FEDERAL RESERVE BANK *of* ATLANTA

Financial Literacy and Subprime Mortgage
Delinquency: Evidence from a Survey
Matched to Administrative Data

Kristopher Gerardi, Lorenz Goette, and Stephan Meier

Working Paper 2010-10
April 2010

WORKING PAPER SERIES

Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data

Kristopher Gerardi, Lorenz Goette, and Stephan Meier

Working Paper 2010-10
April 2010

Abstract: The exact cause of the massive defaults and foreclosures in the U.S. subprime mortgage market is still unclear. This paper investigates whether a particular aspect of borrowers' financial literacy–their numerical ability–may have played a role. We measure several aspects of financial literacy and cognitive ability in a survey of subprime mortgage borrowers who took out mortgages in 2006 or 2007 and match these measures to objective data on mortgage characteristics and repayment performance. We find a large and statistically significant negative correlation between numerical ability and various measures of delinquency and default. Foreclosure starts are approximately two-thirds lower in the group with the highest measured level of numerical ability compared with the group with the lowest measured level. The result is robust to controlling for a broad set of sociodemographic variables and not driven by other aspects of cognitive ability or the characteristics of the mortgage contracts. Our results raise the possibility that limitations in certain aspects of financial literacy played an important role in the subprime mortgage crisis.

JEL classification: R2, D1, D8

Key words: subprime mortgage, delinquency, default, financial literacy, cognitive ability, survey

I. Introduction

The expansion of credit in the early-to-mid 2000s has had a profound impact on real estate and financial markets. In a very short time, it has broadened homeownership, particularly among individuals that had traditionally been shut out of credit markets. This growth occurred despite lagging income growth in these groups over the same time, and also in areas in which little house price growth could be expected (Mian and Sufi, 2009a). Moreover, on top of broadening homeownership rates, the expansion of credit has also led to a substantial increase in borrowing by previous homeowners taking advantage of quickly rising home prices over the same period of time: Evidence suggests that they borrowed as much as 30 cents for every dollar increase in the value of their homes (Mian and Sufi, 2009b).

Improved access to credit as such offers the benefits of consumption smoothing over time. The standard model in economics assumes that agents are rational and fully understand their environment. In such a model, making more credit available will unambiguously increase welfare, and so from that perspective the expansion of credit was viewed by many as welfare-enhancing. However, as house prices leveled off in 2006 and began to decline, this was accompanied by a massive increase in late payments on subprime mortgages, and an explosion of outright defaults (e.g. Foote et al., 2008b; Mayer et al., 2009). This led to a sharp drop in the value of mortgage-backed securities and to the worst financial and macroeconomic crisis since the Great Depression.

In light of these dramatic developments, a debate has started over how to explain the increase in late mortgage payments and defaults, which precipitated the broad economic and financial crisis. Several papers discuss the role of credit supply changes, and in particular the potential role of relaxed underwriting standards in generating an expansion of mortgage credit (e.g., Gerardi et al., 2009; Mian and Sufi, 2009a; Nadauld and Sherlund, 2009). By contrast, this paper examines the borrower side. Many market observers, including Akerlof and Shiller (2009), believe that departures from full rationality are an important factor in explaining the decline of the subprime mortgage market and the subsequent foreclosure

crisis. In part, "irrational exuberance" (Shiller, 2005) — the belief that house prices will just keep rising — may have played a role. But Akerlof and Shiller (2009) and others (e.g., Boeri and Guiso, 2007) argue that individuals' limited ability to make complicated financial decisions contributed importantly to the sharp rise in mortgage defaults.

Our paper is the first, to our knowledge, to directly examine this hypothesis. We combine micro data on subprime mortgage terms and their entire stream of payments with survey data from a telephone interview with a sample of subprime borrowers. A key feature of this survey is a set of questions measuring several aspects of financial literacy, notably numerical ability and economic literacy, in our respondents. Earlier studies have shown that these aspects of financial literacy are poor in large parts of the population (Lusardi and Mitchell, 2009; Banks and Oldfield, 2007; Lusardi and Tufano, 2008; McArdle et al., 2010) and that individuals make systematic mistakes, such as underestimating interest rates from payment streams (Stango and Zinman, 2008). Of even greater concern however, is that differences in financial literacy are correlated with consumption and savings decisions. A low ability to perform simple mathematical calculations, for example, is correlated with lower levels of saving (Banks and Oldfield, 2007; McArdle et al., 2010), less planning for retirement (Lusardi and Mitchell, 2009), poorer comprehension of credit, and the feeling that spending is out of control (Lusardi and Tufano, 2008).[1] Indirect evidence also points to the possibility that cognitive limitations play an important role in the choice of a mortgage instrument. Evidence from micro data on mortgages shows that individuals are confused about important contract terms of their mortgage (Bucks and Pence, 2008), and that individuals who were rated as confused by the interviewer were more likely to have adjustable-rate mortgages (Bergstresser and Beshears, 2009). Agarwal et al. (2010a) show that participants in a voluntary financial education program are less likely to fall behind on their mortgage

---

[1]Bernheim et al. (2001) present evidence that increases in mandatory schooling have positive effects on financial market participation. Mandatory increases in schooling on this scale have been shown not to have any effect on college graduation rates (Lochner and Moretti, 2004). The result is thus suggestive that basic education increases understanding of financial decision making. By contrast, Cole and Shastry (2007), using mandatory introduction of financial literacy education find little evidence that they have much effect on their measure of financial market participation.

payments, potentially indicating that increased financial literacy leads to lower delinquency rates (see also Agarwal et al. (2010b) for the effect of mandatory counseling on mortgage market outcomes). In addition, micro evidence on mortgage delinquency shows a pattern of financial struggle: Many consumers fall behind, then become current again, only for the cycle to repeat itself again and again. By contrast, fewer borrowers show an abrupt stoppage of payments that would be indicative of strategic default (Foote et al., 2009).

We explicitly test for the role of financial literacy and cognitive limitations in the rise of subprime mortgage delinquencies and defaults, and present robust evidence of a correlation between a specific aspect of financial literacy, numerical ability, and mortgage delinquency. We conducted a survey in the summer of 2008 on a sample of subprime borrowers in the states of Connecticut, Massachusetts, and Rhode Island who obtained mortgages in 2006 and 2007. We measured aspects of their financial literacy (numerical ability and economic literacy) and cognitive ability using methods that are standard in the literature (Banks and Oldfield, 2007; Dohmen et al., 2009; Lusardi and Mitchell, 2009). Our sample of subprime mortgage borrowers is taken from a dataset on privately securitized subprime mortgages that the Federal Reserve Bank of Boston purchased from First American LoanPerformance. This data set contains detailed information on mortgage terms and complete payment history streams as reported by the mortgage servicers, which allows us to track the sample over time and follow their subsequent mortgage outcomes. We find a large and statistically significant negative correlation between financial literacy and measures of mortgage delinquency and default, and the finding is robust to the inclusion of controls for income, education, risk aversion, and time preferences, thus ruling out a broad set of potential biases from omitted variables. The point estimates are remarkably robust, and quantitatively important: 20 percent of the borrowers in the bottom quartile of our financial literacy index have experienced foreclosure, compared to only 5 percent of those in the top quartile. Furthermore, borrowers in the bottom quartile of the index are behind on their mortgage payments 25 percent of the time, while those in the top quartile are behind approximately

10 percent of the time.

We include as control variables measures of other aspects of financial literacy and a general measure of cognitive ability, but find that the correlation is highly specific to one aspect of financial literacy: numerical ability. We devote the remainder of the paper to determining the underlying cause of this correlation. There are of course two possibilities. First, there could be a causal effect of numerical ability on mortgage repayment behavior. For example, individuals who cannot perform the simple mathematical calculations that are necessary to maintain a household budget or to calculate whether or not monthly mortgage payments are affordable over a long horizon may be more likely to miss mortgage payments over time. This interpretation is consonant with the picture that emerges from the survey evidence linking poor financial literacy to higher consumption, less saving, and out-of-control credit usage (Banks and Oldfield, 2007; Lusardi and Mitchell, 2009; Lusardi and Tufano, 2008). However, because this is a survey and we are not able to randomize through a controlled experiment there is always the possibility of omitted variable bias – a different factor may be causally responsible for the variation in mortgage repayment behavior in our data, and also correlated with our measure of financial literacy. We take this possibility very seriously in our empirical work, and our rich dataset allows us to narrow down the channels through which numerical ability could be associated with mortgage default. In addition to ruling out the influence of socio-demographic variables (including income), other cognitive skills, and preference parameters, we also find no evidence that the link between financial literacy and delinquency is mediated by differing contract terms or bigger mortgages for a given home value. Thus, individuals with poor numerical ability do not appear to make systematically different choices with respect to their mortgage terms. We also do not find that the link is related to differences in experience with mortgages. While we cannot rule out all other possible explanations, the robustness of our empirical estimates leads us to conclude that limited numerical ability played a non-trivial role in the subprime mortgage crisis. Our result also offers new ways to further test this relationship and points to the

importance of addressing limited financial literacy in the post-crisis reformation of mortgage markets.

The remainder of this paper is structured as follows: section II details the empirical setup and provides an overview of the data. Section III discusses our empirical methodology. Section IV presents the results. Section V discusses some caveats of our study, and concludes the paper.

II. Design of Study

In this section we provide a detailed discussion of our sample and survey design. First, we describe the pool of mortgage borrowers that we chose to draw our survey sample from, and discuss potential sample selection biases. Then, we discuss the survey procedure and the different parts of our survey in detail.

A. The Sample

In order to obtain objective measures of mortgage delinquency and default, we construct our survey sample from data that combines two micro-level mortgage datasets. The first is a loan-level dataset constructed and maintained by FirstAmerican LoanPerformance (LP). LP collects information on individual mortgages that are used as collateral for non-agency, mortgage-backed securities (MBS) and sold to investors on the secondary mortgage market. We use LP data that the Boston Fed purchased in mid-2007. This dataset covers Massachusetts, Connecticut, and Rhode Island from the late-1990s through March 2009. The LP dataset contains extensive loan-level information on mortgage characteristics, including interest rates (initial levels and changes over time), documentation levels, payment histories, loan-to-value ratios, and various other lending terms. It also contains some information regarding borrower characteristics, such as the borrower's credit score and debt-to-income ratio at origination (borrower's monthly debt payment divided by his or her monthly in-

come). Finally, the LP dataset identifies the type of MBS each loan was packaged into — subprime, Alt-A, or prime.[2]

The second dataset we use was supplied by The Warren Group, a private Boston firm that has been tracking real estate transactions in New England for more than a century. The Warren Group collects publicly available real estate transaction records that are filed at Registry of Deeds offices throughout New England, and have maintained an electronic database of these records for the past twenty years. The data that we use includes the universe of purchase-money mortgages, refinance mortgages, home equity loans, home equity lines of credit (only information on capacities and no information on utilization rates), and purchase deeds (including foreclosure deeds) transacted in Massachusetts, Connecticut, and Rhode Island. Unlike the LP data, this data contains the precise location of each property and the exact names of the buyers and sellers of each property as well as the names of the mortgage borrowers. These data allow us to construct a history of mortgage transactions for a household in a given property. In other words, with the Warren Data we are able to follow households in the same house across different mortgages. Since the data include information on all mortgage liens and the sale price for each property, we are able to construct a precise measure of the cumulative loan-to-value ratio at the time of purchase,[3] and to keep track of the total number of mortgages obtained by each homeowner.

