HOUSEHOLD DEBT AND DEFAULTS FROM 2000 TO 2010:
FACTS FROM CREDIT BUREAU DATA

Atif Mian
Amir Sufi

ABSTRACT

We use individual level credit bureau data to document which individuals saw the biggest rise in household debt from 2000 to 2007 and the biggest rise in defaults from 2007 to 2010. Growth in household debt from 2000 to 2007 was substantially larger for individuals with the lowest initial credit scores. However, initial debt levels were lower for individuals in the lowest 20% of the initial credit score distribution. As a result, the contribution to the total dollar rise in household debt was strongest among individuals in the 20th to 60th percentile of the initial credit score distribution. Consistent with the importance of home-equity based borrowing, the increase in debt is especially large among individuals in the lowest 60% of the credit score distribution living in high house price growth zip codes. In contrast, the borrowing of individuals in the top 20% of the credit score distribution is completely unresponsive to higher house price growth. In terms of defaults, the evidence is unambiguous: both default rates and the share of total delinquent debt is largest among individuals with low initial credit scores. The bottom 40% of the credit score distribution is responsible for 73% of the total amount of delinquent debt in 2007, and 68% of the total in 2008. Individuals in the top 40% of the initial credit score distribution never make up more than 15% of total delinquencies, even in 2009 at the height of the default crisis.

Atif Mian
Princeton University
Bendheim Center For Finance
26 Prospect Avenue
Princeton, NJ 08540
and NBER
atif@princeton.edu

Amir Sufi
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
amir.sufi@chicagobooth.edu

1 Introduction

From 2000 to 2007, the United States experienced the most dramatic boom and bust in household debt since the Great Depression. Household debt increased at a steady pace through the 1990s, and then jumped by $7 trillion from 2000 to 2007. The boom in debt ended badly: by 2009 the delinquency rate on debt had reached above 10%, much higher than seen since the Great Depression. Figure 1 shows these patterns. Our previous research has argued that the boom and bust in household debt is crucial for understanding the Great Recession of 2007 to 2009 (Mian and Sufi (2010), Mian, Rao, and Sufi (2013), Mian and Sufi (2014)).

Our goal in this study is largely descriptive. We want to answer the following question: which individuals drove the aggregate patterns in debt and delinquencies from 2000 to 2010? To answer this question, we utilize individual level credit bureau data that tracks a random sample of 300 thousand individuals from 1997 to 2010. We sort individuals into quintiles based on their credit score at the beginning of the sample period, and then we track debt and defaults for each quintile through the sample period.

Using individual level credit bureau data has a number of advantages. First, credit bureau data is accurate and aggregates to total household debt from the U.S. Flow of Funds. It therefore provides the cleanest picture available to us to understand the dynamics of who drove the increase in debt and defaults from 2000 to 2010. Second, individual level credit bureau data consolidates all debts for an individual. Mortgage origination-based data sets such as LPS/McDash often miss second mortgages or home equity-lines of credit, and completely miss auto and credit card debt. For example, a debt to income ratio using an origination-based data set is only the debt to income ratio of the mortgage being originated, which often excludes a substantial amount of debt for the individual. Third, using data at the individual level data avoids concerns about changing composition of individuals within a geographic unit such as a zip code or county.

We find that the *growth* in household debt was strongest for the lowest credit score individuals in the sample. Individuals in the lowest 20% of the 1997 credit score distribution see their debt grow by 175% from 2000 to 2007. There is a monotonic decline in debt growth as credit scores increase, with individuals in the highest 20% of the 1997 credit score distribution seeing debt grow by only 40% from 2000 to 2007. This pattern is not driven by differences in age or initial debt levels:

even among individuals of similar age and similar initial debt levels, low credit score individuals see the largest growth in household debt.

The contribution of any group to the total dollar increase in debt from 2000 to 2007 is a function of both debt growth and initial debt levels of the group. Individuals with low credit scores saw the highest debt growth during the boom, but they also started with the lowest initial debt levels. As a result, individuals in the 20th to 60th percentile of the initial credit score distribution contributed most to the total dollar rise in household debt, having both high initial debt levels and relatively strong growth in debt. More specifically, the 40% of households in the 20th to 60th percentile of the initial credit score distribution accounts for just over 50% of the total dollar rise in household debt from 2000 to 2007. The highest 20% of the initial credit score distribution accounts for only 10% of the total rise.

Recent quantitative macroeconomic models have focused on the importance of loosened borrowing and lending constraints in explaining housing market dynamics during the 2000s (Midrigan and Philippon (2011), Favilukis, Ludvigson, and Van Nieuwerburgh (2013), Justiniano, Primiceri, and Tambalotti (2014)). Debt to income and debt to value ratios are important empirical measures in this literature. We match individuals to average income and home values of the zip codes they live in to examine debt to income and debt to value ratios. We find that debt to income ratios skyrocket by 0.8 during the 2000 to 2006 period among individuals in the bottom 60% of the credit score distribution. For the lowest 20% of credit score distribution, this represents a *doubling* of the debt to income ratio. The top 20% of the credit score distribution saw almost no increase in debt to income ratios from 2000 to 2006.

The evidence is more mixed using housing debt to home value ratios. Housing debt includes both mortgage and home equity-related debt, and we examine the ratio of housing debt to home value conditional on individuals that have some housing debt outstanding. From 2002 to 2005, the debt to value ratio actually *declines* for all but the bottom 20% of the credit score distribution, where it is constant. This reflects the fact that aggressive borrowing by households happened concurrently with strong house price growth. From 2005 to 2007, there is a sharp rise in debt to value ratios, as household borrowing continues to increase while house price growth stagnates.

Given the importance of home-equity driven borrowing during this period, we do a double sort of individuals by credit score and by house price growth of the zip code in which the individual

resides. All of the patterns on growth and the share of debt across the credit score distribution are amplified in zip codes with strong house price growth. Debt growth is strongest among low credit score individuals living in high house price growth zip codes. The contribution to the total dollar rise in household debt is largest for individuals in the 20th to 60th percentile of the credit score distribution living in high house price growth zip codes.

In terms of delinquencies during the bust, the evidence is unambiguous: the mortgage default crisis was driven by low credit score individuals. Default rates for the lowest 20% of the initial credit score distribution reached a stunning 25% by 2009. Default rates for the top 40% of the initial credit score distribution were below 5%. The share of total dollar amounts in delinquency was also largest for the lowest part of the credit score distribution. In 2007, the lowest 20% of the initial credit score distribution accounted for 40% of the dollars in delinquency. The bottom 40% of the initial credit score distribution accounted for almost 70% of the total dollars in delinquency in 2008. Even at the peak of the mortgage default crisis when defaults had spread throughout the country, the top 40% of the credit score distribution never made up more than 15% of total delinquent dollars. Both *default rates* and the *share of delinquent dollars* were highest for low credit score individuals.

In the next section, we present data and summary statistics. Sections 3 and 4 examine the rise in household debt from 2000 to 2007. Section 5 examines delinquencies from 2007 to 2010 and Section 6 concludes. In the conclusion, we relate the facts presented here to our earlier research.

2   Data and Summary Statistics

2.1   Data

Our main data set is based on individual-level credit bureau data from Equifax. This is the same data set used in Mian and Sufi (2011).[1] The initial random sample of individuals was drawn for the year 1997 from a group of 4,025 zip codes with Fiserv Case Shiller Weiss data available. The initial sample contains 320,295 individuals.

We limit the sample to the 288,042 individuals that have credit information available from 1997

---

[1]In Mian and Sufi (2011), we were required to sort individuals into groups of five. After the publication of that study, we were granted permission to use the individual level data. The records do not contain any information to identify individuals.

to 1999. We make this restriction because there is a large amount of attrition in the initial two years, driven by individuals with very few accounts. The attrition rate is 6% in the first two years, but then is reduced to a constant 2% afterward. We include all individuals that have data available for the first three years of our sample. We have data for these individuals through 2010.

We isolated the sample to people living in zip codes covered by Fiserv Case Shiller Weiss because our original research using the data required zip level house price indices. These 4,025 zip codes contain 25% of U.S. population and 40% of total debt. As has been discussed in our previous research (Mian and Sufi (2009), Mian and Sufi (2011)), the main difference between zip codes in the sample and not in the sample is population density. Only zip codes with a large number of households and a large number of housing transactions generate enough data to construct a zip code level house price index. These are mostly urban areas or suburban areas close to urban areas.

There are two potential sources of sample bias from our data. First, we miss new individuals that first enter the Equifax credit bureau system after 1997. Fortunately, this is likely not a major concern given that new entrants typically do not take on large sums of debt, and they therefore are unlikely to change the conclusions of our analysis focusing on aggregate debt patterns. Second, we only have individuals that resided in zip codes in 1997 with FCSW data. It is harder to assess whether this selection would materially change any results. However, Figure 2 gives us comfort that the FCSW criteria is not a major issue. It shows debt growth according the Federal Reserve Flow of Funds and according to our sample. As Figure 2 shows, debt growth from the flow of funds and from our data match closely.

The credit bureau data has excellent information on debt and defaults, but it only has limited information on individual characteristics. The primary measure we have is the Vantage Score, which is a credit score based on creditworthiness of the individual. We will discuss the Vantage Score in more detail below. We also have age for 90% of the sample.

Given the lack of data on income or home values, we supplement the credit bureau data using zip level information on income and home values from other data sources. More specifically, we match each individual to their zip-level average adjusted gross income per tax return and zip-level average home value. Zip level average home value comes from taking the average house price from Zillow in 2000, and then growing the house price by zip-level price indices from CoreLogic.

