Does Greater Inequality Lead to More Household Borrowing? New Evidence from Household Data

Olivier Coibion
UT Austin and NBER
ocoibion@austin.utexas.edu

Yuriy Gorodnichenko
UC Berkeley and NBER
ygorodni@econ.berkeley.edu

Marianna Kudlyak
Federal Reserve Bank of Richmond
marianna.kudlyak@rich.frb.org

John Mondragon
UC Berkeley
jnm@econ.berkeley.edu

This Draft: October 9th, 2014

Abstract: Using household data on debt during 2001-2012 and regional variation in inequality, we show that low-income households in high-inequality regions accumulated *less* debt relative to income than those in lower-inequality regions. These patterns are consistent with inequality affecting the supply of credit to households when lenders use income, combined with local income inequality, to infer the underlying quality of applicants. Higher inequality makes lenders channel relatively more credit toward higher-income applicants. Data on individual mortgage applications supports this mechanism: low-income households were more likely to be denied or pay higher interest rates in high-inequality areas than in low-inequality areas.

JEL: E21, E51, D14, G21
Keywords: inequality, household debt, credit, income, Great Recession

We are grateful to Meta Brown, Donghoon Lee, Daniel Molitar, and seminar participants at the CES-Ifo, Cologne, CREI, Boston College, LBS, NBER SI ME and EFACR, FRB New York, Rice, FRB Richmond, SED, FRB St. Louis, Tinbergen Institute, U. of Houston, and VCU for helpful comments. The views expressed here are those of the authors and do not reflect those of the Federal Reserve Bank of Richmond or the Federal Reserve System or any other institution with which the authors are affiliated. Mondragon thanks the Richmond Fed for their generous support while part of this paper was written as well as support from the NSF. Gorodnichenko thanks the NSF and Sloan Foundation for financial support.

1 Introduction

The financial crisis of 2008-09 was preceded by an exceptional rise in borrowing by U.S. households, accounted for primarily by a rise in mortgage debt. Much of this mortgage debt was securitized and ultimately played a key role in destabilizing the financial system once housing prices began to decline and the associated mortgage-backed securities fell sharply in value. The increase in household leverage over the 2000s continued a broader trend of debt accumulation since the 1980s. Over the same timeframe income inequality in the U.S. increased to the highest levels seen in the post-war period (see Figure 1). These striking movements have led observers to wonder if the rise in income inequality caused some of the increase in household leverage.

There are numerous mechanisms discussed in the literature through which an increase in income inequality might affect household debt accumulation. If the increase in income inequality is due to an increase in the dispersion of temporary shocks then standard incomplete markets models suggest low-income households might borrow more to smooth their consumption (e.g., Iacoviello 2008, Krueger and Perri 2006). Alternatively, there is a long line of behavioral thought suggesting that income dispersion might affect household debt through "keeping up" motives. As inequality increases, high-income households are able to consume relatively more than low-income households. If low-income households experience a disutility from not consuming equivalent amounts as high-income households, they might try to maintain a higher level of consumption (e.g., Kumhof et al. 2013, Bertrand and Morse 2013), potentially funded by debt. Along similar lines Rajan (2010) argues that as inequality rose, credit was made increasingly available to lower-income groups to support their consumption levels in the face of stagnant incomes. While this argument suggests inequality increased the supply of credit, the political demand for this expansion in credit supply was arguably driven by disutility from the dispersion in consumption between the rich and the poor.

In this paper we focus specifically on the potential link between inequality and household borrowing. In particular, we investigate whether borrowing patterns on the part of low-, middle- and high-income households differed depending on the level of local income inequality where we define "local" as ranging from as fine a geographic level as the zip code to as aggregated a level as the state. There are several advantages to exploiting local inequality. First, there is considerably more variation in local inequality relative to the time series. This allows us to do numerous subsample and robustness checks in order to isolate the role of inequality. Second, local inequality is, from a household's point of view, likely to be a relevant metric for "keeping up" effects. Third, by varying the definition of local, we can shed light on the extent to which any observed relationship "aggregates." Finally, much of the rise in income inequality in the U.S. since the 1970s reflects a rise in inequality within regions rather than inequality across regions. If we wish to study the sensitivity of borrowing to inequality then local inequality levels are likely of first order interest.

To assess whether borrowing patterns depended on local inequality levels, we study the changes in debt-to-income ratios at the *household* level over the course of the 2000s and their relationship with households' relative standings in the income distribution and the amount of local income inequality. We use unique data from the New York Federal Reserve Bank Consumer Credit Panel/Equifax (CCP) which provides comprehensive debt measures for millions of U.S. households since 1999, including detailed decompositions of debt by type (i.e. mortgage, auto, credit cards, etc.). Because this dataset does not include a measure of household income, we use the relationship between household debt and income, conditional on observable household characteristics, in the Survey of Consumer Finances to predict initial household income in 2001. This imputation allows us to study the relationship between income and debt in unprecedented detail. We then characterize the evolution of household debt levels, relative to initial income levels, across income groups in areas with different levels of income inequality, akin to a "difference-in-differences" approach across income groups and regional inequality levels.

Our main finding is that high-income households in high-inequality regions accumulated more debt relative to their incomes than did low-income households in the same regions, or equivalently that low-income households in high-inequality regions borrowed relatively *less* than similar households in low-inequality regions. This effect is precisely the opposite of what one would have expected from "keeping up" motives or consumption smoothing driving the rise in household debt during the 2000s. We show that this result is remarkably robust and holds up to an extensive array of robustness checks: e.g. we find these patterns within households with low or high credit scores, within regions which experienced either high or low home price appreciation, within households with either low or high initial debt levels, etc. We measure inequality at the zip code, county, and state and find similar results across levels of aggregation. This robustness to controlling for a wide range of other local factors that are correlated with inequality levels suggests that it is indeed the level of inequality that matters rather than inequality being a stand-in for other economic channels.

Because our data provide disaggregated information on household debt, we assess the link between local inequality and different forms of debt: mortgage debt, auto debt, and credit card debt. Importantly, we find strong evidence that low-income households in high-inequality regions borrowed less in terms of both mortgage and auto debt than those in low-inequality regions, which implies that our results are not driven just by local housing markets, but rather reflect broader borrowing patterns. A unique feature of the data is that we have information on both credit card balances as well as credit card limits. This is particularly useful because we expect credit card balances to be more elastic with respect to demand than credit card limits. We find that low-income households in high-inequality regions saw their credit limits rise by less than those in lower inequality regions as was the case with mortgage and auto debt. At the same time, no economically significant heterogeneity is observed in credit card balances. We interpret this contrast as pointing to supply side factors as being at the root of the differential debt accumulation patterns that we observe in the data.

To illustrate how supply-side factors can explain the differential borrowing behavior tied to regional inequality, we present a model in which each region is composed of two types of households. High-type households have higher income on average than low-type households and are also less likely to (exogenously) default on debt. Banks in each region lend to these households but they do not observe households' types, only their income and another signal correlated with the underlying type. As income inequality rises, banks treat an applicant's income as an increasingly precise signal about their type and therefore target lending toward higher-income households on average. How they do so, however, can vary with the local banking structure. For example, if banks are perfectly competitive and can charge different interest rates to different applicants, then higher-income applicants will on average face lower interest rates than low-income applicants, and this difference will be increasing in the amount of local income inequality. If instead we model the banking system as being monopolistic and forced to charge a common interest rate to all applicants, then this bank will reject low-income applicants more frequently than high-income applicants, and this difference will again be increasing in the amount of local inequality. In both cases, banks will make credit more accessible (or cheaper) to high-income households when local inequality is higher since income is a more precise signal of applicant types.