We matched data from LP to data from the Warren Group, and only used the sample of first-lien mortgages contained in subprime MBS from the LP dataset that were originated in 2006 and 2007. The match is based on the zip code of the property (LP contains only the identity of the zip code where the property is located), the date of mortgage origination, the amount of the mortgage, whether the mortgage was for purchase or refinance, and the identity of the institution that originated the mortgage. The match rate was approximately

---

[2]The sample of prime loans in the LP dataset consists of mortgages with values above the GSE (Government Sponsored Enterprise) conforming loan limits. This segment of the prime market is often referred to as jumbo-prime.

[3]The LP data has only sporadic information on the presence of second liens, and thus does not allow for the construction of accurate cumulative loan-to-value ratios.

45 percent, and left us with a sample of more than 74,000 mortgages.[4]

We randomly selected mortgages from this matched dataset to construct our sample of borrowers for the survey. To contact borrowers we used two different strategies: 1) *Cold-calls* involved calling borrowers by phone. This was possible as we know each borrower's name and address from the Warren dataset, and used this information in an internet search engine (USAPeopleSearch.com) to find each borrower's phone number(s). 2) *Mail-ins* involved mailing invitations to participate in the survey to the addresses listed in the Warren data.

[Table 1 about here.]

Table 1 displays response rates for these two strategies. For the *Cold-call* strategy, we called a total of 3,523 borrowers[5] in the summer of 2008 (June - August). We were unable to reach a working phone line for approximately one-third (1,043) of these calls, while for a little more than one-third (1,366), we were able reach a working line, but unable to verify that the phone number corresponded to the borrower in the data.[6] Finally, we were able to positively identify the borrower in slightly less than one-third of the cases. In half of those cases (559) we were unable to speak to the actual borrower, and thus never received a response to our interview request.[7] In 296 cases we reached the borrower, but he or she refused to participate in the survey,[8] and in 259 cases we reached the borrower and he or she agreed to participate in the survey. Based on these statistics, we report two participation

---

[4]The main issue that contributed to the low match rate was the inconsistent definition of dates between the two datesets. The date listed in LP is the date of origination, while the date listed in the Warren data is the date that the mortgage document was recorded. It usually takes at least a few days for documents to be filed in the Registry of Deeds offices (sometimes a few weeks), and thus, these two dates do not match. Therefore, we were forced to use a date range in our matching algorithm, and consequently often found cases of multiple mortgages of the same amount, originated in the same zip code, in a given date range. We were forced to throw out these cases of multiple matches. The identity of the originating institution often helped us in these cases, but unfortunately the LP data contain only sparse information on this variable.

[5]We often found multiple possible phone numbers for each borrower in the data, so the actual number of phone numbers that we called was much larger than the number of borrowers.

[6]This included cases in which nobody picked up the phone and cases in which we reached an answering machine and left a message, but received no response (and could not identify the borrower from the answering machine message).

[7]In most of these cases we either left a message on an answering machine and never heard back, or spoke to another member of the household, but were not able to reach the actual borrower.

[8]We include cases in which the borrower agreed to participate at a later date, but never followed through on that agreement.

rates for the *Cold-calls* in the first column of Table 1. Of the borrowers that we actually spoke to directly, 46.6 percent agreed to participate in the survey, while 10.4 percent of the borrowers for whom we were able to verify a correct phone number agreed to participate.

We mailed almost 5,000 invitation letters to borrowers for whom we could not find phone contact information (*Mail-ins*). The invitation letter was one page (two-sided) and contained a brief description of the survey and the survey conductors. We also included a small response card that contained a question asking if the borrower would be interested in participating in the survey, and space for the borrowers who agreed to participate to list working phone numbers and times of the day that were best to contact them. We included a response envelope and postage. In the vast majority of cases (97.5 percent), we never received a response. When we did receive a response, we attempted to call the borrower to conduct the interview. Of the borrowers that we were able to reach (93), approximately 97 percent agreed to participate in the survey (74 percent of the borrowers for whom we could verify a correct address).

[Table 2 about here.]

Sample selection bias is always a serious concern in surveys such as this one. As we have information about observable mortgage and borrower characteristics for *all* of the borrowers we contacted, we can test whether there is sample selection on those observable characteristics. Table 2 contains detailed information on the presence of sample selection in observable mortgage and borrower characteristics. The table compares average characteristics between the respondents and non-respondents for both the *Cold-Calls* sample (Panel A) and the *Mail-In* sample (Panel B). There is no evidence of sample selection in the phone call sample. The difference in averages for all variables is never statistically significant at even the 10 percent level. Furthermore, there is very little evidence of sample selection in the mailing sample. The only difference that is statistically significant (at the 10 percent level) is the average mortgage size. Importantly, there is no difference in the probability of foreclosure after the mailing went out between respondents and non-respondents. We also

perform a more rigorous test of potential sample selection bias on observables. For each outcome measure k in Table 2, we estimate

$$y_i^k = \alpha_k + \gamma_k R_i + \beta_k CC_i + \epsilon_i^k \tag{1}$$

where α_k is the constant for outcome k, γ_k is the difference in the outcome if the individual was a respondent (and $R_i = 1$), and β_k is the difference in the outcome if individual i was a cold call (and $CC_i = 1$). Finally, ϵ_i^k is the residual for outcome k. We estimate the k equations in (1) by seemingly unrelated variables, thus allowing the residuals ϵ_i^k to be correlated across outcomes within individuals. We then test the hypothesis $\gamma_k = 0$ for all k outcome measures. The p-value of the corresponding χ^2-test is $p = 0.52$, and thus we find no evidence of selectivity into the survey on these 10 important variables.

While it does not appear that selection into the survey sample is an issue, the timing of the survey raises some important issues. The survey was conducted in the summer of 2008 between June and August, while the borrowers chosen for the survey obtained their mortgages in 2006 and 2007. August 2007 is the last month that a mortgage was originated in the survey sample, quite simply because the subprime mortgage market had completely shut down at that point and no new mortgages were originated. This means that the subprime borrowers taking the survey had been paying their respective mortgages for at least 10 months and up to 32 months (for mortgages originated in January 2006). In addition, one of the requirements that we imposed for inclusion into the sample was that each borrower not be in the foreclosure process at the time that the survey was conducted. Because of this design feature, the results in this study are not necessarily representative of all subprime mortgage borrowers. Many subprime borrowers defaulted on their loans and experienced foreclosure within the first year of origination. The average number of months to default for all subprime mortgages originated in 2006 and 2007 in the LP dataset for which the servicer has initiated foreclosure proceedings is slightly less than 18. More than one-quarter of the defaults occurred within one year of origination. As we will discuss in more detail below,

the effects of numerical ability on delinquency and default in this analysis are likely to be lower bounds, as the composition of mortgages in our sample potentially means that the least sophisticated subprime mortgage borrowers defaulted before we conducted the survey, and thus did not make it into our sample.

B. The Survey

The survey contained four important parts: 1) Measures of two aspects of individuals' financial literacy, numerical ability and basic economic literacy, and a measure of general cognitive ability. 2) Measures of time and risk preferences. 3) Questions about the details of the mortgage contract (we already know much of this information from the micro datasets) and the experience of shopping for the mortgage. 4) An extensive list of socio-demographic characteristics that complements information from the LP dataset.

On average, the survey took about 20 minutes to complete, and individuals were compensated $20 for their participation.

B.1. Financial Literacy: Numerical Ability and Cognitive Ability

The first measure of financial literacy, and the one that we primarily focus on in this study, determines the proficiency of a respondent in solving basic mathematical calculations. We asked participants five questions developed by Banks and Oldfield (2007). The questions are as follows:

1. *In a sale, a shop is selling all items at half price. Before the sale, a sofa costs $300. How much will it cost in the sale?*

2. *If the chance of getting a disease is 10 per cent, how many people out of 1,000 would be expected to get the disease?*

3. *A second hand 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?*

4. *If 5 people all have the winning numbers in the lottery and the prize is $2 million, how much will each of them get?*

5. *Let's say you have $200 in a savings account. The account earns ten per cent interest per year. How much will you have in the account at the end of two years?*

To construct an index of numerical ability, Banks and Oldfield (2007) suggest dividing individuals into four separate groups based on the responses to the five questions. A borrower is placed into the first group corresponding to the lowest level of numerical ability if he answers questions 1, 2, and 3 incorrectly *or* answers question 1 correctly, but gets questions 2, 3, and 4 incorrect. The second group is made up of borrowers who answer at least one of the first four questions incorrectly (the outcome of the fifth question is not considered for this group). The third group contains borrowers who answered questions 1, 2, 3, and 4 correctly, but answered question 5 incorrectly. Finally, borrowers who answered all five questions correctly are placed into the fourth group. Table 3 shows the distribution of the numerical ability index in our sample as well as the distribution from Banks and Oldfield. Approximately 16 percent of borrowers fall into the lowest group, 54 percent into the second group, 17 percent into the third group, and 13 percent into the highest group. Despite being characterized by a very different group of individuals, the distribution of the index in the Banks and Oldfield study is very similar the distribution of our sample.

[Table 3 about here.]

In order to distinguish numerical ability from general cognitive ability, we use a verbal fluency measure that was introduced by Lang et al. (2005). Participants are asked: *" In the next 90 seconds, name as many animals as you can think of. The time starts now."* The number of animals named has been shown to be highly correlated with IQ (e.g. Lang et al., 2005). The reason for this is that intelligence is highly correlated with the ability to retrieve known information. As most people know hundreds, if not thousands of animals, the question reveals how easy it is to retrieve information. Obviously, the ability to name

animals in English also depends on individuals' English language skills, which we elicit separately (see below). In the economics literature, Dohmen et al. (2009) also use this question to measure cognitive ability. Figure 1 compares the distribution of responses in our survey to their study, which used a representative sample of the German population. The shape of the distributions is very similar.

[Figure 1 about here.]

In addition to our measure of financial literacy that focuses on respondents' numerical ability, we measure respondents' basic understanding of economic mechanisms using two questions from Lusardi and Mitchell (2009). Lusardi and Mitchell (2009) refer to these as "basic financial literacy" questions, but in our opinion they measure an individual's understanding of basic economic concepts, and thus we refer to them as questions about "economic literacy."

1. *Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, how much would you be able to buy with the money in this account? More than today, exactly the same as today, or less than today?*

2. *Suppose that in the year 2020, your income has doubled and prices of all goods have doubled too. In 2020, how much will you be able to buy with your income? More than today, exactly the same as today, or less than today?*

In our sample, approximately 79 percent of borrowers answered the first question correctly, and 74 percent answered the second question correctly, while 60 percent answer both questions correctly. These results are very similar to those obtained by Lusardi and Mitchell (2009).