4

2.2 Credit scores

As mentioned above, the credit score we have in the credit bureau data is known as the Vantage Score. Like FICO and other credit scores, it is meant to measure the creditworthiness of a borrower. The Vantage Score varies between 550 and 990, as opposed to 300 to 850 for FICO scores. There is no specific cutoff in the Vantage Score data that indicates a subprime borrower, but 700 is a cut-off widely used to indicate a low credit quality borrower. In our random sample of individuals, about 35% of the sample has a Vantage Score below 700 in 1997.

In the analysis below, we split the sample based on the individual's Vantage Score in 1997. Given that 35% of the sample has a Vantage Score below 700, the subprime cutoff is at the high end of our second quintile. It is important to note that the credit score here applies to *borrowers*, not to the *product* used by the borrowers. People with low credit scores may obtain prime mortgages, and people with high credit scores may obtain subprime mortgages. We are not interested in evaluating the rise in debt across different *products*, but rather across different *people*.

In all of the analysis below, we group individuals by their initial credit score as of 1997 and track them over time. Why are credit scores the most natural characteristic on which to group individuals? The most obvious reason is practical: it is one of the only individual level characteristics we have in the credit bureau data. However, there are deeper economic justifications for grouping individuals in this manner. Credit scores are one of the primary variables used in credit origination decisions, and individuals with low credit scores have been shown in a number of studies to have high denial rates and a high marginal propensity to borrow. Further, credit scores are designed to predict default, which is itself an important outcome to evaluate.

We group individuals based on initial credit scores in 1997; an alternative approach would be a dynamic sort, in which individuals are grouped into credit score bins every year. We prefer the static sort for a number of reasons. First, our goal is to evaluate which individuals contributed to the rise in debt and defaults. A dynamic sort would have different individuals in different groups every year, which makes answering our primary question of which individuals drove the rise in debt more difficult.

Second, and more importantly, credit scores become endogenous to credit outcomes and house price growth over time. For example, Mian and Sufi (2011) show that low credit score homeowners in

inelastic housing supply cities see a decline in default rates during the housing boom relative to low credit score homeowners in elastic cities. The main reason is that low credit score homeowners seeing high house price growth refinance their way out of defaults – i.e., they do a cash-out refinancing if they cannot make a mortgage payment. As a result, low credit score homeowners in high house price growth areas will see a relative improvement in credit scores that is not driven by fundamental improvements in credit quality, but instead by the housing boom. We discuss this in more detail in Section 5.

As another example, an individual that starts in a high credit score group but then sees their credit score drop almost assuredly experienced some kind of financial distress. As a result, we should not be surprised that individuals that enter the low credit score group from a high credit score group see lower debt growth. We want to purge our credit score classification from such endogenous determinants. We use the initial credit score as of 1997, and individuals remain in their group throughout the sample. One disadvantage to this approach is mechanical mean reversion – if low credit score individuals start with low debt, we should expect faster growth. We address this concern in detail in the results below.

2.3   Summary statistics

Table 1 contains summary statistics. Total debt per individual at the beginning of the sample is $49 thousand on average, and mortgage or home equity related debt accounts for $40 thousand. Some of the debt on individual credit reports is held jointly with a spouse, and so debt held by an individual in credit report data is best viewed as something between debt held by an individual and a household. Table 1 also contains summary statistics on zip-level average adjusted gross income, and zip-level house price growth from 2000 to 2006.

In Table 2, we present summary statistics by credit score quintile. Individuals with lower credit scores have lower debt balances in 1997 and are less likely to have housing-related debt. However, the relation between debt balance and credit score quintile is not monotonic – those in the fourth quintile actually have less debt than those in the third quintile. Lower credit score individuals are younger, and live in zip codes with lower adjusted gross income per tax return. Conditional on having housing debt, the lowest 60% of the credit score distribution have debt to value ratios around 71%. Individuals with lower credit scores live in zip codes in 2000 that subsequently experience

6

higher house price growth, which is consistent with the zip code level evidence presented in Mian and Sufi (2009).

3 Rise in Debt, by Credit Score

3.1 Debt growth

The left panel of Figure 3 shows cumulative growth in household debt since 2000 by initial credit score quintile. The right panel plots growth from 2000 to 2007 by initial credit score quintile. Low credit score individuals saw the strongest growth in debt from 2000 to 2007. Debt grew by 175% for the lowest credit score quintile, and only 40% for the highest quintile. The relation between debt growth and credit score quintile is monotonic.

One potential explanation of the pattern in Figure 3 is age cohort effects. We know lower credit score individuals are younger, and perhaps we should expect debt to rise more quickly for younger individuals given standard life-cycle predictions. Before addressing this explanation directly, we want to point out that it is not obvious that younger people show be borrowing more in an environment in which house prices are increasing. Most borrowing is associated with housing, and we know home equity withdrawal was a primary driver of increases in household debt from 2000 to 2007. In most life-cycle models without borrowing constraints, an increase in house prices for older people who plan on reducing their housing services consumption would be a true wealth shock. A standard model without borrowing constraints would predict that older people would borrow more out of a rise in house prices than younger people who likely need to purchase more housing services in the future (see Mian and Sufi (2011) for more on this point).

In Table 3, we sort individuals by both initial credit score and age. As the first column shows, younger individuals did in fact borrow more aggressively during this time period, which is consistent with evidence from Mian and Sufi (2011) that younger households borrowed more against the increase in house prices. However, the stronger debt growth among lower credit score individuals is robust across every age cohort with the exception of individuals over 60. Stronger debt growth by lower credit score individuals is not merely a reflection of cohort effects.

Another concern is mechanical mean reversion. Low credit score individuals start with lower debt (see Table 2), and therefore we should expect debt to increase more rapidly. Before addressing

7

this concern, it is important to remember that the relation between initial debt and credit scores is not monotonic – the fourth quintile actually has less debt than the third quintile in Table 2. Regardless, Table 4 sorts individuals by both initial credit score and debt in 1999. As it shows, the higher debt growth among lower credit score individuals is robust across quintiles of 1999 debt levels. Stronger debt growth by lower credit score individuals is not merely a reflection of mean reversion.

3.2 Share of aggregate increase in debt

The growth in debt from 2000 to 2007 was largest for lower credit score individuals, but who contributed most to the dollar rise in debt? The left panel of Figure 4 shows the average debt level for individuals by 1997 credit score quintile. As it shows, the lowest credit score quintile has substantially less debt in 2000 relative to the other groups. However, all four higher quintiles start with similar debt levels.

From 2000 to 2007, the largest rise in the nominal amount of debt occurred among the second and third quintile of the initial credit score distribution. Individuals in these two groups experienced an increase in debt of about $75 thousand. Individuals in the lowest credit score quintile witnessed a rise in debt of about $50 thousand. Recall that each quintile contains the same fraction of the overall sample (20%). As a result, we immediately know from the left panel of Figure 4 that the total dollar rise was strongest for the 20th to 60th percentile of the credit score distribution.

The right panel of Figure 4 plots the share of the total rise in debt for each quintile. The 20th to 60th percentile of the initial credit score distribution accounts for just over 50% of the overall rise in debt. The lowest and fourth quintile contribute almost 40%, and the highest credit score quintile contributes only 10%.

Even though the lowest credit score individuals did not contribute the most to the total rise in debt, Figure 5 shows a large jump in the fraction of aggregate debt held by these individuals. In 2000, they held just less than 10% of total debt in the sample. By 2007, they held almost 15%. A smaller jump occured for the second quintile. The highest credit score quintile saw the largest decline, going from 25% to 18%.

The uniform message from both debt growth and the share of the aggregate debt increase is that the highest credit score individuals were least important in explaining the explosion of household

8

debt from 2000 to 2007. They both had the lowest growth in debt, and the lowest contribution to the rise in debt. The evidence on lower credit score individuals is mixed. Growth was strongest among the lowest quintile, but the 20th to 60th percentile of the initial credit score distribution contributed most significantly to the dollar rise in debt.

3.3 Debt to income and debt to value ratios

Midrigan and Philippon (2011), Favilukis, Ludvigson, and Van Nieuwerburgh (2013), and Justiniano, Primiceri, and Tambalotti (2014) present quantitative macroeconomic models of the housing boom. A potential driver of the debt expansion in these models is a loosening of *borrowing* constraints, which is modeled as lenders allowing borrowers to borrow a higher fraction of collateral value. Formally, the constraint is:

$$B_{it} \leq \theta_{it} Q_{it} h_{it}$$

where B_{it} is total household borrowing for individual i, Q_{it} is the price of a unit of housing, h_{it} is the quantity of housing, and θ_{it} is the borrowing constraint parameter. In this framework, a loosening of borrowing constraints for individual i over time would be an increased θ_{it}. A higher value of θ_{it} would show up as higher housing debt to home value ratios in the data.

Alternatively, Justiniano, Primiceri, and Tambalotti (2014) argue that the boom in household debt was not driven by a relaxation of borrowing constraints. Instead, it was driven by a loosening of a *lending* constraint, which they model as an increase in the total amount of mortgage lending:

$$B_{it} \leq \bar{L}$$

An increase in $\bar{L}$ is what they refer to as a loosened lending constraint, and they use debt to income ratios as a proxy for $\bar{L}$. Justiniano, Primiceri, and Tambalotti (2014) argue that a loosening of the borrowing constraint alone cannot explain what happened during the housing boom, because a loosening of the borrowing constraint would increase the demand for borrowing putting upward pressure on interest rates. In contrast, they argue that a loosening of a lending constraint is more consistent with falling interest rates and steady housing debt to home value ratios during the boom.

9

A key difference in which constraint was loosened is the behavior of debt to value versus debt to income ratios. If the lending constraint is more important, debt to value ratios would move less and debt to income ratios more.

Figure 6 shows the aggregate debt to income ratio and the housing debt to real estate value ratio for the United States. Debt to income ratios saw a dramatic rise beginning in 2000, whereas the housing debt to home value ratio was constant until 2005. The latter reflects the fact that house prices rose rapidly, which means the ratio of housing debt to home values was constant. However, the housing debt to home value ratio increased in 2006 and 2007.