The credit supply mechanism presented in the model has testable implications for the behavior of the prices and availability of credit. If inequality is affecting credit supply in the manner we suggest, then we expect to see richer households be denied less often when applying for credit in high-inequality regions relative to similar household in low-inequality regions. Similarly, we expect richer households in high-inequality regions to pay lower interest rates on a loan on average as inequality is increasing. In other words, we expect the price of credit, broadly understood, for high-income households to *decline* with inequality as the supply of credit to these households increases. At the same time we expect the price of credit to *increase* with inequality for low-income households as the supply of credit to these households declines.

We test these theoretical predictions using detailed mortgage application information from the publicly available Home Mortgage Disclosure Act data (HMDA). These data track mortgage applications as they go through the origination process and contain information on applicants (including their income, the amount of the loan requested, their locale, and whether the loan is denied or originated). We document that high-income households in high-inequality regions were less likely to be denied for mortgages than their counterparts in low-inequality regions, as predicted by the theory. High-income households in high-inequality regions were also less likely to be charged higher interest rates for their mortgages than similar households in low-inequality regions. Thus, both theoretical predictions from the model are confirmed in the data.

In summary, we document that high-income households borrow more as inequality increases and low-income households borrow less. This finding is in stark contrast to inequality affecting household debt through "keeping up" motives or consumption smoothing. We argue that these results can instead be explained through an information channel: applicants' incomes are a stronger signal of their underlying quality when local inequality is high so banks are likely to channel relatively more credit to high-income applicants. Using information on prices from mortgage data we find evidence consistent with this supply-side story.

This paper is closely related to recent work evaluating the strength of "keeping up with the Joneses" forces. Bertrand and Morse (2013) study whether rising consumption of the rich induces the non-rich to consume more.[1] Using the Consumer Expenditure Survey (CES), they find that, within a state, the consumption of the rich predicts higher consumption for the non-rich, holding everything else constant including own income. Bertrand and Morse interpret their estimates as supporting the view that rising income inequality in a geographic market translates into more demand for credit by low- and middle-income households (e.g. Rajan, 2010). In contrast, by focusing explicitly on the borrowing decisions of households and exploiting a finer level of geographic variation, we document that low- and middle-income households living in high-inequality regions borrowed less than similar households in low-inequality regions. This need not be interpreted as contradicting the results of Bertrand and Morse (2013) since the differences in consumption that they document could have been financed through channels other than debt, e.g. through increased labor force participation, longer working hours, etc. However, our results indicate that "keeping up with the Joneses" forces are unlikely to have played a primary role in accounting for the dramatic rise in household *leverage* during the 2000s.

This paper also relates to a broader line of research investigating the macroeconomic consequences of income inequality, such as whether they are systematically related to financial crises. Kumhofet al. (2013), for example, argue that a rise in inequality driven by an increase in the share of income going to those at the top of the income distribution induces the latter to save more, lowering interest rates and inducing poorer households to borrow more, ultimately leading to more financial fragility and a higher likelihood of a financial crisis. Bordo and Meissner (2012) find little evidence of such a link based on aggregate data since 1920 for fourteen advanced economies, whereas Perugini et al. (2013) find a positive link between income inequality and private sector indebtedness since 1970 across eighteen economies. We contribute to this literature by documenting how, within U.S. regions, debt accumulation patterns across different segments of the population over the course of the 2000s were systematically related to local levels of income inequality. We also provide a novel interpretation for these effects: local income inequality can be used in combination with an applicant's income level to refine inference about borrower types. Higher levels of income inequality then induce banks to reallocate credit toward higher-income applicants and away from lower-income applicants, thereby potentially amplifying the implications of a more unequal income distribution for the distribution of consumption.

The relationship between income inequality and the allocation of credit emphasized in our paper also relates to the literature on consumption and income inequality. Krueger and Perri (2006) and related works argue that consumption inequality during the last decades did not rise with income inequality.[2] Krueger and Perri argue that low-income households have experienced income shocks that increased income inequality, but due to enhanced financial intermediation these households have been able to smooth their consumption such that consumption inequality remained stable. Iacoviello (2008) replicates the trend and cyclicality of household debt since the 1960s and also argues that increased access to credit has allowed households to smooth increasingly

[1] Prior evidence in the same spirit as Bertrand and Morse (2013) includes Neumark and Postlewaite (1998), Zizzo and Oswald (2001), Christen and Morgan (2005), Luttmer (2005), Daly and Wilson (2006), Maurer and Meier (2008), Charles et al. (2009), Kuhn et al. (2010), Heffetz (2011), and Guven and Sorensen (2012).
[2] Related papers are Blundell et al. (2008), Heathcote at al. (2010a), and Heathcote et al. (2010b).

volatile income processes so that the aggregate level of debt increases with inequality. In contrast, Aguiar and Bils (2012) argue that, when one corrects for measurement errors associated with underreporting of consumption expenditures over time and across different goods, consumption inequality has tracked income inequality closely over the last three decades. While this line of research appeals to financial intermediation as a link between consumption and income inequality, it could not measure directly the quantitative importance of formal borrowing for smoothing shocks and its relation to income inequality due to data constraints. We examine this issue directly using household level data on debt accumulation. Our results are consistent with the findings in Aguiar and Bils (2012): if low-income households were smoothing temporary income shocks with debt then low-income households should have accumulated relatively more debt where inequality is higher, but we find the opposite.

We also contribute to the vast literature on household borrowing that covers such diverse topics as pricing of mortgages, optimal portfolios of household debt, risk scoring, and determinants of default probabilities. Our paper is most related to studies of default determinants (e.g., Fay et al. 2002, Gross and Souleles 2002) and lenders' treatment of loan applications (e.g., Tootell 1996, Munnell et al. 1996, Turner and Skidmore 1999) in the sense that we attempt to understand who obtains credit and at what terms. However, while previous research studies these aspects for borrowers without relating a given individual to the pool of borrowers, we explicitly focus on how the relative positions of borrowers in the income distribution as well as the properties of the income distribution can affect the level of debt that households ultimately accumulate. Thus, in contrast to the previous literature, we examine directly the interplay between debt and inequality, which has been the subject of recent policy and academic debates.

This paper is structured as follows. We describe our primary source of data in section 2 as well as our imputation procedure for income. In section 3, we present household-level regressions describing the differential debt accumulation patterns across income levels in regions with different levels of income inequality. Section 4 presents a model that can explain these patterns. In section 5, we test and confirm the additional predictions of the model using data on mortgage applications by individuals in different inequality areas. Section 6 concludes.

2 Data

In this section, we first describe the dataset used to measure household debt accumulation over the course of the 2000s. Second, we discuss how we impute household income based on observed patterns in the Survey of Consumer Finances. Third, we construct local income inequality measures and describe some of their properties.

2.1. The New York Federal Reserve Bank Consumer Credit Panel/Equifax

We measure household debt accumulation using the New York Federal Reserve Bank Consumer Credit Panel/Equifax (CCP) data. The CCP is a quarterly panel of individuals with detailed information on consumer liabilities, delinquency, some demographic information, credit scores, and geographic identifiers to the zip level.[3] The core of the database constitutes a 5% random sample of all U.S. individuals with credit files. The

[3] For complete details on the data set and variables construction, see Appendix B.

database also contains information on all individuals with credit files residing in the same household as the individuals in the primary sample. The household members are added to the sample based on the mailing address in the existing credit files. Using the households' identifiers, we aggregate individual records into households' records and construct measures of households' debt. The resulting sample is a quarterly sample of U.S. households in which at least one member has a credit file. We use 100% of the CCP sample. Lee and van der Klaauw (2010) provide an excellent and detailed description of the database.