As a further measure of cognitive ability, we also include the mean time it took the participants to respond to the Banks and Oldfield (2007) questions. We measure the time from the moment the surveyor has finished reading the question to the moment the person

gives the answer. Table 5 displays the correlations between all the measures of cognitive ability. There is a strong positive correlation between all measures of cogntive ability. Thus, individuals who score well in one measure of cognitive ability also score better in the others. There is also a strong and negative correlation of every measure of cognitive ability with the response time in the numerical ability questions. Thus, individuals who responded more quickly to these questions also had a higher score in them, and in all the other measures of cognitive abilities.[9]

A factor analysis reveals one common factor between the five variables. Only the first eigenvalue is greater than one, while all others are almost exactly equal to zero. Finding one common component to different measures of intelligence is quite common, and found in many other studies (See, e.g., Flynn, 2007; Burks et al., 2009a).

B.2. Time and Risk Preferences

In order to measure time and risk preferences, we asked individuals in the survey to make a number of hypothetical choices that allowed us to calculate their discount factors and risk aversion parameters.

Similar to experimental measures of time preferences (see, e.g., Meier and Sprenger, 2010, 2009), individuals decide on an amount that makes them indifferent between receiving a certain monetary amount now versus waiting x months for a larger monetary amount. This procedure allows us to calculate an individual's discount factor. We asked individuals to make such intertemporal trade-offs either for now versus $x = 6$ months or $x = 12$ months. The two time frames also allow us to construct a measure of whether individuals have dynamically inconsistent time preferences (e.g., Laibson, 1997). In our sample, the average discount factor is 0.97 (over one month) and 81 percent of our sample exhibit dynamically

---

[9]A skeptic may argue that differences in our measure of cognitive abilities rather pick up styles in which individuals answer questions. Some may take the time to think about the question and then answer, while others may have just blurted out the first thing that came to their mind. The negative correlation between response times and the measures of cognitive ability also goes against this interpretation, as it shows that individuals who struggled to answer, also were more likely to get the answer wrong.

consistent time preferences, similar to Meier and Sprenger (2010). In addition, we also asked the borrowers to assess their impatience on a 11-point scale from 0 corresponding to "very impatient" to 10 corresponding to "very patient." In our empirical work we mainly use the measure of impatience that is based on the set of hypothetical choices, but the results do not change if we instead us the subjective scale.

Our measure of risk aversion also followed standard experimental strategies (e.g., Barsky et al., 1997). We asked participants to hypothetically choose between a certain payoff and a 50-50 chance of receiving a good or a bad payoff:

Which would you prefer: A mortgage for which you paid 1000 dollars per month for the next thirty years, or a mortgage, in which, after two years the payment is either $500 or $1100 with equal chance?

If the participant accepted the uncertain lottery, we raised the high mortgage payment of the uncertain mortgage by increments of $100. We use the payoff at which the participant switches to the safe mortgage as our measure of risk tolerance. The mean switching amount was $ 1184, revealing a substantial degree of risk aversion. In addition, we asked participants to assess their level of risk tolerance on a scale from 0 to 10 as in Dohmen et al. (2005). As with the self-assessed impatience measure, the second risk measure does not require any numerical skills. Nevertheless, in our empirical analysis we primarily use the risk measure based on the set of hypothetical choices (most related to experimental risk measures), although the results are robust to using the self-reported scale measure.

B.3. Mortgage Details and Mortgage Experience

The survey contains numerous questions about the characteristics of the mortgage contracts including questions about the size of the mortgage, initial interest rate, whether the mortgage is an adjustable-rate or fixed-rate instrument, whether the mortgage was for the purchase of the home or a refinance of a previous loan, and the existence of a prepayment penalty. These questions complement the mortgage information from the combined Loan-

Perfomance and Warren dataset, which includes the credit score of the borrower, the initial cumulative loan-to-value ratio, the extent of the documentation provided to the lender, and the monthly debt-to-income ratio of the borrower (including other forms of debt).

We obtain information on previous experience in mortgage markets from the Warren dataset. We are able to calculate the number of mortgages obtained since the house purchase (going back to January 1987), which allows us to calculate the number of mortgages taken out by each household before the current one.[10] We supplement this information with additional proxies for borrowers' experience with mortgages and their search behavior prior to obtaining the mortgage. The survey asks participants whether they were first-time homebuyers, whether they had taken a home buying class or had received counseling, if they obtained information about mortgage pricing before obtaining their loan, and if they had, how they obtained the information (internet, relative, friend, etc.).

B.4. Socio-demographics

Finally, the survey contains detailed questions about socio-demographic characteristics and information about household income and employment status. We asked participants about their race and ethnicity, gender, age, place of birth, amount of time spent in the United States, marital status, number of children, education level, and proficiency with the English language (scale from 0 corresponding to a "beginner" to 10 corresponding to a "native speaker"). We included questions on the amount of household income, the number of family members that contribute to household income, and the volatility of household income (on a three-point scale with 1 signifying that "it's been pretty stable"; 2 signifying "it has gone up and down a little over the last few years"; and 3 indicating that "it has gone up and down a lot over the last few years"). Finally, we asked participants about their current employment status and the number of times that they had been out of work over the previous five years.

Table 4 displays sample means and standard deviations of the survey variables for each of

---

[10]We only see mortgage information for the current property, and do not have any information on previous residences.

the four numerical ability groups. The patterns for each variable across the different groups accord well with intuition. Proficiency with the English language, the percent of households born in the United States, cognitive ability, and household income is monotonically increasing in the level of numerical ability as measured by the four groups. In contrast, income volatility, the percentage of black and hispanic households, the percentage of households with a high school diploma or less, the level of the initial contract interest rate, and the percentage of low documentation mortgages is monotonically decreasing in the level of numerical ability. The pattern for the two variables that summarize a household's experience in mortgage markets is very interesting. The percentage of first-time homebuyers is monotonically decreasing in the level of numerical ability, as 70 percent of households in the lowest numerical ability group are first-time homebuyers, while only 33 percent of households in the highest group are first-time homebuyers. This pattern suggests that the least experienced households with respect to buying a home are also the least financially sophisticated. But, the average number of previous mortgages obtained by a household is also monotonically decreasing in the level of numerical ability. This implies that, on average, households with the most experience with mortgages are the most numerically challenged. This seems to contradict the relationship between numerical ability and being a first-time homeowner, but these two observations can be reconciled by noting that a much lower percentage of mortgage originations were for the purchase of a home for the lower numerical ability groups as compared with the higher numerical ability groups. In addition, the lower numerical ability groups are characterized by a disproportionate number of households that have frequently refinanced. We conjecture that this is a result of a greater amount of cash-out refinancing activity on the part of households characterized by lower numerical ability.

There are a few other interesting patterns that are worth mentioning in table 4. Average credit scores (as measured by FICO) are much higher for the most numerically able group of borrowers as compared to the three lower groups, but there is very little difference in average credit scores across groups 1, 2, and 3. The correlation between numerical ability

and higher education appears puzzling in the table, but that reflects the manner in which we constructed the education groups. We constructed the groups to be mutually exclusive, so for example a household that obtained an undergraduate college degree as well as a professional degree would be given a value of 1 in the "Higher degree" group, but a value of 0 in the "College" group (even though that household also obtained an undergraduate college degree). This is the reason for the significant decrease in the "College" variable percentage when moving from the second most numerically able group of households to the most numerically able group. The highest group is characterized by a large percentage of households that obtained a graduate degree (about 40 percent), and by only a small percentage of households that obtained an undergraduate college degree without a graduate degree (25 percent).

[Table 4 about here.]

C. Measures of Mortgage Delinquency

We use three different measures of mortgage delinquency in the empirical analysis. First, we construct a variable that measures the fraction of time a borrower is behind on at least one of his mortgage payments. This measures the time during which a household is unable or unwilling to meet the promised mortgage payments. Imagine a household misses only one payment, and makes all the future payments in time. This first measure counts him as behind in each period until he makes that payment.

The second measure of mortgage delinquency is the fraction of mortgage payments missed. This variable is an explicit measure of the extent of delinquency. For example, a borrower who has had a mortgage for 12 months and who has missed 6 payments would be assigned a value of 50 percent for this measure, while a borrower who has had the mortgage for the same amount of time, but who has only missed 3 payments, would be assigned a value of 25 percent.

Our third measure is a dichotomous variable that takes a value of one if foreclosure

18

proceedings have been initiated by the lender. Normally, foreclosure proceedings are initi-
ated when a borrower is 120 days delinquent on his or her mortgage (or equivalently is 4
payments behind).[11]

[Table 6 about here.]

Table 6 contains information on the distributions of the three delinquency measures in
our sample. The average borrower in our sample is behind on this payments 20 percent of
the time, and has missed 11 percent of his or her mortgage payments. Half of the borrowers
in our sample are delinquent more than 7 percent of the time and have missed more than
5 percent of their mortgage payments, while 10 percent of the borrowers are delinquent
more than 60 percent of the time and have missed more than 30 percent of their payments.
Almost 20 percent of the borrowers in our sample have been in the state of foreclosure at
some point in their mortgage experience.

III. Empirical Specification

Our main empirical specification takes the following form:

$$D_i = \gamma NA_i + \mathbf{x_i}'\beta + \epsilon_i \tag{2}$$

where D_i corresponds to the first two measures of delinquency discussed above, the percent
of time spent in delinquency and the percent of mortgage payments missed, for household
i. The term NA_i represents the numerical ability group of household i, $\mathbf{x_i}$ represents a
vector of control variables, and ϵ_i is the residual. We estimate the equation by ordinary
least squares (OLS)[12], accounting for possible heteroskedasticity in the standard errors. For

---

[11]One of the participation criteria was not being in foreclosure at the time of the survey. But, there are a
few instances in which a borrower had been in foreclosure in the period before the survey was administered,
but then had recovered by the time of the survey. These borrowers were included in the survey sample.

[12]The results are robust to using tobit regressions instead of OLS (see Table A1 in the appendix).

our third measure of delinquency, the initiation of foreclosure proceedings, we estimate a probit model,

$$Pr[F_i = 1|NA_i, \mathbf{x_i}] = \Phi(\gamma NA_i + \mathbf{x_i}'\beta) \tag{3}$$

where F_i takes the value of one if foreclosure proceedings have been initiated on the borrower and zero otherwise, and $\Phi()$ is the cumulative distribution function of the standard normal distribution.

We focus on numerical ability as the independent variable of interest in an initial step, because the strongest evidence from the previous literature comes from studies linking numerical ability to savings (Banks and Oldfield, 2007). In a later step, we then include other measures of financial literacy that are not directly related to the ability to perform mathematical calculations, but more to the ability to understand financial concepts, as well as controls for general cognitive ability (Dohmen et al., 2009; Lusardi and Mitchell, 2009).

We include control variables $\mathbf{x_i}$ to avoid two types of omitted variable biases: there may be some other variable x that has a causal impact on defaults and on numerical ability, but due to the omission of x, that effect is captured by numerical ability. Secondly, it is also possible that numerical ability does not have a direct effect on delinquency, but rather causes x, which in turn affects delinquency and default propensities. In this case, omitting x would lead to the improper conclusion that numerical ability affects delinquency directly, while it really is a determinant of some other variable that in turn affects delinquency. We will not be able to distinguish between these two types of biases, but including controls helps us narrow down the admissible set of interpretations of our result.