We contribute to this debate by showing how debt to value and debt to income ratios changes across the credit score distribution. More specifically, in Figure 7, we use individual level data to see which individuals drove the increase in the debt to income ratio from 2000 to 2007. Debt comes from the individual credit reports, whereas income is measured as the average adjusted gross income per tax return in the zip code in which the individual resides. As the right panel shows, the rise in the debt to income ratio was driven by borrowers in the bottom 60% of the credit score distribution, and the increase was similar across all three groups. In percentage terms, the increase was largest for the lowest credit score quintile, given that individuals within this quintile started from a much lower base. The rise in the debt to income ratio was modest for the top 20% of the credit score distribution.

Figure 8 examines the housing debt to home value ratio conditional on having housing debt. Housing debt to home value ratios actually fell for all groups except for the lowest credit score quintile from 2000 to 2005. However, they rose sharply in 2006 and especially 2007. The rise from 2005 to 2007 was driven by both additional borrowing (see Figure 3) and a stagnation of house prices. The rise from 2001 to 2007 was largest for the lowest 40% of the credit score distribution.

4   Rise in Debt: By Credit Score and House Price Growth

As we show in Mian and Sufi (2011), an important driver of the increase in household debt during the housing boom was existing homeowners borrowing against the rising value of home equity. This finding motivates an additional sorting variable to examine which individuals drove the rise in aggregate debt: house price growth. In this section, we sort individuals into groups based on

house price growth from 2000 to 2006 of the zip code where the individual lived in 2000. We split individuals into five groups of house price growth, where the cutoffs are picked to keep roughly 20% of the sample in each group.

Table 5 shows debt growth from 2000 to 2007 by initial credit score quintile and house price growth group. In every single house price growth group, the growth in debt is larger for lower credit score individuals, which confirms the findings above. However, the new information in this table is the much stronger growth in debt for high house price growth zip codes, especially at the low end of the credit score distribution. For the lowest credit score quintile, debt growth was 107% in zip codes with less than 40% house price growth. In zip codes with greater than 130% house price growth, debt growth was almost twice as large: 207%.

Debt growth was significantly stronger in high house price growth zip codes for the bottom 80% of the initial credit score distribution. However, there is no relation between debt growth and house price growth at the very top of initial credit score distribution. Debt growth is almost constant across house price growth groups for individuals with highest initial credit scores.

The growth in debt was strongest for low credit score individuals living in high house price growth zip codes. But what about the share of the total rise in debt? The double-sort methodology introduces a complication because the share of the total population is no longer constant across each of the 25 cells. In the top panel of Table 6, we show the share of the population in each cell of the two-way sort. If house price growth and initial credit scores were randomly distributed, we would expect 4% of the population in each of the 25 cells. However, house price growth was stronger for lower credit score individuals, and so more of the population is in the top right corner and bottom left corner of the top panel of Table 6.

Initial debt levels are an important determinant of who contributes most to the rise in debt for given debt growth. The middle panel shows initial debt levels for each cell. Initial debt levels are smaller for lower initial credit score individuals, which we know from above. However, the relation between initial debt level and house price growth is not constant across the initial credit score distribution. In the lowest credit score quintiles, it appears that initial debt levels are lowest in zip codes where house prices rose the most.

The bottom panel of Table 6 shows what cells were responsible for the rise in debt in dollar terms from 2000 to 2007. Individuals in the 20th to 60th percentile of the initial credit score

11

distribution that also experienced high house price growth contributed most strongly to the rise in debt. More specifically, the eight cells in credit score quintiles 2 and 3 with house price growth above 40% make up 36% of the population. However, they make of 45% of the rise in total debt from 2000 to 2007. Once again, individuals in the highest credit score quintile contribute very little to the rise in household debt from 2000 to 2007.

5   Delinquencies

5.1   Who drove the household debt default crisis?

From 2000 to 2007, debt growth was strongest among low credit score individuals living in high house price growth areas. The lower-middle part of the credit score distribution with high house price growth was responsible for the largest share of the aggregate rise in household debt. In this section, we examine defaults. We answer two questions: first, who had the highest default rates? And second, who was responsible for the largest share of total dollars in delinquency?

It turns out that the answer to both questions is the same: individuals with low initial credit scores drove the default crisis. The left panel of Figure 9 shows the default rate for individuals based on initial credit score quintile. The default rate jumped from 9% to over 25% for the lowest credit score quintile from 2005 to 2009. The default rate also increased substantially in the second quintile. The rise in the default rate was more modest in the top 40% of the initial credit score distribution.

The right panel examines the share of total dollars in delinquency by quintile. In 2007, the bottom 20% of the credit score distribution accounted for over 40% of the dollars in delinquency in our sample. The bottom 40% made up 73% of the dollars in delinquency in 2007. As Figure 1 shows, the default rate in 2007 was almost 6%, much higher than the United States witnessed in any recent recession. Further, delinquencies on mortgages led to credit market disruptions in the summer and fall of 2007. Regulators were discussing mortgage problems as early as May 17th, 2007 when Chairman Ben Bernanke gave a speech on rising defaults in the subprime mortgage market. Figure 9 shows that the mortgage default crisis was triggered primarily by defaults among low credit score individuals.

By 2008 and 2009, the default crisis spread to higher credit score borrowers. However, the

12

bottom 40% of the initial credit score distribution continued to account for the lion's share of delinquencies: 69% in 2008 and 66% in 2009. Individuals in the top 40% of the initial credit score distribution never accounted for more than 15% of total dollars in delinquency, even at the height of the mortgage default crisis.

The fact that the default rate is so much higher for lower credit score individuals is not surprising. However, it may be surprising that the lowest quintile makes up the largest fraction of total defaults given that the lowest quintile had smaller debt burdens. Recall however that the lowest quintile increased their debt burden substantially from 2000 to 2007 (see left panel of Figure 4). So by 2007, individuals in the lowest quintile of the initial credit score distribution had increased leverage substantially.

We can see this effect in Figure 10. It plots the share of total debt outstanding in 2007 and the share of total delinquencies in 2008 across the initial credit score distribution. By 2007, individuals in the lowest quintile of the initial credit score distribution had 14% of debt outstanding. Individuals in the bottom two quintiles made up 37% of total debt in 2007. Recall that the majority of individuals in the bottom 40% of the initial credit score distribution had a vantage score below 700 at the beginning of the sample, which is considered subprime. So low credit score individuals accounted for a sizable amount of total debt by 2007. As Figure 10 shows, much higher default rates translated into a large fraction of total delinquencies by the bottom 40% of the credit score distribution in 2008.

An alternative approach to constructing credit score bins would be to ensure that each bin contains 20% of the total debt outstanding as of 2006 as opposed to 20% of the individuals. This may be the more relevant sort from an aggregate investor perspective, as it measures the share of defaults for the same principal at risk across the five bins.[2] Figure 11 shows an even higher share of defaults for the low credit score bins using this alternative weighting scheme. Delinquencies on the 20% of total debt in 2006 held by the lowest credit score individuals accounts for over 50% of all delinquencies in 2007, and 48% in 2008.

In Table 7, we examine the share of all delinquencies using the double sort of initial credit score quintile and zip-code level house price growth. As it shows, the default crisis was driven by low

---

[2]Of course, this means there will not be the same number of individuals in each bin – in particular, the lower credit score bin contains more individuals under this alternative weighting scheme.

credit score individuals living in high house price growth areas. The top-right 4 cells in Table 7 include 18% of the sample, but 39% of the defaults.

5.2 Contrasting with other research

The results in Figure 9, Figure 10, and Table 7 stand in contrast to the evidence in Adelino, Schoar, and Severino (2015) who say "...borrowers in the middle and top of the income distribution, as well as those with credit scores above 660, are the ones that contributed most significantly to the increase in the dollar value of mortgages in default after 2007." Why is there a discrepancy? One reason discussed in Mian and Sufi (2015) is that Adelino, Schoar, and Severino (2015) rely on income reported on mortgage applications that is fraudulently overstated in lower income areas. But this concern is less important when sorting by credit scores.

We believe a number of factos explain the discrepancy. First, Adelino, Schoar, and Severino (2015) use a data set that systematically under-represents mortgages to low credit score individuals. The data used by Adelino, Schoar, and Severino (2015) is based on mortgage loan originations for home purchase in 2006 recorded by Lender Processing Services (LPS). As Adelino, Schoar, and Severino (2015) note, LPS "covers approximately 60 percent of the U.S. mortgage market." They do not, however, discuss any potential biases due to the mortgages that are missing in LPS.

Existing research demonstrates that LPS systematically misses mortgages made to low credit score individuals. As Keys, Mukherjee, Seru, and Vig (2010) note, "... the McDash/LPS database significantly under-represents the subprime market before 2005 and even after 2005 covers only about 30% of originated non-agency loans." Keys, Mukherjee, Seru, and Vig (2010) compare a data set that captures 90% of subprime mortgages (Loan Performance) with LPS, and they find that LPS captures only 17% of subprime mortgages with low documentation originated in 2006 (see their Figure 15). Amromin and Paulson also note that the LPS data cover only a fraction of the subprime market. Recall that Adelino, Schoar, and Severino (2015) focus on the share of debt in delinquency, not default rates. If their data set systematically misses mortgages to low credit score individuals, the the fraction of total delinquencies by low credit score individuals will be systematically biased downward.

