Credit Growth and the Financial Crisis: A New Narrative*

Stefania Albanesi, University of Pittsburgh, NBER and CEPR
Giacomo DeGiorgi, GSEM-University of Geneva, ICREA/MOVE, BGSE, CEPR
Jaromir Nosal, Boston College

February 11, 2018

Abstract

A broadly accepted view contends that the 2007-09 financial crisis in the U.S. was caused by an expansion in the supply of credit to subprime borrowers during the 2001-2006 credit boom, leading to the spike in defaults and foreclosures that sparked the crisis. We use a large administrative panel of credit file data to examine the evolution of household debt and defaults between 1999 and 2013. Our findings suggest an alternative narrative that challenges the large role of subprime credit in the crisis.

We show that credit growth between 2001 and 2007 was concentrated in the prime segment, and debt to high risk borrowers was virtually constant for all debt categories during this period. The rise in mortgage defaults during the crisis was concentrated in the middle of the credit score distribution, and mostly attributable to real estate investors. We argue that previous analyses confounded life cycle debt demand of borrowers who were young at the start of the boom with an expansion in credit supply to subprime borrowers over that period.

*We are grateful to Christopher Carroll, Ambrogio Cesa-Bianchi, Gauti Eggertsson, Joel Elvery, Nicola Gennaioli, Richard Harrison, Marianna Kudlyak, Douglas McManus, Virgiliu Midrigan, Giuseppe Moscarini, Ned Prescott, Giorgio Primiceri, Joe Tracy, Eric Swanson, Paul Willen and many seminar and conference participants for useful comments and suggestions. We also thank Matt Ploenzke, Jakob Fabina, Richard Svoboda and Harry Wheeler for excellent research assistance. Correspondence to: stefania.albanesi@gmail.com.

1 Introduction

The broadly accepted narrative about the financial crisis is based on the findings in Mian and Sufi (2009) suggesting that most of the growth in credit during the 2001-2006 boom was concentrated in the subprime segment, despite the fact that income did not rise over the same period for this group of borrowers. The expansion of subprime credit is seen as the leading cause of the rise in mortgage delinquencies and foreclosures, which caused the housing crisis and subsequent the 2007-2009 recession (see Mian and Sufi (2010), Mian and Sufi (2011), Mian, Rao, and Sufi (2013) and Mian, Sufi, and Trebbi (2015)).

This paper studies the evolution of household borrowing and default between 1999 and 2013 using the Federal Reserve Bank of New York Consumer Credit Panel/Equifax data, a large administrative panel of anonymous credit files from the Equifax credit reporting bureau. The data contains information on individual debt holdings, delinquencies, public records and credit scores. We examine the evolution of mortgage debt and defaults during the housing boom and throughout the housing crisis and its aftermath. Our findings suggest an alternative narrative that challenges the view that an expansion of the supply of mortgage credit to subprime borrowers in 2001-2007 played a large role in the housing and financial crisis. Specifically, we show that credit growth between 2001 and 2007 is concentrated in the middle and and at the top of the credit score distribution. Borrowing by individuals with low credit score is virtually constant over this period. We also find that the rise in defaults during the financial crisis is concentrated in the middle of the credit score distribution. Borrowers with subprime credit score typically have higher default rates than those with higher credit scores, however, during the housing crisis and the 2007-2009 recession the fraction of mortgage delinquencies experienced by the lowest quartile of of the credit score distribution dropped from 40% to 30%, and the fraction of foreclosures from 70% to 35%.

Mian and Sufi (2009) and Mian and Sufi (2016) identify subprime individuals based on their credit score in 1996 and 1997, respectively. We show that, since low credit score individuals at any time are disproportionally young, this approach confounds an expansion of the supply of credit to low credit score borrowers with the life cycle demand for credit of borrowers who were young at the start of the boom. To avoid this pitfall, our approach estimates future growth in mortgage balances and mortgage delinquencies based on the borrowers' recent lagged credit score. This is closer to industry practices and prevents joint endogeneity of credit scores with borrowing and delinquency behavior, but ensures that the ranking best reflects the borrower's likely ability to repay debt at the time of borrowing. Using payroll data for 2009, we show that the cross sectional variation of credit scores is mostly explained by variation in labor income, conditional on age. Moreover, the life cycle growth in credit scores is tightly related to the life cycle growth in income.

Our finding that the rise in defaults during the housing crisis and subsequent recession was greatest for borrowers in the middle and at the top of the credit score distribution is puzzling, as these borrowers historically exhibit very low default rates on any type of debt, as well as very low foreclosure rates. To gain insight on what may have driven defaults by borrowers with relatively high credit scores, we explore the role of real estate investors. Using our data, we can identify investors as borrowers who hold 2 or more first mortgages, following Haughwout et al. (2011). There are four main reasons that may lead real estate investors to display higher default rates than other borrowers with similar credit scores. First, mortgages for non owner occupied properties must meet stricter credit standards and are usually charged an additional premium to qualify for GSE insurance. This makes it more likely for real estate investors to contract non-standard mortgages, which are intrinsically more risky.[1] Second, if investors are motivated by the prospect of capital gains,[2] they are more likely to default if the value of the mortgage is higher than the value of the property, especially in states in which foreclosure is non recourse.[3] Third, only the primary residence is protected in personal bankruptcy (see Li (2009)). Thus, a financially distressed borrower whose primary residence satisfies the homestead exemption could potentially file for Chapter 7 bankruptcy and discharge unsecured debt to avoid missing payments on the mortgage.[4] Finally, the financial and psychological costs of default for resident owners are typically quite substantial, including moving and storage costs, longer commute times. Real estate investors are not subject to these costs.

We find that real estate investors play a critical role in the rise in mortgage debt only for the middle and the top of the credit score distribution. The share of mortgage balances of real estate investors rose from 21% to 33% between 2004 and 2007 for quartiles 2-4 of the credit score distribution. Most importantly, we find that the rise in mortgage delinquencies

[1] Agarwal et al. (2016) document clear patterns of product steering by mortgage brokers, who directed borrowers eligible for conventional fixed interest rate mortgages to riskier products with higher margins, increasing default risk even for standard borrowers.

[2] Case, Shiller, and Thompson (2012) show using survey evidence that long-term home price expectations reached abnormally high levels relative to rental rates during the housing boom. Foote, Gerardi, and Willen (2012) and Adelino, Schoar, and Severino (2015) also emphasize the role of overoptimistic house price expectations.

[3] Ghent and Kudlyak (2011) show that foreclosure rates are 30% higher in non-recourse state during the crisis.

[4] Albanesi and Nosal (2015) provide empirical evidence on the relation between consumer bankruptcy, delinquency and foreclosure, while Mitman (2016) develops a quantitative model of bankruptcy where default on unsecured debt is prioritized over mortgage default.

and foreclosures is virtually *exclusively* accounted for by real estate investors. The fraction of borrowers with delinquent mortgage balances grew by 30 percentage points between 2005 and 2008 for the lowest three quartiles of the credit score distribution, and by 10 percentage points for borrowers in the top quartile, while it was virtually constant for borrowers with only one first mortgage. This striking result provides guidance to policy makers interested in understanding the cause of the housing crisis and designing interventions to mitigate and prevent future such episodes.[5]

We also explore the broader macroeconomic implications of our findings, linking them to the literature that emphasizes the role of the collateral channel in the transmission of financial shocks to real economic activity. There is a large theoretical literature on the role of collateral constraints in causing or amplifying swings in economic activity, following the pioneering work of Kiyotaki and Moore (1997). This literature proliferated in response to the financial crisis, leading to numerous theoretical and quantitative contributions.[6] Following the 2007-2009 recession, a large empirical literature also developed. The empirical literature exploits geographical variation to relate mortgage debt growth to the severity of the recession at a regional level, linking the size of the credit boom and the depth of the recession in different geographical areas.[7]

We also examine the behavior of debt and defaults at the zip code level. Because we also have access to individual data, our analysis can provide important insights into the relation between individual and geographically aggregated outcomes, shedding light on the mechanism through which credit growth affects other economic outcomes.[8]

Following Mian and Sufi (2009), we rank zip codes by the initial fraction of subprime borrowers, identifying subprime borrowers as those with Equifax Risk Score below 660. Based on our data, zip codes in the top quartile of the distribution of the fraction of subprime borrowers exhibit larger growth in per capita mortgage balances, confirming previous findings. However, in all quartiles *prime* borrowers are responsible for most of the credit growth. The growth in mortgage debt by subprime borrowers during the boom is modest in terms of balances, and even weaker in terms of number of mortgages and originations. We also show

[5] One implication of our findings is that many renters were displaced as their landlords defaulted on their mortgages, leading to foreclosure of the home. See Bazikyan (2009) and Robinson and Todd (2010) for a discussion.

[6] Some recent contributions include Iacoviello (2004), Guerrieri and Lorenzoni (2011), Berger et al. (2015), Corbae and Quintin (2015), Mitman (2016), Justiniano, Primiceri, and Tambalotti (2016), Kaplan, Mitman, and Violante (2017).

[7] Some examples include Mian and Sufi (2011), Mian, Sufi, and Trebbi (2015), Mian, Rao, and Sufi (2013), Mian and Sufi (2010), Midrigan and Philippon (2011), Kehoe, Pastorino, and Midrigan (2016), Keys et al. (2014).

[8] Most existing analyses have access to either geographically aggregated data or individual data.

that irrespective of the fraction of subprime borrowers, the rise in defaults during the crises is mostly driven by prime borrowers.

Based on our findings with individual level data, we examine the role of the age distribution in different quartiles of the fraction of subprime borrowers. The median age declines by quartile of the fraction of subprime, while the proportion of borrowers younger than 35 rises. We conduct counterfactuals to quantify the role of the age distribution, and find that 83% of the difference in mortgage balance growth between the top and bottom quartile of the fraction of subprime borrowers is accounted for by differences in the age distribution in these zip codes. These results confirm our findings with individual data on the effect of life cycle demand for credit on the observed borrowing by initial credit score during the boom.

The empirical papers that exploit geographical variation to link the size of mortgage debt growth during the credit boom to the depth of the recession (measured in terms of consumption drop or unemployment rate increase) attribute this correlation to the tightening of collateral constraints during the crisis, resulting from mortgage defaults by high risk/low income borrowers with high marginal propensity to consume. Our findings are not consistent with this causal mechanism. We therefore explore additional characteristics of these geographical areas that may explain this correlation. We show that several indicators that are critical to business cycle sensitivity are systematically related to the fraction of subprime borrowers. Zip codes with higher fraction of subprime borrowers are younger, as previously noted, have lower levels of educational attainment and have a disproportionately large minority and African American share of the population. It is well known that younger, less educated, minority workers suffer larger and more persistent employment loss during recessions (see Mincer (1991) and Shimer (1998)). Zip codes with a large fraction of subprime borrowers also exhibit more income inequality. It follows that the aggregation bias that is generated by the fact that, within zip code, prime borrowers experience larger credit growth than subprime borrowers is accentuated.[9] We also examine real estate investor behavior at the zip code level. We find investor activity is mainly accounted for by prime borrowers and that the distribution of number of mortgages is very similar across quartiles. However, in areas with large subprime population (quartiles 3 and 4, specifically), investors exhibit a much larger growth in mortgage balances and a much more pronounced rise in

[9] The distribution of the fraction of subprime borrowers is quite stable at the zip code level, and this is also true for other characteristics salient to business cycle sensitivity, as shown in Section 8. Therefore, the timing of the ranking by fraction of subprime does not change zip code level patterns. However, some aggregate trends, such as the historical decline in wages, labor force participation and employment rates for unskilled, young and minority workers, and the rise in income inequality may influence economic outcomes at the zip code level over time.

foreclosures. These areas are also disproportionately urban and exhibit larger home prime increases during the boom and more pronounced drops during the crisis. The urban nature of these areas may have jointly contributed to the rise in home prices and the intensity of investor activity, resulting from gentrification (see Guerrieri, Hartley, and Hurst (2013)).

Taken together, our findings suggest that using geographically aggregated data does not provide an accurate account of the patterns of borrowing at the individual level. Moreover, the positive correlation between credit growth during the boom and the depth of the recession may be due to other geographical characteristics, such as the prevalence of young, minority or low education workers.

Our findings confirm and expand those in Adelino, Schoar, and Severino (2015) and Adelino, Schoar, and Severino (2017), who show that the growth in mortgage balances during the boom are concentrated in the middle of the income distribution. We show that the large contribution of middle and upper credit score (and income) households to credit growth during the 2001-2007 boom and the stark rise in defaults and foreclosures for these households is primarily driven by real estate investors. Moreover, we explain the role of the positive relation between credit score and age in generating the discrepancy in distribution of debt based on initial and recent credit scores.[10] Our results are also consistent with Foote, Loewenstein, and Willen (2016), who find that the geographical relation in mortgage debt growth and income does not change relative to previous periods during the 2001-2006 credit boom, and there is no relative growth in debt for low income households. Our analysis also reconciles the pattern of borrowing at the individual level and at the zip code level, showing that though mortgage balances grow more in areas with a larger fraction of subprime borrowers, within those areas, debt growth is driven by high credit score borrowers. The fact that zip codes with high fraction of subprime borrowers are associated with low income levels and growth during the boom may be due to demographics, specifically the high fraction of young, low education minority borrowers. High population density and very extreme levels of income inequality in these zip codes exacerbates the aggregation bias associated with using geographically aggregated data.

The rest of the paper is organized as follows. Section 2 describes the data used in this analysis. Section 3 reports the existing evidence on credit growth and default behavior by credit score. Section 4 examines the role of life cycle factors for credit demand and credit scores. Section 5 explores the relation between credit score and income. Section 6 examines

[10]Ferreira and Gyourko (2015) also find that default activity by prime borrowers intensifies during the crisis, however, their definition of prime/subprime borrowers is based on lender characteristics, not on the individual characteristics of the borrower.

the behavior of debt and defaults by recent credit score and Section 7 discusses the role of investors. Section 8 presents the zip code level analysis and Section 9 concludes.

2 Data

We use the Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data (CCP), which is an anonymous longitudinal panel of individuals, comprising a 5% random sample of all individuals who have a credit report with Equifax. Our quarterly sample starts in 1999Q1 and ends in 2013Q3. The data is described in detail in Lee and van der Klaauw (2010). We use a 1% sample for the individual analysis, which includes information for approximately 2.5 million individuals in each quarter. We use the full 5% sample for the zip code level analysis.

The data contains over 600 variables, allowing us to track all aspects of individuals' financial liabilities, including bankruptcy and foreclosure, mortgage status, detailed delinquencies, various types of debt, with number of accounts and balances. Apart from the financial information, the data contains individual descriptors such as age, ZIP code and credit score. For 2009, we also have access to payroll data for a subset of aproximately 11,000 borrowers. The data is described in detail by the Center for Microeconomic Data at the Federal Reserve Bank of New York.[11]

3 Existing Evidence

The credit score is a summary indicator intended to predict the risk of default by the borrower and it is widely used by the financial industry. For most unsecured debt, lenders typically verify a perspective borrower's credit score at the time of application and sometimes a short recent sample of their credit history. For larger unsecured debts, lenders also typically require some form of income verification, as they do for secured debts, such as mortgages and auto loans. Still, the credit score is often a key determinant of crucial terms of the borrowing contract, such as the interest rate, the downpayment or the credit limit.

The most widely known credit score is the FICO score, a measure generated by the Fair Isaac Corporation, which has been in existence in its current form since 1989. Each

[11] A technical note with a description of the dataset is available here: `https://www.newyorkfed.org/medialibrary/interactives/householdcredit/data/pdf/Technical_Notes_HHDC.pdf`. The data dictionary is available at `https://www.newyorkfed.org/medialibrary/interactives/householdcredit/data/pdf/data_dictionary_HHDC.pdf`.

of the three major credit reporting bureaus– Equifax, Experian and TransUnion– also have their own proprietary credit scores. Credit scoring models are not public, though they are restricted by the law, mainly the Fair Credit Reporting Act of 1970 and the Consumer Credit Reporting Reform Act of 1996. The legislation mandates that consumers be made aware of the 4 main factors that may affect their credit score. Based on available descriptive materials from FICO and the credit bureaus, these are payment history and outstanding debt, which account for more than 60% of the variation in credit scores, followed by credit history, or the age of existing accounts, which explains 15-20% of the variation, followed by new accounts and types of credit used (10-5%) and new "hard" inquiries, that is credit report inquiries coming from perspective lenders after a borrower initiated credit application.

U.S. law prohibits credit scoring models from considering a borrower's race, color, religion, national origin, sex and marital status, age, address, as well as any receipt of public assistance, or the exercise of any consumer right under the Consumer Credit Protection Act. The credit score cannot be based on information not found in a borrower's credit report, such as salary, occupation, title, employer, date employed or employment history, or interest rates being charged on particular accounts. Finally, any items in the credit report reported as child/family support obligations are not permitted, as well as "soft" inquiries[12] and any information that is not proven to be predictive of future credit performance.

We have access to the Equifax Risk Score, which is a proprietary measure designed to capture the likelihood of a consumer becoming 90+ days delinquent within the subsequent 24 months. The measure has a numerical range of 280 to 850, where higher scores indicate lower default risk. It can be accessed by lenders together with the borrower's credit report. Mian and Sufi (2009) rank MSA zip codes by the fraction of residents with Equifax Risk Score below 660 in 1996, and Mian and Sufi (2016) rank individuals by their 1997 Vantage Score, the credit score produced by the Experian credit bureau. Based on this approach, they show that zip codes and individuals with lower credit scores exhibit stronger credit growth during the credit boom. We will show that this result is a consequence of the fact that low credit score individuals are disproportionately young and zip codes with a high share of subprime borrowers have a younger population. Individuals who are young exhibit subsequent life cycle growth in income, debt and credit scores. Hence, the growth in borrowing by individuals who have low credit score at some initial date does not necessarily reflect an expansion in the supply of credit, but simply the typical life cycle demand for borrowing.

[12]These include "consumer-initiated" inquiries, such as requests to view one's own credit report, "promotional inquiries," requests made by lenders in order to make pre-approved credit offers, or "administrative inquiries," requests made by lenders to review open accounts. Requests that are marked as coming from employers are also not counted.

To illustrate the results associated with ranking borrowers by their initial credit score, we consider data at the individual and at the zip code level and, following Mian and Sufi (2016) and Mian and Sufi (2009), we rank them by the earliest available date. For individuals, we consider quartiles of the Equifax Risk Score distribution in 1999. For the zip code level analysis, we rank zip codes by the fraction of individuals with Equifax Risk Score lower than 660 in 2001.[13] The 660 cutoff is a standard characterization for subprime individuals, and mirrors the approach in Mian and Sufi (2009).

Figure 1 displays the growth of per capita mortgage debt balances relative to 2001Q3, which is the last quarter of the 2001 recession, according to the NBER business cycle dates. The left panel displays the individual data, where borrowers are ranked based on their average credit score in 1999. The first quartile contains the individuals with the lowest credit score.[14] The right panel presents zip code level evidence. Here, quartile 1 corresponds to the zip codes with the *lowest* fraction of subprime borrowers in 2001, where subprime borrowers are identified as having an Equifax Risk Score lower than 660. The median fraction of subprime borrowers in 2001 is 19% in quartile 1, 32% in quartile 2, 44% in quartile 3 and 60% in quartile 4.[15] All statistics are computed for the population of 20-85 year old individuals.

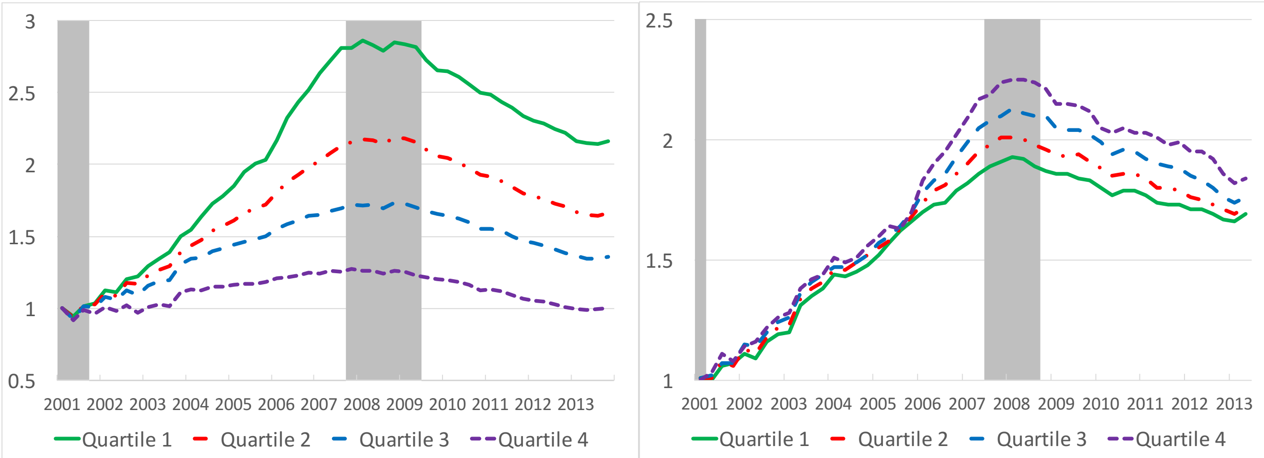
For the individual data, the net growth in per capita mortgage balances between 2001Q3 and 2007Q4 by initial credit score is 146% for quartile 1, 121% for quartile 2, 74% for quartile 3, and 20% for quartile 4. The expansion of mortgage balances continues well into and past the 2007-2009 recession, reaching a peak of 255% for quartile 1, 188% for quartile 2, 111% for quartile 3, and 38% for quartile 4 in 2010Q2. The drop in mortgage balances in the aftermath of the crisis is very dramatic for quartiles 1 and 2, approximately one third from the peak, whereas it is considerably smaller for quartiles 3 and 4, approximately 10% and 5% from the peak.