The data cover all major categories of household debt including mortgages, home equity lines of credit (HELOC), credit cards, and student loans. Because of the large sample size, the breadth of variables observed, detailed location, and the ability to construct a quarterly household panel these data provide the most detailed picture of household debt available.

2.2. Income Rank Imputation

While the CCP provides detailed records of household debt and geographical location, it does not include information on household income. To address this issue, we impute income for the households in the CCP using information from the Survey of Consumer Finances (SCF). The SCF is a household-level survey that contains information on debt balances and income as well as a rich set of demographic characteristics. However, the SCF does not provide geographic identifiers in the publicly available data. We use the SCF to estimate how household income relates to debt and demographic characteristics available in both the CCP and SCF data sets. We then use these estimates to impute household income in the CCP data. Finally, we use the imputed income and the estimated error terms from the SCF to impute the household's income rank in the household's geographical area and the distribution of income in that area.

In our analysis, we restrict the sample to households for whom the household head's age is between 20 and 65 to minimize potential age-related selection effects. The data in the CCP are updated quarterly. We use data from the third quarter of the CCP for years 2001 - 2012. We follow Brown et al. (2011) and choose the third quarter to maximize the match with the SCF survey (typically administered between April and December), which we use to impute the initial income distribution as described below. For consistency, we then use the third quarter of each subsequent year to generate annual measures of household debt.

Table 1 contains the summary statistics from the CCP and SCF samples from the third quarter of 2001. The statistics from the CCP and SCF are similar for most categories with the exception of credit card balances. This finding is consistent with Brown et al. (2011) reporting that overall and in the majority of disaggregated debt categories (mortgages, auto loans, and HELOCs), borrower characteristics and debt levels reported in the CCP and SCF are similar. Brown et al. (2011) suggest that some of the discrepancy between the credit card balance statistics in the two datasets might come from the way credit card balances are recorded: the CCP contains records of all credit card balances, whereas the households in the SCF might only report the fraction of the balance they intend to roll over.[4] The mortgage balance and HELOCs in the CCP are slightly higher than in the SCF because the CCP measure includes secondary/investment properties, while in the SCF it does not (see Brown et al. 2011). The auto debt balance is also slightly higher in the CCP because the CCP includes auto leases, while in the SCF respondents usually do not necessarily report car leases as auto debt. The bankruptcy rates are very similar between the two samples. The tables also show some differences between the delinquency statistics in the two datasets. It is possible that SCF households only report severe delinquencies on large quantities of debt and do not report delinquencies that they regard as temporary or small.[5]

To impute the rank in the income distribution for a household in the CCP, we first estimate the following relationship between the household's gross income and observable characteristics in the 2001 SCF,

$$\log\left(Y_{i,SCF}\right) = \beta f(\, X_{i,SCF}) + \, \epsilon_{i,SCF}, \tag{1}$$

where $Y_{i,SCF}$ is the income of household i, and $X_{i,SCF}$ is the vector of the household's characteristics that include (logs of) mortgage balance, credit card balance, credit card limit, an indicator for positive credit card limit, the credit card utilization rate conditional on positive credit card limit, auto loan balance, HELOC balance, student loan balance, an indicator for bankruptcy, an indicator of 60 days or more past due on any loan, the age of the head of the household and the household size. $f(.)$ is a vector-valued function that includes polynomials, interaction terms, and dummy variables. Appendix F provides more information on the specification and variables. We estimate equation (1) using OLS (with the SCF sampling weights) and eliminate outliers using Cook's distance.[6] The unadjusted R^2 for this regression is 0.55.

Using the estimated β, we construct the expected imputed (log) income for each household i in the third quarter of 2001 in the CCP data:

$$\mathrm{E}[\log(Y_i)] = \hat{\beta} f\left(X_{i,CCP}\,\right),$$

and the expected imputed income (in levels)

$$\mathrm{E}[\,Y_i] = \exp[\mathrm{E}[\log(Y_i)] + 0.5\sigma^2_{\hat{\epsilon}_{i,SCF}}\,],$$

where $\sigma^2_{\hat{\epsilon}_{i,SCF}} = 0.423$ is the variance of $\epsilon_{i,SCF}$ estimated in equation (1).

Having imputed households' income in the CCP, we then estimate the household's rank in the local income distribution. For each household i in area c we construct its income rank in 2001, $R_{i,c,2001}$, as the rank

[4] In the CCP, the credit balance is recorded on some date during the quarter. For some individuals, this can be the date right before they pay off most of their credit balance, and the balance might largely reflect the transaction use of the credit cards. For other individuals, the date might be the date after they pay off the intended balance and the remaining amount reflects the carry-over balances. In the SCF, the credit balance reported likely does not reflect the use of credit card for transactions, but rather the debt that the household does not plan to repay in the current period. In addition, the households in the SCF might forget older balances.

[5] In the SCF data, the 60DPD indicator is the indicator of whether a household has ever been delinquent on any loan for 60 days or longer. In the CCP data, the 60DPD indicator is the indicator of whether a household is delinquent on any loan for 60 days or longer in the current quarter.

[6] Equation (1) is estimated only for observations with positive values of income. We also restrict our analysis to the 50 U.S. states and the District of Columbia, dropping the observations from Puerto Rico and U.S.-owned territories.

of the household's expected imputed income, E[$\log(Y_{i,2001})$], in the imputed income distribution for location c. We approximate the local income distribution through a simple resampling procedure. In particular, we assume that the distribution of income residuals estimated in the SCF is the same across all locations. Note that if this assumption is not appropriate, we will tend to bias our results against finding any role for inequality in accounting for debt dynamics. However, our results are robust to using alternative measures of inequality that do not rely on this imputation procedure. After drawing a household from location c in the CCP and calculating its expected income, we add a randomly drawn residual estimated on the SCF sample to obtain a simulated household income:

$$\log(Y_{i,c,CCP}) = \hat{\beta} f(X_{i,c,CCP}) + \hat{\epsilon}_{SCF}\,.$$

By repeating the process 50,000 times, with draws done with replacement, we approximate the local income distribution. We then calculate each household's percentile rank ($R_{i,c,2001}$) using their expected income relative to the simulated distribution of incomes from that region. The higher the value of $R_{i,c,2001}$, the relatively richer is household i in its geographical location c in 2001.

We separately construct the rank of the household by the household's location at the three different levels of aggregation: zip code, county, and state. When the measure is constructed at the zip code level, we restrict the analysis to zip codes with at least 100 households in our CCP sample. This gives us 14,529 distinct zip codes in 2001. At the county level, we restrict the analysis to counties with at least 300 households in our CCP sample. This procedure gives us 2,303 counties in 2001, covering over 35,000 zip codes.