There is another reason for the discrepancy based on changes in credit scores over time. Adelino, Schoar, and Severino (2015) sort individuals into groups based on the credit score at origination

14

on home purchase mortgages in 2006. We sort individuals into groups based on their credit score before the boom in house prices and household debt. We sort on credit scores before the boom because, as Mian and Sufi (2011) show, low credit score individuals living in high house price growth areas saw a sharp relative decline in defaults from 1997 to 2005 (see the middle panel of their Figure 5). In that study, we argue lower default rates were due to the ability of homeowners in high house price growth areas to extract equity to avoid default in case of a negative shock such as unemployment (e.g., Hurst and Stafford (2004)). When house prices crashed, the pattern reverses and default rates increase by far more for the same households that saw the largest drop in default rates during the boom.

Because of the lower default rates during the boom, low initial credit score homeowners in high house price growth areas saw the largest increase in their credit scores from 2000 to 2006. Table 8 shows that the largest increase in credit scores occurred among the lowest 40% of the credit score distribution living in high house price growth areas (the top right four cells). We know from Table 7 that these individuals made up the largest share of defaults in 2008. In other words, conditional on the 2006 credit score, the increase in credit scores from 2000 to 2006 predicts higher defaults in 2008, a result we confirm in a regression framework.

We sort individuals in the top right cells of Table 8 into low credit score bins based on their 1997 credit score, whereas the methodology of Adelino, Schoar, and Severino (2015) would sort them into higher credit score bins based on their 2006 credit score. Their sort mechanically shows higher default rates among high credit score individuals because of the endogeneity of credit scores with respect to the housing boom. This is why our research always sorts on *initial* credit scores before the housing boom when examining the expansion in debt and the subsequent default crisis (Mian and Sufi (2009), Mian and Sufi (2011)).

Finally, Adelino, Schoar, and Severino (2015) focus only on defaults on home purchase mortgages originated in 2006. We focus on the universe of all defaults on any debt outstanding. There may be systematic differences in default rates or originated amounts on home purchase mortgages in 2006 versus all debt outstanding. For example, it may be that LPS does not include second mortgages or home equity lines, which have higher default rates and originated amounts for lower credit score individuals. Or alternatively, it could be that the highest share of mortgages for low credit score borrowers was originated in years other than 2006. The use of credit bureau data avoids these

problems, as it includes all debt outstanding and all defaults.

6   Conclusion

Our purpose in this study is to present facts on which individuals drove the rise in household debt from 2000 to 2007 and the subsequent default crisis during the Great Recession. Using individual level credit bureau data, we show the following four facts:

- Individuals with low credit scores as of 1997 had the largest growth in debt from 2000 to 2007, and this fact is not driven by differences in age or initial debt levels

- Individuals in the 20th to 60th percentile of the initial credit score distribution contributed most to the dollar rise in household debt from 2000 to 2007; individuals in the highest 20% of the credit score contributed the least

- Low credit score individuals saw the largest increase in debt in high house price growth zip codes, consistent with the importance of home-equity based borrowing. Borrowing by the highest credit score individuals was completely unresponsive to higher house price growth

- The household debt default crisis from 2007 to 2010 was driven primarily by individuals in the lowest 40% of the credit score distribution, who had much higher default rates and contributed most to the total debt amount in delinquency

We have purposefully avoided causal statements up to this point. However, we believe all of the facts presented here are consistent with the narrative of our research that more carefully considers causation and alternative hypotheses. In Mian and Sufi (2009), we examined zip-code level data on new mortgage originations for the purpose of home purchase, and we found that mortgage origination growth was strongest in low credit score zip codes. This is consistent with the evidence presented here that low credit score individuals had the largest growth in debt from 2000 to 2007. We also argued that the expansion of credit to low credit score individuals triggered the mortgage default crisis in 2007 and early 2008. The results on delinquencies and default rates presented here support that view.

In Mian and Sufi (2011), we argued that rising house prices had a causal effect on borrowing by existing homeowners, and this home-equity based borrowing was responsible for a large fraction of the overall rise in household debt. We also argued that the elasticity of borrowing with respect to house price growth was strongest among low credit score individuals, but it was only zero for the highest credit score individuals. As we said, "for a consumer one standard deviation above the mean 1997 credit score, the elasticity of debt with respect to house prices is 0.35. For a consumer one standard deviation below the mean 1997 credit score, the elasticity is 0.76." The facts presented here are consistent with these findings. Only the top quintile of the credit score distribution sees no relation between house price growth and debt growth. The relation is present for the bottom 80% of the credit score distribution. We also argued that aggressive borrowing against home equity was responsible for higher default rates among homeowners during the Great Recession. In this study, we find that low credit score individuals living in high house price growth zip codes contributed most to the overall rise in defaults.

References

Manuel Adelino, Antoinette Schoar, and Felipe Severino. Changes in buyer composition and the expansion of credit during the boom. Technical report, National Bureau of Economic Research, 2015.

Gene Amromin and Anna Paulson. Comparing patterns of default among prime and subprime mortgages. *Federal Reserve Bank of Chicago Economic Perspectives*, 2009.

Jack Favilukis, Sydney C Ludvigson, and Stijn Van Nieuwerburgh. The macroeconomic effects of housing wealth, housing finance, and limited risk-sharing in general equilibrium. *SSRN*, 2013.

Erik Hurst and Frank Stafford. Home is where the equity is: Mortgage refinancing and household consumption. *Journal of Money, Credit and Banking*, pages 985–1014, 2004.

Alejandro Justiniano, Giorgio E Primiceri, and Andrea Tambalotti. Credit supply and the housing boom. 2014.

Benjamin J Keys, Tanmoy Mukherjee, Amit Seru, and Vikrant Vig. 620 fico, take ii: Securitization and screening in the subprime mortgage market. *Working Paper*, 2010.

Atif Mian and Amir Sufi. The consequences of mortgage credit expansion: Evidence from the us mortgage default crisis. *The Quarterly Journal of Economics*, 124(4):1449–1496, 2009.

Atif Mian and Amir Sufi. Household leverage and the recession of 2007–09. *IMF Economic Review*, 58(1):74–117, 2010.

Atif Mian and Amir Sufi. House prices, home equitybased borrowing, and the us household leverage crisis. *The American Economic Review*, pages 2132–2156, 2011.

Atif Mian and Amir Sufi. What explains the 2007–2009 drop in employment? *Econometrica*, 82 (6):2197–2223, 2014.

Atif Mian and Amir Sufi. Fraudulent income overstatement on mortgage applications during the credit expansion of 2002 to 2005. Technical report, National Bureau of Economic Research, 2015.

Atif Mian, Kamalesh Rao, and Amir Sufi. Household balance sheets, consumption, and the economic slump. *The Quarterly Journal of Economics*, 128(4):1687–1726, 2013.

Virgiliu Midrigan and Thomas Philippon. Household leverage and the recession. 2011.

Figure 1: Aggregate Household Debt and Defaults

The left panel of this figure plots nominal household debt according to the Federal Reserve Flow of Funds. The right panel plots the default rate on household debt according to our sample of credit reports.
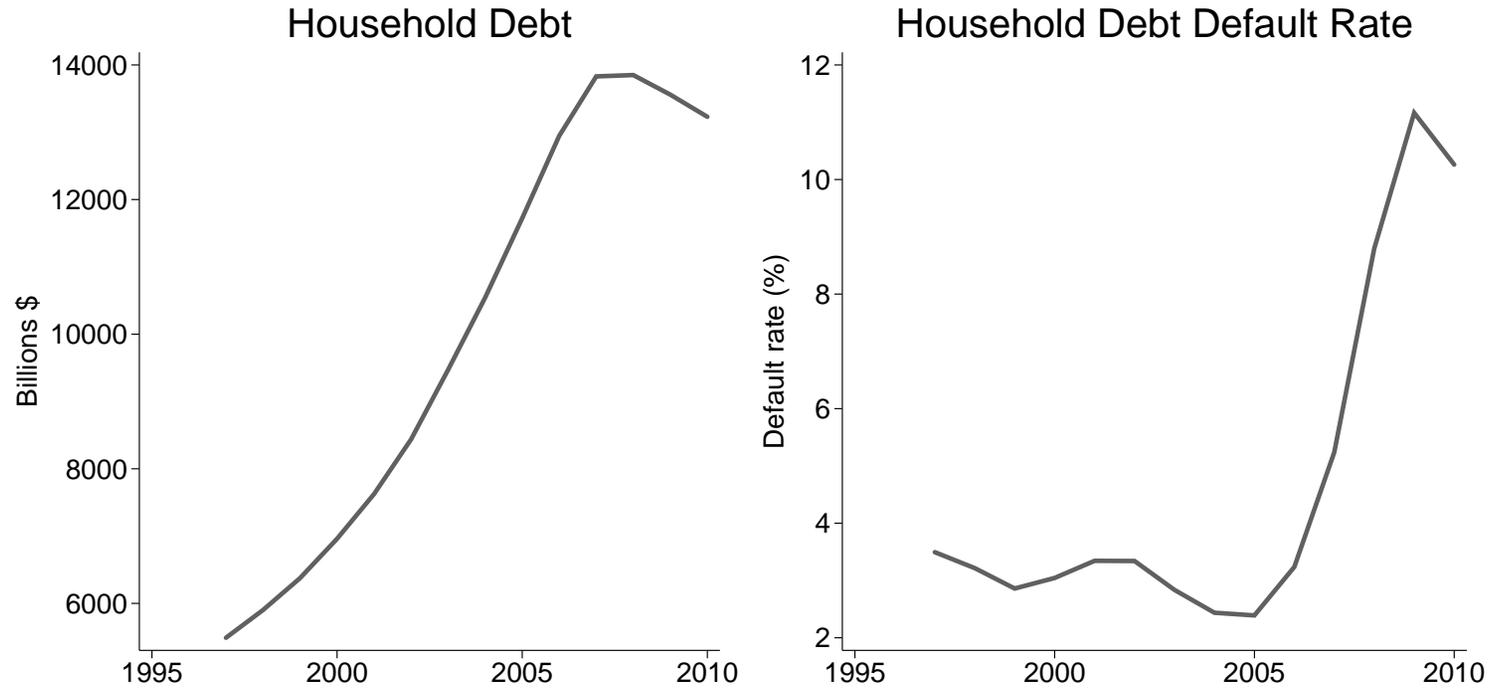
Figure 2: Debt Growth: Sample Matches Aggregate