At the zip code level, the growth of per capita mortgage balances by the fraction of subprime borrowers during the expansion is 58% for quartile 1 (lowest fraction), 64% for quartile 2, 70% for quartile 3, and 77% for quartile 4 (highest fraction). For quartile 4, mortgage balances grow by an additional 5 percentage points during the recession, while they are approximately stable for the other quartiles. Between 2009Q2 and the end of the sample, mortgage balances drop from 19% for quartile 1 to 24% for quartile 4. While at the

[13] We use 2001 rather than 1999 as an initial year to avoid problems relating to missing credit scores for certain zip codes in 1999. The findings using the 1999 ranking are virtually identical.

[14] The cut-off for the individual ranking are 615 for quartile 1, 710 for quartile 2, 778 for quartile 3, and 836 for quartile 4. The cut-off used to identify subprime borrowers with the Equifax Risk Score is 660, therefore, quartile 1 comprises only subprime borrowers, while quartile 2 contains mainly prime individuals and a small subset of subprime.

[15] Section 8 presents more detailed summary statistics at the zip code level.



(a) Individuals: Ranked by 1999 Equifax Risk Score (b) Zip Codes: Ranked by Fraction of Subprime in 2001

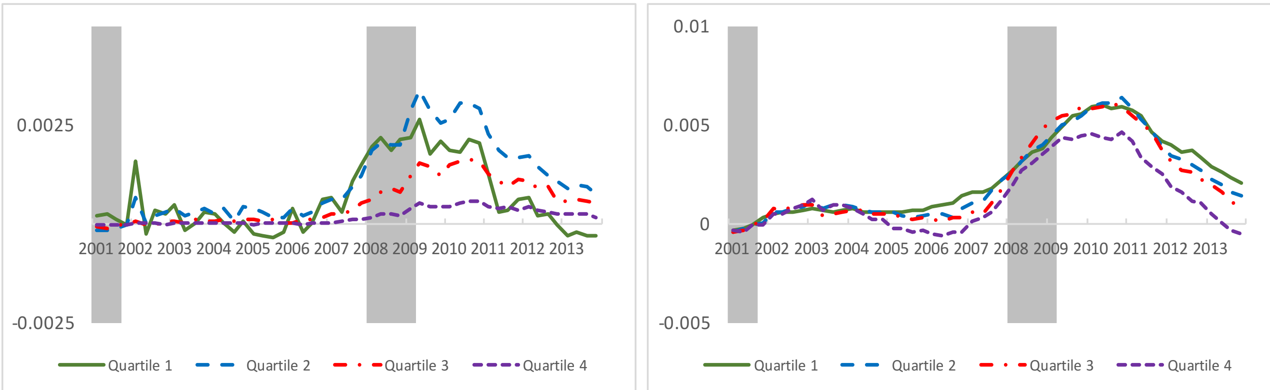
Figure 1: Per capita real mortgage balances, ratio to 2001Q3. Deflated by CPI-U. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

individual level there is much more dispersion across quartiles in mortgage debt growth, both the individual and the zip code level data suggest a stronger growth in mortgage balances for borrowers with low initial credit score in 1999 and zip codes with a large initial fraction of subprime borrowers.[16]

Another basic tenet of the commonly accepted view of the financial crisis is that the growth in credit extended to subprime borrowers during the boom led to a rise in mortgage delinquencies and foreclosures for those borrowers during the crisis. We examine this premise in the next two charts, which display the per capita default rate at the individual and at the zip code level, based on the initial credit score and fraction of subprime ranking. Figure 2 presents the per capita foreclosure rate, specifically the difference in this variable relative to the 2001Q3 value. For individuals (left), the foreclosure rate is virtually constant until the end of 2006, however, during the crisis there is a notable rise in foreclosure rates, especially for borrowers in quartiles 1 and 2 for the 1999 credit score distribution, and to some degree for quartile 3. At the zip code level, the foreclosure rate is constant during the boom and rises during the crisis. The rise is virtually identical for zip codes in quartiles 1-3 of the 2001

[16] The growth in mortgage balances mostly involves intensive margins. If we consider mortgage originations, displayed in Appendix A, the growth is limited only to individuals with 1999 credit scores in quartiles 2-4, and occurs only in the period between 2001Q3 and the end of 2004. A similar pattern prevails at the zip code level, where the growth in originations is negatively related to fraction of subprime borrowers, and there is virtually no growth in the fraction with new mortgage originations in the last year for quartile 4, the zip codes with the largest fraction of subprime borrowers displays the behavior of originations.

fraction of subprime, and slightly lower for zip codes in quartile 4, which have the highest share of subprime borrowers.



(a) Individuals: Ranked by 1999 Equifax Risk Score (b) Zip Codes: Ranked by Fraction of Subprime in 2001

Figure 2: Per capita foreclosure rate, difference from 2001Q3. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

In Section 6, we take a lenders' perspective and estimate mortgage credit growth and defaults at various horizons based on a recent lagged credit score. This approach prevents joint endogeneity between credit score and borrowing behavior, and at the same time provides a more accurate description of borrowers creditworthiness as perceived by lenders at the time in which the loans are extended. The use of a recent credit score to rank individuals, in addition to being closer to industry practices, also better reflects the probability of default at the time of borrowing. In the next section, we examine the link between age, debt and credit scores in detail. This analysis illustrates the issues associated to using initial credit scores to rank individuals and rationalizes the use of recent credit scores by showing that the most important determinant of credit score variation, in addition to age, is income, which is closely related to a borrower's ability to remain current on debt payments.

4 The Role of Age

We now explain why ranking individuals by their credit score 15 years prior, as in Mian and Sufi (2016) and Mian and Sufi (2009) magnifies credit growth for low credit score individuals. Specifically, we will show that low credit score individuals are disproportionately young, and they experience future credit growth, as well as income and credit score growth, due to life cycle factors. As a consequence, their credit score at the time of borrowing is considerably higher than when young. On this basis, we will argue that using a recent lagged credit score provides a better assessment of a borrower's default risk. We will also show that a recent lagged credit score is closely related to income at time of borrowing.

We begin by showing that low credit score individuals are disproportionately young. Table 1 reports the median age by quartile of the credit score distribution, which varies from 39 for quartile 1 to 58 for quartile 4. Figure 3 displays the entire age distribution by credit score quartile. Quartile 1 has the highest share of borrowers between the age of 25 and 40, and the mass shifts right for higher credit score quartiles. For quartile 4, most of the mass is concentrated on borrowers older than 60.

Table 1: Median Age by Credit Score Quartile

Quartile 1	Quartile 2	Quartile 3	Quartile 4
39	44	48	58

Source: Authors' calculation based on Experian Data.

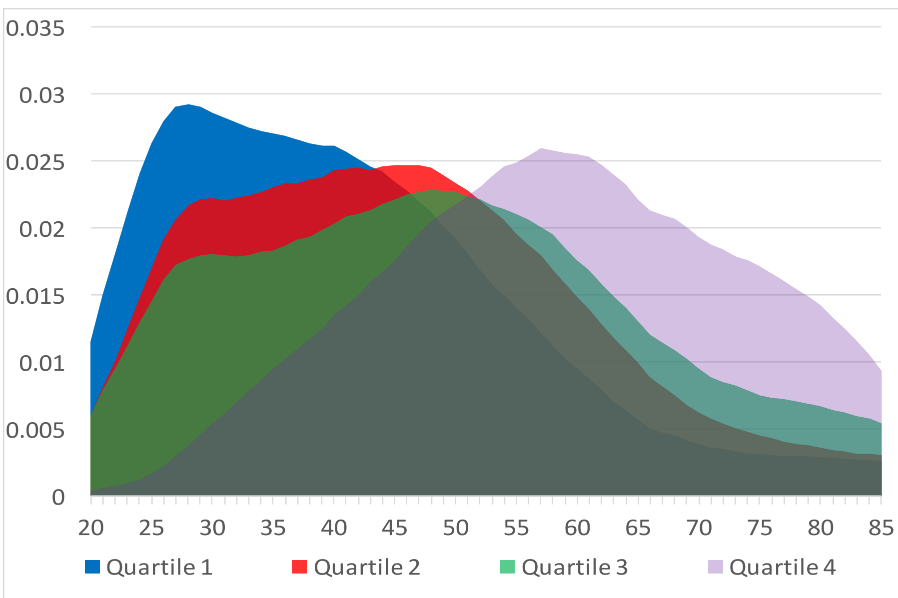


Figure 3: Age distribution by credit score quartile. Source: Authors' calculation based on Experian Data.

Given their relatively young age, and correspondingly short credit history, low credit score individuals in 1999 exhibit credit score growth over time. This is illustrated in figure 4, which plots the current over the 2001 credit score ratio over the sample period by 1999 credit score quartile. For individuals in quartile 1, the credit score grows by more than 10%

between 2001 and the end of 2013. The credit score grows by about 2% for individuals in the second quartile, and is essentially flat for quartiles 3 and 4 of the 1999 credit score distribution.

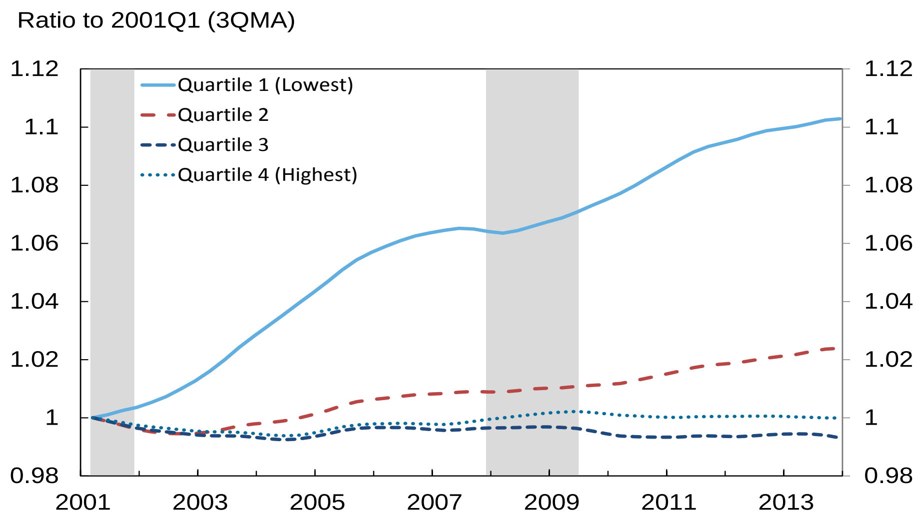


Figure 4: Current credit score as ratio to 2001, by Equifax Risk Score quartile in 1999. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

To more precisely assess the relation between age and the credit score, we estimate age effects for the Equifax Risk Scorein a specification that includes time effects and state fixed effects.[17] Figure 5 plots the estimated age effects between age 20 and 85. The growth in credit score as a function of age is strongest between age 25 and 35, and weakest after age 65. Between the age of 25 and 35, credit score rise by approximately 40 points, and by 60 points between the age 25 and 45. Therefore, an individual in the first quartile of the credit score distribution at age 25 would typically be in the second quartile at 35 and in the third at 45.

[17] U.S. legislation prevents credit scoring agencies to use location as a factor in their models, even if location may affect default behavior. However, we include state effects due to the sizable cross state variation in important regulations regarding foreclosure, bankruptcy, wage garnishment and other factors that could affect the incidence of financial distress and the resulting credit score distribution.

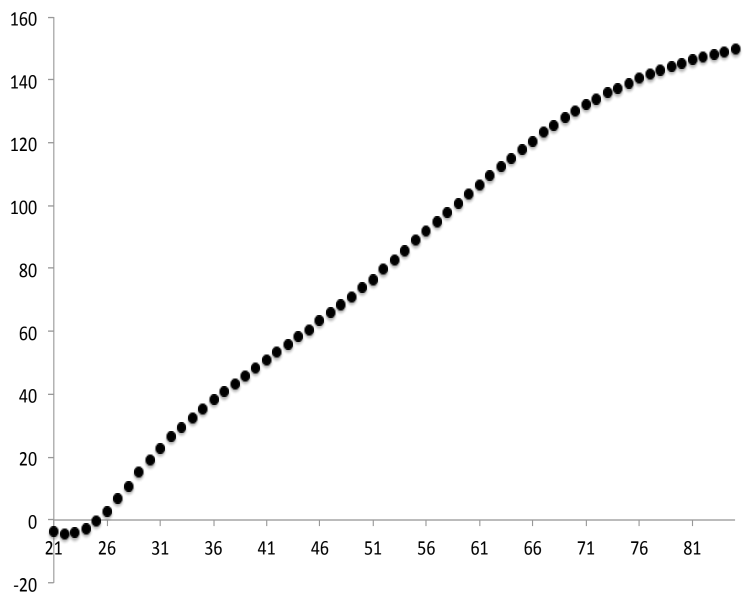


Figure 5: Estimated age effects for the Equifax Risk Score. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

4.1 Counterfactuals

To further illustrate the role of the life cycle in mortgage borrowing by initial credit score, we construct two counterfactuals using the individual data.[18] The first counterfactual eliminates differences in the age distribution across quartiles, while the second removes life cycle effects on the growth in mortgage balances.

Let $\pi^{i,j_{1999}}$ be the fraction of individuals in age bin $i = 1, 2, ...$ and Equifax Risk Score quartile $j = 1, 2, 3, 4$ in 1999. We consider the following age bins: $1 = [20, 35)$, $2 = [35, 45)$, $3 = [45, 55)$, $4 = [55, 64)$ and $5 = [65, 85]$. Further, let $\overline{m}_t^{j_{1999}}$ be per capita mortgage balances of borrowers in quartile $j = 1, 2, 3, 4$ of the 1999 credit score distribution in quarter t and $m_t^{i_{1999},j_{1999}}$ be per capita mortgage balances for borrowers in age bin i and Equifax Risk Score quartile $j = 1, 2, 3, 4$ in 1999 at quarter t. Then:

$$\overline{m}_t^{j_{1999}} = \sum_i \pi^{i_{1999},j_{1999}} \times m_t^{i_{1999},j_{1999}}. \tag{1}$$

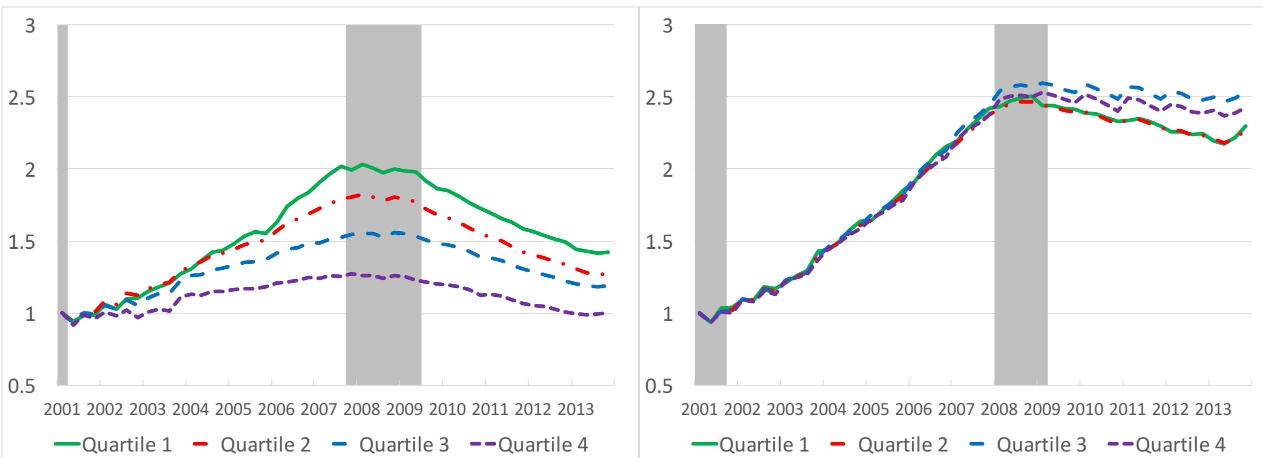
Counterfactual 1: Age Distribution We first calculate a counterfactual designed to isolate the role of differences in the age distribution of across the 1999 credit score quartiles. To do so, we impose the quartile 4 age distribution on quartiles 1-3. That is, for each

[18]We will consider similar counterfactuals at the zip code level in Section 8.

$j = 1, 2, 3$, we compute:

$$\tilde{m}_t^{j_{1999}} = \sum_i \pi^{i_{1999}, 4_{1999}} \times m_t^{i_{1999}, j_{1999}}. \tag{2}$$

Panel (a) in figure 6 plots the resulting counterfactual growth in per capital mortgage balances relative to 2001. Compared to the actual growth rate of mortgage balances displayed in figure 1, mortgage balance growth is much weaker for the counterfactual series than for the actual for quartiles 1-3. However, even in the counterfactual, mortgage balance growth is inversely related to the initial quartile of the credit score distribution, consistent with Mian and Sufi (2016). Based on this approach, we can compute the fraction of the difference between quartile 1 to 3 and quartile 4 in cumulative 2001Q3-2007Q4 growth in mortgage balances accounted by the difference in the age distribution relative to quartile 4. This amounts to 26% for quartile 1, 20% for quartile 2 and 14% for quartile 3.



(a) Counterfactual 1

(b) Counterfactual 2

Figure 6: Per capita real mortgage balances, ratio to 2001. Deflated by CPI-U. Counterfactual 1 attributes to all quartiles the age distribution of quartile 4. Counterfactual 2: Attributes to borrowers in a given age bin in 1999 the mortgage balances of borrowers in that age bin in the current quarter. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

Counterfactual 2: Life Cycle Effects The second counterfactual is designed to isolate the impact of life cycle factors for borrowers in different quartiles of the 1999 Equifax Risk Score distribution. We remove life cycle effects by maintaining borrowers at their age in 1999. This is achieved by attributing to borrowers in age bin i and credit score quartile j in 1999 the debt balances of individuals in age bin i and credit score quartile j in each subsequent quarter t. That is:

$$\hat{m}_t^{j_{1999}} = \sum_i \pi^{i_{1999}, j_{1999}} \times m_t^{i_t, j_t}. \tag{3}$$

Panel (b) in figure 6 displays the resulting counterfactual mortgage balance growth relative to 2001. Based on this counterfactual, there is virtually no difference in mortgage balance growth across quartiles, which is consistent with differences in life cycle effects accounting for most of the difference in borrowing across 1999 credit score quartiles. The counterfactual remove life cycle effects but captures time effects in mortgage borrowing by age. The strong growth in the counterfactual balances for all quartiles suggests that there was a generalized growth in mortgage borrowing for all age groups. This may be driven by an increase in the supply of credit or by the rise in housing values, which necessarily increases the size of the typical mortgage. However, the difference across quartiles of the initial credit score distribution is mostly accounted for by life cycle factors.

Taken together, these results suggest that life cycle effects in borrowing are very strong and sizably affect mortgage debt growth especially for individuals at the bottom of the 1999 credit distribution.

5 Credit Scores, Income and Debt Over the Life Cycle

This section documents the life cycle relation between income, credit score and borrowing. Based on this analysis, we argue that a recent lagged credit score should be used to assess a borrower's probability of default, as this measure better reflects default risk at the time of borrowing. In addition, we show that the life time evolution of credit score and debt is closely related to the lifetime evolution of income. Since the ability to make timely payments on outstanding debt critically depends on income at the time of borrowing and throughout the life of the loan, the tight relation between a recent credit score and contemporaneous income conditional on age supports the notion that it should be used as an indicator of default risk.

To study the relation between credit scores, borrowing and income over the life cycle, we use payroll information- so called Worknumber data- for 2009 from a large income verification firm, linked to the Equifax credit files. The income data is available for a nationally representative subsample of over 11,000 individuals in the credit panel. We construct a total labor income measure using information on pay rate and pay frequency. Appendix B reports detailed information on the construction of this income measure, and shows that the distribution of our income measure is comparable by age and location to that of similar measures obtained from the CPS.

5.1 Life-Cycle Relation

The availability of labor income data for a subsample of borrowers in 2009 and their full credit profile enables us to assess the lifecycle relation between income, credit score and debt.

We begin by relating the debt and credit score evolution from 1999 to 2009, by 2009 total labor income and 1999 age. We find that young borrowers in 1999 with high income in 2009 exhibit the largest growth in mortgage and total balances, and credit score between 1999 and 2009. Figure 7 illustrates this pattern for the 25-34 year olds in 1999 that are in our Worknumber Data sample for 2009. The charts clearly show that 25-34 year olds in 1999 who are in the top quintile of the labor income distribution in 2009 exhibit a much stronger growth in credit scores and mortgage balances. For those in the bottom quintile, the credit score rises by only 10 points between 2001 and 2009, while it grows by 40 points for those in the top quintile. Similarly, (real) mortgage balances grow by a factor of 3.3 between 2001 and and the start of the recession for the top quartile, and by a factor of 2.4 for the bottom quintile. The growth in both credit scores and mortgage debt balances is monotonically increasing in 2009 income quintile. We report only quintile 1 and 5 for clarity.

Figures 8 and 9 present the same variables for 35-44 year olds in 1999 and 45-54 year olds in 1999. The same qualitative patterns apply, however, the magnitude of the increase in both credit score and mortgage balances between 2001 and 2009 is much smaller, as credit demand is much smaller for these age groups.

Our second exercise relates credit score growth between 1999 and 2009 to income levels and debt levels in 2009 for borrowers in the bottom quartile of the credit score distribution in 1999. Table 2 summarizes these results. The columns correspond to the quartiles on the 2009 credit distributions for borrowers (of any age) that were in the first quartile of the credit score distribution in 1999. We report mean income and mean total debt balances. Clearly, 2009 income and total debt balances are increasing in the 2009 credit score, even if all these borrowers begin in the bottom quartile of the credit score distribution in 1999.

