and the period from 2005 to 2007 (Column (2)) respectively. We report odds ratio of each variable to measure the marginal effects. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable: Delinquency dummy		
	(1) full documentation	(2) after 2005
Literacy dummy	0.849** (-2.03)	0.863** (-2.07)
Control Variables	Yes	Yes
MSA fixed effects	Yes	Yes
Constant	Yes	Yes
Property type fixed effects	Yes	Yes
Origination year fixed effects	Yes	Yes
Servicer fixed effects	Yes	Yes
Pseudo R-squared	0.239	0.235
Observations	53178	81625

Table 9 Compare Mark-to-Market LTV

This table compares the summary statistics of the mark-to-market LTV of financial industry borrowers and non-financial industry borrowers when loans are delinquent. Since the loans can be delinquent in multiple periods, the dataset is a panel dataset.

Subsample	Obs	Mean	Std. Dev.	Min	Max
Non-financial industry	2018	63.900	23.533	0	97.315
Financial Industry	63897	67.139	21.040	0	104.679

Table 10 Strategic Default Incentives

This table presents the generalized least squares regression of the mark-to-market LTV of all delinquent mortgages on financial literacy and other information. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable	Mark-to-Market LTV
	(1)
Literacy dummy	-0.032 (-0.07)
Control Variables	Yes
MSA fixed effects	Yes
Constant	Yes
Property type fixed effects	Yes
Origination year fixed effects	Yes
Servicer fixed effects	Yes
Pseudo R-squared	0.947
Observations	65915

Table 11 Falsification I

We randomly assign high financial literacy to observations and run regression in equation (1) and (3), using delinquency dummy as dependent variable. For simplicity we only report the odds ratio of the financial literacy dummy. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

	(1)	(2)	(3)	(4)	(5)
Delinquency dummy (Odds ratio)	0.965 (-0.59)	1.057 (0.97)	1.120** (2.36)	1.016 (0.35)	1.041 (0.78)
	(6)	(7)	(8)	(9)	(10)
Delinquency dummy (Odds ratio)	1.029 (0.46)	1.049 (0.89)	1.026 (0.57)	1.108* (1.85)	0.971 (-0.60)
	(11)	(12)	(13)	(14)	(15)
Delinquency dummy (Odds ratio)	0.969 (-0.66)	1.013 (0.30)	1.032 (0.66)	0.959 (-0.87)	0.956 (-0.77)
	(16)	(17)	(18)	(19)	(20)
Delinquency dummy (Odds ratio)	0.98 (-0.44)	1 (-0.01)	1.024 (0.52)	1.082 (1.64)	1.063 (1.17)
	(21)	(22)	(23)	(24)	(25)
Delinquency dummy (Odds ratio)	1.032 (0.63)	0.989 (-0.20)	1.01 (0.20)	0.984 (-0.28)	0.924 (-1.50)
	(26)	(27)	(28)	(29)	(30)
Delinquency dummy (Odds ratio)	1.05 (0.90)	0.998 (-0.03)	1.011 (0.19)	0.99 (-0.18)	1.003 (0.06)

Table 12 Falsification II

We assign high financial literacy to two professions: engineers, and servicemen and police officers. We report the odds ratio of the financial literacy dummy. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable: Delinquency Dummy		
	(1)	(2)
	Servicemen and Police Officers	Engineers
Literacy dummy	1.120 (1.08)	0.976 (-0.27)
Control Variables	Yes	Yes
MSA fixed effects	Yes	Yes
Constant	Yes	Yes
Property type fixed effects	Yes	Yes
Origination year fixed effects	Yes	Yes
Servicer fixed effects	Yes	Yes
R-squared	0.866	0.866
Observations	136619	136619

Table 13 Comparison between Borrowers from the Financial and the Non-Financial Industry: Continuous Measures

This table presents the summary statistics for the two delinquency measures for borrowers from the financial industry and the non-financial industry respectively. *delinq_ratio* is the fraction of time when the mortgage is in delinquency and *missed_ratio* is the number of missed mortgage payments divided by the number of total payments.

Variable	Industry	Obs	Mean	Std. Dev.	Min	Max
delinq_ratio	Non-financial industry	130591	0.030	0.106	0	1
	Financial Industry	4849	0.025	0.094	0	0.857
missed_ratio	Non-financial industry	130591	0.027	0.094	0	1
	Financial Industry	4849	0.023	0.086	0	0.857

Table 14 Alternative Delinquency Measures

This table presents the logistic regression of two alternative dependent variables on financial literacy and other information. We use the two other delinquency measures *delinq_ratio* and *missed_ratio* introduced in 6, and estimate the coefficients using generalized least-squares method. Robust standard errors are clustered at the MSA level. Robust z-statistics are reported in the parentheses. *, **, and *** indicate the estimates are significant at 10%, 5%, and 1% levels respectively.

Dependent variable	delinq_ratio	missed_ratio
	(1)	(2)
Mean Dependent variable	0.029	0.027
Literacy dummy	-0.0026** (-2.03)	-0.0021* (-1.68)
Control Variables	Yes	Yes
MSA fixed effects	Yes	Yes
Constant	Yes	Yes
Property type fixed effects	Yes	Yes
Origination year fixed effects	Yes	Yes
Servicer fixed effects	Yes	Yes
Pseudo R-squared	0.351	0.330
Observations	136619	136619