This figure plots the growth in household debt in the Federal Reserve Flow of Funds and growth in total debt for our sample of individual credit reports.
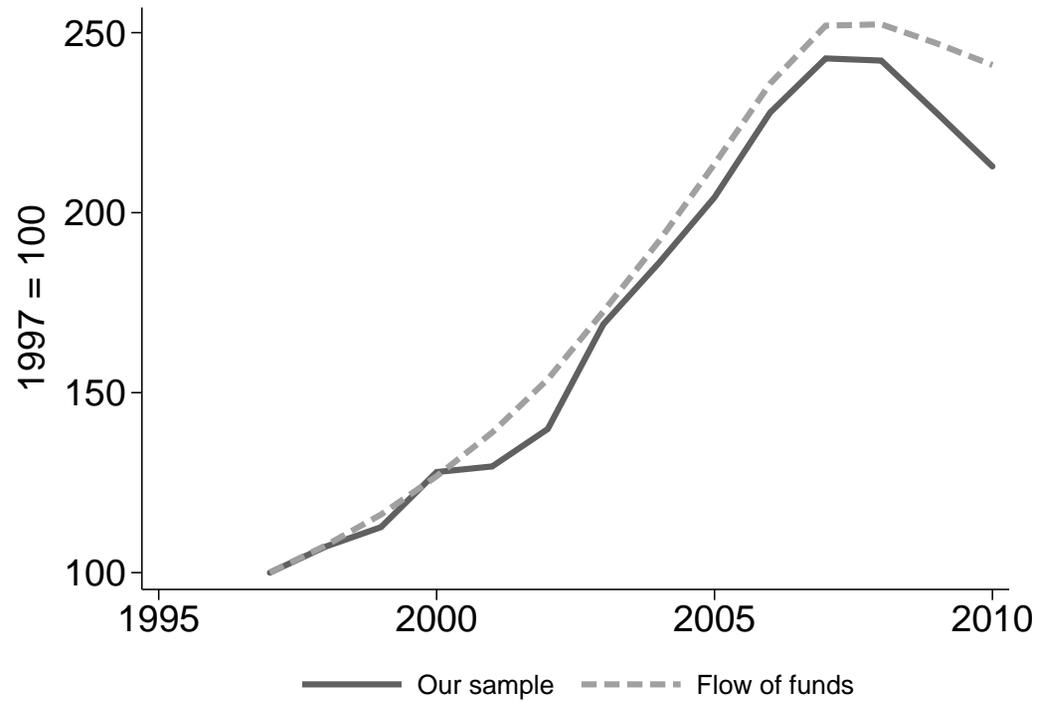
## Figure 3: Growth in Debt, by 1997 Credit Score

This figure plots the growth in debt for individuals sorted into quintiles by their 1997 credit score. Each quintile contains 20% of the sample. The left panel shows cumulative growth since 2000, and the right panel shows growth from 2000 to 2007.

Debt growth by 1997 credit score

Debt growth by 1997 credit score
2000 to 2007
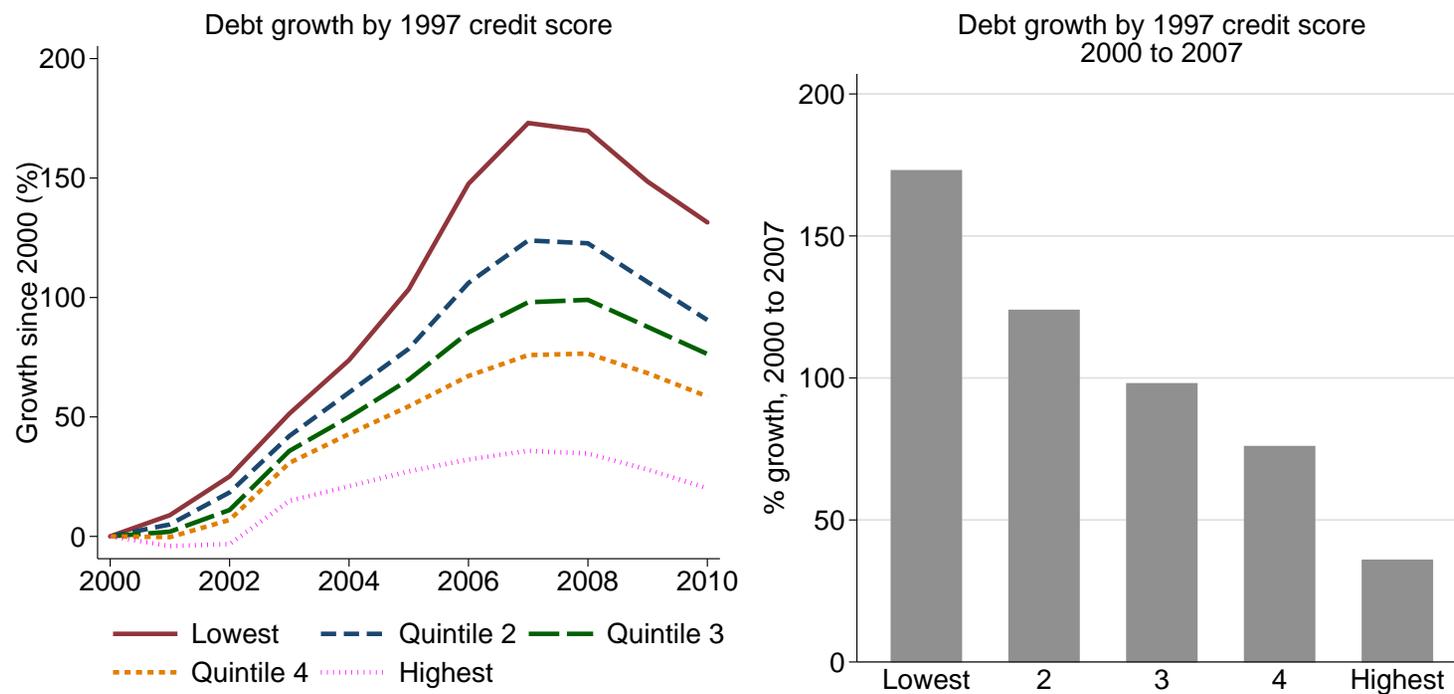
Figure 4: Share of Aggregate Rise in Debt, by 1997 Credit Score

The left panel plots average debt for individuals in each credit score quintile. The right panel shows the share of the aggregate dollar rise in debt by credit score quintile. Each quintile contains 20% of the sample.
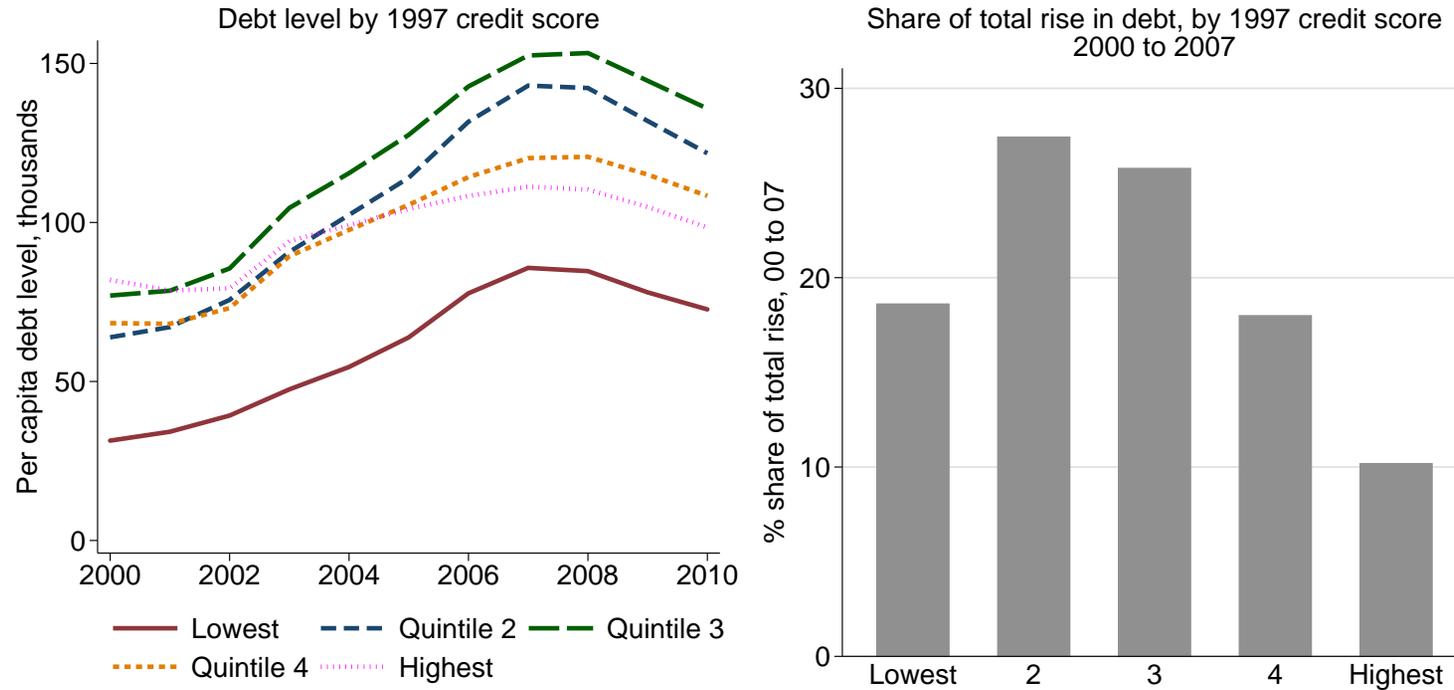
Figure 5: Share of Debt in 2000 and 2007, by 1997 Credit Score