We check the quality of our imputation in a number of ways. First, we can easily check the quality of the rank imputation within the SCF itself, although this does not speak to the quality of the imputation across geographies. Regressing the true percentile rank on the imputed rank and a constant gives us a coefficient of 0.69 with a robust standard error of 0.004, extremely significant. To test that the imputation provides a useful measure of income in the CCP, Table 2 presents the moments of the income distribution imputed in the CCP and the same moments calculated from the SCF. The two sets of moments are very similar especially in the middle of the distribution, suggesting that our imputation function is sensible. We also check the quality of our income imputation procedure by bringing income information to the CCP data from an alternative source. We merge the CCP data with the data from a proprietary database. This database has detailed mortgage-level panel data that contain information on a majority of mortgages originated in the U.S. Critically, these data include the debt-to-income ratio associated with each mortgage at the time of origination. We use information on the mortgage origination month, location (zip code) and balance from this proprietary database and the same attributes from the mortgage trade-line data in the CCP to match households in the two datasets as in Elul et al. (2010). The earliest year when the debt-to-income variable is available in the proprietary dataset and when the SCF is available is 2007; thus we merge the data using the first mortgages originated in 2007 and re-estimate our imputation equation for 2007. Prior to the merge, we eliminate all cases of multiple mortgages with the same combination of open month, initial balance and zip code in both datasets to ensure that the match is unique. For the sample of matched households we use the debt-to-income ratio from the proprietary database and the debt in the CCP to estimate the income. For this subset of matched households we compare the income rank derived from the proprietary data with the income rank derived from the SCF-CCP imputation. The two measures of rank are highly positively correlated (Spearman correlation is 0.55). Regressing the imputed CCP income measure on the actual measure of income yields a slope estimate which is practically one suggesting a classical measurement error relationship between the two measures of income.

2.3. Local Inequality Measures

Having imputed income in the CCP, we construct the local inequality measures for 2001 ($I_{c,2001}$). Our preferred measure of inequality is the difference between expected log income at the 90th percentile and expected log income at the 10th percentile, i.e.,

$$I_{c,2001} = p90_c[\, E\, \{ \log(Y_{i,c,2001})\}\,] \; - \; p10_c[\, E\, \{ \log(Y_{i,c,2001})\}\,] \; .$$

We then compare this measure to inequality measures constructed from alternative sources. At the zip code level, we use data from the IRS on household adjusted gross income (AGI) drawn from the 2001 tax returns. At the county level, we use the Census data on household income from 2000. Both of these sources provide income bins and the fraction of the population within each bin. Using this information, we construct a simple approximation to the Gini coefficient. The CCP measure constructed from imputed incomes is highly correlated with Gini coefficients based on Census or IRS data. For example, the correlation between Gini coefficients from the 2000 Census and 90-10 differences in the CCP data at the county level is 0.59.

Figure 2 plots a map of U.S. inequality at the county level. Inequality is on average highest in the southern states, as well as California and the Pacific Northwest. Midwestern states, in contrast, stand out for having some of the lowest levels of inequality on average. The map also shows that inequality tends to be higher in large cities than in more rural areas. The map masks even greater regional heterogeneity in inequality at the zip code level. Figure 3 plots histograms of our CCP inequality measure at each level of aggregation. Average inequality is higher at lower levels of aggregation with a mean across zip codes of 2.24 and a mean of 1.68 across states. The standard deviation of inequality is twice as high (0.15) at the zip level compared to the state level (0.07).

We focus on local income inequality for a number of reasons. First, this is likely to be the most relevant metric when households compare themselves to others. Second, it avoids measurement issues associated with comparing incomes across very different areas (e.g. $100K in New York vs. Tulsa). Third, much of the rise in aggregate inequality in the U.S. reflects rising inequality within regions rather than across regions.[7] Finally, there is much more variation in income inequality across regions than in aggregate inequality over time, which is necessary for isolating any potential effects of inequality on household behavior.

[7] In Appendix C, we describe in detail a decomposition of aggregate income inequality in the U.S. from 1970 to 2000 measured using Census income data. When we measure the relative importance of differences in mean incomes across regions ("between" inequality) versus the dispersion of incomes within regions ("within" inequality) for each Census, we find that "between" inequality has consistently accounted for less than two percent of total inequality and that this share has, if anything, been declining over time.

3 Empirical Analysis of Debt and Inequality

In this section, we investigate whether households' borrowing patterns from 2001 to 2012 varied with local inequality. We do so using household level regressions of debt-to-income changes over time as a function of household characteristics, their position in the local income distribution, and interactions of the latter with local inequality measures. While the evidence supports the notion that local inequality affected debt accumulation patterns across income groups, the direction of the effect is opposite to what one would expect from "keeping up" or consumption smoothing effects. We document the robustness of this result along a variety of dimensions.

3.1. Baseline Results

We are interested in estimating the role of initial local income inequality on the relationship between a household's debt accumulation and their rank in the initial local income distribution. In particular, we estimate the change in each household's debt between 2001 and year t, $2002 \leq t \leq 2012$, as a function of their income rank in the 2001 local income distribution, conditional on local income inequality in 2001. The benchmark specification is

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \beta I_{c,2001} + \gamma R_{ic,2001} \times I_{c,2001} + c^{+} + \epsilon_{ict}, \qquad (2)$$

where $\frac{\Delta D_{ict}}{E[Y]_{ic,2001}}$ is the change from year 2001 to year t in the debt of household i that resides in location c relative to the household's (imputed expected) income in 2001 (in levels), i.e., $\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} \equiv \frac{D_{ict} - D_{ic,2001}}{E[Y]_{ic,2001}}$, where D_{ict} is deflated by the CPI-U and expressed in 2001 dollars. c^{+} is a fixed effect of the geographical location that is at one level of aggregation higher than the geographic area used to construct the income distribution and the income inequality measure.[8] We use the 2001 measure of local income inequality because it is predetermined relative to subsequent household debt accumulation decisions, although inequality is highly persistent over time (see Appendix D).

Parameters α, β and γ describe the relationship between a household's debt accumulation and local inequality. If $\alpha < 0$, low-rank households within an area accumulate relatively more debt than high-rank households. If $\beta = \gamma = 0$, then local inequality is irrelevant for household debt accumulation. This case is shown in Panel A of Figure 4. Panel B of Figure 4 illustrates the case when $\alpha < 0, \beta > 0, \gamma < 0$. If $\beta > 0$, an area with higher inequality is associated with higher debt accumulation. If $\gamma < 0$, this effect weakens as household rank increases, which is an example of the "keeping up" hypothesis. The final panel illustrates a case where $\gamma > 0$. In this case there is a crossing point such that to the right high-income households accumulate more debt as inequality increases. To the left of this crossing point low-income households accumulate less debt as inequality increases. The aggregate effect depends on the exact crossing point and relative slopes.

[8] For example, in the regressions with zip code-level distribution of income and inequality, we control for county-level fixed effects. In the regressions with county-level rank and inequality, we control for state-level fixed effects. We do not control for the geographical fixed effects in the regressions with state-level income rank and inequality.

We estimate equation (2) separately for each year t, $2002 \leq t \leq 2012$. In each year t, we follow Guerrieri et al. (2013) and restrict the sample to households that reside in the same geographical area c in 2001 and in t. In each regression, we exclude the observations below the 2nd and above the 98th percentile of the distribution of $\frac{\Delta D_{ict}}{E[Y]_{ic,2001}}$ in year t. The standard errors are clustered by geographic location c.[9]

Our baseline estimates of equation (2), estimated at the zip code level with county fixed effects for years ranging from 2002 to 2012, are reported in Panel A of Table 3.[10] Our first finding is that the coefficient on a household's rank in the income distribution (α) is consistently negative, with a peak absolute value in 2007. Hence, debt accumulation over the course of the early to mid-2000s was, on average, greater for lower-income households. Second, the estimated coefficient on the inequality level of the zip code is systematically negative, again peaking in absolute value in 2007. This implies that, holding everything else constant, households living in the more unequal areas within a county accumulated *less* debt over the early to mid-2000s than did those in lower inequality areas in the same county.