This evidence speaks directly to the relation between income and debt during the credit boom. Using zip code level data, Mian and Sufi (2009) show that during the period between 2001 and 2006, the zip codes that exhibited the largest growth in debt were those who experiences the smallest growth in income. They argue that the negative relation between

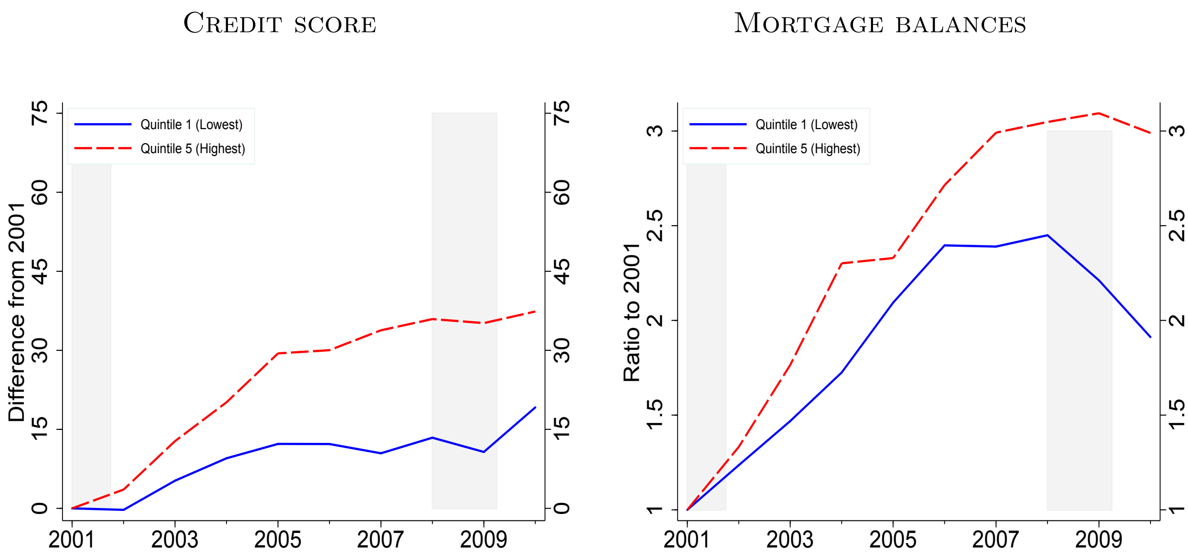


Figure 7: Equifax Risk Score and mortgage balances for 25-34 yo in 1999 by their 2009 Worknumber total annual labor income quantile. Difference with 2001 (credit score) and ratio to 2001 (balances). Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Table 2: Relation between Credit Score, Income and Debt Balances

2009 credit score	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Debt balances	\$38k	\$74k	\$126k	\$213k
Income	\$39k	\$47k	\$57k	\$62k

Mean income and total debt balances by 2009 Equifax Risk Score quartile for individuals in the first quartile of the 1999 Equifax Risk Score distribution. Worknumber total annual labor income for restricted Worknumber sample. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

debt growth and income growth at the zip code level over that period is consistent with a growth in the supply of credit to high risk borrowers. We show that this negative relation does not hold for individual data. The differences in credit growth between 2001 and 2009 are positively related to life cycle growth in income and credit scores. Moreover, debt growth for young/low credit score borrowers at the start of the boom occurs primarily for individuals who have high income by 2009, and the growth in income is associated in a growth in credit score. Older individuals in 1999 exhibit much lower subsequent debt and credit score growth, still positively related to their income in 2009. The strong correlation between recent

Credit score | Mortgage balances

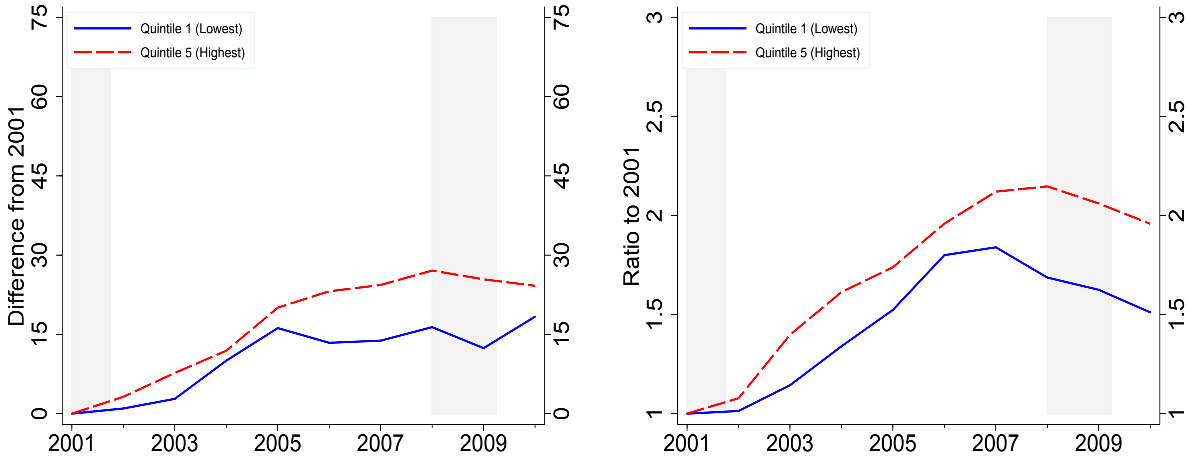


Figure 8: Equifax Risk Score and mortgage balances for 35-44 yo in 1999 by their 2009 Worknumber total annual labor income quantile. Difference with 2001 (credit score) and ratio to 2001 (balances). Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Credit score | Mortgage balances

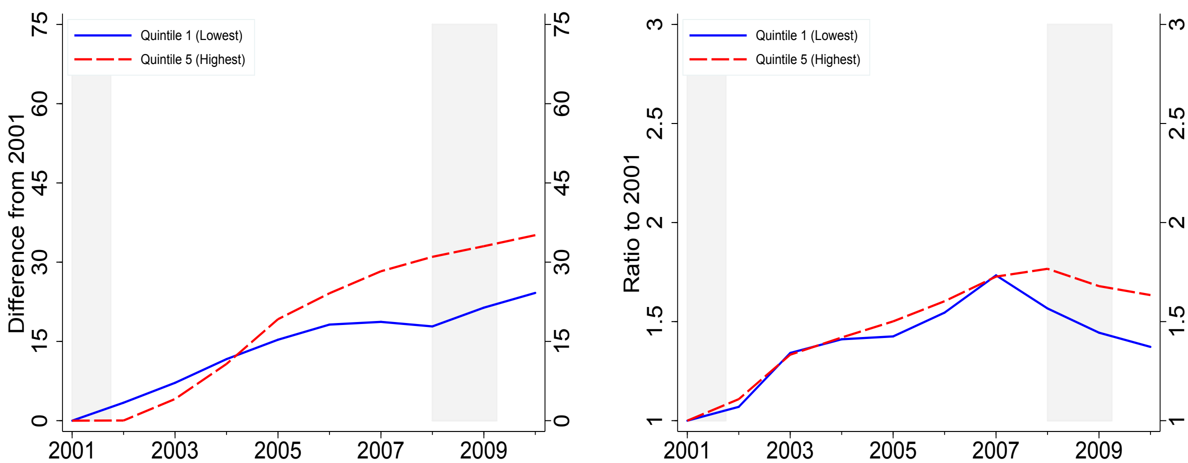


Figure 9: Equifax Risk Score and mortgage balances for 45-54 yo in 1999 by their 2009 Worknumber total annual labor income quantile. Difference with 2001 (credit score) and ratio to 2001 (balances). Source: Authors' calculations based on FRBNY CCP/Equifax Data.

credit scores and income suggests recent credit scores are better indicator of default risk. Appendix C reports estimates of the relation between the growth in total debt balances and total income using the PSID over the 1999-2007 period. The PSID analysis confirms the positive relation between income growth and growth in debt balances in 2001-2006.

The positive relation between income growth and debt growth during the credit boom casts doubt on the notion that there was an increase in the supply of credit, especially to high risk borrowers. Instead, it is more likely that the rise in house prices caused an increase in mortgage balances. This is confirmed by the fact that the fraction of borrowers with mortgages did not rise for any quartile of the credit score distribution, as we show in Section 6.1.1 below.

5.2 Cross-Sectional Relation

We also estimate the cross-sectional relation between credit scores and income, conditional on age. To evaluate the relation between income and credit score, we regress the 8 quarter lagged credit score on income, income square, age, age square, and interactions between age, income and state fixed effects.[19] Specifically, we estimated the following:

$$CS^i_{2009-h} = \alpha + \beta_1 y^i_{2009} + \beta_2 \left(y^i_{2009}\right)^2 + \gamma_1 \text{age}^i_{2009} + \gamma_2 \left(\text{age}^i_{2009}\right)^2 + \text{interactions} + \varepsilon^i_{2009} \quad (4)$$

where i denoted individual borrowers, CS^i_{2009-h} is a borrower's credit score in quarter $2009 - h$, and h denotes the leads/lags in the credit score relative to income, with $h \in \{-8Q, -4Q, 0, 4Q, 8Q\}$. The coefficient α corresponds to the constant and y^i_{2009} is a borrower's total labor income in 2009.

Figure 10 displays the in sample projected relation between the 8 quarter lagged credit score and income for different age levels. The range of income levels varies by age as they do in our sample. Clearly, credit scores are strongly positively related to income given age. The intercept of the relation increases with age while the slope declines with age. We estimate the same specification for the 4 quarter lagged, current, and 4 quarter and 8 quarter ahead credit score, with very similar results.

The strong and positive relation between recent credit scores to to income, given age, provides a rationale for considering recent credit scores as an indicator of default risk at the time of borrowing, since income is a key determinant of a borrower's ability to make timely

[19]Since the credit score is bounded above, we use a truncated regression approach. Standard errors are clustered at the state level.

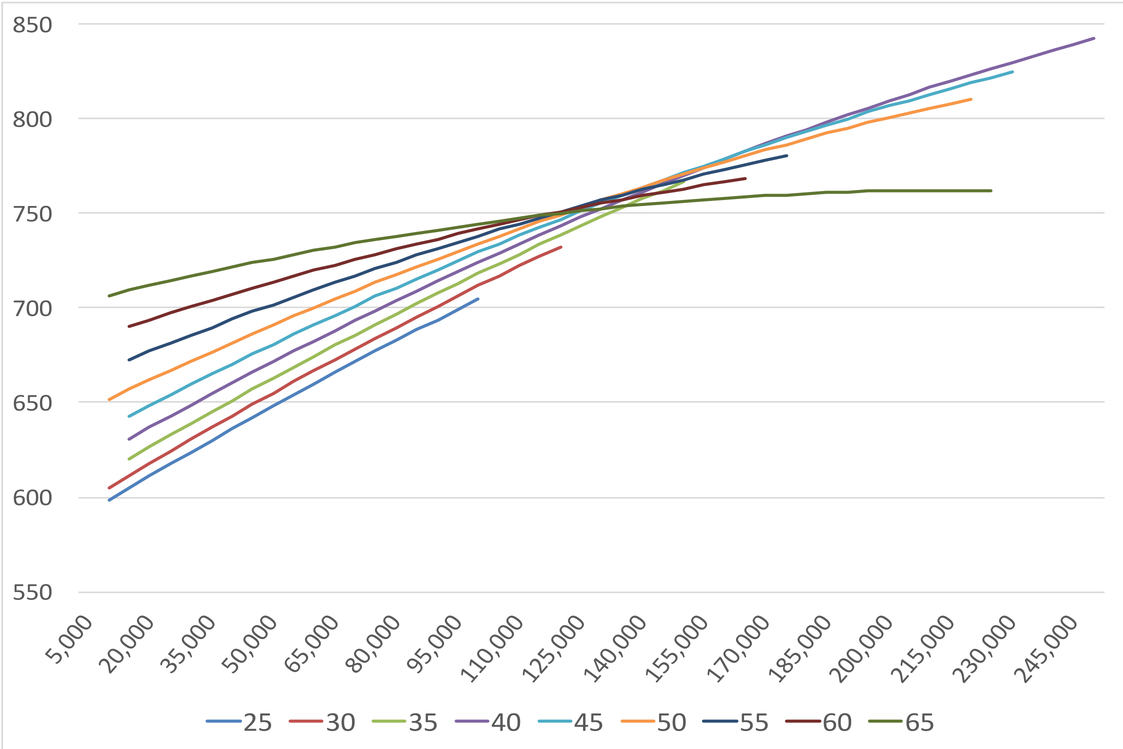


Figure 10: Predicted 8Q lagged Equifax Risk Score by age and 2009 Worknumber total annual labor income, for age specific 1-99 percentile of income range. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

payments.

6 Debt and Defaults by Recent Credit Score

We now present our approach to characterizing the distribution of debt growth during the boom and defaults during the crisis based on recent credit scores. We adopt a lender's perspective, and relate future credit growth at various horizons to a recent lagged credit score to capture the credit score at the time of borrowing. This strategy is based on the observed patterns of credit extension in the U.S. An increase in debt balances between two time periods, say one year, would arise due to either a new loan or credit line, or to an increase in the maximum balance on an outstanding loan or credit line. In most cases, the borrower would have applied for the loan or the balance increase, leading the lender to check the borrower's credit score. Given that our data is quarterly and for most types of debt such requests are processed in a matter of days, the credit score in the quarter before the increase in debt balances is the best proxy of the one that would be available to the lender at the time of application.

Lenders often may also check some other variables in an applicant's credit history, such as the number of missed payments or credit utilization in the last 1-2 years. These factors would be reflected in changes in the credit score in the corresponding period. Changes in the credit score before the application date may also be motivated by the intention to borrow. For example, individuals intending to finance a car purchase may be motivated to improve their credit score in the period leading up to their purchase or to delay the purchase until their credit score has improved- for example by paying down credit card balances- in order to secure better terms. For these reasons, we also include the change in the credit score as an explanatory variable. For most unsecured debt and auto loans, lenders would not typically verify a borrower's income. For mortgage loans, lenders typically also verify a lender's recent income history. We do not have access to income, therefore, we only use the credit score in the last quarter and the change in the score between the last quarter and some previous dates as our main explanatory variables. As we have shown, income and recent credit score are positively related, conditional on age.

Our baseline specification is:

$$\Delta B^i_{t,t+h} = \sum_{j=1,2,3,4} \alpha(j_{-1}) + \eta \Delta CS^i_{t-1,t-1-k} + \text{ time fe} + \text{ age fe} + \text{ interactions} + \varepsilon^i_t, \tag{5}$$

where i denotes and individual, t denotes a quarter, $\Delta B^i_{t,t+h}$ is the change in balances between quarters t and $t+h$, and $h \in \{4, 8, 12\}$ is the horizon. The explanatory variables are $\alpha(j_{-1})$ which is a fixed effect for the 1 quarter lagged quartile of the credit score distribution and $\Delta CS^i_{t-1,t-1-k}$, which represents the change in credit score between $t-1$ and $t-1-k$, with $k \in \{4, 6\}$ length of the credit score history considered. The baseline specification includes quarter effects, age effects and their interaction with the 1 quarter lagged credit score quartile.

Our estimates show that during the boom credit growth was highest for borrowers in the middle and top quartiles of the one quarter lagged credit score distribution, at all horizons. We find that past changes in the credit score also have a sizable effect on subsequent balance growth. Consistent with our analysis in Section 4, we find strong age effects in balance growth but *only* for individuals in quartile 2-4 of the 1 quarter lagged credit score distribution. We also find that the growth in delinquent balances during the crisis is concentrated in the middle of the credit score distribution.

In the rest of this section we report our findings. We complement our regression based evidence with an analysis of extensive margins, such as mortgage originations, first mortgages, foreclosures by eight quarter lagged credit scores. We find there is no growth in the fraction with first mortgages or with new mortgage originations for borrowers in the first quartile of the eight quarter lagged credit score distribution. Additionally, consistent with Adelino, Schoar, and Severino (2015), we find that the distribution of credit scores at originations is virtually constant throughout the boom. Further, we show that the rise in mortgage defaults and foreclosures is greatest for borrowers in quartiles 2 and 3 of the eight quarter lagged credit score distribution.

6.1 Mortgage Debt Growth

This section presents our regression results for mortgage balances. In Appendix D, we report results for total debt balances, as well as some robustness analysis.

Our baseline specification uses the 8 quarter ahead change in mortgage balances as the dependent variable and includes the 4 quarter change in credit score as a regressor. Table 3 reports the fixed effects estimates, and figure 11 presents the age adjusted interactions between the time effects and each quartile of the 1 quarter lagged credit score distribution. The credit score quartile fixed effects show a non-monotone pattern, with quartile 2 and 3 estimates of the average eight quarter ahead mortgage balance change above $9,000, approximately three times as large as the value for the first quartile, and approximately double the value for quartile 4. The coefficients on the change in the credit score distribution are $50 for the 4 quarter lag and $51 for the 6 quarter lag, and highly significant. To understand the economic impact of these estimates, it is useful to consider how common credit events may affect the credit score over 4 or 6 quarters. Based on commonly used credit score simulators, a borrower with credit score of 610 (in quartile 1) with a $15,000 balance on revolving trades (credit cards) and no mortgages can increase her credit score by 30 points over a one year period by paying down 5% of her balance every month for 12 months. Paying her bills on time over the same period may improve her credit score by only 5 points, while taking out a mortgage would not affect her credit score. The simulated changes in credit score corresponding to these events have smaller effects for borrowers in quartile 2, and negligible effects for borrowers in quartiles 3 and 4, based on the same simulators. Missing a payment reduces the credit score by at least 35 points instantly, irrespective of initial score. Based on these simulations, common positive credit event inducing a 5-30 point increase in the credit score over a 4-6 quarter period could change the predicted eight quarter ahead change in balances by $250-1,530 for low credit score borrowers, while a missed payment could reduce the predicted change in balances by $1,750-1,785 for all borrowers. These magnitudes are sizable, especially for borrowers in quartile 1, for whom the changes are approximately equal to half of their eight quarter ahead predicted change in balances.

Table 3: Mortgage Balance Growth

Dependent Variable: 8Q Ahead Mortgage Balance Change					
1Q lagged CS Quartile Effects				Credit Score Change	
1	2	3	4	4Q	6Q
221	7540	7451	3297		
3,182	9,559	9,291	4,803	50	
4,129	10,164	9,787	5,173		51

Estimated 1Q lagged Equifax Risk Score quartile effects and coefficients for 4Q, 6Q past change from 1Q lagged score in balance change regressions, in USD. Baseline specification. All estimates significant at 1% level. Sample period 2001Q1-2011Q4. Number of obs. 64,588,488. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 11 presents the predicted age adjusted time effects for each quartile of the 1 quarter lagged credit score distribution. The estimated age effects, presented in figure 12 and the quartile fixed effects are used to adjust the level of the raw quartile time effects, using the quartile specific age distribution. According to these calculations, quartile 1 borrowers experience a steady $5,000 8 quarter ahead change in balanced throughout the credit boom. The change in balanced for quartile 4 is about twice than the change in quartile 1 between 2001 and late 2003, when it start to rapidly decline. Quartiles 2-3 show a very similar increase in balances throughout the sample period. Their 8 quarter ahead change in balances rises from $10,000 in 2001 to a peak of approximately $17,000 at the end of 2005. Starting at the end of 2005, all quartiles experience a sharp decline in the eight quarter ahead growth in mortgage balances, which bottoms out to approximately zero in 2009Q1 for quartile 2-4. For quartile 1, the growth in balances continues to decline until 2009Q4, when it reaches a minimum of -$8,000. This finding is particularly striking, since quartile 1 borrowers experienced very modest mortgage balance growth during the boom, suggesting the the costs in terms of credit contraction were mainly borne by borrowers who reaped little benefit from the previous boom. Part of the decline in mortgage balances for quartile 1 may also be driven by charge offs by borrowers who had higher credit scores during the boom and drop into quartile 1 during and after the housing crisis as a consequence of mortgage defaults.[20]

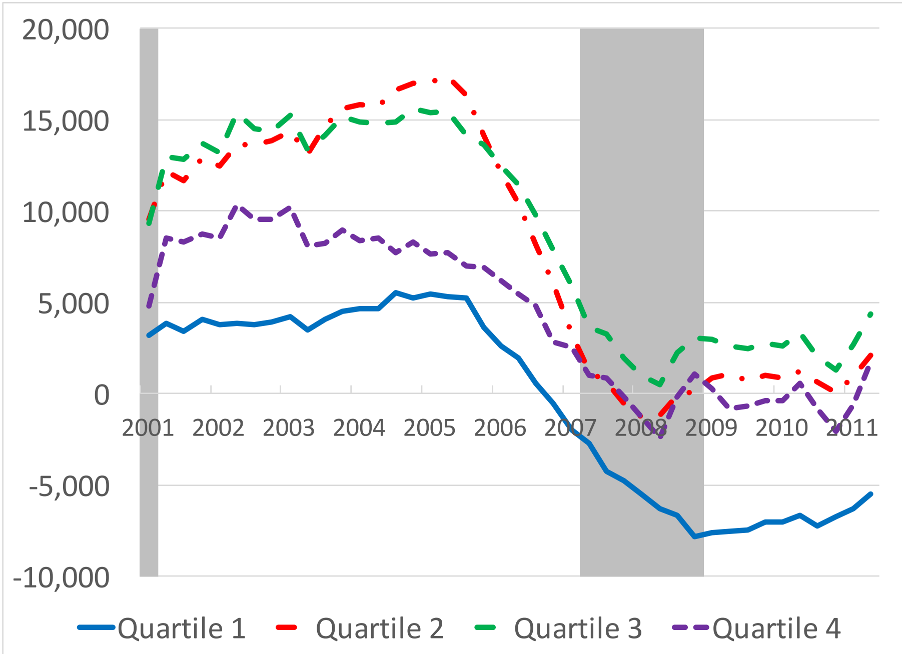


Figure 11: Predicted age adjusted time effects by 1Q lagged Equifax Risk Score quartile from balance change regressions obtained from the time×credit score quartile interactions. Dependent variable is the 8Q ahead change in per capita mortgage balances in USD. Sample period 2001Q1-2011Q4. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 27 in Appendix D presents the difference between the time×quartile effect interactions for quartiles 2-4 relative to quartile 1, with 5% confidence intervals. These charts clarify that the difference in time effects across quartiles is sizable and highly significant throughout the sample period.