In choosing the appropriate set of controls, we use guidance from the literature on models of mortgage default. Economic models of mortgage default emphasize the role of liquidity shocks and differences in household financial situations that make borrowers differentially vulnerable to those shocks (Gerardi et al., 2009; Pennington-Cross and Ho, 2010; Sherlund, 2008; Demyanyk and Van Hemert, 2010). If the shocks or exposure to the shocks are

20

correlated with, or caused by, numerical ability, then they may act as omitted variables and thus, would need to be taken into account. For example, numerical ability may affect the financial situation of the household in general, thus putting them at a disadvantage from the beginning of the mortgage. But in order to understand the contribution of numerical ability on the decisions made during the mortgage tenure, we need to control for initial differences in financial situations.

Models of mortgage default also predict that preference parameters, such as time preferences and risk preferences should be related to default (Foote et al., 2009). For example, more impatient individuals may be more likely to default on a mortgage, all else being equal. Furthermore, a recent literature suggests that these preference parameters may be correlated with certain aspects of cognitive abilities (Burks et al., 2009b; Benjamin et al., 2006; Dohmen et al., 2009). Thus, including measures of time preference and risk preference parameters allows us to examine the impact of numerical ability on delinquency separate from a potential correlation with those preference parameters.

The empirical mortgage literature also documents large differences across borrowers in the extent of delinquency and default that are due to differences in mortgage characteristics and house price movements. Gerardi et al. (2009), Foote et al. (2008b), Mayer et al. (2009) and Foote et al. (2009) show that there are large differences in default rates between fixed-rate and adjustable-rate mortgages, and that local house price movements are strongly correlated with default rates. However, from a theoretical view, the type of mortgage, or the location of a house are among the most important choices of any home purchase. Therefore, these differences should be regarded as differences in choices of the individuals. To the extent that these are correlated with numerical ability, differences in these choices may plausibly be a consequence of limitations in numerical ability. It is therefore not clear that differences in these choices should be treated as an omitted variable. We return to this issue below.

IV. Results

A. Baseline Findings

Figure 2 displays the relationship between the numerical ability index and our three measures of delinquency. There are three bar graphs corresponding to each of the three delinquency measures. There is a monotonically decreasing relationship between the percent of time delinquent and numerical ability (Panel A). Borrowers in the lowest numerical ability group on average spend almost 25 percent of the time in delinquency, while those in the highest group spend on average only 12 percent of the time in delinquency. In Panel B we also see a similar relationship between the percent of missed mortgage payments and numerical ability. The lowest group has missed almost 15 percent of mortgage payments on average, while the highest group has missed only 6 percent of payments on average. Foreclosure also appears to be negatively related to numerical ability (Panel C). While there is a small difference in the percentage of foreclosure between the first and second numerical ability group, the third group is characterized by a significantly lower percentage of foreclosures than the first two groups (15 percent versus more than 20 percent), while the fourth and highest group is characterized by a significantly lower percentage of foreclosures than the third group (7 percent versus 15 percent).

Table 7 displays the coefficient estimates from the linear regressions (columns (1) and (4)) and the estimated marginal effects from the probit model of foreclosure starts (column (7)). They indicate that, as suggested by the figure, the correlations between numerical ability and the delinquency measures are positive and statistically significant. As the figure already suggested, magnitudes are also quantitatively important: A borrower in the lowest category of the numerical ability index spends, on average, approximately 15 percent more time in delinquency than a borrower in the highest category. The differences in foreclosure rates across numerical ability groups are also very large. According to the estimates, the difference in foreclosure rates between the bottom quartile and top quartile is approximately

18 percentage points. To put this number in perspective, it is similar to the difference in foreclosure rates in the subprime mortgage sector when house prices were rapidly rising in 2005 (roughly 5 percent) compared to when they were rapidly falling in 2007 (roughly 20 percent, see Foote et al. (2009)).[13]

[Figure 2 and Table 7 about here.]

While quantitatively important, there are many possible interpretations of the estimated correlation between numerical ability and mortgage delinquency. There could be a causal effect of numerical ability on mortgage repayment behavior. The inability to perform simple mathematical calculations is likely to negatively impact a borrower's ability to manage a household budget. In addition, such an inability may adversely affect the borrower's ability to choose the appropriate type of mortgage given his or her current financial status and expected future financial situation. Both of these scenarios would likely put a borrower at risk of falling behind on his or her mortgage.

Alternatively, the correlation could be the result of omitted variables that are related to certain socio-economic characteristics of the borrowers. For example, poor numerical ability in the form of poor numerical ability could lead to less success in the labor market, which in turn could adversely impact income levels that could result in a higher incidence of mortgage delinquency and default. Lower numerical ability may also simply be correlated with lower levels of education, with other aspects of education other than numerical ability contributing to the higher delinquency levels.[14] An emerging literature also documents that preferences are correlated with cognitive ability (Burks et al., 2009a; Dohmen et al., 2009; Benjamin et al., 2006), so it is possible that differences in preferences could account for the correlation between numerical ability and mortgage delinquency. Finally, there could be

---

[13]The extent of delinquency and foreclosure is also estimated to be monotonically increasing in numerical ability when we employ a specification that includes a separate dichotomous variable for each numerical ability group (see Table A2 in the Appendix).

[14]For example, in Burks et al. (2009a), smoking and cognitive abilities are highly correlated in a univariate regression. But once one controls for education, cognitive abilities are no longer significant.

omitted variables related to the borrower's financial situation at the time of mortgage origination that might lead to delinquency at some later date and happen to be also correlated with our measure of numerical ability. For example, individuals with poor numerical ability may be more burdened with debt before they obtain their mortgage.

In the remainder of the paper, we use the richness of our survey dataset and mortgage datasets to try to distinguish between these alternative explanations in order to determine the channel through which numerical ability affects mortgage repayment behavior.

B. Socio-economic Characteristics, Preferences, and Household Financial Status

In columns (2), (5) and (8) of Table 7, we include the socio-economic variables and preference parameters that we collected in the survey to the regressions as control variables. They include variables for age, gender, ethnicity, education, the size of the household, time and risk preference parameters, labor market status over the previous five years, the household's income, and the subjective measure of income volatility. As can be seen in the first row of Table 7, the inclusion of these control variables does not significantly alter the point estimates or the standard errors associated with our numerical ability index. Numerical ability remains significantly correlated with mortgage delinquency, and the point estimates remain large and virtually unchanged. The control set, however, does contain important predictors of delinquency, as can be seen in the increase in the R^2 from roughly 2 to 14 percent. In particular, variables related to labor market success, such as income and income stability, as well as the number of times out of work over the previous five years, have a significant impact on delinquency.

In columns (3), (6) and (9), we also include controls for certain aspects of the household's financial situation at the time of origination. We include the FICO score, and dummy variables for whether the borrower is an investor (owner occupant as the reference group), as well as whether the mortgage is for a "purchase" ("refinance" as the reference group).

Again, the coefficient estimates are unaffected, and remain statistically significant in all specifications. The inclusion of these controls also significantly increases the R^2 of the regression from around 15 percent to approximately 25 percent. The FICO score, in particular, is an important determinant of delinquency and default. The fact that the correlation between numerical ability and delinquency does not change when we include the FICO score at origination is an important finding.[15] It implies that our measure of numerical ability is not just capturing the fact that borrowers who have defaulted on previous debts are more likely to default on their mortgage compared to borrowers with good credit histories.[16]

C. Different Measures of Financial Literacy and Cognitive Ability

The next step in our analysis is to pin down the particular aspect of financial literacy that affects mortgage repayment behavior. In addition to the education variable, we include as control variables an additional aspect of financial literacy and a measure of cognitive ability that is unrelated to financial literacy. Our measure of cognitive ability is a verbal IQ measure that is related to information processing, while our second measure of financial literacy includes two questions taken from Lusardi and Mitchell (2009) that are meant to measure basic economic literacy. We also include the response times to the numerical ability questions as an explanatory variable. Table 8 displays the results. In columns (1), (3) and (5), we include only the verbal IQ measure. The inclusion of the verbal IQ measure does not affect the magnitude or statistical significance of the estimated coefficient associated with financial literacy. The verbal IQ measure, conditional on our numerical ability measure, is not correlated with the first two measures of delinquency (percent of time behind, and percent of payments behind). However, it does enter significantly into the probit model for foreclosure. An increase of one standard deviation in the verbal IQ measure (8 points),

---

[15]Notice also that the inclusion of the FICO score renders most labor market controls that were significant in columns (2), (5), and (8), insignificant, with the exception of the volatility of income. Since the FICO score is constructed to be a catch-all predictor for delinquency, this is not entirely surprising.

[16]The estimated correlation between numerical ability and delinquency is not affected by the inclusion of debt-to-income ratios at origination, which capture other types of debt in addition to mortgage debt (see the discussion below and Table 9).

is associated with a 4.8 percentage point decrease in the foreclosure rate. An important difference between foreclosure and the other two delinquency measures is that foreclosure is initiated by the lender. One possible interpretation of this finding is that lenders may be less likely to foreclose on an intelligent person who is behind, and that this is picked up by our measure of IQ.

Columns (2), (4) and (6) display the results when we also include the measures of economic literacy and the response times in the set of control variables. They are not correlated with any of the three measures of delinquency, and do not affect the point estimate of our numerical ability measure. These findings lead us to conclude that the correlation between financial literacy and mortgage repayment behavior is specific to borrowers' numerical ability. The addition of both a verbal IQ measure, the response time and a different aspect of financial literacy, economic proficiency, does not explain differences in mortgage delinquency and default, and does not affect the the correlation between our numerical ability index and mortgage delinquency.

[Table 8 about here.]

D. Mortgage Terms and Prior Experience in Mortgage Markets

An important potential channel through which financial literacy could affect mortgage delinquency is in leading individuals to obtain mortgages with unfavorable terms, because they may be more likely to make mistakes in assessing the financial consequences of a particular contract. To examine this possibility, we add two sets of control variables to our basic specification. The first set controls for the contract terms of the mortgage, such as whether the mortgage has a fixed interest rate, and the initial interest rate of the mortgage. In a second step we also include choice variables with respect to the size of the mortgage. We include the loan-to-value ratio (LTV), and the debt-to-income ratio (DTI) to examine whether individuals with poor numerical ability take out loans that are significantly larger than those with higher literacy levels, and whether this causes their repayment problems.

Table 9 displays the results. Columns (1), (3) and (5) display the results when we control for differences in contract terms. The control variables do not add to the explanatory power of our baseline specification and, consequently, leave the point estimate of the impact of numerical ability and its standard error, essentially unchanged. In columns (2), (4) and (6), we also include LTV and DTI as control variables. The two variables have no impact on delinquency or foreclosure. Again, the inclusion of the variables does not affect the magnitude or statistical significance of the correlation between numerical ability and delinquency. LTV is statistically significant in the probit model of foreclosures. According to the estimates, a 10 percentage point increase in LTV at origination is associated with a 5.3 percentage point increase in the probability of foreclosure. There are likely two explanations for this finding. First, all else equal, a higher LTV at origination, implies a worse equity position at each future date, and thus a higher probability of foreclosure (see for example (Foote et al., 2008a)). Second, there is likely a selection effect, whereby borrowers that are more likely to default, perhaps because they have less wealth, choose to produce lower down payments at the time of purchase.

[Table 9 about here.]