The key parameter for us is γ, which captures the interaction of household rank in the local income distribution and local inequality. Our main finding is that γ is positive over this time period. This implies that debt accumulation was relatively higher for (sufficiently) high-income households in high-inequality regions than in low-inequality regions, or equivalently that *lower-income households in high-inequality regions borrowed relatively less than their counterparts in lower inequality regions*. This result is precisely the opposite of what one would have expected from "keeping up" effects. Panel C of Figure 4 describes our results qualitatively. Households with rank to the right of the crossing accumulate more debt on average as inequality increases. Households to the left of the crossing accumulate relatively less debt as inequality increases.

To give a sense of the economic magnitudes, we calculate the change in debt accumulation in response to a one standard deviation increase in local inequality for households of several different ranks. Panel A of Figure 5 plots these calculated effects at the 80th, 50th, and 20th percentiles for each time sample. At the 80th percentile a one standard deviation increase in inequality implies an increase in household debt over expected income of more than nine percentage points in 2007. At the 20th percentile we estimate that households decreased debt relative to income by a little over seven percentage points in 2007. In the same year, the median household saw an increase in debt-to-income of little more than one percentage point.

3.2. Specifications with Additional Controls

Our baseline specification does not include any household-specific controls other than their rank in the income distribution. To control for potentially confounding household characteristics, we consider an expanded specification augmented to include a vector of household-specific regressors:

---

[9] Each specification below is estimated using household sampling weights from 2001, as described in Appendix B.

[10] In general we report standard errors uncorrected for the fact that rank and inequality are generated regressors. The standard errors are very similar but computationally burdensome when we use a bootstrap to correct for the generated regressor.

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \beta I_{c,2001} + \gamma R_{ic,2001} \times I_{c,2001} + \psi X_{ic} + c^{+} + \epsilon_{ict}, \qquad (3)$$

where X_{ic} is the set of household-specific controls. The latter include the age of the head of the household, household size, (logarithm of) the level of household's mortgage debt, (logarithm of) the level of household's auto debt, (logarithm of) the level of household's HELOC debt, (logarithm of) the level of household's student loan debt, an indicator for a non-zero credit card debt limit, (logarithm of) the level of household's credit card debt, (logarithm of) the level of household's credit card limit, the credit card utilization rate conditional on non-zero credit card limit, default indicators, and the average of household members' credit scores. All controls are from 2001, with the exception of credit scores for which we include both 2001 values (to control for initial access to credit) as well as year *t* values (to control for access to credit in subsequent years). Results from this augmented specification are presented in Panel B of Table 3. The results for the estimated effects of rank, inequality, and the interaction of the two are almost identical to those from the parsimonious specification.

A second concern one might have is that regional inequality is correlated with other regional economic characteristics and that it is the latter that are most relevant for household debt accumulation decisions. We control for this possibility in several ways. First, we include an additional vector of zip-level control variables:

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \beta I_{c,2001} + \gamma R_{ic,2001} \times I_{c,2001} + \psi X_{ic} + \kappa W_c + c^{+} + \epsilon_{ict}, \qquad (4)$$

where W_c is the set of location-specific controls. The set of location-specific controls includes the median expected income in the zip code in 2001, the median of (log of) the household's total debt in 2001, and the median of (log of) the household's mortgage debt in 2001. Results are presented in Panel C of Table 3. Again, our baseline estimates of the effects of household rank, local inequality and their interaction are almost unchanged. This is also illustrated graphically in Panel B of Figure 5: our estimates with both household and regional controls suggest that increasing inequality by one standard deviation is associated with households at the 80th percentile increasing borrowing relative to income by almost 11 percentage points, at the 50th percentile households increase borrowing over income by over one percentage point, and at the 20th percentile households decrease borrowing over income by about eight percentage points. The difference between high- and low-rank households is essentially identical.

Another way to control for regional characteristics is to estimate our baseline specification with fixed effects at the level of the zip code rather than the county:

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \gamma R_{ic,2001} \times I_{c,2001} + \psi X_{ic} + \delta_c + \epsilon_{ict}. \qquad (5)$$

With zip code-specific fixed effects δ_c, we can no longer separate the effect of local inequality from other regional characteristics, but we can still estimate the coefficient on the interaction term between the household's income rank and local inequality, γ. The results from estimating equation (5) are presented in Panel D of Table 3: the estimate of γ is again almost unchanged relative to those from our parsimonious specification (2) or specifications augmented with household (3) and regional controls (4).

We also check for omitted variable bias in the interaction term by adding the interaction of the household credit risk score with local inequality to the specification in equation (3). Specifically, this deals with the concern that income might be a proxy for some other variable actually driving debt accumulation. If the measure of income rank primarily picked up the relative importance of the household's credit risk score, the estimate of γ should differ significantly after including this interaction. We estimated the following modification of specification (3):

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \beta I_{c,2001} + \gamma R_{ic,2001} \times I_{c,2001} + \psi X_{ic}$$
$$+\phi Risk_{ic,2001} + \sigma Risk_{ic,2001} \times I_{c,2001} + c^{+} + \epsilon_{ict}, \qquad (3')$$

The estimates of γ across all years (Panel A, Table 4) are robust to the inclusion of the interaction term.

Similarly, we check whether the results are sensitive to including an interaction of the household's initial debt level with local inequality in specification (3):

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha R_{ic,2001} + \beta I_{c,2001} + \gamma R_{ic,2001} \times I_{c,2001} + \psi X_{ic}$$
$$+\phi Debt_{ic,2001} + \sigma Debt_{ic,2001} \times I_{c,2001} + c^{+} + \epsilon_{ict}, \qquad (3'')$$

Our baseline findings are unchanged with these additional controls (Panel B of Table 4).

We verify that our results do not hinge on the CCP measure of income inequality. We replicate our results from Table 3 in Appendix Table A1 using the measure of inequality constructed from IRS data and described in section 2.3 and find almost identical results. Finally, we also check that we are not mechanically inducing any spurious correlation between the interaction term and our outcome by using the imputed income on the left hand side and imputed rank in the interaction. To check this we estimate two additional specifications. The first replaces rank with the inverse of imputed income

$$\frac{\Delta D_{ict}}{E[Y]_{ic,2001}} = \alpha \frac{1}{E[Y]_{ic,2001}} + \beta I_{c,2001} + \gamma \frac{1}{E[Y]_{ic,2001}} \times I_{c,2001} + \psi X_{ic} + \kappa W_{c} + c^{+} + \epsilon_{ict}. \qquad (6)$$

By including the inverse of imputed income on the right hand side, we are inherently removing any first-order correlation between the outcome and variables on the right hand side. Thus, any higher-order correlation must be a feature of the data. The results of this estimation are found in Appendix Table A2 Panel A and show that with this specification we get qualitatively the same results since now the signs are reversed. In Appendix Table A2 Panel B we also estimate a specification where the outcome variable is the log difference of total debt keeping the baseline regressors and controls as in (4). We again find qualitatively similar results: low-income households saw their debt grow by less in high inequality areas than similar households in less unequal areas.

In short, the differential debt-accumulation patterns by households of differing income levels across inequality regions are a robust feature of the data.