The findings are very similar using the 4 quarter ahead and 12 ahead change in mortgage balances, including additional robustness results for the 8 quarter ahead change in mortgage balances. Table 13 in Appendix D summarizes the implied cumulated mortgage debt balance change for 2002-2006 and 2007-2010 by sub period using the 4, 8 and 12 quarter ahead mortgage balance change regressions.

Role of Age Figure 12 presents the estimated age effects, obtained from the interaction with the quartile fixed effects. The estimates are consistent with those in Section 4, since the cumulated growth in mortgage debt balances between age 20 and age 30, which corresponds

[20]This is consistent with the behavior of delinquent balances, described in Section 6.2.1.

to peak growth over the life cycle, based on the 8 quarter ahead change is approximately \$35,000.[21] However, the interactions between the quartile and age effects suggest that only borrowers in quartiles 2-4 of the 1 quarter lagged credit score distribution experience a life cycle growth in mortgage balances, and the size of this growth is increasing with the credit score quartile. This result is consistent with our findings in Section 5, where we show that the life cycle growth in mortgage balances is closely related by the life cycle growth in income and credit scores.

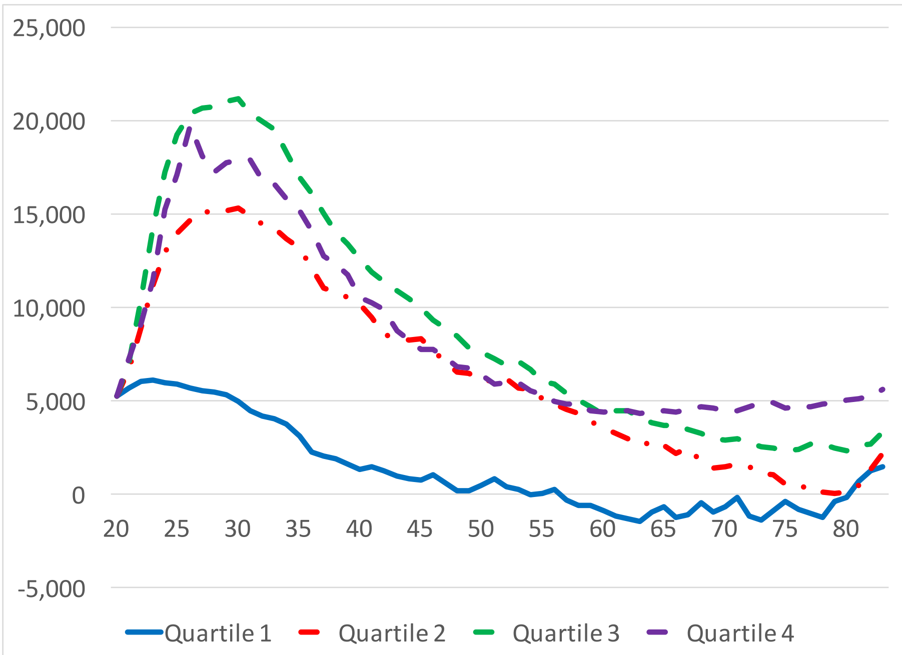


Figure 12: Estimated age effects from balance change regressions obtained from age×credit score quartile interactions. Dependent variable is the 8Q ahead change in per capita mortgage balances in USD. Sample period 2001Q3-2011Q4. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

6.1.1 Homeownership and Originations

To corroborate the regression analysis on mortgage balances, we also examine borrowing behavior by recent credit score on the extensive margin. Consistent with our baseline regression specification, we rank borrowers by their 8 quarter lagged credit score. Our findings are robust to alternative recent rankings, such as 4 quarter lagged credit score.

[21] This estimate is obtained by averaging out the quartile fixed effects and adding them to the age effects.

Figure 13 presents the fraction with first mortgages, which can be taken to correspond to the home ownership rate in these data, and the fraction with new originations by 8 quarter lagged credit score. Both the fraction with first mortgages and the fraction with new mortgage originations are virtually constant for quartile 1 during the boom. The fraction with first mortgages grows by approximately 10 percentage points between 2001Q3 and 2007Q4 for quartiles 2 and 3, and by about 6 percentage points for quartile 4. Quartiles 2-4 experience a boom in new originations between 2001 and 2004Q1. The fraction with new mortgage originations rises from just below 20% in 2001Q1 to 23% and 27% at the peak for quartiles 2 and 3, respectively. For quartile 4, it rises from 12% in 2001 to 22% in 2004Q1. The sizable rise in mortgage originations for prime borrowers early in the boom combined with the modest rise in the fraction of borrowers with first mortgages for that period suggests that most of the originations reflect refinancing activity[22] or real estate investing, as we document in Section 7.1 below. The fraction with new mortgage originations drops thereafter for quartiles 2-4, reaching lows of 6-8% in 2009Q2, when it starts to slowly recover. For quartile 1, the fraction of borrowers with new originations declines between 2001 and 2006, reaching 8% in 2006, and then stabilizes between 2006Q1 and 2007Q1. It then decline to close to zero by the end of 2009.

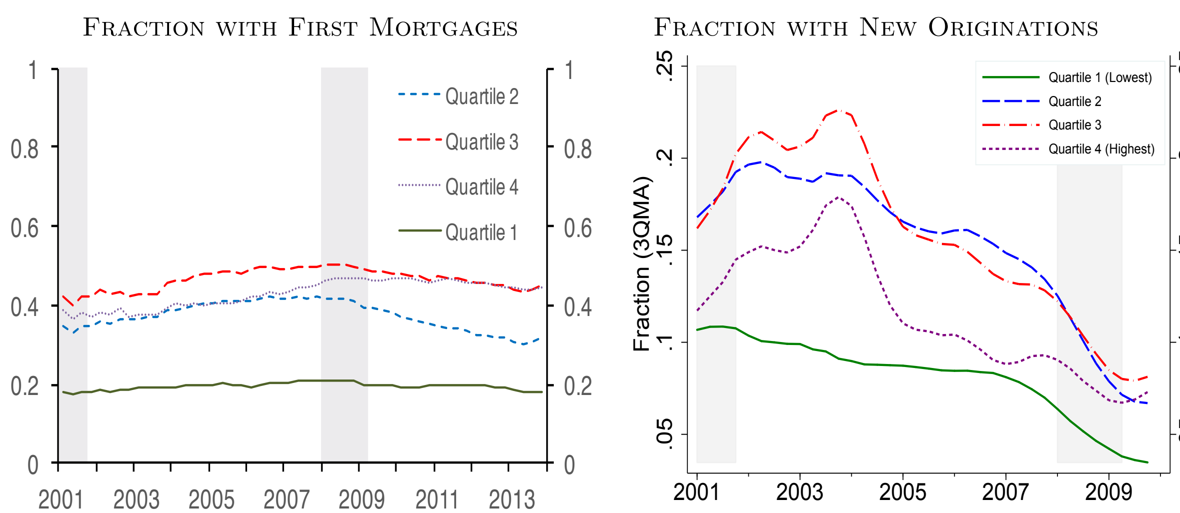


Figure 13: Fraction with first mortgages and fraction with new mortgage originations by 8Q lagged Equifax Risk Score quartile . Quartile cutoffs: 615, 720, 791, 840.
Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 14 presents the distribution of credit scores at originations for each quarter of

[22]Chen, Michaux, and Roussanov (2013) and Bhutta and Keys (2016) document the rise of refinancing activity during the credit boom and argue that in 2001-2004 it was mainly driven by lower mortgage rates.

our sample period. The fraction of new mortgage originations attributable to borrowers in quartiles 1 and 3 of the credit score distribution remains virtually constant throughout the sample period. There is a modest rise in the fraction of originations to borrowers in quartile 2, from 23% in 2003Q4 to a peak of 30% in 2006Q4, after which they drop to a low of 20% in 2011Q2. The fraction of new originations to borrowers in quartile 4 of the credit score distribution peaks at 28% in 2003Q3 during the boom, but rises during the crisis from 20% in 2006Q4 to 31% in 2011Q2 and then stabilizes. This rise reflects the tightening of lending standards during the crises.[23]

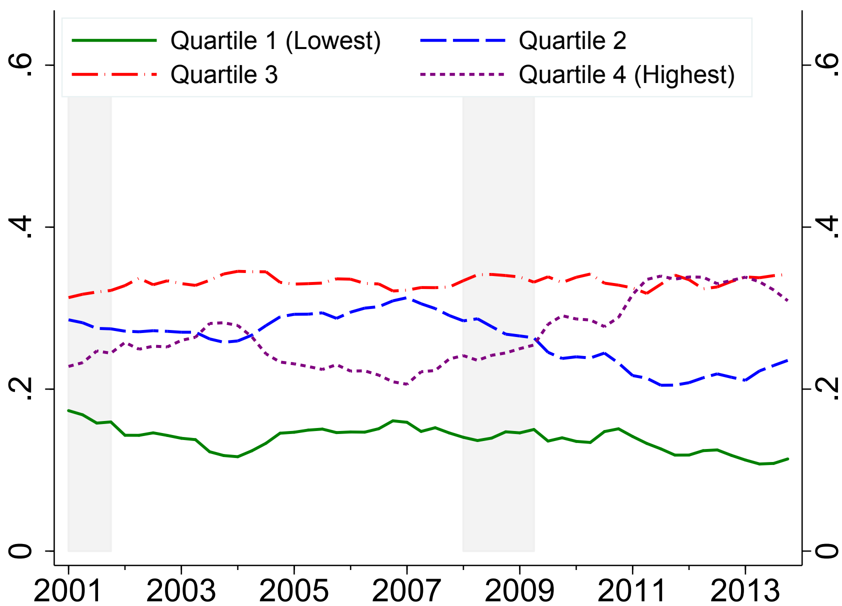


Figure 14: Individuals with a new mortgage origination. Fraction in each quartile of the 4Q lagged Equifax Risk Score distribution. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

6.2 Defaults

We now examine default activity by recent credit score. As for debt growth, we use regression analysis to examine the behavior of delinquent balances over the sample period, and then use a recent credit score ranking to the examine the distribution of mortgage delinquencies and foreclosures.

[23]See Brown et al. (2014) for a discussion.

6.2.1 Delinquent Balances

We follow the same regression specification described in Section 6.1 for the 8 quarter ahead change in 90+ days delinquent mortgage balances. The estimated quartile fixed effects are presented in Table 4. The average 8 quarter ahead change in delinquent balances falls with the 1 quarter lagged credit score, with the estimated effects for quartiles 3-4 about half as large as for quartiles 1-2. As for debt growth, the contribution of past credit score changes to the growth in delinquent balances is non negligible especially for low credit score borrowers. For example, a 30 point increase in the credit score over a 4-6 quarter period corresponds to an increase average delinquent balances of approximately $1,000. The fact that a past increase in the credit credit score corresponds to an increase in predicted delinquent balances is due to the fact that such an increase is associated in an increase in average balances over the same horizon, as shown in Table 3.

Table 4: Delinquent Mortgage Balance Growth

Dependent Variable: 8Q Ahead 90+ Days Delinquent Mortgage Balance Change (USD)					
1Q lagged CS Quartile Effects				ΔCS_{-1}	
1	2	3	4	4Q	6Q
-3378	-435	-254	-1367		
371	91	-3	41	26	
781	287	162	166		27

Estimated 1Q lagged Equifax Risk Score quartile effects and coefficients for 4Q, 6Q past change from 1Q lagged Riskscore in balance change regressions. Baseline specification. All estimates significant at 1% level. Sample period 2001Q3-2011Q4. Number of obs. 64,588,488. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 18 presents the age adjusted 8 quarter ahead change in delinquent mortgage balances by 1 quarter lagged credit score. The growth in delinquent balances is very close to zero between 2001 and the end of 2004 for quartiles 1-3, and hovers around -$1,000 for quartile 4. There is a very large rise in the 8 quarter ahead change in delinquent balances for all quartiles starting in late 2004, but the growth in delinquent balances is considerably larger for quartiles 2-4. For quartiles 2 and 4, the growth in delinquent balances peaks at the start of 2007, when it reaches $5,500 and $4,200, respectively. For quartile 3, the peak occurs in early 2008, at $3,000. Quartile 1 exhibits the smaller growth in delinquent balances during the crisis. The peak in delinquent balance growth for these borrowers occurs at the end of

2006, with a growth of approximately \$2,000. The growth in delinquent balances declines for all borrowers during the 2007-09 recession and for about a year after. For quartiles 2-4, the growth in delinquent balances goes back to pre crisis values by 2011, whereas it hovers around -\$6,000 in 2009 and 2010 for quartile 1. This pattern is driven by the large decline in mortgage balances for borrowers in the first quartile, discussed in Section 6.1, and may in part be driven by charge offs. We find similar results for the change in delinquent balances at 4 and 12 quarter ahead horizon.[24]

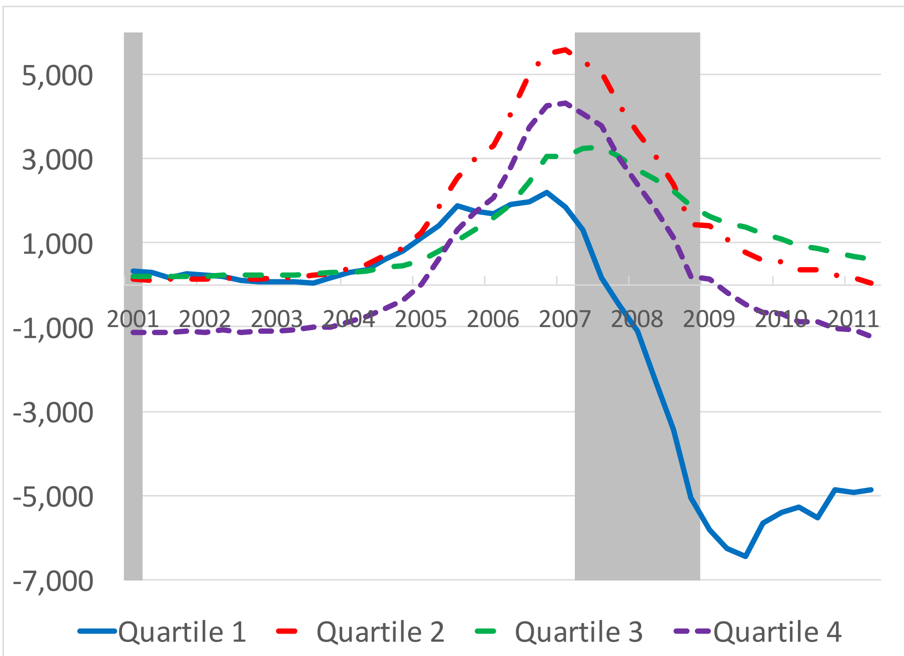


Figure 15: Predicted age adjusted time effects by 1Q lagged Equifax Risk Score quartile from balance change regressions. Baseline specification. Dependent variable is the 8Q ahead change in per capita 90+ days delinquent debt balances in USD. Sample period 2001Q3-2011Q4. Number of obs. 64,588,488. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

6.2.2 Defaults by Recent Credit Score

We now examine default behavior on the extensive margin by recent credit score, and again we present results by 8 quarter lagged credit score as a baseline. Results are very similar for 4 quarter lagged credit score.

[24]Appendix D.2 reports additional results for delinquent balances, including the estimated age affects.

Figure 16 presents the distribution of new mortgage delinquencies. The fraction of borrowers with a new 90+ days mortgage delinquency in the last 4 quarters (left panel) is highest for borrowers in quartile 1 in 2001-2004. During this period, it drops from 1.8% to 1%, and by 2004Q1, the fraction with a new mortgage delinquency in quartile 1 is very similar to the fraction for quartile 2. The delinquency rate starts rising for both quartile 1 and 2 in 2005Q2, though the rise for quartile 2 is much bigger than for quartile 1, so that the fraction with new delinquencies peaks at 1.3% in 2007Q2 for quartile 1 and at 1.7% in 2009Q1 for quartile 2. The fraction with new delinquencies hovers around 0.3% for quartile 3 and 0.15% for quartile for during the boom. During the crisis, it rises to a peak of 0.45% in 2009Q3 for quartile 3, with a very modest rise for quartile 4 over the same period. As a result of the large rise in the fraction of new delinquencies for borrowers in quartile 2 and 3, the quartile 1 share of new delinquencies (right panel) falls by 10 percentage points during the crisis. The share of delinquencies for quartile 2 borrowers rises by 8 percentage points during the crisis and by 11 percentage points for quartile 3.

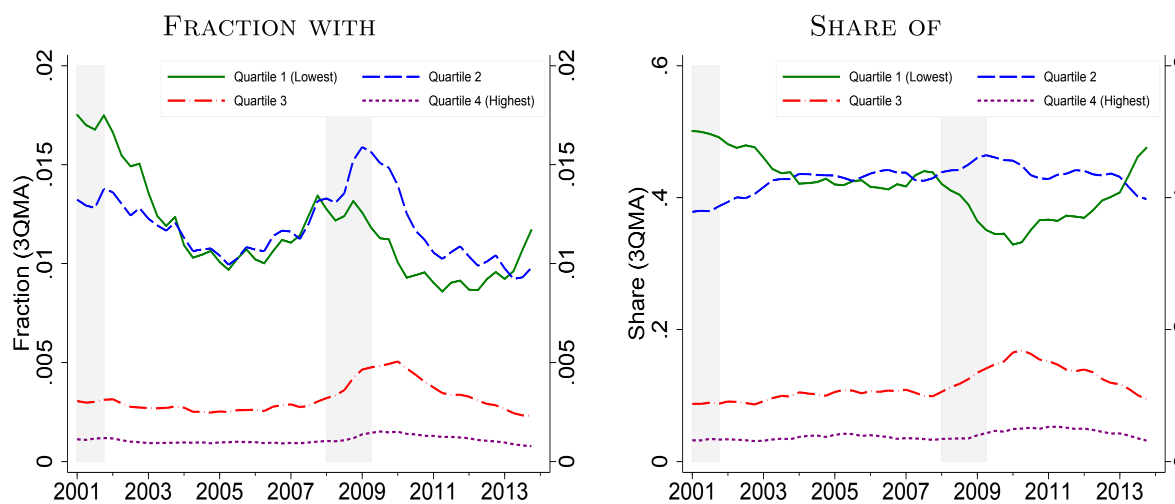


Figure 16: New 90 days+ delinquencies by credit score quartile, 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 17 presents the same statistics for new foreclosures. The quartile 1 and 2 fraction with new foreclosures in the last 4 quarters (left panel) average to 0.26% and to 0.1%, respectively, for the period ending in 2005Q2. For quartile 3 and 4, this fraction is very close to zero until 2006Q3. In mid 2006, new foreclosures start rising for all quartiles, and the rise is particularly pronounced for borrowers in quartile 2 and 3 of the 8 quarter lagged credit score distribution. As a result, the share of new foreclosures (right panel) for quartile

1 borrowers drops from 73% during the boom to a low of 39% in 2009Q1. By contrast, the share of new foreclosures to quartile 2 borrowers rises from 21% during the boom to a peak of 38% in 2009Q1. The share of foreclosures to quartile 3 also rises noticeably from around 4% during the boom, to a peak of 13% in 2009Q2, and the share for quartile 4 also experiences a 5 percentage point rise over the same period.

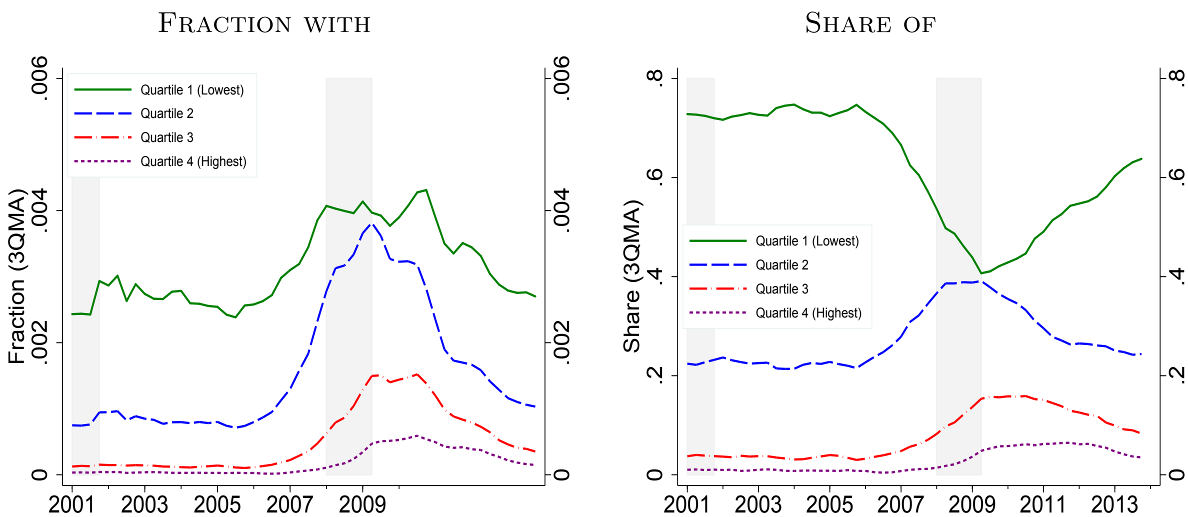


Figure 17: New foreclosures by credit score quartile, 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

In summary, using a lender's approach based on recent credit scores, we find that credit growth during the boom is concentrated in the middle and the top of the credit score distribution and that the rise in defaults during the crisis is concentrated in the middle of credit score distribution. The share of new mortgage delinquencies and foreclosures to low credit score borrowers drops considerably during the crisis, challenging the notion the increase debt and defaults by low credit score borrowers was the main determinant of the housing crisis.