This figure plots the share of total household debt in 2000 and 2007 for each quintile of the 1997 credit score distribution. Each quintile contains 20% of the sample.
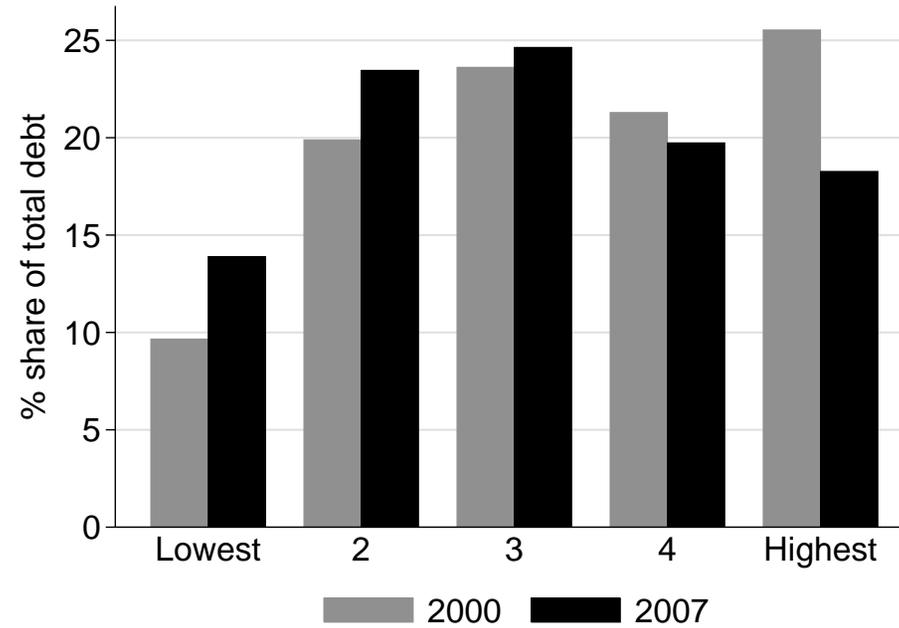
Figure 6: Aggregate Debt to Income and Debt to Value Ratios

This figure plots the household debt to income ratio, where household debt is from the Flow of Funds and income is from NIPA and is measured as compensation from wages. It also plots the mortgage-related household debt to real estate asset ratio, both variables are from the Flow of Funds.
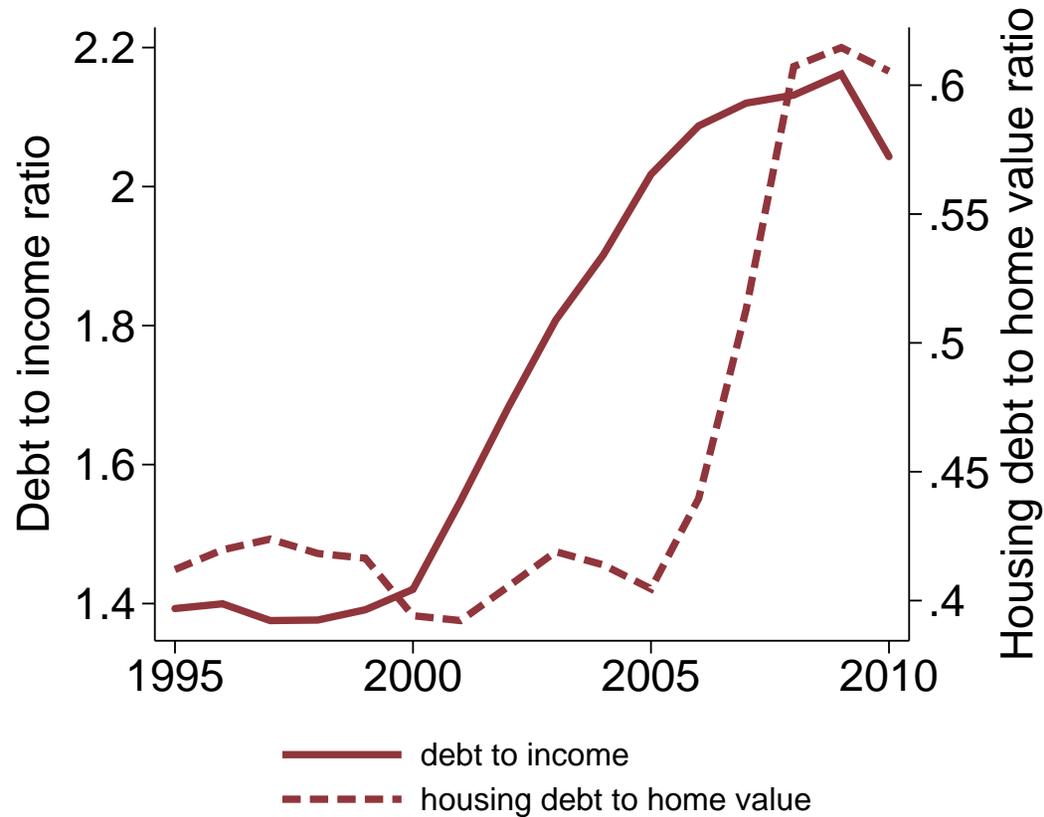
Figure 7: Debt to Income, by 1997 Credit Score

This figure plots the debt to income ratio for individuals based on their 1997 credit score. Income is measured as average adjusted gross income per tax return in the zip code in which the individual resides. Each quintile contain 20% of the sample.
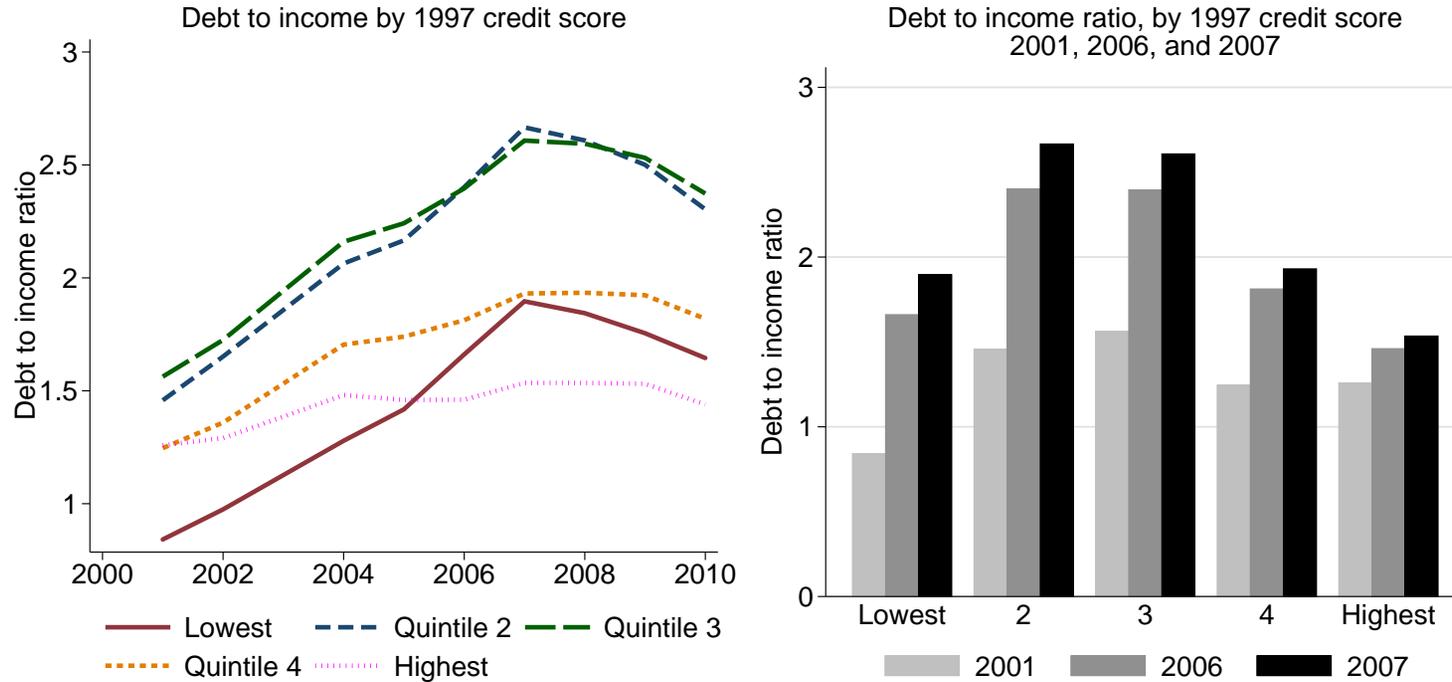
Figure 8: Housing Debt to Home Value Ratio, by 1997 Credit Score