3.3. Subsample analysis

Our finding that debt accumulation was higher for poorer households in low-inequality regions than high-inequality regions is robust to controlling for a wide variety of household and regional observables. One may be concerned however that our interaction effect is capturing some other nonlinear characteristic of household borrowing, which need not be captured by linear controls. To address this possibility, we consider an additional set of robustness checks in which we verify that our results still obtain within subsets of the data. Specifically, we break our regions along four dimensions: geographic areas, initial debt burdens, credit scores and house price growth. Note that in each of the subsample regressions we do not normalize inequality so that differences in magnitude are not necessarily the result of differences in economic effects.

For geographic areas, we estimate our specification with household and regional controls (equation (4)) separately for each of the four Census regions: Midwest, Northeast, South and West. We present the results of the household level regressions of debt accumulation from 2001 to 2007 (the main period over which household debt increased sharply) for each region in Panel A of Table 5, with the full set of yearly regressions by region available in Appendix Table A3. For each region, the coefficients are of the same sign as before and of approximately the same order of magnitude. Hence, our baseline results are confirmed within each region of the country.

Second, we decompose zip codes by the average level of credit scores among households in each locale in 2001. Specifically, we group zip codes into three bins: low credit scores (below the 33rd percentile of average credit score distribution), medium (between the 33rd and 67th percentiles) and high credit scores (above the 67th percentile of the average credit score distribution). We then rerun our specification with household and regional controls within each of these three credit score areas. The results for 2001-2007 are presented in Panel B of Table 5, with all yearly regressions by credit score grouping available in Appendix Table A4. Again, the results are qualitatively similar across credit score groups, although they are somewhat smaller in high credit score regions.

Third, we split zip codes according to median debt-to-income ratios in 2001. Specifically, we construct median initial debt-to-income ratios across all households in a zip code, then split zip codes into three groups based on these median ratios: low initial debt levels (below the 33rd percentile of the debt-to-income distribution), medium (between the 33rd and 67th percentiles) and high debt-to-income ratios (above the 67th percentile of the debt-to-income distribution). We then estimate our specification with household and regional controls within each of these three subsets of zip codes. We again present results for 2001-2007 in Panel C of Table 5, with the full set of yearly regressions by initial debt-to-income ratio available in Appendix Table A5. We find that our qualitative result holds across zip codes of different initial debt-to-income ratios but that the differential effects of inequality on household borrowing across income groups were largest in regions with higher initial debt-to-income ratios.

Finally, we assess whether our results are sensitive to either the growth in house prices or the initial level of house prices relative to income. We measure house prices for each zip code using data from the Core Logic index. These data are only available for a subset of our zip codes (about 6,600) which constitutes about 70% of our original sample. We split zip codes according either to their growth rates in house prices between 2001 and 2005 or according to their initial (2001) ratio of average house price to median income. In each case, we group zip codes into three bins: low (below the 33rd percentile), medium (between the 33rd and 67th percentiles), and high (above the 67th percentile). We re-estimate the specification with household and regional controls within each sub-grouping of zip codes and present results from 2001-2007 in Panels D (for house price growth) and E (for initial levels of house prices relative to income) of Table 5, with the full set of yearly regressions in Appendix Tables A6 and A7 respectively. The interaction of household rank and local inequality remains statistically significant within each subset of the data, with the results varying little depending on initial relative house price levels or subsequent house price appreciation.[11]

3.4. Results from a Nonparametric Specification

The specification in equation (2) assumes a linear relationship between debt accumulation, income and rank and local inequality. In this section, we relax this assumption and estimate a nonparametric specification. Specifically, we first split the sample of households into three bins according to the level of local inequality. In particular, each location (zip code) is assigned to one of the three bins based on the location's level of inequality in the distribution of inequality across locations in 2001, i.e., low-inequality bin (less than the 20th percentile of the distribution of local inequality levels), mid-level inequality bin (between the 20th and 80th percentile), and high-inequality bin (above the 80th percentile). The assignment of locations to inequality bins remains constant through 2002-2012. We similarly group households into bins based on income ranks (below 20th percentile, above 80th percentile, and between 20th and 80th percentiles). We then run a regression of households' relative debt accumulation on dummies for each income rank category and inequality bin, with regional controls and the county-specific fixed effects for each year separately. The omitted category is the dummy for low-rank households in low-inequality regions.

Figure 6 shows the estimated coefficients for low- and high-rank households in each type of region.[12] The differences across inequality regions for high-ranked households (i.e. those above the 80th percentile) are small throughout the time sample. In contrast, low-ranked households display much larger differences in debt accumulation patterns across low- and high-inequality regions, with differences in debt accumulation reaching over 50 percent of initial income levels by 2008. Hence, the link between inequality and debt accumulation was relatively more important for low-income households than for high-income households.

3.5. Results with County- and State-Level Income Distribution and Inequality Measures

Previous work on inequality and consumption has used measures of inequality at the state level (see Bertrand and Morse, 2013) and most discussion of inequality and debt has focused on measures of inequality at the national level, as in Figure 1. We explore how our results vary as we increase the level of geographic aggregation for inequality by estimating equation (4) using the income distribution at the county and state level.

---

[11] Another way to characterize the insensitivity of our results to housing is to split the sample into households who had mortgage debt in 2001 vs. those who did not. As we document in Appendix Table A8, we find the same qualitative results for both groups: debt accumulation of low-income households was more pronounced in low-inequality regions than high-inequality regions regardless of whether individuals already had a mortgage in 2001.

[12] Results for mid-rank households are included in Appendix Figure 1. They display no meaningful differences across areas of high or low-inequality.

We construct the area income distribution using the same resampling procedure we used for zip codes and now we compute a household's percentile rank within the larger area (e.g. county) income distribution and inequality statistics of that distribution. We keep all household and regional-level controls that we used before except now we include state fixed effects for county-level regressions and no fixed effects for state-level regressions.

Panels A and B of Table 6 report the results with county- and state-level income distribution and inequality measures, respectively. At the county level, we find very similar results to our zip code regressions once we consider that the standard deviation of inequality is smaller at the county level. We also find very similar estimates of the interaction term when inequality is measured at the state level, although there is some loss of precision in our estimates due to the aggregation. These results indicate that the effects we measure at the zip-level are also apparent at higher levels of aggregation. Also noteworthy is that the estimate of β is positive at the state level, implying that households on average accumulated relatively more debt in states with higher levels of inequality. This is similar to the result obtained by Bertrand and Morse (2013) that typical households consumed more in states where consumption of the rich was higher.

3.6. Results by Form of Debt

We now consider debt accumulation patterns along different dimensions of debt: mortgages, auto loans and credit cards. For each, we reproduce our household-level regressions with household and regional controls and county fixed effects and report yearly results in Table 7. Panel A documents that the results for mortgages are almost identical to those found for total debt. Because mortgage debt on average accounts for two-thirds of total debt, it is likely the primary driver of total debt patterns described above. Panel B documents that very similar qualitative results obtain for auto loans: both α and β are estimated to be negative while the interaction term γ is positive. However, the interaction effects are significantly smaller for auto loans than for mortgages, even if we adjust for the relative magnitudes of each form of debt (i.e. convert to growth rates). For example, the peak interaction effect on auto loans is about 0.05, which when adjusted by the average ratio of auto debt to mortgage debt (mortgage debt is almost eight times as large as auto debt on average) becomes 0.4, one-third to one-fourth of the mortgage interaction effect. Though auto loans display the same qualitative patterns, the mapping from local inequality to differential borrowing patterns across households is quantitatively weaker than for mortgages.