7 Explaining Defaults by Prime Borrowers

The findings presented in the previous section are puzzling given the typically very low default rates for high credit score borrowers. It is then natural to ask why did individuals with good credit histories experience defaults during the crisis. In this section, we document the rise in real estate investors in the prime segment, and we show the increase in mortgage defaults among prime borrowers is primarily driven by real estate investors. We also consider the rise in non-conforming loans. We focus specifically on jumbo loans, which are not eligible for GSE insurance. We find that jumbo loans rise modestly only for prime borrowers.[25]

7.1 Role of Investors

We follow Haughwout et al. (2011) and define investors as borrowers who hold 2 or more first mortgages. Real estate investors are particularly interesting as they may be more prone to default than mortgage borrowers who reside in the property that secures the mortgage, as we discuss below. Moreover, conventional GSE sponsored mortgages are only available for owner occupied properties, which implies investors are more likely to use alternative products, such as Alt-A mortgages, adjustable rate mortgages and other non-standard products.[26] Additionally, if investors are motivated by the prospect of capital gains, they have an incentive to maximize leverage, as this strategy increases potential gains, while the potential losses are limited, especially in states in which foreclosure is non recourse.

Table 5 presents the fraction of borrowers with 2 or more first mortgages and their share of total balances, among all first mortgage holders, by 8 quarter lagged credit score quartile. The fraction of investors (bottom panel) is very stable between 2001Q3 and 2004Q3 for all credit score quartiles, and increases by quartile. The fraction of investors starts increasing rapidly in 2004Q4 and peaks in 2007Q4. Most notably, quartiles 2-4 experience a 50% increase in the fraction of investors between early 2004 and the start of the 2007-09 recession. For quartiles 2-3, the fraction of investors drops to pre boom levels by 2011, but it settles at the 2007 peak for borrowers in quartile 4. By contrast, the fraction of investors for quartile 1 is about half of the fraction for higher quartiles, and rises only modestly during the boom.[27] The time path of the investor share of mortgage balances (bottom panel) is very similar to the path of the fraction of investors, but this share is sizably larger than the fraction of

[25] Another possibility is that the 2007-2009 was so severe that it affected relatively high income individuals and led to a rise in mortgage defaults in populations that are not usually affected. Indeed, Foote, Gerardi, and Willen (2008) argue that negative equity is a necessary but not sufficient condition for mortgage default, and show that negative income shocks may be the ultimate trigger for defaults.

[26] Keys et al. (2012) document the sizable increase of Alt-A mortgages, that have low standard for income documentation and would be particularly appropriate for real estate investors who have variable and hard to document income. Further, Foote and Willen (2016) also discuss the role of alternative mortgage products and the fact that their structure may increase the risk of default. However, Elul and Tilson (2015) present evidence of substantial misrepresentation of home purchases as primary residences, for the purpose of qualifying for GSE sponsored mortgages.

[27] Ferreira and Gyourko (2015) find that the fraction of investors is very similar for prime and subprime borrowers. However, their definition of investors includes only businesses and borrowers with a tax address different from their mortgage address. Chinco and Mayer (2014) also identify real estate investors using the difference between property address and tax address.

investors, as balances are substantially larger for investors. At the beginning of the sample, the share of mortgage balances held by investors is stable. The average over this period is about 12% for quartile 1, and varies between 20 and 23% for quartile 2-4. The investor share of balances rises by just under 50% for quartile 1 by the end of 2007, while it increases by approximately 65% for those in quartile 2 and 3, starting in 2004, and by 35% for borrowers in quartile 4. During and after the recession the investor share of mortgage balances drops, reaching pre-boom levels for quartile 2 and 3, and stabilizing at the peak level for quartiles 1 and 4. Appendix E presents the fraction of investors and the share of balances by specific number of first mortgages (only 2, only 3 and 4+), and shows that both these statistics are increasing with credit score quartile and display the same overall patterns as the combined statistics.[28]

Table 5: Investor Activity by Credit Score Quartile

Fraction of Investors				
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2001Q3-2004Q3 mean	0.063	0.103	0.110	0.107
2004Q4-2007Q4 change	0.020	0.056	0.055	0.037
2007Q4 peak	0.082	0.156	0.162	0.142
Investor Share of Mortgage Balances				
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2001Q3-2004Q3 mean	0.123	0.196	0.212	0.226
2004Q4-2007Q4 change	0.056	0.129	0.125	0.083
2007Q4 peak	0.183	0.333	0.350	0.317

Borrowers with 2 or more first mortgages, fraction (top panel) and share of mortgage balances (bottom panel) by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

For all quartiles, investor activity sizably intensifies at the end of 2004. While identifying the cause of this rise is beyond the scope of this paper, the timing of the rise in investor activity coincides with the decline in the spread for Alt-A and other unconventional mortgages discussed in Justiniano, Primiceri, and Tambalotti (2017). Investors would typically opt to contract such mortgages, which are issued by lenders that rely on private label securitization,

[28]Bhutta (2015) also finds that new mortgages to real estate investors grew markedly during the housing boom, but he does not examine the differentiation by credit score and the default performance.

since GSE sponsored mortgages for investment properties typically feature very restrictive conditions, such as high loan to value ratios and interest rates, in order to compensate for investors' higher default risk.

Figure 18 and 19 report the fraction of borrowers with mortgage delinquencies and foreclosures, respectively, by number of first mortgages. Figure 18 reports the fraction of borrowers with a 90+ day mortgage delinquency by number of first mortgages. Between 2002 and 2006, delinquency rates are similar for investors and non investors for borrowers in quartiles 2-4, but more than twice as high for investors relative to non-investors for borrowers in quartile 1. For non investors, the fraction of borrowers with mortgage delinquencies approximately doubles between 2007 and 2009 for quartiles 1-3 of the credit score distribution, and rises very modestly for borrowers in quartile 4, returning close to pre-crisis levels rises by 2012. Strikingly, the fraction with new delinquencies rises much more for investors than for non-investors over the same period. It roughly doubles for quartile 1, and exhibits a more than 5 fold increase for higher quartiles.

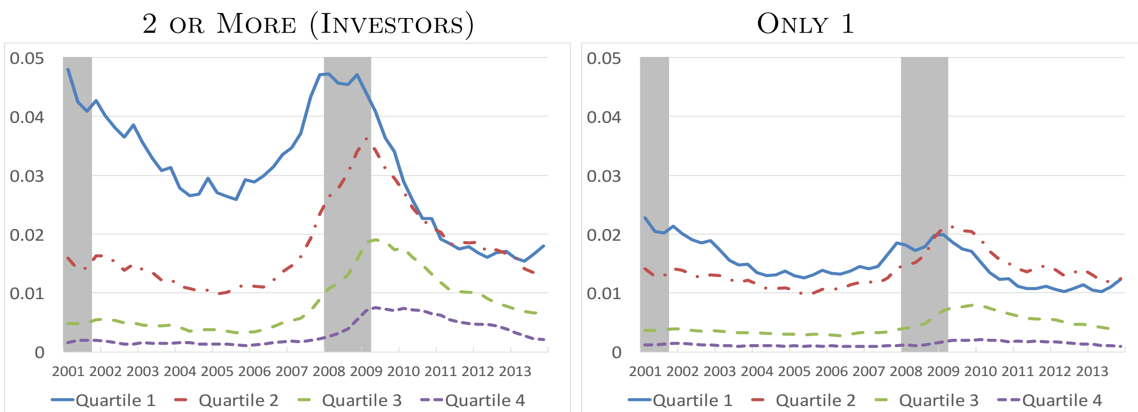


Figure 18: Fraction with new 90+ days mortgage delinquency in the last 4 quarters for borrowers with 2 or more (left panel) and only 1 (right panel) first mortgages by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Figure 19 presents the fraction of borrowers with new foreclosures in the last 4 quarters. As for delinquencies, during the 2002-2006 housing boom the foreclosure rate is very similar for investors and non-investors for all quartiles. However, during the crisis the foreclosure rate diverges, with investors experiencing much higher foreclosure rates than non-investors, especially for higher credit score quartiles. For investors, foreclosure increases by a factor of 4 for the lowest quartile, and by more than a factor of 10 for quartiles 2-4. For non-investors, the foreclosure rate roughly doubles in quartile 1-2, and rises very modestly for quartiles 3-4. Appendix E reports delinquency and foreclosure rates by specific number of first mortgages, showing that the the rise is delinquency and foreclosure rates during the crisis is monotonically increasing in the number of first mortgages for all quartiles.

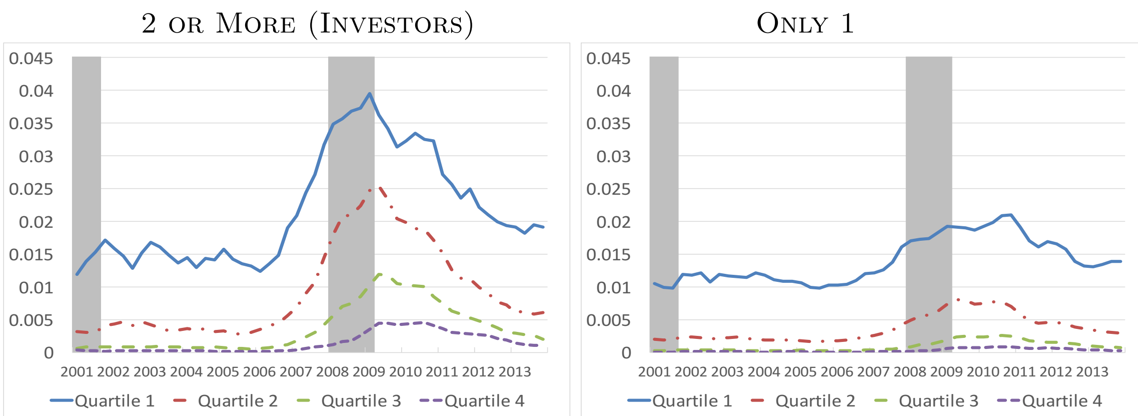


Figure 19: Foreclosure rates for 2 or more (left panel) and only 1 (right panel) first mortgages by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

As a consequence of the greater rise of default rates for investors relative to non-investors, the share of investor defaults rises during the crisis. Figure 20 presents the investor share of 90+ days mortgage delinquencies and foreclosures. The delinquency share of investors is about 10% for all quartiles until mid 2006. This is similar to the share of investors for quartiles 2-4, but about twice the share of investors for quartile 1 over that period. The foreclosure share of investors is about 20% on average during the 2002-2006 boom for quartiles 2-4, which is about twice the fraction of investors for those groups, whereas the investor share of foreclosures for quartile 1 is close to 10%. At the onset of the crisis, there is a sharp rise of the investor share of delinquencies, and especially foreclosures, for borrowers in quartiles 2-4 of the credit score distribution. The share of investor delinquencies rises from 10% to 17% for quartile 1, to 20% for quartile 2, to 30% for quartile 3 and to 40% for quartile 4, with the peak for quartiles 1-3 occurring at the start of the 2007-09 recession, and the peak for quartile 4 at the end of the recession. The investor share of delinquencies subsequently declines, reaching pre-crisis levels by 2012 for quartiles 1-2, but remaining much higher relative to pre-recession levels for quartiles 3-4. The pattern is similar but more dramatic for foreclosures. The investor share of foreclosure rises from 20% to approximately 60% for quartiles 3 and 4, to 40% for quartile 2 and only to 15% for quartile 1 between early 2006 and the start of 2008. For quartiles 1-2, the investor share of foreclosures converges back to pre-crisis levels by the end of 2011, while it remains at more than twice the pre-crisis levels for quartiles 3-4.

The differential default behavior of investors is reflected in their delinquency and foreclosure shares, displayed in figure 20. The investor share of both delinquencies and foreclosures rises dramatically between early 2006 and the end of 2009 for quartiles 2-4. The foreclosure share is particularly large and reaches a peak of over 55% by 2008 for quartiles 3 and 4.

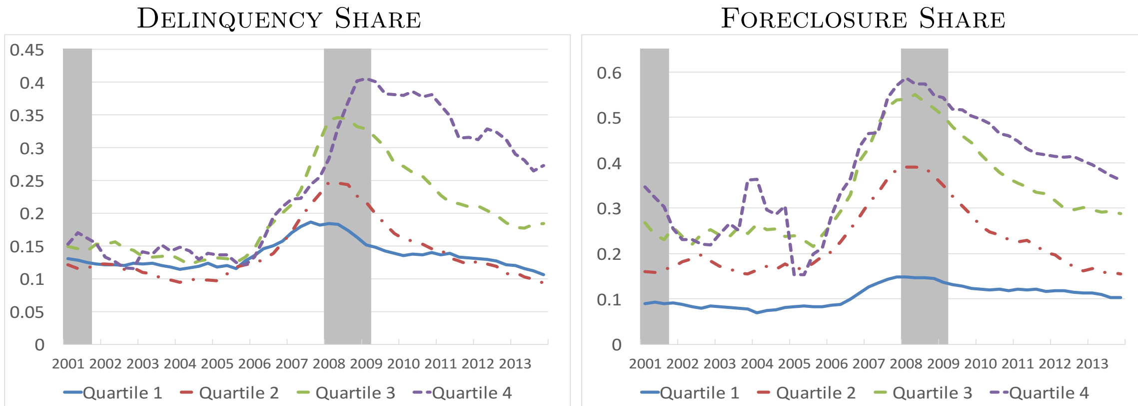


Figure 20: Investor share of 90+ days delinquencies (left panel) and foreclosures (right panel) by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

To more precisely quantify the role of investors in the growth in mortgage balances and defaults, we run a number of counterfactuals presented in Table 6. The top panel considers the growth in mortgage balances in 2001Q3-2006Q4. The first row presents the dollar change in per capita mortgage balances by quartile of the 8 quarter lagged credit score distribution over that period. The second row computes the same change in balances maintaining the distribution of the fraction of first mortgages at the 2001Q3 values, allowing balances per capita by number of first mortgages to take their historical value. Both measures are reported as a fraction of the total change.[29] The growth of mortgage balances per capita with a fixed distribution of number of first mortgages accounts for 77% of the total for quartile 1, 62% of the total for quartiles 2 and 3, and 69% for quartile 4. The third row presents the growth in mortgage balances per capita keeping mortgage balances per capita constant by number

[29] Total balances per capita for quartile $i = 1, 2, 34$ is $B_t^i = \sum_j \pi_t^{i,j} B_t^{i,j}$, where j denotes the number of first mortgages (0,1,2,3,4+) and $t = \left\{\underline{t}, ..., \overline{t}\right\}$ is the quarter. The variables $\pi_t^{i,j}$, $B_t^{i,j}$ are the corresponding fraction of borrowers and the per capita value of balances for that category of borrowers. In the first counterfactual, we set $\pi_t^{i,j} = \pi_{\underline{t}}^{i,j}$ for all periods, and in the second counterfactuals we set $B_t^{i,j} = B_{\underline{t}}^{i,j}$ for all periods. we then consider the change in these 3 statistics between $\underline{t}$ and $\overline{t}$, and report the ratio of the change in the counterfactual value to the total value. These two statistics need not add up to 1 as interactions are not included. The same counterfactuals are computed for delinquency and foreclosure rates, where we only include two groups (j), borrowers with only 1 first mortgages or borrowers with 2 or more.

of first mortgages, but allowing the distribution of the number of first mortgages to follow its historical path. The change in the distribution of the number of first mortgages accounts for only 13% of the total for quartile 1, 20% for quartiles 2 and 4, 22% for quartile 3. This pattern confirms that the the rise in the fraction of borrowers with 2 or more first mortgages is more important for borrowers in higher quartiles.

The second and third panel of Table 6 report similar calculations for the change in delinquency and foreclosure rates in 2006Q3-2009Q4. Here, we report the log change in the rates, since the base delinquency and foreclosure rates vary substantially across quartiles.[30] For delinquency rates, the log change for quartile 1 is 0.12, while it is respectively 0.32, 0.54 and 0.44 for quartiles 2, 3 and 4. The log change with fixed investor share accounts for 97% of the total change in delinquency rates for quartile 1, 99% for quartile 2, 98% for quartile 3 and 95% for quartile 4. The change with fixed delinquency rate for investors is 91% of the total change for quartile 1, 82% for quartile 2, 76% for quartile 3 and 63% for quartile 4. These results confirm the large role of both the rising investor delinquency rate and the rising share of investors in the increase in delinquencies for high credit score borrowers during the crisis. A similar but heightened pattern holds for foreclosures. In this case the log change in foreclosure rates is 0.27 for quartile 1, approximately 2.5 larger for quartile 2, and approximately 5 and 6 times higher for quartiles 3 and 4, respectively. The log change in foreclosure rates with fixed investor share is 98-99% as large as the total change, while the change with constant investor foreclosure rates is 85% of the total change for quartiles 1 and 2, and 77% and 66% of the total change for quartiles 3 and 4, respectively. These results suggest that the increase in foreclosure rates for investors account for larger fraction of the total rise in foreclosure rates for higher quartiles of the credit score distribution.[31]

Real estate investors are particularly likely to contract non-conventional mortgages that are intrinsically more risky and they are also likely to prefer highly leveraged products, as discussed above. An additional factor that may increase the default rate for investors is that only the primary residence is protected in personal bankruptcy, via the homestead exemption.Thus, a borrower who is experiencing difficulties in making payments could potentially file for Chapter 7 bankruptcy and discharge unsecured debt using non exempt assets, and avoid a mortgage delinquency. Perhaps more importantly, the financial and psychological

[30] Similar results obtain using the simple difference in delinquency and foreclosure rates. We select the 2006Q3-2009Q4 time period as it comprises the trough and peak of the delinquency and foreclosure rates for all quartiles of the credit score distribution.

[31] For both delinquencies and foreclosures, the interaction term is negative and large, suggesting a negative relation between the rise in the investor share and the rise in the investor default rates.

Table 6: Role of Investors in Mortgage Balance, Delinquency and Foreclosure Growth

	2001Q3-2006Q4 change in mortgage balances[a]			
total (USD)	8,478	27,608	28,538	20,063
with constant distribution of number of first mortgages[c]	0.7684	0.61594	0.61554	0.6909
with constant balances by number of first mortgages[c]	0.13423	0.2013	0.2180	0.1961
	2006Q3-2009Q4 change in delinquency rates[d]			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
total[b]	0.1175	0.3149	0.5426	0.4373
with constant investor share[c]	0.9706	0.9929	0.9758	0.9497
with constant investor rate[c]	0.9113	0.8177	0.7634	0.6261
	2006Q3-2009Q4 change in foreclosure rate[d]			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
total[b]	0.2649	0.6338	1.0622	1.2854
with constant investor share[c]	0.9892	0.9934	0.9742	0.9676
with constant investor rate[c]	0.8854	0.8535	0.76535	0.6595

Contribution of changing fraction of investors and changing behavior of investors by quartiles of the 8Q lagged Equifax Risk Score distribution. Delinquency rate is defined as fraction with new 90+ day delinquency in last 4 quarters. Foreclosure rate is fraction with new foreclosure in last 4 quarters. Source: Authors' calculations based on FRBNY CCP/Equifax Data.
a. Includes all borrowers.
b. Log difference.
c. Ratio to total. Ratios need not add up to 1 as interaction terms are not reported.
d. Includes only borrowers with at least 1 first mortgage.

costs of default for mortgage borrowers who reside in the home are typically quite substantial, including moving and storage costs, increased commuting costs, and so on. Our results suggest that these factors may have been quite prevalent during the housing crisis.

In Appendix **??**, we also examine the size of the average mortgage. As house prices were rising between 2001-2007, some borrowers were taking on increasingly larger mortgages. Some of these mortgages satisfy the criteria of jumbo loans,[32] which do not qualify for GSE insurance and therefore typically charge higher interest rates. We find a very small rise in the fraction of jumbo mortgages, for borrowers with above median credit score. The fraction of jumbo mortgages rose from 0.7% in 2001 to 1.5% in 2007 for quartile 3 borrowers, and from 1.1% to 1.9% over the same period for borrowers in quartile 4. The rise in the fraction of jumbo loans seems too small to account for the rise in mortgage delinquencies for this group.

8 Interpreting Zip Code Level Evidence

Starting with the seminal work of Mian and Sufi (2009), the macroeconomic literature has used geographical variation to link mortgage debt growth to the severity of the housing crisis and of the ensuing 2007-2009 recession. As shown in figure 1, ranking zip codes by the fraction of subprime borrowers in 2001, suggests that mortgage debt growth in 2001-2007 is stronger in zip codes with high fraction of subprime borrowers at the starting date. However, there is no difference in the growth in total debt balances across quartiles of the fraction of subprime borrowers, as shown in figure **??** reported in Appendix F. In this section, we explore the link between the fraction of subprime borrowers at the zip code level and other population characteristics.

Figure 21 presents zip code level mortgage balance growth since 2001Q3 for prime and subprime borrowers by quartile of the fraction of subprime borrowers. It is clear that prime borrowers experience much higher growth in mortgage balances during the boom relative to subprime borrowers, in all zip codes. However, in zip codes with the highest fraction of subprime borrowers, mortgage balances grow more than in other zip codes for *both* prime and subprime borrowers. As we show in Section **??**, subprime borrowers are disproportionately young and have high demand for credit due to life cycle considerations. Based on this observation, we explore the role of the age distribution at the zip code level.