As a next step, we ask whether borrowers with poor numerical ability are less experienced with mortgages, which may have an independent effect on delinquency.[17] We add as a control variable the number of previous mortgages obtained by the borrower (from the Warren dataset). In addition we include an indicator for first-time homebuyers, as well as a number of variables collected in the survey pertaining to the amount of information the individual collected before signing the mortgage contract. Table 10 displays the results. The correlation between numerical ability and delinquency is not affected. Experience *per se* does not seem to have a strong effect on delinquency. There is, however, some evidence that individuals who purchased a house for the first time are more likely to experience

---

[17]Agarwal et al. (2008) show that people are learning to avoid making mistakes in the credit card market.

foreclosure, though this effect is difficult to interpret.[18]

[Table 10 about here.]

E. Geographic Area and Mortgage Lenders

We explore two additional channels through which numerical ability could indirectly affect mortgage repayment behavior. The first is related to the decline in house prices. Declining house prices play an important role in explaining the rise in foreclosures during the recent housing crisis (e.g., Foote et al., 2008b,a; Gerardi et al., 2007). Individuals with poor numerical ability may have been less prudent in choosing the location of their property because they may not have fully understood the financial ramifications of declining house prices for refinancing. Thus, the correlation between numerical ability and delinquency may be modulated by a poor choice of location. We address this issue by including a full set of town/city fixed effects into our specifications. The results are displayed in columns (1), (4) and (7) of Table 11 for each of our measures of delinquency, respectively. The inclusion leads to a large increase in the R^2, confirming that regional variation is important in explaining variation in mortgage delinquency, as found in many other studies (Foote et al., 2008b,a; Gerardi et al., 2007). However, with 175 town fixed effects, the large increase reflects the fact that in many towns, we observe few borrowers. Yet, most importantly, the correlation between numerical ability and delinquency remains significant, and for all three measures, the point estimate increases. Thus, numerical ability does not appear to operate on delinquency through poor choice of location.[19]

[Table 11 about here.]

---

[18]It may also be that individuals who have purchased a house before have more assets, as they benefitted from increasing house prices. Since this is not the focus of the paper, we do not explore this topic in further detail.

[19]We also estimated a specification in which we included the cumulative amount of house price appreciation experienced between the time the mortgage was originated and the time the survey was conducted. This controls for some of the cross-sectional dispersion in house prices that had developed over the course of the financial crisis. The results are robust to such a specification.

We also explore the possibility that individuals with poor numerical ability may choose mortgage companies that provide poor support for mortgage borrowers. For example, lenders and (and servicers to the extent that they differ from the lender) may differ in how diligent they are about reminding borrowers when payments are due. Such differences may explain variation in delinquency. Thus, in the remaining columns of Table 11, we add originator (42) and servicer (27) fixed effects to our baseline specification. The additional controls increase the R^2, but again leave the coefficient estimate associated with numerical ability unchanged.

V. Interpretation and Conclusion

This paper investigates whether subprime borrowers with limited financial literacy are more likely to be delinquent on their mortgage and more likely to default. We conducted an extensive survey to measure subprime borrowers' financial literacy and cognitive ability in the summer of 2008 and matched the individual-level measures to micro-level datasets that contain extensive information on mortgage characteristics and payment histories. Our estimation results show a significant and quantitatively large association between one aspect of financial literacy, numerical ability, and mortgage delinquency. In addition, we find that foreclosure starts are two-thirds lower in the highest numerical ability group compared to the lowest group. The correlation is robust to several measures of delinquency and to the inclusion of a wide set of socio-economic and demographic control variables. The correlation appears to be specific to numerical ability and is not driven by general cognitive skills or economic literacy. Our results therefore show that limitations in numerical ability are common and that there is a strong and quantitatively important link to subprime mortgage defaults.

The results suggest that the correlation between mortgage delinquency and financial literacy is not due to financially illiterate borrowers taking on too much debt, or choosing

excessively risky mortgages. We are able to control for many details of the mortgage contracts, but find that the correlation is not sensitive to their inclusion in the econometric models. This suggests that limited numerical ability might lead to other mistakes over the course of time, like too much spending, too little savings, or inappropriate reaction to income and/or consumption shocks. Such an interpretation is consistent with results using the same measure of numerical ability for savings (Banks and Oldfield, 2007), and related measures (Stango and Zinman, 2008; Lusardi and Mitchell, 2009; Lusardi and Tufano, 2008). This result also suggests that subprime mortgage borrowers with limited numerical ability were no more likely than others to have been steered into unfavorable contract terms, although there is an important caveat: We surveyed individuals between 1 and 2 years after their mortgages had been originated, and many subprime mortgage defaults (about 60 percent, see, e.g. Foote et al., 2009) happen within two years of origination. Thus, our results do not completely rule out the possibility that limitations in financial literacy led to unfavorable mortgage terms or contracts that contribute to unfavorable mortgage outcomes.

We believe that our results also have several implications for future research and applications. First, the results show that a normally unobservable characteristic/ability can explain part of the heterogeneity in default behavior. This finding provides insights to lending firms on designing contract terms and default reduction strategies. Individuals who have difficulties dealing with numbers seem to be riskier, controlling for usual indicators like FICO scores. In order to better assess the risk of its customers, financial institutions may therefore have an interest in applying tests of numerical ability to screen loan applicants.

Second, one could ask whether the role of financial literacy was particularly important in this environment of rapidly falling house prices. Even among academic economists, the views sharply diverged. Many economists thought that the rise in house prices reflected fundamentals (e.g., Himmelberg et al., 2005), while others saw it as a giant bubble (e.g., Shiller, 2005). Thus, many individuals may have assumed that house prices would keep rising at the previous high rates into the foreseeable future, and thus, may have relied

more heavily on the accumulation of future equity in their homes to refinance into a larger mortgage, or to add a second lien to extract the equity. As the prospects for refinancing darkened, individuals with high financial literacy may have found it easier to adjust their consumption and savings decisions in order to continue making their mortgage payments. Exploring this hypothesis more fully requires data from an episode in which house prices were rising, and our data does not allow us to examine this in more detail.

Finally, our results suggest as a policy implication that more intensive financial education could substantially improve financial decisions later in life, and, in fact, have a profound impact on financial markets as suggested by the evidence in Agarwal et al. (2010a) and Bernheim and Garrett (2003). But it is important to remember that while our data show a strong and robust correlation that is highly specific and robust to a wide set of controls, it is not a setting in which financial literacy has been explicitly randomized in some way. The next logical, but ambitious step, is to randomize financial education and then track the financial decisions of these individuals over time.

References

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, Souphala Chomsisengphet, and Douglas Evanoff, "Learning to Cope: Voluntary Financial Education Programs and the Housing Crisis," *American Economic Review, Papers and Proceedings*, 2010, *100* (2), Forthcoming.

_ , _ , _ , _ , and _ , "The Effects of Financial Education on Household Financial Decision Making: Evidence from a Natural Experiment of Mortgage Advice," *Working Paper*, 2010.

_ , **John Driscoll, Xavier Gabaix, and David Laibson**, "Learning in the credit card market," *NBER Working paper*, 2008.

Akerlof, George and Robert Shiller, *Animal spirits: how human psychology drives the economy, and why it matters for global capitalism*, Princeton University Press, 2009.

Banks, J. and Z. Oldfield, "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving," *Fiscal Studies*, 2007, *28* (2).

Barsky, Robert B., F. Thomas Juster, Miles S. Kimball, and Matthew D. Shapiro, "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *The Quarterly Journal of Economics*, 1997, *112* (2), 537–579.

Benjamin, Daniel, Sebastian Brown, and Jesse Shapiro, "Who is 'Behavioral'? Cognitive Ability and Anomalous Preferences," *Working Paper*, 2006.

Bergstresser, Daniel and John Beshears, "Who selected Adjustable-Rate Mortgages? Evidence from Surveys of Consumer Finances," Working Paper, Harvard Business School 2009.

Bernheim, B. Douglas and Daniel M. Garrett, "The effects of financial education in the workplace: evidence from a survey of households," *Journal of Public Economics*, 2003, *87* (7-8), 1487–1519.

Bernheim, Douglas B., Daniel M. Garrett, and Dean M. Maki, "Education and saving: The long-term effect of high school financial curriculum mandates," *Journal of Public Economics*, 2001, *80* (3), 435–465.

Boeri, T. and L. Guiso, "Subprime crisis: Greenspan's Legacy," *VoxEU.org* 2007.

Bucks, Brian and Karen Pence, "Do borrowers know their mortgage terms?," *Journal of Urban Economics*, 2008, *64*, 218 – 233.

Burks, Stephen, Jeffrey Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive skills affect economic preferences, strategic behavior, and job attachment," *Proceedings of the National Academy of Science*, 2009, *106* (19), 7745–7750.

Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini, "Cognitive skills affect economic preferences, strategic behavior, and job attachment," *Proceedings of the National Academy of Sciences*, 2009, *106* (19), 7745–7751.

Cole, Shawn and Gauri Kartini Shastry, "Smart Money: The Effects of Education, Financial Literacy and Cognitive Ability on Financial Market Participation," *Working Paper*, 2007.

Demyanyk, Y. and O. Van Hemert, "Understanding the Subprime Mortgage Crisis," *The Review of Financial Studies*, 2010, p. Forthcoming.

Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde, "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 2009, p. forthcoming.

_ , _ , _ , _ , **Juergen Schupp, and Gert G. Wagner**, "Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey," *Journal of the European Economic Association*, 2005, p. Forthcoming.

Flynn, James R., *What Is Intelligence?: Beyond the Flynn Effect*, Cambridge and New York: Cambridge University Press, 2007.

Foote, Christopher, Kristopher Gerardi, Lorenz Goette, and Paul S. Willen, "Reducing Foreclosures: No Easy Answers," *NBER Macroeconomics Annual*, 2009, p. Forthcoming.

Foote, Christopher L., Kristopher Gerardi, and Paul S. Willen, "Negative Equity and Foreclosure: Theory and Evidence," *Journal of Urban Economics*, 2008, *64* (2), 234–245.

_ , _ , **Lorenz Goette, and Paul S. Willen**, "Just the Facts: An Initital Analysis of the Subprime Crisis," *Journal of Housing Economics*, 2008, *17* (4), 1–24.

Gerardi, Kristopher, Adam Hale Shapiro, and Paul S. Willen, "Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures," *Federal Reserve Bank of Boston Working Paper 07-15*, 2007.

___, **Adam Shapiro, and Paul Willen**, "Decomposing the Foreclosure Crisis: House Price Depreciation versus Bad Underwriting," *Federal Reserve Bank of Atlanta Working Paper 2009-25*, 2009.

Himmelberg, Charles, Christopher Mayer, and Todd Sinai, "Assessing high house prices: bubbles, fundamentals and misperceptions," *The Journal of Economic Perspectives*, 2005, *19* (4), 67–92.

Laibson, David, "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 1997, *112* (2), 443–477.

Lang, Frieder R., Dorothee Hahne, Stefanie Gymbel, Stefan Schroepper, and Katharina Lutsch, "Erfassung des kognitiven Leisungspotenzials und der "Big Five" mit Computer-Assisted-Personal-Interviewing (CAPI): Zur Reliabilitaet und Validitaet zweier ultrakurzer Tests und des BFI-S," *DIW Research Notes, Berlin* 2005.