Panels C and D report equivalent results for credit card balances and credit card limits. The distinction between credit card balances and limits is useful because the former can be expected to be very elastic with respect to the demand for credit while credit limits should be significantly less elastic with respect to household demand.[13] Strikingly, we find very different results for the two measures. With credit card limits, we recover the same qualitative features as in our baseline estimates for total debt, α and β are both estimated to be systematically negative while the interaction term γ is positive. With credit card limits being approximately half of mortgage debt on average, the estimated peak level of γ of around 0.6 is approximately one-third as large as

[13] This distinction is somewhat offset by the fact that households can endogenously raise their credit limits by applying for more credit cards or requesting higher limits from their current credit card providers.

the peak interaction effect estimated for mortgages in terms of implied growth rates of each form of debt. In contrast, we find no consistent or economically significant relationship between local inequality and the credit card balances of households across different income groups: both β and γ are estimated to be very small (in some years becoming statistically insignificant) and the sign of γ unstable across years. Thus, to the extent that we can interpret credit card balances and limits as reflecting credit demand and supply respectively, these results suggest that the differential borrowing patterns of lower- and higher-income households across regions of different inequality reflect differential credit supply conditions, not differential credit demand.

In section 4, we propose one channel through which credit supply can vary with local inequality in a way that can account for these patterns, namely if lenders use an applicant's income in combination with local inequality to make inferences about the applicant's underlying type. This interpretation of the data would be consistent not just with the difference in our findings for credit card limits and credit card balances, but also with the quantitative differences in the size of estimated effects of inequality across other forms of debt. Mortgages, for example, represent much larger loan amounts than other forms of debt and, although collateralized, it is often difficult and expensive for financial institutions to recover the property associated with the loan in case of default. Auto loans, on the other hand, are much smaller in size and banks face fewer hurdles to repossessing a car. Hence, the incentive of financial institutions to devote resources toward identifying applicants' underlying credit-worthiness should be lower for auto loans than mortgages, leading to weaker utilization of the information provided by local income inequality as found in Table 7. While credit card debt is of the same order of magnitude on average as auto debt in the CCP, credit card debt is unsecured so that financial institutions bear more risk than they do with automobiles. One would therefore expect stronger incentives to utilize available information in extracting credit risk for credit cards than autos, which is again consistent with what we observe in the data.

3.7. Credit Demand

One potential demand-side explanation for our findings is that high- and low-income households' income expectations or growth vary systematically with inequality. If high-income households expect a relatively larger increase in permanent income growth in areas where inequality is high then we might expect them to borrow more. To the extent there is excess sensitivity in consumption even temporary differences in income changes that are correlated with inequality could affect borrowing behavior in the way we observe.

While we do not have the expectations data necessary to test this channel directly, we can test implications of this alternative explanation. One such implication is that if high-income households' incomes were growing faster relative to low-income households in more unequal areas, then we would expect to see *divergence* in income inequality across regions. Specifically, areas with higher initial levels of inequality should experience rising levels of inequality relative to other regions in subsequent years so that β in the following cross-sectional regression should be greater than one

$$Inequality_{it+1} = \alpha + \beta_t \, Inequality_{it} + e_i.$$

We test this implication using data from the Integrated Public Use Microdata Series (IPUMS) on household incomes in metro areas. We restrict the sample to the set of metro areas identified consistently from 1970 to 2000, to households where the respondent's age is between 25 and 65, and where the respondent is the head of the household or the spouse of the head of the household. To calculate income we use total family income. This leaves us with a sample of 117 metro areas covering roughly 60% of the U.S. population. See Appendix C for more details on the data. We measure inequality in each period as log of the p90/p10 ratio although results for other measures are similar. Table 8 provides the OLS coefficients from these regressions using base years of 1970, 1980, and 1990 and inequality levels for 2000 as the dependent variable. For all years the estimated coefficient is positive but significantly below one, suggesting that income distributions are stable on average. Estimates using quantile and robust regression give nearly identical results.

We also test if income growth by income decile varies with local inequality. For the decile j in area I, the average income is $\bar{Y}_{ij}$ so that we estimate

$$\log(\bar{Y}_{ijt+1}) - \log(\bar{Y}_{ijt}) = \alpha + \beta_j \, Inequality_{it} + e_i \, .$$

Figure 7 plots these coefficients along with 95% confidence intervals measuring income growth from 1970 to 2000 and 1990 to 2000. While the bottom decile appears to be a strong outlier, the pictures do not suggest that high-income deciles experienced higher income growth in areas that were more unequal. In fact, the graphs appear to have a downward slope, which suggests a convergence in the income distributions across regions over time. This is consistent with the results in Table 8.

Neither of these exercises suggests that income growth for high-income households was relatively higher in high-inequality areas. We find that lower-income households living in high-inequality regions have tended to experience relatively higher income growth, leading to convergent dynamics in regional inequality over time. These results suggest that differential income growth is not likely to be driving our results.

Another potential demand-side mechanism that could explain our findings is if households try to segregate themselves more when local inequality levels are higher. For example, as high-income households become increasingly richer than low-income households, then high-income individuals may have a greater desire to live with other high-income individuals. One immediate limitation of this story is that it only has implications for mortgage debt while Table 7 documents the qualitative consistency of our results across auto debt and credit card limits. Additionally, in Appendix Table A9, we introduce the interaction of several local observables likely to be correlated with the motivation for economic segregation. We separately interact rank with the share of homeowners, the share of nonwhite residents, the county-level crime rate (computed from the Uniform Crime Reporting Statistics), and the dispersion of housing quality (measured as the log ratio of average house prices at the top and bottom third from Zillow). Our results not only carry through but are essentially unchanged even though a number of these additional interactions are economically and statistically significant.

4 Model

In this section, we develop a stylized model in which lenders use local inequality to extract information about applicant types in order to differentiate between borrowers of varying credit quality. Intuitively, as inequality increases it becomes easier for the lender to tell applicants of different quality apart and so price credit more efficiently, which results in borrowing patterns similar to those we find in the CCP data. We demonstrate these results under two types of market structure: perfect competition and monopoly.

Suppose there are two types of households: High (H) and Low (L). To simplify algebra, we assume that High type households never default on debt while Low type households default with probability d and that the share of High type households is 0.5.[14] The income for an individual i of type $j \in \{H, L\}$ is given by $y_{i,j} = \mu_j + e_i$ where $\mu_H > \mu_L$ are constants and $e_i \sim N(0, \sigma^2)$. Hence, $y_H \sim N(\mu_H, \sigma^2)$ and $y_L \sim N(\mu_L, \sigma^2)$. Denote the pdfs for each distribution with ϕ_H and ϕ_L. The average income in this economy is $\bar{y} = \frac{1}{2}\mu_H + \frac{1}{2}\mu_L$.

We also assume banks observe s, another signal about the quality of borrowers that can incorporate other information about borrowers and is not observed by the econometrician, to capture the idea that loan officers have more information than econometricians. Similar to the income signal, $s_{i,j} = \rho_j + \eta_i$ where $\rho_H > \rho_L$ are constants and $\eta_i \sim iid\ N(0, \omega^2)$. Denote the pdfs for each distribution with q_H and q_L. To simplify algebra, we assume without loss of generality that idiosyncratic shocks to income and signal s are independent.