8.1 The Role of Age

Table 7 reports the age distribution by fraction of subprime borrowers. Not surprisingly, based on our results with individual data, zip codes in quartile 4 of the fraction of subprime

[32] First mortgages above $417,000 were classified as jumbo for the 2001-2007 period. The Obama administration increased this threshold for selected metropolitan areas in 2010 to adjust for regional variation in housing values.

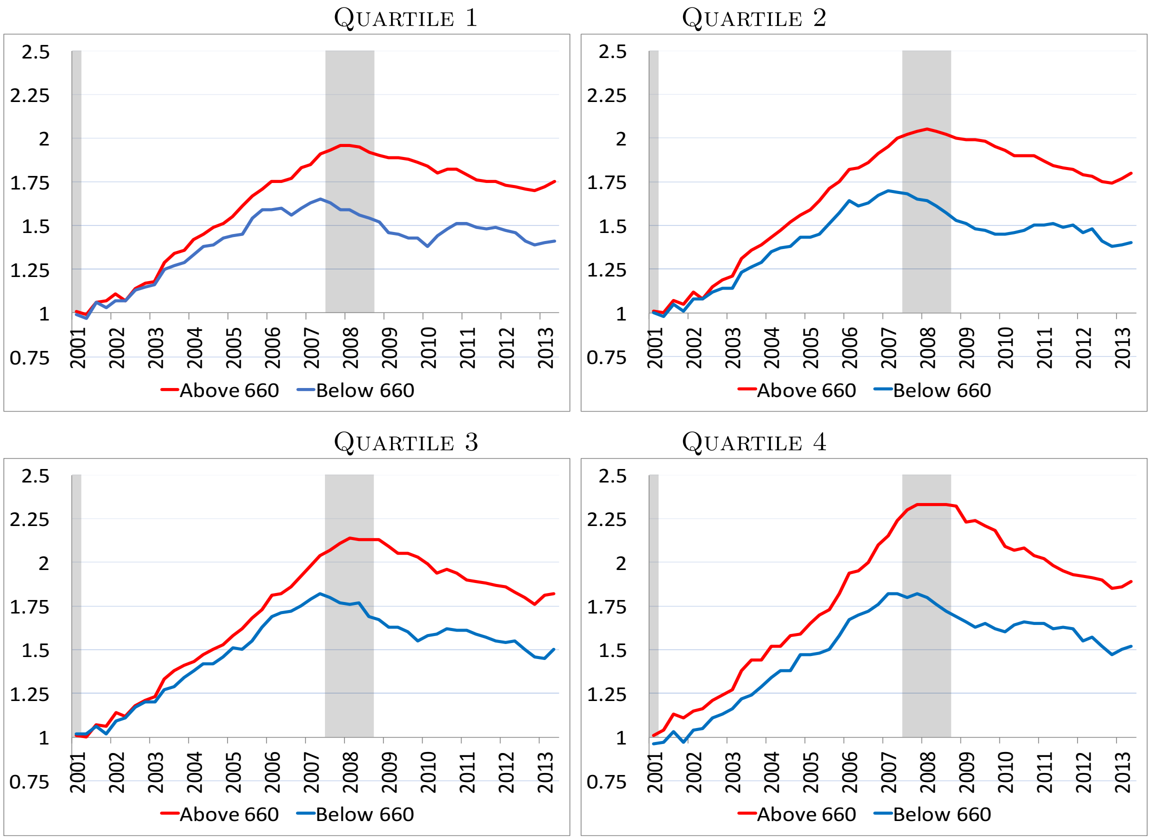


Figure 21: Zip code level per capita mortgage debt growth for prime (Equifax risk Score above 660) and subprime (Equifax risk Score below 660) borrowers by quartile of share of subprime in 2001. Based on 8Q lagged individual credit scores. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

borrowers exhibit a much larger share of borrowers younger than 35.

To quantify the role of the age distribution, we construct counterfactual mortgage balance growth with the age distribution set equal to the age distribution for quartile 1 for all quartiles. We then use this counterfactuals to calculate the contribution of the differences in age distribution across quartiles of the fraction of subprime borrowers to the difference in 2001Q1-2007Q4 (trough to peak) mortgage debt growth relative to quartile 1. These results are reported in Table 8. We find that for zip codes in quartiles 2 and 3 of the fraction of subprime borrowers in 2001, respectively 44% and 43% of the additional cumulative growth in mortgage debt balances relative to quartile 1 is accounted for by differences in the age distribution. This statistic is 84% for zip codes in quartile 4. These findings suggest that even at the zip code level, the age structure is an important determinant of borrowing demand, and strongly affects the observed pattern of debt growth during the 2001-2007 credit boom.

Table 7: Age Distribution by Fraction of Subprime Borrowers

Fraction in each age bin, 2001Q1-2013Q4						
	20-24	25-34	35-44	45-54	55-64	65-85
Quartile 1	0.063	0.157	0.200	0.218	0.171	0.192
Quartile 2	0.070	0.184	0.200	0.205	0.161	0.181
Quartile 3	0.074	0.201	0.206	0.200	0.152	0.168
Quartile 4	0.081	0.212	0.210	0.199	0.145	0.153

Average age distribution in 2001Q1-2013Q4 by quartile of fraction of subprime in 2001. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Table 8: Contribution of Age Distribution to Mortgage Balance Growth

Mortgage Balances		
Quartile 2	Quartile 3	Quartile 4
0.44	0.43	0.84

Contribution of differences in the age distribution to differences in mortgage balance growth 2001Q1-2007Q4. Counterfactuals computed by attributing to each quartile the age distribution of quartile 1 of the fraction of subprime borrowers in 2001. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

8.2 Defaults

We now examine the behavior of defaults by zip code. Figure 22 presents the 90+ mortgage delinquency rate and the foreclosure rate by quartile of fraction of subprime borrowers in 2001. Not surprisingly, zip codes with higher fraction of subprime borrowers exhibit higher delinquency and foreclosure rates, throughout the sample period, with a 2-3 percentage point difference between adjacent quartiles for most of the sample period. The delinquency rate increases modestly during the housing crisis, mostly for zip codes in quartiles 2-4 of the fraction of subprime borrowers in 2001. Foreclosure rates display a similar pattern across quartiles for the entire sample period, despite modest differences in levels. The increase in the foreclosure rate during the crisis is sizable for all quartiles. Foreclosure rates for quartiles 2-4 converge during the crisis, whereas the rate for quartile 1 remains lower, despite its increase. Figure 23 presents the share of 90+ days delinquencies and foreclosures of *prime* borrowers by quartile of the 2001 distribution of the fraction of subprime borrowers. Clearly, prime borrowers contribute more to the growth in delinquent mortgage balances and foreclosures during crisis in all zip codes. The share of prime borrowers' delinquent mortgage balances rises by approximately 30 percentage points between 2006Q2 and 2009Q4, while the share of prime borrowers' foreclosures rises approximately by 40 percentage points over the same period.

DELINQUENCIES FORECLOSURES

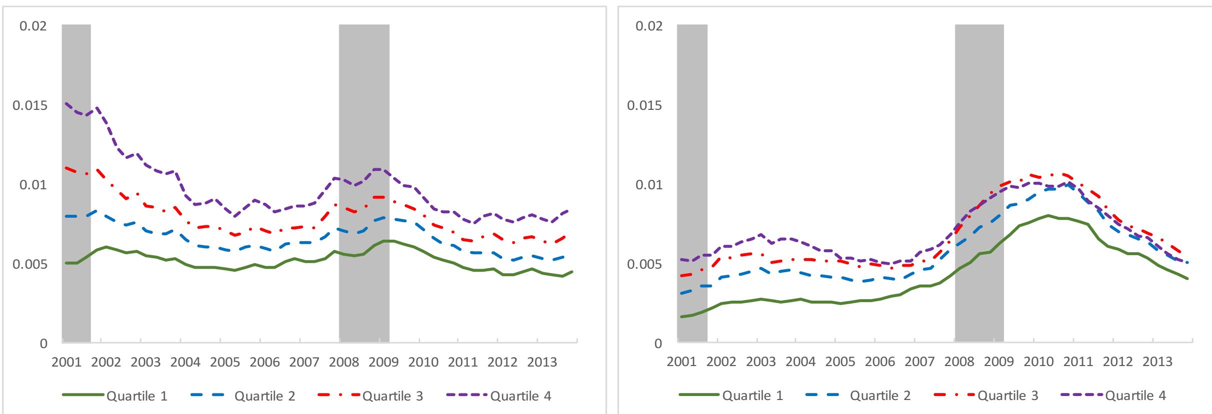


Figure 22: Fraction with 90+ days delinquencies and foreclosures. Zip code level average by quartile of the fraction of subprime share in 2001. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

DELINQUENCIES FORECLOSURES

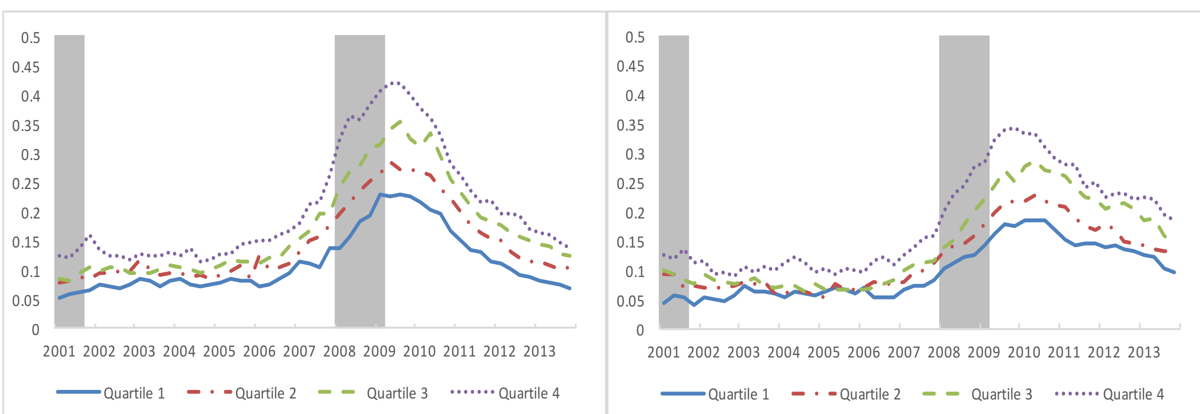


Figure 23: Share of 90+ days delinquencies and foreclosures for prime borrowers, based on 8Q lagged individual credit score. Quartiles of subprime share in 2001. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Interestingly, the prime share of mortgage delinquencies and foreclosures are higher in zip codes with *high* fraction of subprime borrowers, despite the fact that prime borrowers account for a smaller fraction of the population. This suggests that prime borrowers in zip codes with high fraction of subprime borrowers are more vulnerable to financial distress. Though other zip code level characteristics may contribute to this pattern, as we discuss in Section 8.3, here we focus on the role of investors, based on our findings using individual data. Figure 24 presents the fraction of investors at the zip code level for prime and subprime borrowers. There is virtually no difference across quartiles in the fraction of investors for prime borrowers. It starts at approximately 10% in 2001, rises by 5 percentage points between 2005Q1 and 2007Q4, with the average for 2005-2007, reported in Table 9 equal to 12-13%. It then drops during and after the recession, though still remaining above pre-boom levels by the end of 2013. For subprime borrowers, the fraction of investors is decreasing in the quartile of the fraction of subprime in 2001. At the beginning of the sample it is 10% for quartile 1, with a 1-3 percentage point difference across quartiles throughout the sample. The 2005Q1-2007Q4 rise in the fraction of investors is more modest for subprime borrowers, and also decreasing with the quartile of the subprime distribution in 2001. The 2005-2007 average of the fraction of investors among subprime borrowers is 11% and 10% for quartile 1 and 2, and 8% and 7% for quartiles 3 and 4.

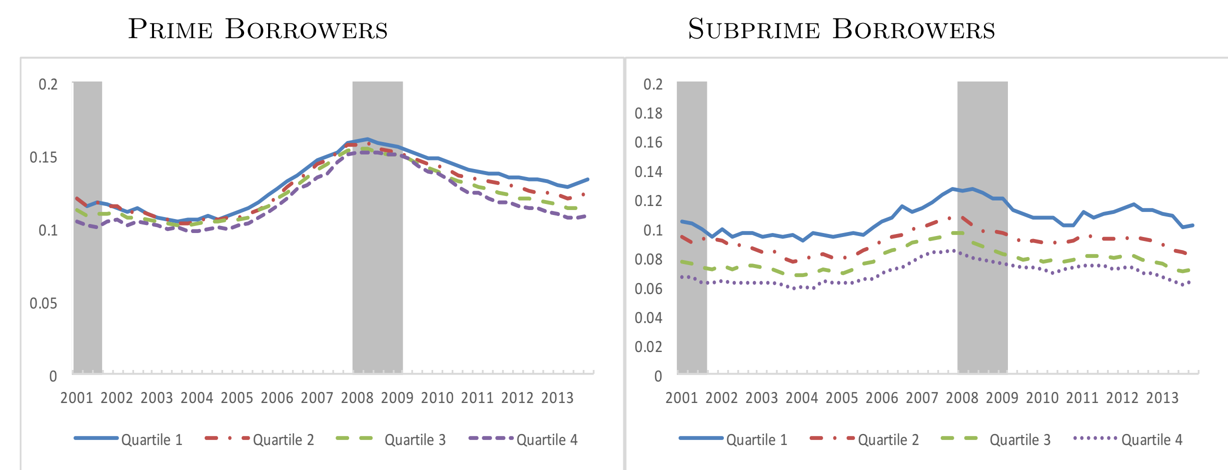


Figure 24: Fraction with 2 or more first mortgages for prime borrowers (left) and subprime borrowers (right), by quartile of fraction of subprime borrowers in 2001. Subprime/prime based on 8Q lagged credit score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Given the large rise in the share of defaults to prime borrowers and the link between investor activity and foreclosures we established using individual data, we examine more detail on investor activity for prime borrowers in Table 9. The distribution of investors across the number of first mortgages is very similar across quartiles, with 79-80% of investors holding 2 first mortgages, 13-14% holding 3 and 7-8% holding 4 or more. However, the growth in average mortgage balances for investors during the boom varies substantially by quartile, and is significantly higher in quartiles with a large fraction of subprime borrowers. The growth in per capita mortgage balances for investors is around 20 percentage points higher for prime borrowers in quartile 4 relative to quartile 1. The growth in mortgage balances for investors with 4 or more first mortgages is particularly high in quartiles 2-4, ranging from 122% to 133%. Turning to defaults, we see that foreclosure rates are sizably higher for investors relative to non-investors, as we found in the individual data. This difference is increasing in the fraction of subprime borrowers. The rise in the foreclosure rate during the crisis for investors in quartiles 3-4 is nearly double the rise in quartiles 1-2, reaching a high of 15% for investors with 4 or more first mortgages in quartile 4.

This pattern in investor borrowing and default behavior may explain why despite large regional variation in predictable default risk, GSE mortgage rates for otherwise identical loans do not vary spatially, while the private market does set interest rates that vary with local risk, as shown in Hurst et al. (2016). GSE mortgages are mostly available for owner occupied properties and default rates among borrowers with only one first mortgage are low in all zip codes. By contrast, default rates on private market products would reflect the geographical variation in investor activity, and corresponding default propensity.

Summarizing, though the fraction of investors with prime credit score is very similar across quartiles, in quartiles with high share of subprime, investors exhibit larger increases in mortgage balances during the boom and a more severe increase in foreclosures during the crisis. This difference in behavior for prime investors may be driven by the behavior of real estate values. As reported in Table 10, the average growth house price index in 2001-2007 varies from 29% in quartile 1 to 47% in quartile 4. The total decline in housing values in 2007-2010 is also increasing in the fraction of subprime, ranging from 21% in quartile 1 to 36% in quartile 4. This suggests that investor activity by *prime* borrowers is associated with a more pronounced house price boom and bust and a more severe foreclosure crisis.

8.3 Zip Code Characteristics

Several studies find a positive relation between the size of the increase in mortgage debt growth or house price debt growth during the 2001-2006 credit boom, often instrumented

Table 9: Investor Activity

	2005-2007 fraction of investors			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
prime borrowers	13%	13%	13%	12%
subprime borrowers	11%	10%	8%	7%
	Prime Borrowers			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2007Q4 fraction of investors with				
2 first mortgages	80%	80%	80%	79%
3 first mortgages	13%	13%	14%	14%
4+ first mortgages	7%	7%	7%	8%
2001Q3-2007Q4 mortgage balance growth				
1 first mortgage	59%	62%	66%	69%
2 first mortgages	86%	85%	97%	104%
3 first mortgages	94%	104%	117%	118%
4+ first mortgages	102%	122%	133%	125%
boom average-peak change in foreclosure rate				
1 first mortgage	0.008	0.012	0.016	0.017
2 first mortgages	0.023	0.027	0.045	0.053
3 first mortgages	0.040	0.063	0.087	0.115
4+ first mortgages	0.076	0.096	0.123	0.151

Selected zip code level indicators of investor activity by quartile of the fraction of subprime borrowers in 2001. The boom average for the foreclosure rate corresponds to the 2002Q1-2005Q4 average. The peak of the foreclosure rate varies by group, with 2007Q4 the most common date. Source: Authors' calculations based on FRBNY CCP/Equifax Data, IPUMS, IRS, BLS, ACS data.

with Saiz (2010) house price elasticities, and the severity of the 2007-2009 recession.[33] These

[33]For example, Mian, Sufi, and Trebbi (2015) find that states with higher foreclosure rates experienced a larger decline in consumption, while Mian and Sufi (2014) use county level data and show that a larger decline in household net worth during the crisis experience a more pronounced decline in non-tradable employment. Mian, Rao, and Sufi (2013) exploit geographic variation in house price declines over the period 2006-2009 and household balance sheets in 2006, to estimate the elasticity of consumption expenditures to changes in the housing share of household net worth, and find a positive and sizable elasticity. Kaplan, Mitman, and Violante (2016) refine this analysis and find that, once the direct effect of the fall in local house prices has been controlled for, household balance sheets do not have an effect on durable consumption.

studies attribute this correlation to the tightening of collateral constraints during the crisis, driven by mortgage defaults and the resulting decline in housing values.[34] Since this causal mechanism is not consistent with our findings, we explore additional economic indicators at the zip code level to shed light on this correlation.

Table 10 reports several economic indicators by quartile of the fraction of subprime borrowers in 2001. Several indicators that are critical to business cycle sensitivity are systematically related to the fraction of subprime borrowers. Zip codes with higher fraction of subprime borrowers are younger, as previously noted, have lower levels of educational attainment and have a disproportionately large minority and African American share in the population. It is well known that younger, less educated, minority workers suffer larger employment loss during recessions.

Zip codes with a large fraction of subprime borrowers also exhibit lower per capita income levels in both the boom and the recession. In 2001-2007, the average real per capita income was \$41,045 in quartile 1 and only \$21,019 in quartile 4, whereas in 2007-2010 it was \$46,341 for quartile 1 and \$21,898 for quartile 4. Consistent with Mian and Sufi (2009), income growth during the boom was lower in zip codes with higher fraction of subprime. Average per capita income grew by 35% between 2001 and 2007 for quartile 1 and only 4% for quartile 4. Similarly, zip codes with large subprime population have higher unemployment rates both during the boom and during the crisis. The average unemployment rate for 2001-2007 was 4.94% in quartile 1 and 5.72% in quartile 4. In 2007-2010, the average unemployment rate rose to 6.93% in quartile 1 and 7.81% in quartile 4. Zip codes with a large subprime population also exhibit higher income inequality. We measure this with the ratio of average income for individuals with incomes above \$200,000 over average income for the entire population, based on IRS data. Higher inequality implies that the aggregation bias generated by the fact within each zip code prime borrowers experience more credit growth than subprime borrowers is accentuated.

Zip codes with high fraction of subprime borrowers experience higher house price growth in 2001-2007, as previously noted. This may be related to the their higher population density, suggesting the prevalence of urban areas for this group. Endogenous gentrification, as described in Guerrieri, Hartley, and Hurst (2013), exerted particularly high pressure on housing values in urban areas over this period, and may have encouraged real estate investor

[34]However, Ferreira and Gyourko (2011) show that local income is the only potential demand shifter found that also had an economically and statistically significant change around the time that local housing booms began. Liebersohn (2017) also shows that the share of growing industries drives the size of housing demand shocks, the magnitude of the housing price increase and household consumption variation between 2000-2006.

activity.[35] The distribution of zip codes with low housing supply elasticity, as captured by the Saiz (2010) index, is fairly even across quartiles. However, 16% of zip codes in quartiles 3 and 4 are in sand states[36], whereas only 11% and 13% of zip codes in quartiles 1 and 2 are in those states.

The distribution of the fraction of subprime borrowers is quite stable at the zip code level, and this is also true for other characteristics salient to business cycle sensitivity, as we show in Appendix F. Therefore, the timing of the ranking by fraction of subprime does not change the patterns at the zip code level. However, some aggregate trends, such as the historical decline in wages, labor force participation and employment rates for unskilled, young and minority workers, and the rise in income inequality may influence economic outcomes at the zip code level over time. One motivation for considering zip code level evidence is the scarcity of information on individual borrowers in credit file data. Geographical aggregation provides access to a number of additional indicators, such as income, housing values and so on. Very often, geographical patterns are interpreted as reflecting individual behavior. For example, differences in debt growth across two zip codes with different fraction of subprime borrowers are assumed to be similar to differences in debt growth across individuals with different credit scores. Our findings suggest that using geographically aggregated data does not provide an accurate account of the patterns of borrowing at the individual level. Moreover, the positive correlation between credit growth during the boom and the depth of the recession may be due to other characteristics at the zip code level, such as the prevalence of young, minority or low education workers.