This figure plots the housing debt to home value ratio for individuals based on their 1997 credit score. The sample is limited to individuals that have some housing debt outstanding, so it is the housing debt to home value ratio conditional on having some housing debt. Housing debt includes both mortgage and home-equity debt. The home value is measured using median home value in a zip code in 2000 from the Census, and then growing the home value using CoreLogic zip-level house price indices. Each quintile contain 20% of the sample.
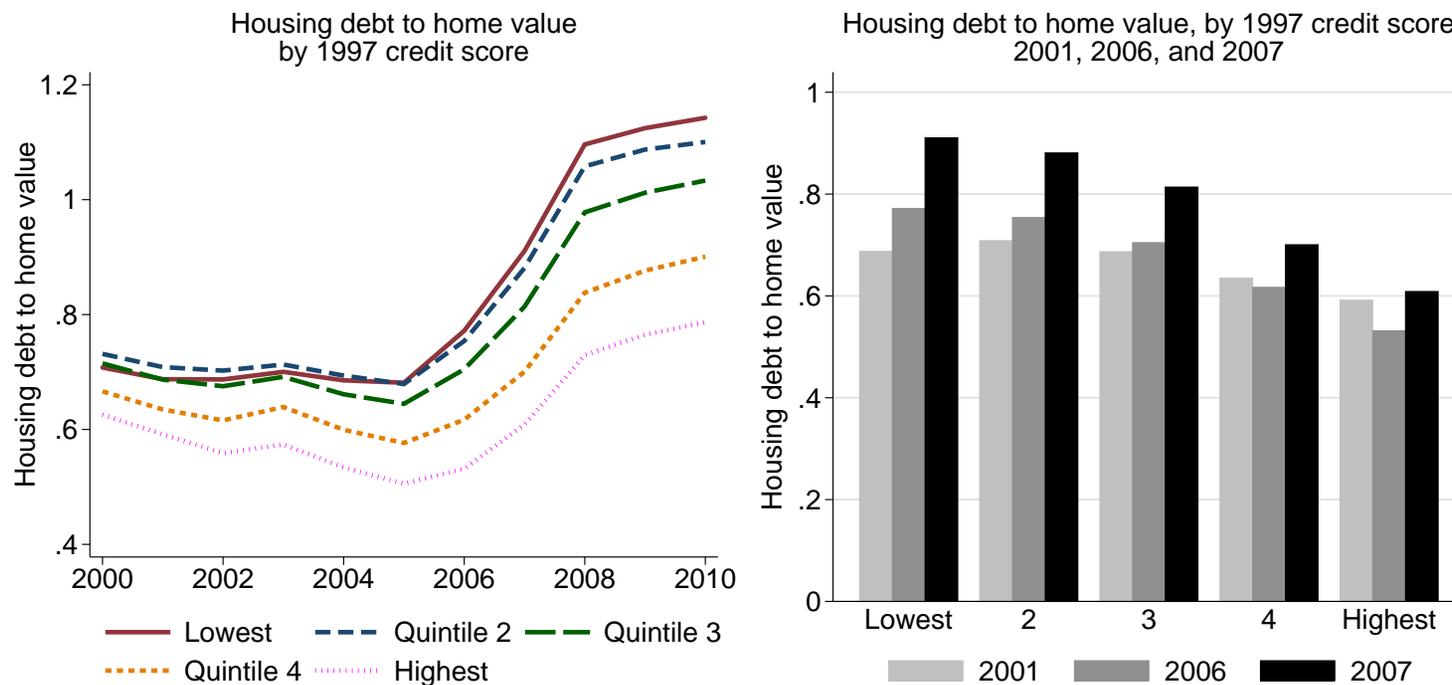
Figure 9: Delinquencies, by 1997 Credit Score

The left panel plots the default rate by 1997 credit score quintile, and the right panel plots the share of total dollars in delinquency by 1997 credit score quintile. Each quintile contain 20% of the sample.

Default rate by 1997 credit score

Share of delinquent debt, by 1997 credit score
2007 to 2010
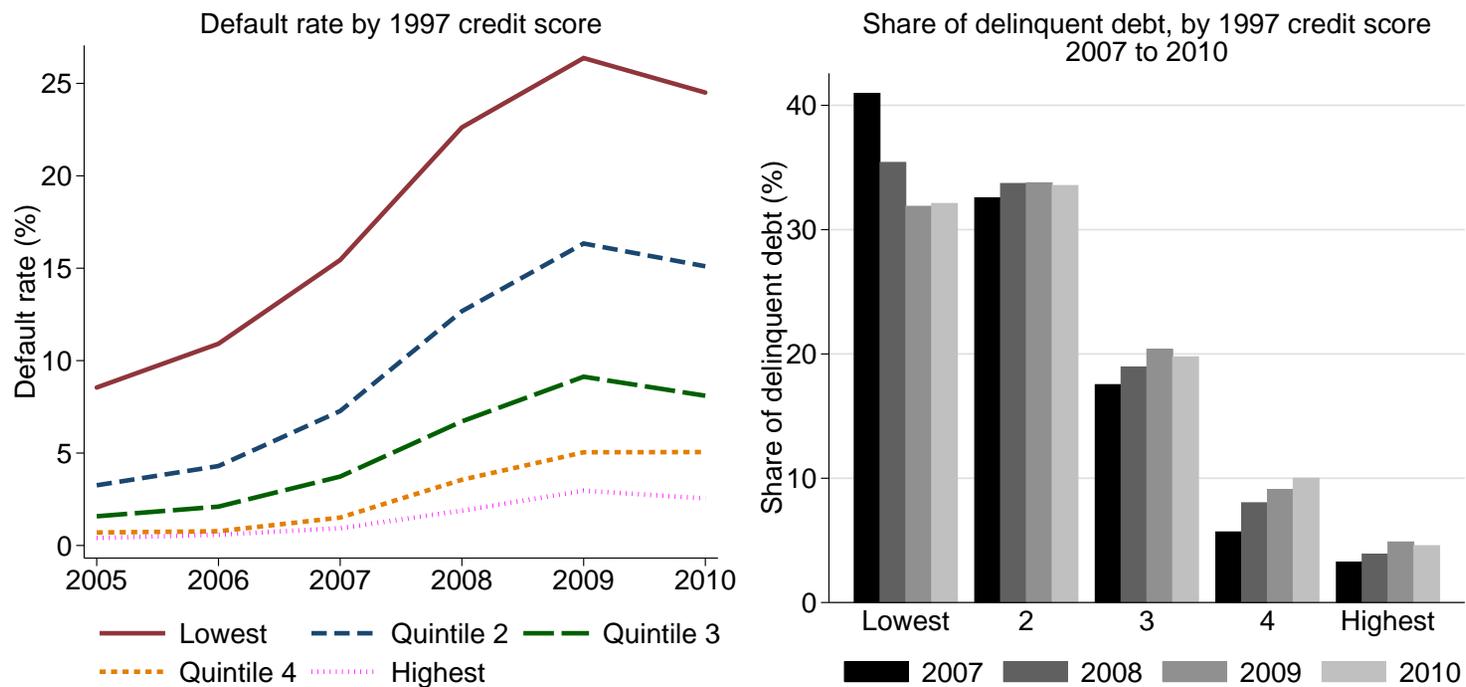
Figure 10: Share of Total Debt and Delinquencies, by 1997 Credit Score

This figure plots the share of total debt in 2007 and the share of total dollars in delinquency in 2008 by 1997 credit score. Each quintile contain 20% of the sample.
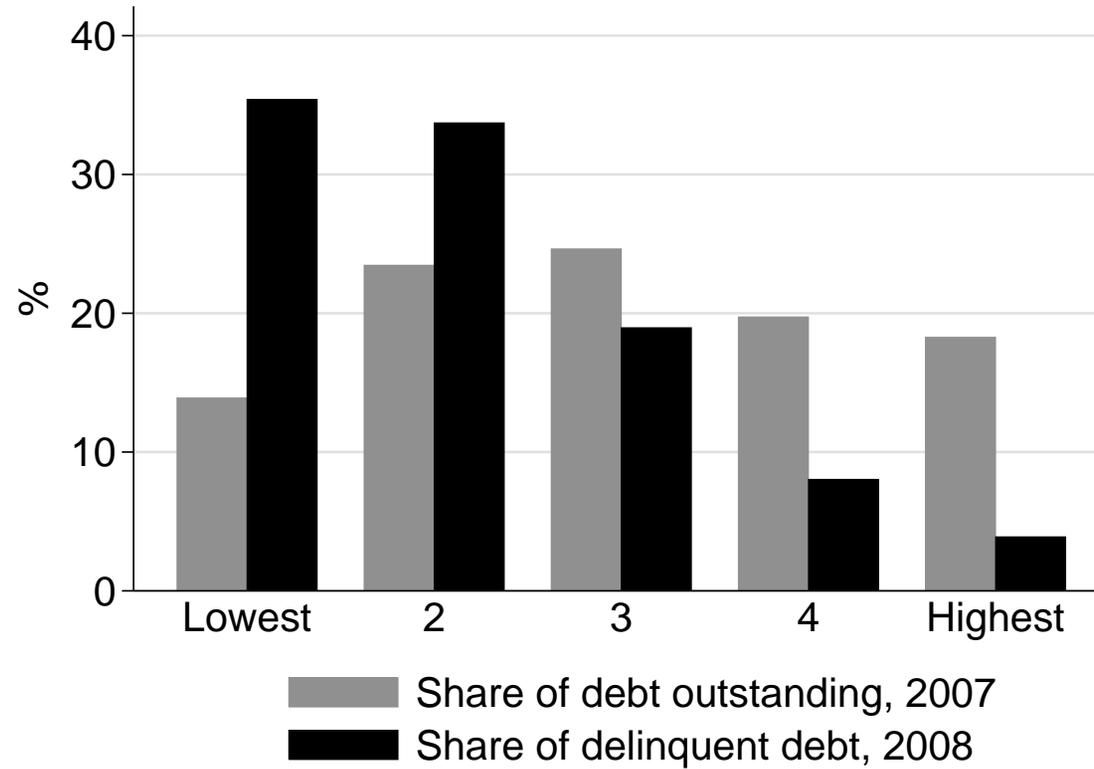
Figure 11: Share of Total Delinquencies, by 1997 Credit Score, Bins Contain 20% of Total Debt in 2006