Lochner, Lance and Enrico Moretti, "The effect of education on crime: Evidence from prison inmates, arrests, and self-reports," *The American Economic Review*, 2004, *94* (1), 155–189.

Lusardi, Annamaria and Oivia Mitchell, "How Ordinary Consumers Make Complex Economic Decisions: Financial Literacy and Retirement Readiness," *NBER Working Paper No. 15350*, 2009.

___ **and Peter Tufano**, "Debt Literacy, Financial Experience, and Overindebteness," Working Paper, Harvard University 2008.

Mayer, Christopher, Karen Pence, and S.M. Sherlund, "The rise in mortgage defaults," *Journal of Economic Perspectives*, 2009, *23* (1), 27–50.

McArdle, John J., James P. Smith, and Robert Willis, "Cognition and Economic Outcomes in the Health and Retirement Survey," Technical Report 15266, NBER Working Paper 2010.

Meier, Stephan and Charles Sprenger, "Stability of Time Preferences," *Working Paper*, 2009.

___ **and** ___ , "Present-Biased Preferences and Credit Card Borrowing," *American Economic Journal: Applied Economics*, 2010, *2* (1), 193–210.

Mian, Atif and Amir Sufi, "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis," *Quarterly Journal of Economics*, 2009, (124), Forthcoming.

__ **and** __ , "House Prices, Home Equity-Based Borrowing, and the U.S. Household Leverage Crisis," Working Paper, University of Chicago 2009.

Nadauld, Taylor D. and Shane M. Sherlund, "The Impact of Secondary Mortgage Market Activity on the Expansion of Subprime Credit," Working Paper, Brigham Young University 2009.

Pennington-Cross, Anthony and Giang Ho, "The Termination of Subprime Hybrid and Fixed Rate Mortgages ," *Real Estate Economics*, 2010, p. Forthcoming.

Sherlund, Shane, "The Past, Present, and Future of Subprime Mortgages," *Finance and Economics Discussion Series Federal Reserve Board 2008-63*, 2008.

Shiller, Robert, *Irrational Exuberance*, Princeton University Press, 2005.

Stango, Victor and Jonathan Zinman, "Exponential Growth Bias and Household Finance," *Journal of Finance, Forthcoming*, 2008.

Figure 1. Distribution of Verbal IQ Scores

Panel A: Distribution in this study
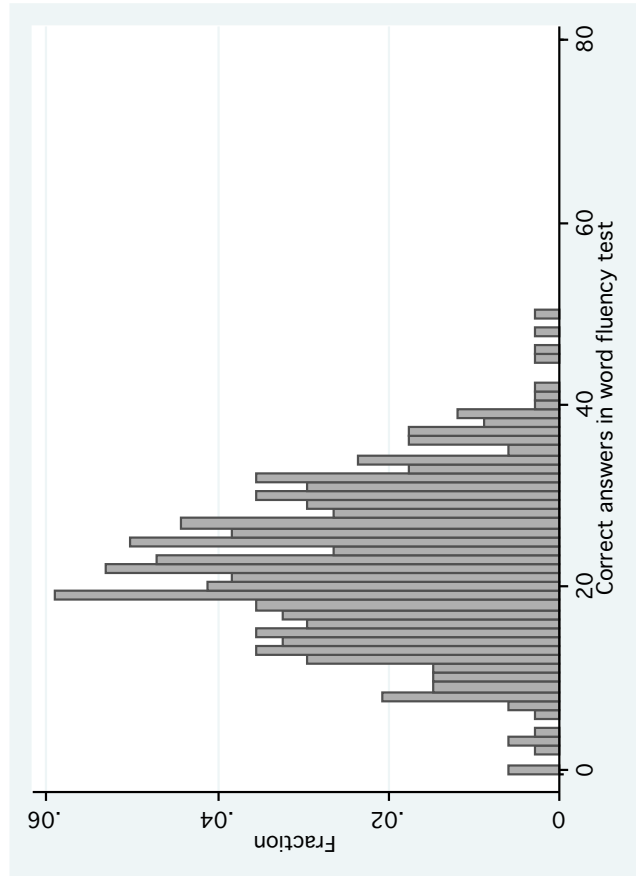
Panel B: Distribution in Dohmen et al. (2009)
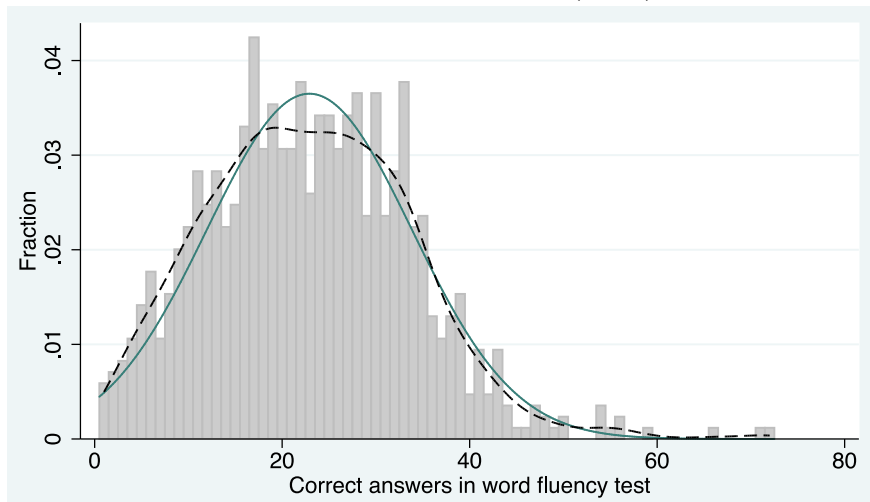
Figure 2. Delinquency and Numerical Ability Histograms

Panel A: Percent of Time Delinquent

Panel B: Percent of Payments Behind

Panel C: Frequency of Foreclosure Petitions
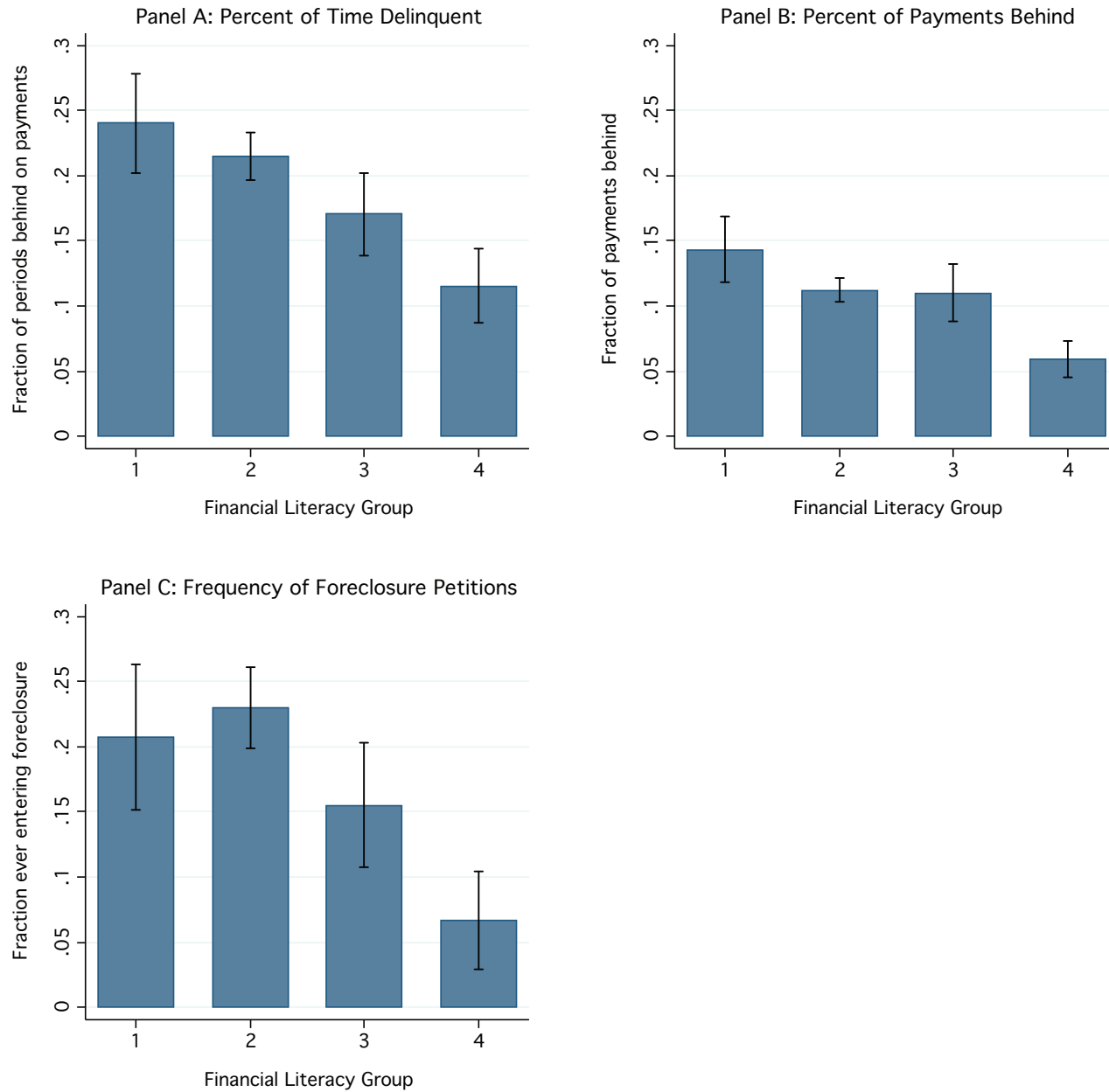
Table 1

Response Rate for Different Contact Strategies

	Cold-calls	Mail-Ins
Dead phone # or address	1,043	3
Unknown identity (not reached)	1,366	4,871
Known identity (not reached)	559	29
Refused to response	296	3
Responded	259	90
Total	3,523	4,996
Response Rate		
of working phone #s	10.4 %	–
of individuals answering phone	46.6 %	96.7 %

Notes: Dead means that none of the phone numbers were working or mail got "returned to sender". *Unknown* means that a phone was ringing, but the subject could not be identified (or no response to the mailing). *Known* means that the phone number did belong to the target subject, but could never be reached in person. *Refused* means that the subject was reached, but refused to participate. *Responded* means that the subject was reached and participated in the survey.