9 Conclusion

Our analysis suggests a reassessment of the role of growth in the supply of subprime credit in the 2001-2006 housing boom and in the 2007-2009 financial crisis. We find that most of the increase in mortgage debt during the boom and of mortgage delinquencies during the crisis is driven by mid to high credit score borrowers, and it is these borrowers who disproportionately default on their mortgages during the crisis. The growth in defaults is

[35] However, the more sizable housing boom in some zip codes with large subprime population may have masked negative employment growth over this period, as shown by Hurst et al. (2016), and increased income and reduced unemployment rates in those areas above what would have been consistent with their industry and demographic composition.

[36] These are Arizona, California, Colorado, Florida, and Nevada. These states exhibit the largest swings in housing values during the housing boom and the subsequent foreclosure crisis. Chinco and Mayer (2014) show that in Phoenix, Las Vegas and Miami out of town second home buyers may have contributed to an inflation in housing values.

Table 10: Zip Code Level Indicators

	Demographics			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Median age	50	49	48	46
Associate+ degree (2012)	45%	31%	23%	17%
Percent white	93%	90%	83%	63%
Percent black	1.7%	3.6%	7.6%	24.6%
	Economy			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Average UR 2001-2007	4.94%	5.19%	5.38%	5.72%
Average UR 2007-2010	6.93%	7.30%	7.51%	7.81%
Average PDI 2001-2007	\$41,045	\$30,442	\$25,692	\$21,019
Average PDI 2007-2010	\$46,341	\$33,224	\$27,491	\$21,898
PDI Growth 2001-2007	25%	16%	10%	4%
PDI Growth 2007-2010	10%	10%	11%	10%
$\frac{\text{Mean Income} \geq \$200K}{\text{Mean Income}}$ (2006-11)	6.4	7.9	9.4	11.8
	Mortgage Markets			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2001 fraction subprime (med)	19%	32%	44%	60%
HPI Growth 2001-2007	29%	37%	42%	47%
HPI Growth 2007-2010	-21%	-30%	-27%	-36%
Low Saiz elasticity	17%	13%	11%	12%
	Geography			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
In sand states	11%	13%	16%	16%
Pop per sq mile	1214	1380	1386	2322
Percent never moved	53%	53%	51%	51%

Selected zip code level indicators by quartile of the fraction of subprime borrowers in 2001. PDI (personal disposable income) and HPI (housing price index) expressed in 2012 USD, adjusted by CPI-U. UR (unemployment rate) is the U3 official rate. Source: Authors' calculations based on FRBNY CCP/Equifax Data, IPUMS, IRS, BLS, ACS data.

mostly accounted for by real estate investors. The disproportionate role of investors in the foreclosure crisis are consistent with the findings in Ospina and Uhlig (2016), who show that non-agency residential mortgage backed securities, which would have been used for investor mortgages, overall exhibited very modest losses during the crisis. However, the bulk of the losses were concentrated among a small set of securities, especially those issued after 2004, when investor activity started surging.

At the zip code level, we show that prime borrowers experience a larger rise in debt and defaults than subprime borrowers, irrespective of the size of the subprime population in the zip code, however, zip codes with a large fraction of subprime do experience stronger mortgage credit growth during the boom. We show that these zip codes have a high share of young residents and are predominantly urban and that they exhibit the largest rise in prime defaults and the most dramatic house price fluctuations between 2001 and 2010. Zip codes with a large share of subprime borrowers also exhibit a high fraction of low education, high minority residents that may be particularly sensitive to business cycle shocks. These new findings should inform discussions of the causes and consequences of the 2007-2009 financial crisis and of the appropriate policy responses.

References

Adelino, Manuel, Antoinette Schoar, and Felipe Severino. 2015. “Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class.” Working paper 20848, National Bureau of Economic Research.

———. 2017. “Dynamics of Housing Debt in the Recent Boom and Great Recession.” Working paper 23502, National Bureau of Economic Research.

Agarwal, Sumit, Gene Amromin, Itzhak Ben-David, and Douglas D. Evanoff. 2016, September. “Loan Product Steering in Mortgage Markets.” Working paper 22696, National Bureau of Economic Research.

Albanesi, Stefania, and Jaromir Nosal. 2015. “Insolvency After the 2005 Bankruptcy Reform.” Cepr discussion papers 10533, CEPR Discussion Papers.

Bazikyan, Armine. 2009. “Renters: The Innocent Victims of the Foreclosure Mortgage Crisis.” *Sw. L. Rev.* 39:339.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra. 2015, October. “House Prices and Consumer Spending.” Working paper 21667, National Bureau of Economic Research.

Bhutta, Neil. 2015. "The ins and outs of mortgage debt during the housing boom and bust." *Journal of Monetary Economics* 76:284 – 298.

Bhutta, Neil, and Benjamin J Keys. 2016. "Interest rates and equity extraction during the housing boom." *The American Economic Review* 106 (7): 1742–1774.

Brown, Meta, Andrew Haughwout, Donghoon Lee, and Wilbert van der Klaauw. 2014. "The Financial Crisis at the Kitchen Table: Trends in Household Debt and Credit." *Current Issues in Economics and Finance* 19, no. 2.

Case, Karl E., Robert J. Shiller, and Anne Thompson. 2012, September. "What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets." Working paper 18400, National Bureau of Economic Research.

Chen, Hui, Michael Michaux, and Nikolai Roussanov. 2013, September. "Houses as ATMs? Mortgage Refinancing and Macroeconomic Uncertainty." Working paper 19421, National Bureau of Economic Research.

Chinco, Alex, and Christopher Mayer. 2014, January. "Misinformed Speculators and Mispricing in the Housing Market." Working paper 19817, National Bureau of Economic Research.

Corbae, Dean, and Erwan Quintin. 2015. "Leverage and the Foreclosure Crisis." *Journal of Political Economy* 123 (1): 1–65.

Elul, Ronel, and Sebastian G Tilson. 2015. "Owner occupancy fraud and mortgage performance."

Ferreira, Fernando, and Joseph Gyourko. 2011, August. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009." Working paper 17374, National Bureau of Economic Research.

———. 2015, June. "A New Look at the U.S. Foreclosure Crisis: Panel Data Evidence of Prime and Subprime Borrowers from 1997 to 2012." Working paper 21261, National Bureau of Economic Research.

Foote, Christopher L., Kristopher S. Gerardi, and Paul S. Willen. 2012, May. "Why Did So Many People Make So Many Ex Post Bad Decisions? The Causes of the Foreclosure Crisis." Working paper 18082, National Bureau of Economic Research.

Foote, Christopher L., Lara Loewenstein, and Paul S. Willen. 2016. "Cross-sectional patterns of mortgage debt during the housing boom: evidence and implications." Technical Report 22985, National Bureau of Economic Research.

Foote, Christopher L., and Paul S. Willen. 2016. “The Subprime Mortgage Crisis.” In *Banking Crises*, 324–336. Springer.

Foote, Christopher L, Kristopher Gerardi, and Paul S Willen. 2008. “Negative equity and foreclosure: Theory and evidence.” *Journal of Urban Economics* 64 (2): 234–245.

Ghent, Andra C, and Marianna Kudlyak. 2011. “Recourse and residential mortgage default: evidence from US states.” *Review of Financial Studies*, p. hhr055.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. “Endogenous gentrification and housing price dynamics.” *Journal of Public Economics* 100:45–60.

Guerrieri, Veronica, and Guido Lorenzoni. 2011. “Credit crises, precautionary savings, and the liquidity trap.” Technical Report 17583, National Bureau of Economic Research.

Haughwout, Andrew, Donghoon Lee, Joseph S Tracy, and Wilbert Van der Klaauw. 2011. “Real estate investors, the leverage cycle, and the housing market crisis.” *Federal Reserve Bank of New York Staff Reports*, no. 514.

Hurst, Erik, Benjamin J. Keys, Amit Seru, and Joseph Vavra. 2016. “Regional Redistribution through the US Mortgage Market.” *American Economic Review* 106 (10): 2982–3028 (October).

Iacoviello, Matteo. 2004. “Consumption, house prices, and collateral constraints: a structural econometric analysis.” *Journal of Housing Economics* 13 (4): 304–320.

Justiniano, Alejandro, Giorgio E. Primiceri, and Andrea Tambalotti. 2017, September. “The Mortgage Rate Conundrum.” Working paper 23784, National Bureau of Economic Research.

Justiniano, Alejandro, Giorgio E Primiceri, and Andrea Tambalotti. 2016. “Quantitative Modeling of the Financial Crisis A Simple Model of Subprime Borrowers and Credit Growth.” *The American Economic Review* 106 (5): 543–547.

Kaplan, Greg, Kurt Mitman, and Giovanni L. Violante. 2016, May. “Non-durable Consumption and Housing Net Worth in the Great Recession: Evidence from Easily Accessible Data.” Working paper 22232, National Bureau of Economic Research.

———. 2017, August. “The Housing Boom and Bust: Model Meets Evidence.” Working paper 23694, National Bureau of Economic Research.

Kehoe, Patrick, Elena Pastorino, and Virgiliu Midrigan. 2016. “Debt constraints and employment.” Technical Report, National Bureau of Economic Research.

Keys, Benjamin J, Tomasz Piskorski, Amit Seru, and Vikrant Vig. 2012. "Mortgage financing in the housing boom and bust." In *Housing and the Financial Crisis*, 143–204. University of Chicago Press.

Keys, Benjamin J, Tomasz Piskorski, Amit Seru, and Vincent Yao. 2014. "Mortgage rates, household balance sheets, and the real economy." Technical Report, National Bureau of Economic Research.

Kiyotaki, Nobuhiro, and John Moore. 1997. "Credit Cycles." *Journal of Political Economy*, no. 2:211–248.

Lee, Donghoon, and Wilbert van der Klaauw. 2010. "An Introduction to the FRBNY Consumer Credit Panel." FRBNY Staff Report 479.

Li, Wenli. 2009. "Residential housing and personal bankruptcy." *Business Review* 2:19–29.

Liebersohn, Jack. 2017. "Housing Demand, Regional House Prices and Consumption." *Manuscript, MIT.*

Mian, Atif, Kamalesh Rao, and Amir Sufi. 2013. "Household Balance Sheets, Consumption, and the Economic Slump." *The Quarterly Journal of Economics* 128 (4): 1687–1726.

Mian, Atif, and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *The Quarterly Journal of Economics* 124 (4): 1449–1496.

———. 2010. "The Great Recession: Lessons from Microeconomic Data." *The American Economic Review* 100 (2): 51–56.

———. 2011. "House Prices, Home Equity–Based Borrowing, and the US Household Leverage Crisis." *The American Economic Review* 101 (5): 2132–2156.

———. 2014. "What explains the 2007–2009 drop in employment?" *Econometrica* 82 (6): 2197–2223.

Mian, Atif, Amir Sufi, and Francesco Trebbi. 2015. "Foreclosures, house prices, and the real economy." *The Journal of Finance* 70 (6): 2587–2634.

Mian, Atif R, and Amir Sufi. 2016. "Household debt and defaults from 2000 to 2010: The credit supply view." *Manuscript.*

Midrigan, Virgiliu, and Thomas Philippon. 2011. "Household Leverage and the Recession." *NBER Working Paper 16965.*

Mincer, Jacob. 1991. "Education and unemployment." Technical Report, National Bureau of Economic Research.

Mitman, Kurt. 2016. “Macroeconomic effects of bankruptcy and foreclosure policies.” *The American Economic Review* 106 (8): 2219–2255.

Ospina, Juan, and Harald Uhlig. 2016. “Mortgage-Backed Securities and the Financial Crisis of 2008: A Post Mortem.” Technical Report, University of Chicago.

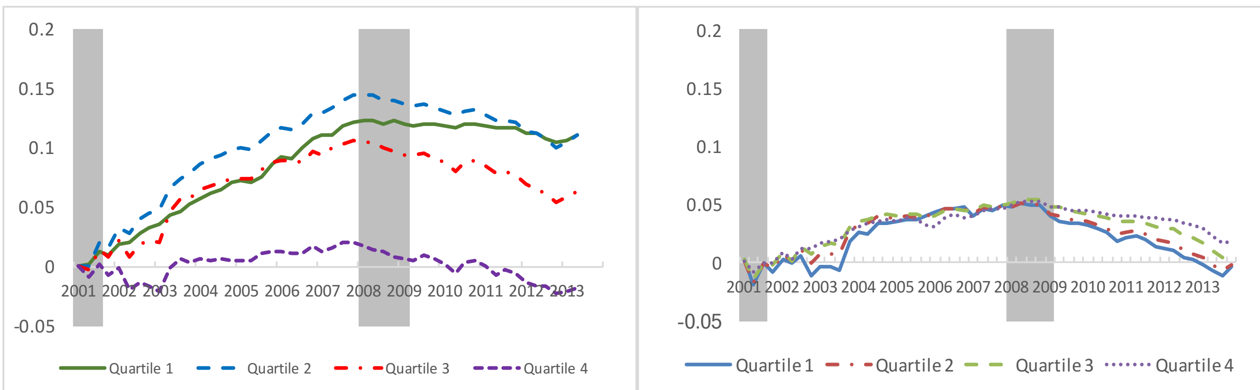
Robinson, Break L., and Richard M. Todd. 2010. “The Role of Non-Owner-Occupied Homes in the Current Housing and Foreclosure Cycle.” Working paper 10-11, Federal Reserve Bank of Richmond.

Saiz, Albert. 2010. “The geographic determinants of housing supply.” *The Quarterly Journal of Economics* 125 (3): 1253–1296.

Shimer, Robert. 1998. “Why is the US unemployment rate so much lower?” *NBER Macroeconomics Annual* 13:11–61.

A Initial Credit Score Ranking: Additional Results

Figure 25 displays the fraction with first mortgages by 1999 credit score ranking for individuals (left panel) and by fraction of subprime borrowers in 2001 for zip codes (right panel). Based on the individual level data, the fraction with first mortgages growth by 10-13 percentage points between 2001Q3 and the start of the recession for quartiles 1-3, and only by 2 percentage points for borrowers in quartile 4. At the zip code level, there is little difference in the change in the fraction with first mortgages across zip codes during the boom, though the decline during and after the recession is more pronounced for lower quartiles.



(a) Individuals: Ranked by 1999 Equifax Risk Score (b) Zip Codes: Ranked by Fraction of Subprime in 2001

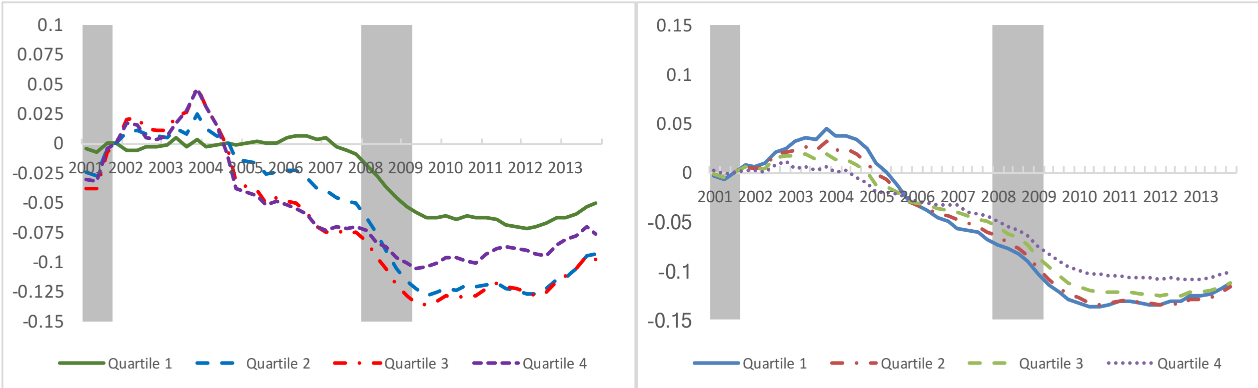
Figure 25: Fraction with first mortgages, difference from 2001Q3. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

Figure 26 displays the fraction with new mortgage originations. Even with the 1999 quartile ranking, borrowers in quartile 1 do not exhibit any growth in new mortgage originations during the boom, and most of the growth in new mortgage originations occurs between 2001 and 2004 for borrowers in quartiles 2-4 (left panel). At the zip code level (right panel), zip codes with the lowest fraction of subprime borrowers exhibit stronger growth in mortgage originations between 2001 and 2004, and this fraction declines for all quartiles after 2004.

B Income Data

In this section, we describe the supplementary payroll data used for the income imputation procedure. This data is merged with our credit panel data, allowing us to map individuals' incomes for 2009 to their credit files.

The Equifax Workforce Solutions data provided by Equifax is a nationally-representative random sample of individuals containing employment and payroll verification information provided directly from the employers. The information provided for each employee includes the last three years of total income, the date of first hire, tenure, and for the current year status (part time/full time), weekly hours, pay rate and pay frequency.



(a) Individuals: Ranked by 1999 Equifax Risk Score (b) Zip Codes: Ranked by Fraction of Subprime in 2001

Figure 26: Fraction with new mortgage originations in the last 4 quarters, difference from 2001Q3. Source: Authors' calculation based on Federal Reserve Bank of New York's Consumer Credit Panel/Equifax Data.

Income Measure Description There are various income measures provided in the Worknumber dataset. For each year of data available variables are given for the total 12-month base, bonus, overtime, and commission compensation in year t, $t-1$, and $t-2$. This information however is only available for a little over $\frac{1}{3}$ of the sample. The other measure of income, which is widely available across the sample, is rate of pay and pay frequency. We therefore impute total income using a simple *rate* $\times$ *frequency* approach to account for the lack of representation found in the sample regarding the total 12-month income variables. This yields about 11,000 observations for 2009. The sample of records is nationally representative, both in terms of geographical and age distribution.

Comparison with the CPS To gauge the accuracy of the imputed income measure in our data, we performed a simple comparison with the income levels reported in the Consumer Population Survey. We present results based on income quintiles below.

Table 11: Income Distribution Comparison by Quintile

Calculation	Dataset	1	2	3	4	5
Mean	CPS	11058.67	24791.32	36584.61	51872.45	110192.2
	Worknumber	17078.07	26565.46	39589.76	58510.22	117260.1
Median	CPS	12000	25000	36000	50000	85000
	Worknumber	16640	27040	39520	57512	99990

Source: IPUMS, Equifax Worknumber. Worknumber income calculations made using proxied income from pay periods and pay rate. CPS income calculations made using total wage and salary income.

We conduct a similar analysis, comparing the distribution of income and age by state in the Worknumber sample and compare it to the American Community Survey. We also find that the sample is consistent with this survey. These results are available upon request.

C PSID Evidence on Income and Debt

To assess the generality of the relation between income, age and debt described in Section 5.2, we use the PSID to estimate the relation between debt growth and income during the boom period. Using zip code level data, Mian and Sufi (2009) show that during the period between 2001 and 2006, the zip codes that exhibited the largest growth in debt were those who experiences the smallest growth in income. They argue that the negative relation between debt growth and income growth at the zip code level over that period is consistent with a growth in the supply of credit, via a relaxation of lending standards. Using the panel stricture of the PSID, we can directly assess the relation between income and debt growth at the individual data. While debt is poorly measured in the PSID relative to the Consumer Credit Panel that we use for our main analysis, we have income at a yearly or bi-yearly frequency.

The estimates for various specifications are displayed in Table 12. The dependent variable is the change in real log total debt between 2007 and 1999, and the baseline specification includes the change in log income over the same period as a dependent variable. The coefficient is positive and highly significant, with a 1 log point change in income corresponding to a 0.066 log point increase in the change in debt over the period. This coefficient implies that 1 10,000$ increase in income from a value of 50,000$ in 1999 is associated with a 1$ increase in debt. The second column includes 1999 age and 1999 age squared. The coefficient on the change in income changes little, and the coefficient on age is negative and significant, consistent with our previous finding on the fact that debt accumulation slows with age, and debt accumulation is strongest for borrowers who are young in 1999. The third column includes a an interaction between 1999 age and the change in income, log income in 1999 and no squared age term. In this case the coefficient on the change in log income is positive but much smaller and not significant, while the coefficient on age is still negative and significant, but smaller in magnitude. The coefficient on log income in 1999 is positive but not significant. The last column also adds an interaction between log income in 1999 and age in 1999. In this case the coefficient on the change in income is positive and larger in magnitude relative to previous specifications, but not significant. The other coefficients are similar, with a larger magnitude of the negative coefficient on age. The interaction between age and log income in 1999 is positive and significant, suggesting that higher initial income is associated with larger growth in debt conditional on age. These results confirm our findings based on the Equifax data, suggesting that income growth and debt growth are positively related over the 2001-2006 boom.

Table 12: Relation Between Debt Growth and Income Growth

Dependent Variable: 2007-1999 change in log total debt (real USD)				
Δlog(income)	0.066**	0.068**	0.21	0.081
1999 age		-0.064***	-0.01***	-0.070**
1999 age sq		0.001***		
1999 age $\times \Delta log$(income)			-0.003	-0.001
log(income$_{1999}$)			0.001	-0.270
1999 age $\times$ log(income$_{1999}$)				0.006*

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$ No. obs. 1,395. Source: Authors' calculations based on PSID Data.