This figure plots the share of total dollars in delinquency by 1997 credit score. In contrast to the previous figures and tables, each quintile in this figure contains 20% of *total debt in 2006* as opposed to 20% of *individuals*. The mean Vantage score in 1997 for each bin moving from left to right is 603, 696, 760, 827, and 894.
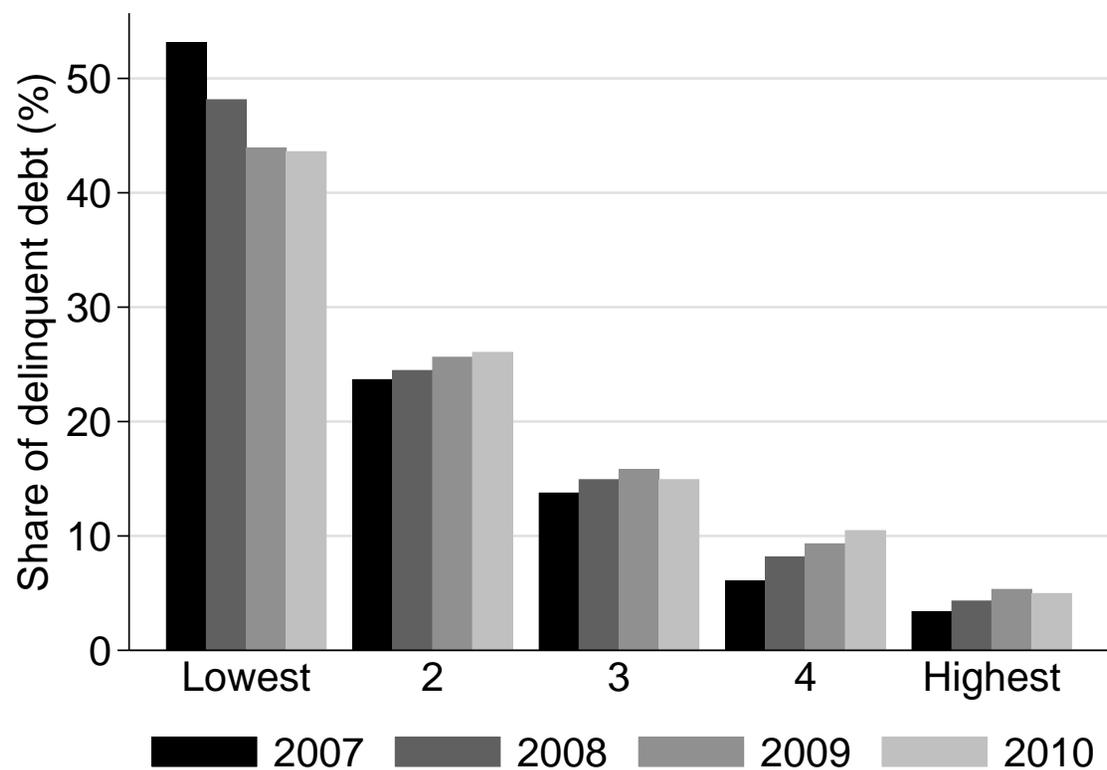
Table 1: Summary Statistics

This table present summary statistics for our sample of 288,042 individuals from Equifax. Housing debt to home value is measured only for individuals with some housing-related debt outstanding as of 2000.

	N	Mean	SD	10th	90th
Total debt, 1997, thousands	288042	49.81	94.24	0.00	148.94
Housing debt, 1997, thousands	288042	40.43	88.31	0.00	134.00
Has housing debt, 1997	288042	0.36	0.48	0.00	1.00
Credit score (Vantage), 1997	288042	747.99	113.83	586.00	888.00
Age, 1997	257338	45.58	15.85	26.00	70.00
Zip average AGI, 1998, thousands	286799	49.80	30.35	26.97	77.18
Debt to income ratio, 1998	284164	1.13	1.77	0.00	3.39
Zip average home value, 2000	237358	183.07	106.17	87.70	309.30
Housing debt to home value, 2000	93995	0.68	0.50	0.16	1.23
Zip house price growth (%), 2000 to 2006	237883	89.94	50.36	19.67	157.40

Table 2: Averages by 1997 Credit Score Quintile

This table presents averages by 1997 credit score quintile. Each quintile contains 20% of the sample. Housing debt to home value is measured only for individuals with some housing-related debt outstanding as of 2000.

	Means by credit score quintile				
	1	2	3	4	5
Credit score (Vantage), 1997	581.0	678.8	756.7	830.5	894.5
Total debt, 1997, thousands	25.44	42.02	54.04	51.23	76.60
Housing debt, 1997, thousands	14.49	28.77	43.23	44.22	71.76
Has housing debt, 1997	0.164	0.299	0.401	0.392	0.544
Age, 1997	37.23	40.53	44.05	50.23	55.35
Zip average AGI, 1998, thousands	41.83	45.76	49.98	53.17	58.28
Debt to income ratio, 1998	0.646	1.152	1.341	1.146	1.355
Zip average home value, 2000	158.5	171.9	184.3	189.7	210.5
Housing debt to home value, 2000	0.708	0.732	0.715	0.666	0.626
Zip house price growth (%), 2000 to 2006	97.07	93.36	89.13	84.94	85.33

Table 3: Growth in Debt, by Credit Score and Age Cohort

This table shows the growth in debt from 2000 to 2007 by credit score quintile and age in 1997.

| Credit Score | Debt growth, 2000 to 2007 (%) | | | | |
| Quintile | Age in 1997 | | | | |
	lt 30	30-40	40-50	50-60	gt 60
1	366.6	171.4	114.9	71.0	5.6
2	277.5	126.1	88.5	51.8	6.0
3	258.1	107.1	71.3	44.3	8.0
4	197.4	90.3	63.3	39.9	4.1
5	135.8	58.2	40.6	22.2	-11.1

Table 4: Growth in Debt, by Credit Score and Initial Debt

This table shows the growth in debt from 2000 to 2007 by 1997 credit score quintile and initial debt quintile in 1999.

| Credit Score | Debt growth, 2000 to 2007 (%) | | | | |
| Quintile | Debt in 1999, Quintile | | | | |
	1	2	3	4	5
1	618.4	490.0	302.4	94.6	48.5
2	485.1	421.9	289.2	97.9	53.6
3	447.0	401.7	241.4	81.2	52.1
4	342.7	295.6	196.7	56.7	43.4
5	193.1	150.2	112.8	29.2	22.7

Table 5: Growth in Debt, by Credit Score and House Price Growth

This table shows the growth in debt from 2000 to 2007 by 1997 credit score quintile and by house price growth from 2000 to 2007. Each individual is assigned the house price growth from 2000 to 2007 of the zip code in which they reside in 2000.

Credit Score Quintile	Debt growth, 2000 to 2007 (%) House Price Growth Category				
	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	106.7	175.7	181.5	194.5	207.1
2	83.5	126.8	133.2	138.2	142.4
3	76.2	102.7	107.1	100.4	109.6
4	61.3	74.0	78.8	80.1	93.2
5	33.0	38.7	35.1	36.4	38.9

**,* Coefficient statistically different than zero at the 1% and 5% confidence level, respectively.

Table 6: Share of Rise in Debt, by Credit Score and House Price Growth

This table shows means by 1997 credit score quintile and by house price growth from 2000 to 2007. Each individual is assigned the house price growth from 2000 to 2007 of the zip code in which they reside in 2000. The bottom panel shows the share of total debt increase from 2000 to 2007 for each cell.

| Credit Score | Share of population, 1999 (%) | | | | |
| | House Price Growth Category | | | | |
Quintile	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	3.7	3.2	3.7	3.7	5.5
2	3.8	3.8	4.0	3.6	4.9
3	3.9	4.3	4.0	3.5	4.2
4	4.1	4.7	4.3	3.4	3.5
5	3.7	5.0	4.5	3.6	3.4

| Credit Score | Debt level, 2000 (thousands) | | | | |
| | House Price Growth Category | | | | |
Quintile	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	32.2	33.9	32.3	33.5	29.3
2	63.6	69.0	65.9	67.6	60.1
3	75.9	83.4	76.7	82.6	74.0
4	66.1	77.4	66.4	69.0	63.6
5	76.7	88.8	77.5	84.8	76.1

| Credit Score | Share of Debt Increase, 2000 to 2007 (%) | | | | |
| | House Price Growth Category | | | | |
Quintile	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	2.1	3.2	3.6	4.0	5.6
2	3.4	5.6	5.9	5.6	7.1
3	3.8	6.1	5.5	4.8	5.7
4	2.8	4.5	3.8	3.2	3.5
5	1.6	2.9	2.0	1.9	1.7

Table 7: Share of Delinquencies, by Credit Score and House Price Growth

This table shows the share in delinquencies in 2008 by 1997 credit score quintile and by house price growth from 2000 to 2007. Each individual is assigned the house price growth from 2000 to 2007 of the zip code in which they reside in 2000.

| Credit Score | Share of delinquent debt, 2008 (%) | | | | |
| | House Price Growth Category | | | | |
Quintile	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	3.7	4.9	6.4	7.8	12.5
2	3.2	5.6	6.6	6.8	11.6
3	1.8	3.4	4.0	3.7	6.2
4	1.0	1.5	1.5	1.7	2.4
5	0.3	0.8	0.6	0.7	1.3

Table 8: Change in Credit Scores, by Credit Score and House Price Growth

This table shows the change in the Vantage Score from 2000 to 2006 by 1997 credit score quintile and by house price growth from 2000 to 2007. Each individual is assigned the house price growth from 2000 to 2007 of the zip code in which they reside in 2000.

| Credit Score | Change in Vantage Score, 2000 to 2006 | | | | |
| | House Price Growth Category | | | | |
Quintile	lt 40%	40-75%	75-105%	105-130%	gt 130%
1	30.2	39.3	38.7	44.3	43.2
2	31.4	41.4	45.4	47.7	48.6
3	32.2	37.7	41.7	43.5	43.1
4	22.6	26.8	24.9	26.2	24.5
5	11.2	10.1	8.9	8.5	7.7