Table 2
Comparing Characteristics of Responders and Non-Responders

	# of Obs		Means		t-test of Differences	
	NR	R	NR	R	NR-R	p-value
Panel A: Cold-Calls						
FICO Score	2,346	242	632.3	638.7	-6.4	0.116
Fixed-Rate Mortage (=1)	2,346	242	0.410	0.397	0.014	0.678
Interest-Only (=1)	2,346	242	0.082	0.095	-0.013	0.479
Balloon Payment (=1)	2,346	242	0.203	0.227	-0.024	0.372
Refinance (=1)	2,346	242	0.529	0.492	0.037	0.275
Loan-to-Value Ratio	2,346	242	78.264	77.706	0.559	0.520
Amount of Mortgage	2,346	242	237,215	250,294	-13,079	0.124
Initial interest rate	2,346	242	8.003	7.938	0.065	0.408
Debt-to-Income Ratio	2,153	227	41.666	41.348	0.318	0.619
Full-Doc Status (=1)	2,346	242	0.725	0.723	0.002	0.949
Foreclosure after mailing went out (=1)	2,017	217	0.105	0.092	0.013	0.553
Panel B: Mail-Ins						
FICO Score	4,902	90	621.3	612.9	8.4	0.173
Fixed-Rate Mortage (=1)	4,902	90	0.161	0.178	-0.017	0.659
Interest-Only (=1)	4,902	90	0.079	0.056	0.023	0.421
Balloon Payment (=1)	4,902	90	0.303	0.344	-0.042	0.394
Refinance (=1)	4,902	90	0.781	0.778	0.003	0.943
Loan-to-Value Ratio	4,902	90	81.200	80.530	0.670	0.556
Amount of Mortgage	4,902	90	257,982	235,381	22,601	0.080
Initial interest rate	4,902	90	8.201	8.000	0.200	0.103
Debt-to-Income Ratio	4,537	86	42.226	43.756	-1.530	0.116
Full-Doc Status (=1)	4,902	90	0.653	0.600	0.053	0.294
Foreclosure after mailing went out (=1)	3,779	76	0.160	0.145	0.015	0.718

Notes: Table shows number of observations, mean of various mortgage and borrower characteristics of responders (R) and non-responders (NR). To test whether the difference of the various means is statistically significant between R and NR (NR-R), the table shows *p*-values of *t*-tests. The information about *Debt-to-Income Ratio* is missing for a few observations in the Warren data set. To compare *Foreclosure after mailing went out* we focus on individuals who were never in foreclosure between the origination and the date we contacted them. For some of the borrowers who were "current" on their mortgage when we contacted them (one criteria for being in the sample), a foreclosure petition had been filed before and they may have already been in the process of moving out.

Table 3
Distribution of Numerical Ability Index

	Numerical Ability Group			
	1	2	3	4
This study:	15.6%	53.9%	17.1%	13.3%
Banks and Oldfield (2007):	16.2%	46.6%	26.8%	11.1%

Table 4
Summary Statistics by Numerical Ability Group

	Numerical Ability Group					Numerical Ability Group			
	1	2	3	4		1	2	3	4
English Fluency	9.30	9.78	9.81	9.98	Risk tolerance	1157	1171	1202	1256
	(1.87)	(0.92)	(0.66)	(0.15)		(119)	(146)	(168)	(208)
Income Volatility	1.98	1.87	1.84	1.80	Discount Factor	0.97	0.96	0.96	0.97
	(0.77)	(0.81)	(0.77)	(0.79)		(0.03)	(0.03)	(0.03)	(0.03)
Born in U.S. (dv)	0.74	0.84	0.86	0.91	Present Bias (dv)	0.19	0.21	0.14	0.24
# Years spent in U.S.	43.5	43.1	41.1	45.2	Duration of mortgage	28.4	28.6	29.4	28.3
	(15.8)	(14.0)	(11.8)	(12.6)		(4.2)	(4.9)	(5.0)	(4.7)
Asian (dv)	0.02	0.02	0.00	0.00	# Months already in home	50.7	52.6	38.1	35.8
						(57.9)	(61.3)	(58.6)	(55.4)
Black (dv)	0.32	0.20	0.14	0.07	2007 cohort (dv)	0.19	0.19	0.16	0.20
Hispanic (dv)	0.17	0.06	0.05	0.04	Owner non-occupant (dv)	0.04	0.04	0.02	0.04
Native American (dv)	0.04	0.02	0.00	0.07	Cognitive ability score	17.5	21.5	26.9	27.1
						(7.9)	(8.2)	(8.6)	(7.6)
Other race (dv)	0.00	0.04	0.00	0.04	Fixed-rate loan (dv)	0.36	0.31	0.38	0.38
High school or less (dv)	0.49	0.29	0.10	0.04	Initial interest rate	8.2	8.0	7.9	7.8
						(1.1)	(1.1)	(1.1)	(1.2)
Some college (dv)	0.34	0.36	0.28	0.31	Low documentation (dv)	1.40	1.33	1.22	1.22
College (dv)	0.11	0.25	0.52	0.24	Cumulative LTV	0.86	0.83	0.85	0.85
						(0.17)	(0.16)	(0.15)	(0.17)
Higher degree (dv)	0.06	0.11	0.10	0.40	Back-end DTI ratio	43.1	42.5	39.9	41.3
						(6.8)	(8.1)	(9.3)	(8.6)
Employment status (dv)	0.75	0.86	0.84	0.87	# Previous mortgages	2.1	1.9	1.5	1.3
						(2.6)	(2.2)	(2.2)	(2.0)
Age of borrower	50	46	44	47	First-time Homebuyer (dv)	0.70	0.58	0.50	0.33
	(10)	(11)	(8)	(10)					
# Children	2.5	2.1	1.8	1.8	Home counseling (dv)	0.06	0.11	0.05	0.09
	(1.5)	(1.5)	(1.3)	(1.4)					
Credit score (FICO)	625	632	624	650	Shop around (dv)	0.47	0.56	0.76	0.71
	(49)	(62)	(59)	(72)					
Purchase mortgage (dv)	0.40	0.39	0.53	0.62	Income ($ thousands)	51.6	69.4	100.6	127.1
						(29.9)	(33.3)	(63.3)	(98.9)

Notes: Table shows means of variables for the four numerical ability groups. Standard deviations in parentheses.

Table 5
Correlation Between Measures of Cognitive Ability

	Numerical ability group	Verbal IQ measure	Savings scenario	Inflation scenario
Verbal IQ measure	0.356 (0.000)	1		
Savings scenario correct (DV)	0.236 (0.000)	0.153 (0.005)	1	
Inflation scenario correct (DV)	0.273 (0.000)	0.251 (0.000)	0.093 (0.087)	1
Reaction time in numerical ability questions	-0.279 (0.000)	-0.303 (0.000)	-0.157 (0.004)	-0.207 (0.000)

Notes: $N = 339$. p-values in parentheses. A factor analysis performed on these correlations reveals one common factor ($\lambda = 1.17$), while all other eigenvalues are less than 0.005.

Table 6
Distribution of Delinquency Measures

	Mean	Std. Dev.	Percentiles 10	25	50	75	90
Fraction of periods during which household is behind on at least one payment	0.198	0.247	0	0	0.077	0.367	0.621
Fraction of missed payments	0.110	0.143	0	0	0.056	0.167	0.304
Foreclosure	0.192	.	.	.	.	.	.

Notes: $N = 339$ observations.

Table 7: The Baseline Result

	Fraction of Time in Delinquency			Fraction of Payments Missed			Foreclosure Initiated (=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Numerical Ability Index	− 0.043***	− 0.039**	− 0.052***	− 0.024***	− 0.024**	− 0.031***	− 0.059**	− 0.065**	− 0.082***
	(0.014)	(0.020)	(0.019)	(0.009)	(0.011)	(0.010)	(0.026)	(0.030)	(0.028)
Fluency in English		0.001	0.005		0.005	0.007		0.013	0.014
		(0.016)	(0.014)		(0.008)	(0.008)		(0.028)	(0.026)
Born in USA (DV)		− 0.034	− 0.053		− 0.020	− 0.027		− 0.115	− 0.108
		(0.070)	(0.076)		(0.042)	(0.043)		(0.162)	(0.155)
Years lived in US		0.004	0.005		0.002	0.002		0.003	0.002
		(0.003)	(0.003)		(0.002)	(0.002)		(0.005)	(0.005)
Asian (DV)		− 0.158**	− 0.275***		− 0.103**	− 0.163***			
		(0.075)	(0.082)		(0.047)	(0.048)			
African American (DV)		0.109**	0.095**		0.070**	0.062**		0.153**	0.120*
		(0.043)	(0.039)		(0.029)	(0.026)		(0.071)	(0.067)
Hispanic (DV)		0.013	0.025		0.017	0.021		0.018	0.006
		(0.057)	(0.055)		(0.029)	(0.027)		(0.101)	(0.089)
Native American (DV)		− 0.054	− 0.032		− 0.032	− 0.024		0.026	0.010
		(0.110)	(0.102)		(0.053)	(0.040)		(0.179)	(0.181)
Other Ethnicity (DV)		0.111	0.101		0.084	0.079		0.166	0.172
		(0.131)	(0.099)		(0.089)	(0.076)		(0.232)	(0.235)
Age		− 0.003	− 0.003		− 0.001	− 0.001		0.001	0.002
		(0.003)	(0.003)		(0.001)	(0.002)		(0.005)	(0.005)
Some High School (DV)		− 0.112	− 0.078		− 0.047	− 0.029		− 0.029	0.012
		(0.096)	(0.100)		(0.055)	(0.055)		(0.155)	(0.167)
High School Degree (DV)		− 0.081	− 0.069		− 0.028	− 0.021		0.069	0.101
		(0.071)	(0.073)		(0.046)	(0.047)		(0.175)	(0.176)
Some College (DV)		− 0.032	− 0.007		− 0.002	0.012		0.204	0.243
		(0.070)	(0.073)		(0.045)	(0.046)		(0.177)	(0.175)
College Degree (DV)		0.019	0.039		0.017	0.029		0.211	0.249
		(0.073)	(0.075)		(0.046)	(0.046)		(0.192)	(0.193)
Professional Degree (DV)		0.014	0.047		0.009	0.029		0.292	0.384
		(0.075)	(0.076)		(0.047)	(0.047)		(0.231)	(0.244)
Number of Children		0.009	0.003		0.003	− 0.000		0.024	0.021
		(0.010)	(0.010)		(0.005)	(0.005)		(0.015)	(0.014)
Male (DV)		0.017	0.050		0.016	0.033*		0.116**	0.147***
		(0.032)	(0.031)		(0.018)	(0.017)		(0.048)	(0.045)
Estimated δ		− 0.665	− 0.261		− 0.455	− 0.257		− 1.116	− 0.556

Table 7: (continued)

	Fraction of Time in Delinquency			Fraction of Payments Missed			Foreclosure Initiated (=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		(0.732)	(0.701)		(0.424)	(0.403)		(1.117)	(1.003)
Present-Biased (DV)		0.023	0.010		0.015	0.010		-0.033	-0.051
		(0.037)	(0.036)		(0.020)	(0.019)		(0.061)	(0.060)
Risk preference		0.000	0.000		0.000	0.000		0.000	0.000
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
Log household income		-0.054**	-0.026		-0.023	-0.008		-0.071	-0.044
		(0.025)	(0.025)		(0.015)	(0.015)		(0.043)	(0.040)
Volatility of HH Income		0.040**	0.042**		0.025**	0.026***		0.056*	0.052**
		(0.018)	(0.016)		(0.010)	(0.009)		(0.029)	(0.026)
Number of times out of		0.017**	0.009		0.006	0.002		-0.020	-0.026*
work in last five years		(0.008)	(0.008)		(0.004)	(0.004)		(0.016)	(0.014)
Employed (DV)		-0.006	-0.006		0.001	0.002		0.034	0.044
		(0.046)	(0.044)		(0.026)	(0.025)		(0.056)	(0.045)
FICO Score / 10			-0.015***			-0.008***			-0.020***
			(0.002)			(0.001)			(0.004)
Home Purchase (DV)			-0.030			-0.006			0.058
			(0.035)			(0.019)			(0.061)
Months since home purchased			-0.000			-0.000*			-0.000
			(0.000)			(0.000)			(0.000)
Originated in 2007 (DV)			-0.050			-0.021			-0.044
			(0.034)			(0.019)			(0.043)
Investor (DV)			-0.012			-0.004			-0.007
			(0.060)			(0.034)			(0.099)
Constant	0.296***	0.838	1.351*	0.164***	0.467	0.782*			
	(0.037)	(0.762)	(0.703)	(0.023)	(0.428)	(0.402)			
R^2	0.023	0.148	0.260	0.022	0.138	0.245			
F-test of H_0: All coefficients are equal to zero	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
N	339	322	322	339	322	322	339	318	318

Notes: Robust standard errors in columns (1) - (6). ***, **, * indicate significance at the 1, 5, 10 percent level, respectively. Regression coefficients are reported in columns (1) - (6). Marginal effects from probit model are reported in columns (7) - (9).