D Balance Change Regressions: Additional Results

D.1 Mortgage Balances

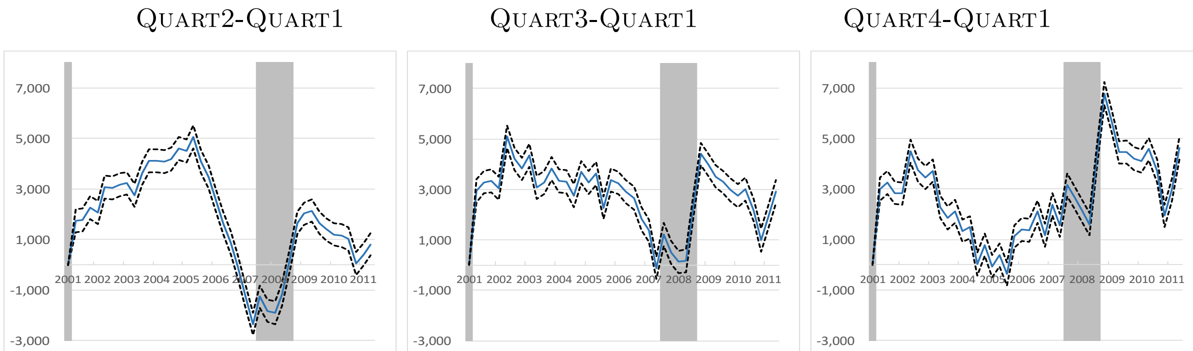


Figure 27: Estimated time effects by 1Q lagged Equifax Risk Score quartile from balance change regressions. Baseline specification. Dependent variable is the 8Q ahead change in per capita mortgage balances in USD. Dashed lines denote 5% confidence intervals. Sample period 2001Q3-2011Q4. Number of obs. (baseline) 64,588,488. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Table 13 presents the estimated cumulative change in mortgage balances for different time periods in our sample based on the balance change regressions. These are derived from the 8 quarter ahead change specification, discussed in detail in Section 6.1, as well as for the 4 quarter ahead and 12 quarter ahead change. The variation across quartiles for the 4 and 12 quarter ahead change is consistent with the findings for the 8 quarter ahead change.

D.2 Delinquent balances

We report additional results for the estimates for delinquent balances described in Section 6.2.1. Figure 28 reports the differences in the estimated time effects for quartiles 2-4 relative to quartile 1. As for debt balances, there is a sizable and highly significant difference in time effects across quartiles. Figure 29 reports the estimated age effects.The age effects for dellinquent balances largely reflect the age pattern of total debt balances.

E Investors

We report additional detail about investor activity. Figure 30 report the fraction of borrowers with only 2, only 3 and 4 or more first mortgages among all borrowers with first mortgages, by quartile of the 8 quarter lagged credit score distribution. These fractions are lowest for quartile 1 for all categories. There is modest rise is fraction with two mortgages in 2005-2007 for quartile 1, whereas the fraction with 3 or 4+ are stable throughout the period for this quartile. Quartiles 2-4 exhibit strong rise in the number of first mortgages, as previously

Table 13: Mortgage Balances: Summary Regression Results

	Average 4 Quarter Ahead Change in Mortgage Balances			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2002-03	-206	1,998	2,329	1,269
2003-04	235	3,622	4,355	3,182
2004-05	173	3,173	3,150	1,702
2005-06	936	5,316	4,996	2,209
2006-07	798	4,937	3,854	1,115
2007-08	-1,700	1,732	2,505	818
2008-09	-4,690	-2,691	-1,724	-2,355
2009-10	-6,463	-3,075	-1,964	-2,369
2010-11	-5,538	-2,670	-1,269	-1,591
2011-12	-5,189	-3,020	-2,281	-2,921
2002-2006	1,138	14,108	14,831	8,361
2007-2010	-12,853	-4,035	-1,183	-3,905
	Average 8 Quarter Ahead Change in Mortgage Balances			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2002-04	1,202	7,760	9,663	6,745
2004-06	2,449	10,696	10,657	6,351
2006-08	1,175	8,397	8,896	4,260
2008-10	-7,459	-4,732	-2,192	-2,864
2010-12	-9,053	-3,413	-1,276	-2,515
2002-2006	5,304	27,419	30,604	20,248
2007-2010	-9,689	3,835	9,500	2,002
	Average 12 Quarter Ahead Change in Mortgage Balances			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
2002-05	2,073	9,390	10,605	7,132
2005-08	3,256	12,123	11,470	5,417
2008-11	-10,039	-9,680	-7,010	-6,390
2002-2006	5,589	22,032	23,454	15,411
2007-2010	-3,583	8,870	12,561	2,476

Cumulative change in mortgage balances at various horizons in USD. Based on 1 quarter lagged credit score quartile fixed effects and time effect from balance change regressions. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

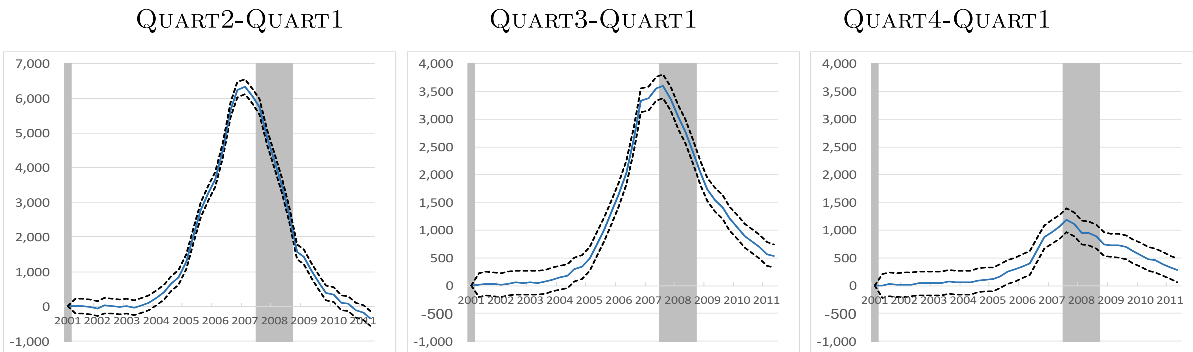


Figure 28: Difference in estimated time effects by 1Q lagged Equifax Risk Score quartile relative to quartile 1. Baseline specification. Dependent variable is the 8Q ahead change in per capita 90+ days delinquent debt balances in USD. Sample period 2001Q3-2011Q4. Number of obs. 64,588,488. Dashed lines denote 5% confidence intervals. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

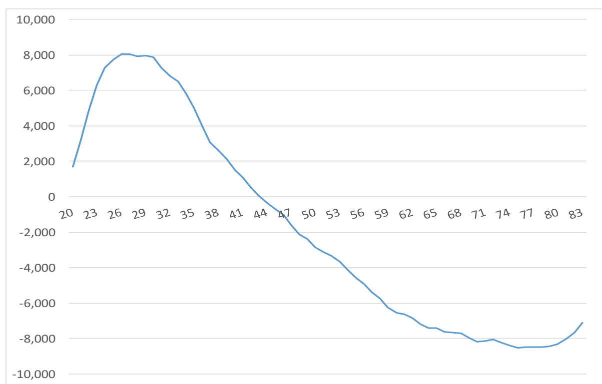


Figure 29: Estimated age effects from balance change regressions. Baseline specification. Dependent variable is the 8Q ahead change in per capita 90+ days delinquent debt balances in USD. Sample period 2001Q3-2011Q4. Number of obs. 64,588,488. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

noted. The fraction with 2 rises by about a third of the 2004 value, the fraction with 3 and 4 or more than double, though they start from a much lower level. For quartiles 2-3, these fractions return to pre-boom levels in the aftermath of the recession, while they remain at peak levels for borrowers in quartile 4. Figure 31 reports the share of balances held by borrowers with only 2, only 3 and 4 or more first mortgages. The behavior of balances broadly reflects the pattern of the fractions.

Figure 32 reports the fraction with a new 90+ days mortgage delinquency in the last 4 quarters by number of first mortgages, whereas figure 33 presents the fraction with a new foreclosure in the last 4 quarters by number of first mortgages. There is a sharp rise in the delinquency and foreclosure rates for borrowers with 2 or more first mortgages, with

Only 2 | Only 3 | 4 or More

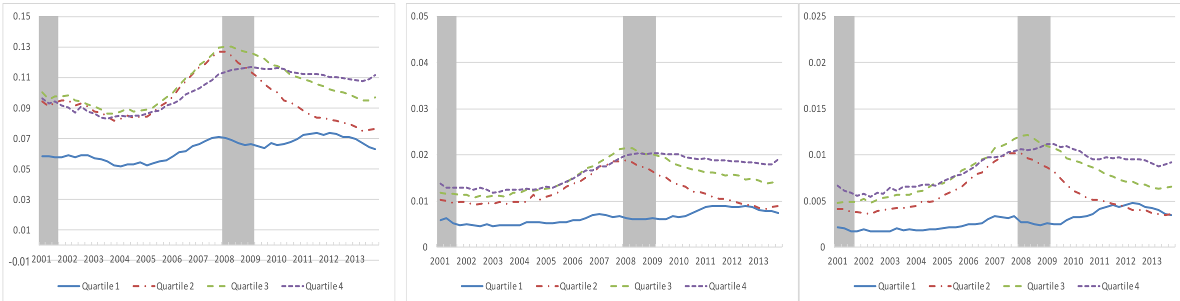


Figure 30: Fraction of borrowers with only 2 (left), only 3 (center) and 4 or more (right) first mortgages by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

Only 2 | Only 3 | 4 or More

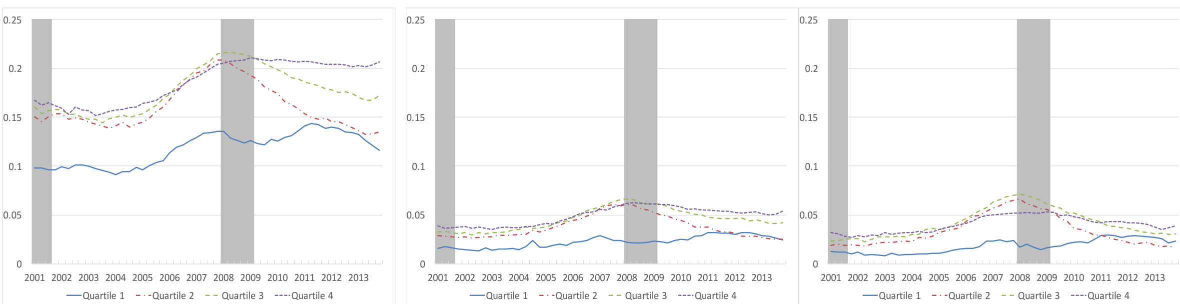


Figure 31: Share of mortgage balances held by borrowers with only 2 (left), only 3 (center) and 4 or more (right) first mortgages by quartile of the 8Q lagged Equifax Risk Score. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

the rates increasing in the number of firts mortgages. This is true for all quartiles of the 8 quarter lagged credit score distribution, though for delinquencies the rise is particular sharp for borrowers in quartiles 2 and 4, compared to both the pre-crisis rates and the rates of borrowers with only first mortgage in the same quartiles. The maximum delinquency rate is registered for borrowers with 4 or more first mortgages in quartile 3 of the credit score distribution, which reaches a rate of 3.5% in 2009Q2. For foreclosures, there is a sharp rise for all quartiles, and again the increase in rates is increasing in the number of first mortgages, with the maximum foreclosure rate attained by borrowers with 4 or more first mortgages in quartile 3 of the credit score distribution, at 3% in 2009Q3.

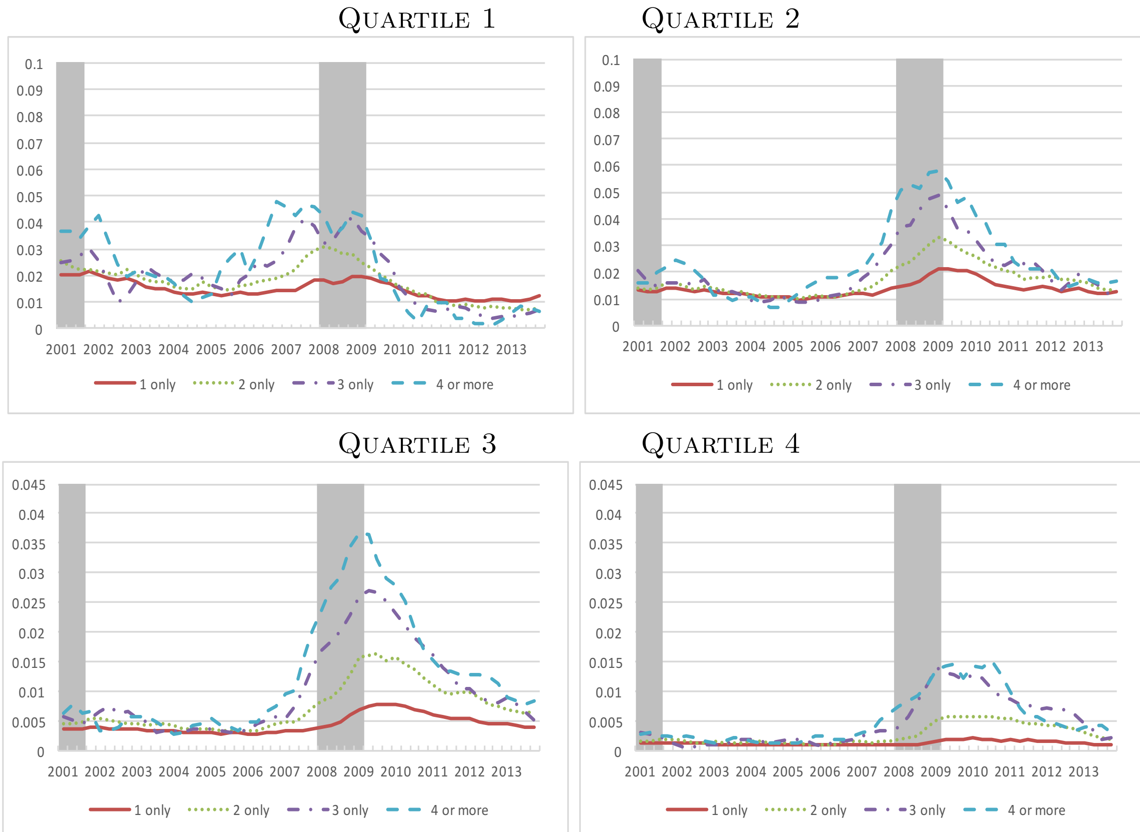


Figure 32: 90+ days mortgage delinquency rates by quartile of 8Q lag Equifax Risk Score quartile by number of first mortgages. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

F Additional Zip Code Level Evidence

F.1 Stability and Consistency of Zip Code Rankings

Mian and Sufi (2009) ranks zip codes by the fraction of subprime in 1996. Mian and Sufi (2011) ranks zip codes by initial personal disposable income or initial leverage, which they define as total debt balances per capita over average personal disposable income. Mian and Sufi (2014) rank counties by the decline in household net worth during the crisis, which is instrumented by the Saiz (2010) house prime elasticities to capture the rise in house prices during the boom and the associated rise in leverage. Here, we examine the relation between these measures at the zip code level.

We first consider the stability of each ranking. Table 14 reports the fraction of zip codes that remain in the same quartile of each ranking in the subsequent year. We consider three indicators: the fraction of subprime borrowers, average personal disposable income (PDI) and average leverage, defined as total balances per capita over average personal disposable income. All rankings are very stable, with approximately 70% of all zip codes remaining in the same quartile of the fraction of subprime borrower distribution year to year, over 90%

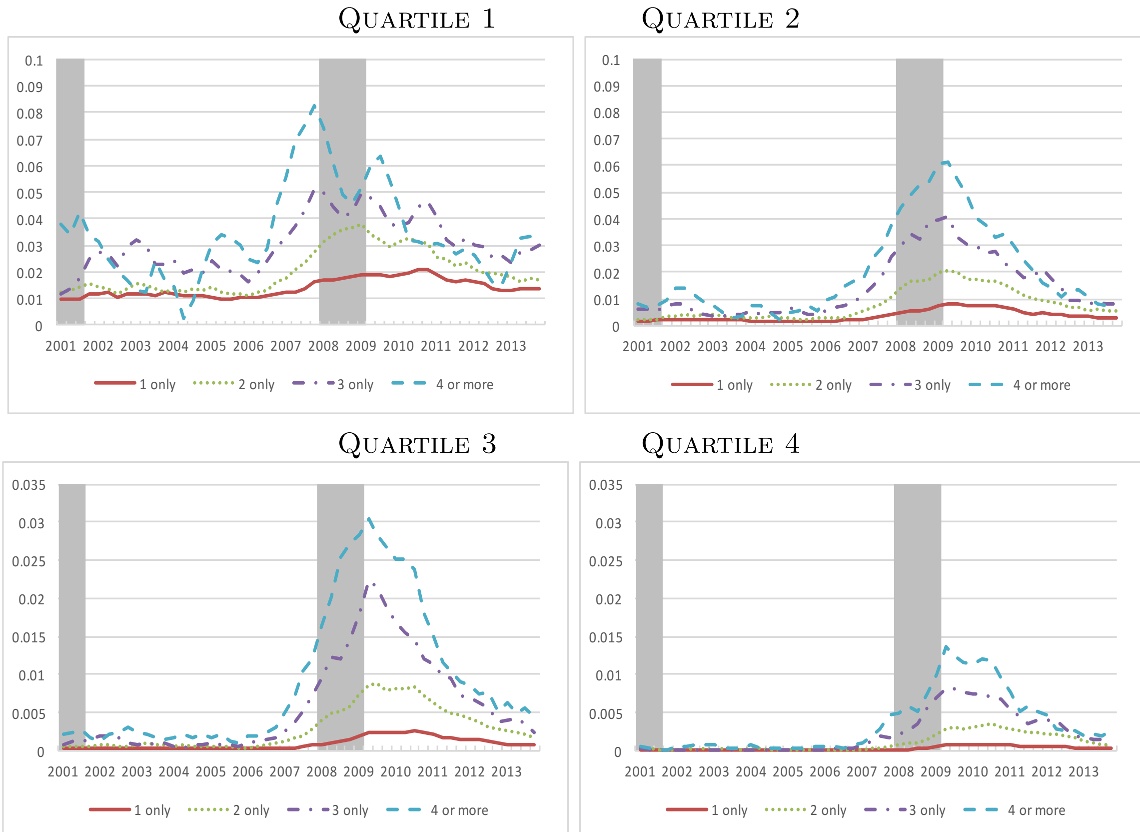


Figure 33: Foreclosure rates by quartile of 8Q lag Equifax Risk Score quartile by number of first mortgages. Source: Authors' calculations based on FRBNY CCP/Equifax Data.

for personal disposable income and 59-75% for leverage. We also examine the correlation between various rankings. The Spearman correlation between fraction of subprime and PDI ranges from -0.46 and -0.58, and decreases over the sample period. The Spearman correlation between fraction of subprime and leverage is negative, ranging between -0.03 at the end of the sample and -0.15 at the height of the credit boom. This is consistent with a greater growth in leverage for zip codes with low fraction of subprime during the boom.

We now concentrate on quartile 4 by fraction of subprime on 2001. We examine their income and leverage ranking throughout the sample period. The results are reported in Table 15. Depending on the sample year, 51-58% of the zip codes in quartile 4 of the fraction of subprime borrowers in 2001 are in the lowest PDI quartile in 2001-2011. Moreover, the fraction of subprime zip codes in higher PDI quartiles declines later in the sample period. The distribution of zip codes with high fraction of subprime borrowers across the leverage distribution is more even, however, in all years more than 50% are in the first 2 quartiles of the leverage distribution, confirming the negative relation between fraction of subprime borrowers and leverage.

Table 14: Stability and Correlation of Zip Code Rankings

	Fraction in same quartile			Correlation with % subprime	
	% subprime	PDI	Leverage	PDI	Leverage
2001	0.68	0.88	0.59	-0.46 ***	-0.04 ***
2002	0.71	0.91	0.62	-0.50 ***	-0.05 ***
2003	0.73	0.92	0.66	-0.51 ***	-0.06 ***
2004	0.70	0.90	0.63	-0.53 ***	-0.10 ***
2005	0.71	0.90	0.67	-0.53 ***	-0.15 ***
2006	0.72	0.89	0.67	-0.55 ***	-0.15 ***
2007	0.72	0.87	0.69	-0.58 ***	-0.09 ***
2008	0.72	0.92	0.73	-0.58 ***	-0.11 ***
2009	0.72	0.95	0.74	-0.58 ***	-0.04 ***
2010	0.73	0.95	0.75	-0.58 ***	-0.03 ***
2011	0.72			-0.57 ***	-0.03 ***

Fraction of zip codes in same quartile in subsequent year, by fraction of subprime borrowers, PDI and leverage. Correlation (Spearman ρ) of fraction of subprime borrowers in 2001 and PDI or leverage in each sample year. Leverage is the ratio of total debt balances to PDI. *** denotes significance at the 1% level. Source: Authors' calculations based on FRBNY CCP/Equifax Data, IPUMS, IRS, BLS, ACS data.

Table 15: Zip Codes in Quartile 4 by % of Subprime Borrowers in 2001

	PDI Quartile				Leverage Quartile			
	1	2	3	4	1	2	3	4
2001	0.51	0.27	0.14	0.07	0.28	0.27	0.23	0.22
2002	0.54	0.27	0.13	0.07	0.29	0.26	0.23	0.21
2003	0.55	0.26	0.13	0.07	0.29	0.27	0.22	0.21
2004	0.57	0.24	0.12	0.07	0.31	0.28	0.21	0.19
2005	0.59	0.23	0.11	0.07	0.35	0.27	0.21	0.17
2006	0.57	0.25	0.12	0.07	0.35	0.28	0.20	0.17
2007	0.58	0.25	0.11	0.06	0.33	0.28	0.20	0.19
2008	0.58	0.26	0.11	0.06	0.34	0.27	0.20	0.19
2009	0.58	0.25	0.11	0.06	0.31	0.26	0.20	0.23
2010	0.58	0.25	0.11	0.06	0.32	0.25	0.20	0.23
2011	0.58	0.26	0.11	0.06	0.31	0.24	0.20	0.24
2002-06 average	0.56	0.25	0.12	0.07	0.32	0.27	0.21	0.19

Fraction of zip codes in quartile 4 of the fraction of subprime borrowers in 2001 in various quartiles of the PDI and leverage distribution in each sample year. Leverage is the ratio of total per capital debt balances to average PDI. Source: Authors' calculations based on FRBNY CCP/Equifax Data, IPUMS, IRS, BLS, ACS data.