Table 8
Controlling for General Cognitive Skills and Economic Literacy

	Fraction of Time in Delinquency		Fraction of Payments Missed		Foreclosure Initiated (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Numerical Ability Index	− 0.047**	− 0.051***	− 0.031***	− 0.033***	− 0.065**	− 0.061**
	(0.019)	(0.019)	(0.011)	(0.011)	(0.027)	(0.028)
Verbal IQ measure	− 0.001	− 0.002	0.000	0.000	− 0.006**	− 0.006**
	(0.002)	(0.002)	(0.001)	(0.001)	(0.003)	(0.003)
Savings Scenario correct (DV)		0.002		− 0.000		− 0.051
		(0.036)		(0.021)		(0.058)
Inflation scenario correct (DV)		0.006		0.011		− 0.016
		(0.033)		(0.018)		(0.047)
Reaction time in NA questions		− 0.003		− 0.000		− 0.001
		(0.002)		(0.001)		(0.003)
Control variables?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.262	0.268	0.242	0.244		
F-test of H_0: All coefficients are equal to zero.	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
N	322	322	322	322	318	318

Notes: Regression coefficients are reported in columns (1) - (4). Marginal effects from probit models are reported in columns (5) - (6). Robust standard errors in parentheses in columns (1) - (4). All specifications contain the full set of control variables as in Table 7.

Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 9
Controlling for Mortgage Attributes

	Fraction of Time in Delinquency		Fraction of Payments Missed		Foreclosure Initiated (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)
Numerical Ability Index	-0.048**	-0.039*	-0.030***	-0.029**	-0.078***	-0.064**
	(0.019)	(0.020)	(0.011)	(0.012)	(0.027)	(0.026)
Fixed-Rate Mortgage	0.027	0.035	0.010	0.014	0.023	0.038
	(0.028)	(0.029)	(0.016)	(0.017)	(0.043)	(0.043)
Initial Interest Rate	0.022	0.025	0.007	0.008	0.021	0.009
	(0.016)	(0.017)	(0.009)	(0.009)	(0.021)	(0.020)
Low-Doc Loan (DV)	0.027	0.011	0.009	-0.001	0.014	0.003
	(0.032)	(0.034)	(0.018)	(0.019)	(0.045)	(0.044)
Loan-to-Value Ratio		0.095		0.084		0.537***
		(0.096)		(0.054)		(0.162)
Debt-to-Income Ratio		0.002		0.001		0.003
		(0.002)		(0.001)		(0.002)
Control variables?	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.272	0.278	0.249	0.259		
F-test of H_0: All coefficients are equal to zero.	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
N	321	293	321	293	317	291

Notes: Regression coefficients are reported in columns (1) - (4). Marginal effects from probit models are reported in columns (5) - (6). Robust standard errors in parentheses in columns (1) - (4). All specifications contain the full set of control variables as in Table 7.
Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 10
Controlling for Previous Homeownership Experience

	Fraction of Time in Delinquency	Fraction of Payments Missed	Foreclosure Initiated (=1)
Numerical Ability Index	– 0.049***	– 0.030***	– 0.075***
	(0.019)	(0.010)	(0.028)
Number of prev. mortgages	0.001	0.005	0.019
	(0.009)	(0.005)	(0.014)
First home purchase (DV)	0.046	0.023	0.089**
	(0.031)	(0.017)	(0.043)
Shopped around before getting mortgage (DV)	0.029	0.017	0.017
	(0.029)	(0.016)	(0.040)
Sought counseling for home buyers (DV)	– 0.026	– 0.003	– 0.068
	(0.050)	(0.029)	(0.053)
Attended home owner classes (DV)	– 0.009	– 0.010	0.127
	(0.047)	(0.024)	(0.102)
Control variables?	Yes	Yes	Yes
R^2	0.270	0.255	
F-test of H_0: All coefficients are equal to zero.	$p < 0.01$	$p < 0.01$	$p < 0.01$
N	322	322	318

Notes: Regression coefficients are reported in columns (1) and (2). Marginal effects from probit models are reported in column (3). Robust standard errors in parentheses in columns (1) and (2). All specifications contain the full set of control variables as in Table 7.
Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table 11
Including Town, Servicer, and Originator Fixed Effects

	Fraction of Time in Delinquency			Fraction of Payments Missed			Foreclosure Initiated (=1)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Numerical Ability Index	− 0.081**	− 0.055***	− 0.044***	− 0.045**	− 0.037***	− 0.026***	− 0.105*	− 0.106***	− 0.065**
	(0.033)	(0.020)	(0.015)	(0.022)	(0.011)	(0.009)	(0.056)	(0.032)	(0.025)
Town Fixed Effects?	Yes	No	No	Yes	No	No	Yes	No	No
Originator Effects?	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Servicer Effects?	No	No	Yes	No	No	Yes	No	No	Yes
Control Variables?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.735	0.361	0.350	0.690	0.358	0.337	0.668	0.318	0.298
F-test of H_0: All coefficients are equal to zero	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$	$p < 0.01$
N	319	307	293	319	307	293	319	307	293

Notes: Regression coefficients are reported in columns (1) - (6). Marginal effects from probit models are reported in columns (7) - (9). Robust standard errors in parentheses in columns (1) - (6). All specifications contain the full set of control variables as in Table 7.
Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

A. Appendix

A. Appendix Tables and Figures

Table A1: The Baseline Result in Tobit Models

	Fraction of Time in Delinquency		Fraction of Payments Missed	
	(1)	(2)	(3)	(4)
Numerical Ability Index	− 0.063***	− 0.067***	− 0.035***	− 0.041***
	(0.023)	(0.026)	(0.013)	(0.015)
Fluency in English		0.004		0.007
		(0.021)		(0.012)
Born in USA (DV)		− 0.042		− 0.018
		(0.115)		(0.067)
Years lived in US		0.007		0.004
		(0.005)		(0.003)
Asian (DV)		− 0.556**		− 0.334**
		(0.227)		(0.133)
African American (DV)		0.140***		0.090***
		(0.049)		(0.028)
Hispanic (DV)		0.050		0.036
		(0.078)		(0.045)
Native American (DV)		− 0.046		− 0.033
		(0.128)		(0.074)
Other Ethnicity (DV)		0.103		0.088
		(0.141)		(0.081)
Age		− 0.006		− 0.003
		(0.005)		(0.003)
Some High School (DV)		− 0.154		− 0.074
		(0.132)		(0.076)
High School Degree (DV)		− 0.138		− 0.060
		(0.111)		(0.065)
Some College (DV)		− 0.055		− 0.016
		(0.107)		(0.062)
College Degree (DV)		0.028		0.023
		(0.110)		(0.064)
Professional Degree (DV)		0.030		0.017
		(0.116)		(0.067)
Number of Children		0.014		0.007
		(0.014)		(0.008)
Male (DV)		0.070*		0.046*
		(0.041)		(0.024)

	Fraction of Time in Delinquency		Fraction of Payments Missed	
	(1)	(2)	(3)	(4)
Estimated δ		-0.350		-0.346
		(0.951)		(0.551)
Present-Biased (DV)		0.036		0.027
		(0.050)		(0.029)
Risk preference		0.000		0.000
		(0.000)		(0.000)
Log household income		-0.067^*		-0.030
		(0.038)		(0.022)
Volatility of HH Income		0.059^{**}		0.036^{**}
		(0.024)		(0.014)
# of times out of work		0.016		0.007
in last five years		(0.012)		(0.007)
Employed (DV)		0.018		0.017
		(0.054)		(0.031)
FICO Score / 10		-0.023^{***}		-0.013^{***}
		(0.003)		(0.002)
Home Purchase (DV)		-0.067		-0.025
		(0.054)		(0.031)
Months since home purchased		-0.001^*		-0.001^{**}
		(0.000)		(0.000)
Originated in 2007 (DV)		-0.063		-0.027
		(0.048)		(0.028)
Investor (DV)		0.014		0.012
		(0.098)		(0.056)
Constant	0.249^{***}	2.022^{**}	0.137^{***}	1.186^{**}
	(0.056)	(0.987)	(0.032)	(0.572)
σ	0.350^{***}	0.294^{***}	0.201^{***}	0.170^{***}
	(0.018)	(0.016)	(0.011)	(0.009)
N	339	322	339	322

Notes: Coefficients of tobit models. Robust standard errors in columns (1) - (4). σ is the estimated standard deviation of the residual.

Level of significance: $^*p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$

2

<div style="text-align:center">

Table A2

Dummy Variables of NA Categories Instead of Linear Term

</div>

	Fraction of Time in Delinquency		Fraction of Payments Missed		Foreclosure Initiated (=1)	
	(1)	(2)	(3)	(4)	(5)	(6)
NA Index = 2 (DV)	− 0.025	− 0.037	− 0.031	− 0.040	0.020	− 0.016
	(0.042)	(0.042)	(0.027)	(0.027)	(0.058)	(0.054)
NA Index = 3 (DV)	− 0.070	− 0.108**	− 0.033	− 0.056	− 0.050	− 0.084*
	(0.050)	(0.053)	(0.033)	(0.035)	(0.065)	(0.048)
NA Index = 4 (DV)	− 0.125***	− 0.142**	− 0.084***	− 0.104***	− 0.142***	− 0.145***
	(0.048)	(0.060)	(0.029)	(0.034)	(0.052)	(0.032)
Control variables?	Yes	Yes	Yes	Yes	Yes	Yes
F-Test: all coefficients of NA are zero	$p = 0.01$	$p = 0.04$	$p < 0.01$	$p = 0.01$	$p = 0.07$	$p = 0.02$
F-Test: Relationship is linear	$p = 0.86$	$p = 0.81$	$p = 0.6$	$p = 0.7$	$p = 0.28$	$p = 0.31$
R^2	0.024	0.261	0.026	0.247		
N	339	322	339	322	339	318

Notes: Regression coefficients are reported in columns (1) - (4). Marginal effects from probit models are reported in columns (5) - (6). Robust standard errors in parentheses in columns (1) - (4). All specifications contain the full set of control variables as in Table 7.

Level of significance: *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$