Banks do not observe household types directly but they observe applicants' incomes and signal s.[15] They can then infer the probability of a given type conditional on observed income. Specifically, using Bayes law, the posterior probability of being High type for a household i with signals y_i and s_i is given by

$$\Pr(H|y_i, s_i) = \frac{\Pr(y_i|H)\Pr(s_i|H)\Pr(H)}{\Pr(y_i|H)\Pr(s_i|H)\Pr(H) + \Pr(y_i|L)\Pr(s_i|L)\Pr(L)}$$

$$= \frac{\phi_H(y_i) q_H(y_i)\frac{1}{2}}{\phi_H(y_i) q_H(y_i)\frac{1}{2} + \phi_L(y_i) q_L(y_i)\frac{1}{2}} = \frac{\Phi(y_i)Q(s_i)}{\Phi(y_i)Q(s_i)+1} \qquad (6)$$

where $\Phi(y_i) \equiv \phi_H(y_i)/\phi_L(y_i)$ and $Q(s_i) \equiv q_H(s_i)/q_L(s_i)$ are the likelihood ratios. Given our assumptions, we have $\Phi' > 0$ and $Q' > 0$, that is, High type households are monotonically more likely to be observed as income y or signal s increase. Since there are only two types, it follows that

$$\Pr(L|y_i, s_i) = 1 - \Pr(H|y_i, s_i) = \frac{1}{\Phi(y_i)Q(s_i)+1}. \qquad (7)$$

Clearly, $\frac{\partial \Pr(L|y_i,s_i)}{\partial y_i} < 0$, $\frac{\partial \Pr(L|y_i,s_i)}{\partial s_i} < 0$, $\frac{\partial \Pr(H|y_i,s_i)}{\partial s_i} > 0$, and $\frac{\partial \Pr(H|y_i,s_i)}{\partial y_i} > 0$.

Banks potentially have two margins to determine which borrowers obtain loans: 1) price of loans; 2) loan denial probability. While in reality banks are likely to use both margins, we consider polar cases to illustrate the workings

[14] We document in Appendix F that high-income households are indeed less likely to default than low-income households.

[15] Obviously, banks observe many other characteristics of households. We abstract from this additional information available to banks to simplify derivations. One may interpret this approach as partialling out these other characteristics. Typically, one of the important indicators of individual's risk is individual's credit score. In the analysis in section 3, we show that the household's income rank has explanatory power for the household's debt even after we control for the credit score.

of each margin separately. For the price margin, we will assume that banks can price discriminate borrowers perfectly, banks compete in all population segments, and banks can freely obtain resources at rate R_0 ("perfect competition"). For the loan denial probability, we assume that there is only one bank serving the market but this bank is threatened by entry of other banks if this bank makes a profit ("monopoly").

4.1 Perfect Competition

With perfect competition and free entry in each lending segment, banks can have only one interest rate for a borrower of a given quality. Since there is a continuum of borrower quality, there is also a continuum of markets where each market is indexed by borrower quality. Consider a set of households with income y_i and signal s_i. Given by the zero profit condition, the interest rate is set to

$$R^*\{(1-d)\Pr(L|y_i,s_i)+\Pr(H|y_i,s_i)\}=R_0 \Longrightarrow$$

$$R^*=\frac{R_0}{(1-d)\Pr(L|y_i,s_i)+\Pr(H|y_i,s_i)}=R_0\frac{\Phi(y_i)Q(s_i)+1}{\Phi(y_i)Q(s_i)+(1-d)}=R^*(y_i,s_i) \tag{8}$$

Note that households with other levels of y and s pay the same interest rate as long as $\Phi(y_i)Q(s_i)=\Phi(y)Q(s)$. That is, each lending segment is characterized by a pair of signals

$$\mathcal{S}(R^*)=\left\{(y,s)\colon R_0\frac{\Phi(y)Q(s)+1}{\Phi(y)Q(s)+(1-d)}=R^*\right\}.$$

where R^* is a sufficient statistic for the quality of borrowers. Because the quality of borrowers is the same in $\mathcal{S}(R^*)$, every borrower in $\mathcal{S}(R^*)$ obtains a loan at the interest rate R^*. Borrowers of a worse quality are offered loans at higher interest rates while borrowers of better quality can obtain a loan with a lower interest rate.

Clearly, $\frac{\partial R^*}{\partial y}<0$ and $\frac{\partial R^*}{\partial s}<0$ so that households with high income y and strong signal s pay lower rates because banks believe that these applicants are more likely to be of the High type. To see the tradeoff between y and s, one can fix $R^*(y,s)$ at level $R^\#$ and find the required signal s to allow a household to borrow at rate $R^\#$ given that this household has income y:

$$s^*(y)=Q^{-1}\left\{\frac{1}{\Phi(y)}\times\frac{R_0-R^\#(1-d)}{R^\#-R_0}\right\} \tag{9}$$

where Q^{-1} is the inverse function of Q. Given that $Q'>0$ and $\Phi'>0$, it follows that $\frac{\partial s^*(y)}{\partial y}<0$.

Although we (unlike loan officers) do not observe signal s in the data, we can still calculate the interest rate paid on average by households with income y, which is observed by the econometrician:

$$R^*(y)=\int R^*(y,s)\left\{q_H(s)\frac{1}{2}+q_L(s)\frac{1}{2}\right\}ds \tag{10}$$

Given that $R^*(y,s)$ is differentiable and otherwise well behaved as well as $\frac{\partial R^*(y,s)}{\partial y}<0$, we have that

$$\frac{\partial R^*(y)}{\partial y}=\int\frac{\partial R^*(y,s)}{\partial y}\left\{q_H(s)\frac{1}{2}+q_L(s)\frac{1}{2}\right\}ds<0. \tag{11}$$

Hence, the model predicts that the interest rate decreases in household income.

One can then consider a thought experiment of raising the income inequality in this economy without changing the mean level of income. Specifically, we increase the distance between μ_H and μ_L but the average income $\bar{y}$ is held constant.[16] Because income levels are now a stronger signal of an applicant's type, banks put a higher weight on signal y, hence the slope of the tradeoff becomes steeper as it takes a larger change in signal s to justify lending at a given interest rate (see Panel A of Figure 8). This will lead to higher borrowing on the part of low-income households in low-inequality regions than in high-inequality regions because, in the former, banks are less sure about the underlying type of the applicant based on income and therefore are more willing to lend to households of different incomes. In other words, $R^*_{equal}(y) < R^*_{unequal}(y)$ when $y < \bar{y}$ where "equal" and "unequal" denote the level of inequality, captured by mean-preserving changes in μ_H and μ_L, and $R^*_{equal}(y) > R^*_{unequal}(y)$ when $y > \bar{y}$. Panel B of Figure 8 illustrates this point. In short, banks charge lower interest rates to high-income households than to low-income households and the difference in the interest rates across income groups rises as the difference between these groups widens.[17]

We also study the effects of an increase in the supply of credit. Since perfect competition prices each borrower type fairly, we can only increase the supply of credit by reducing the cost of funds rate R_0. Equation (9) shows that a decrease in R_0 shifts schedule $s^*(y)$ down and hence all borrowers enjoy a lower cost of credit.

A combination of a positive credit supply shock (R_0 decreases) and an increase in inequality ($\mu_H - \mu_L$ increases) can reconcile how all types of households increased their borrowing on average over the course of the mid 2000s with the cross-sectional variation in debt-accumulation patterns across income groups at different levels of local inequality documented in section 3. The supply shock by itself can explain the former while the increased inequality by itself can explain only the latter.

4.2. Monopoly

In practice, regulatory or informational constraints limit the ability of banks to charge different prices to different borrowers and therefore they often can charge only one rate or a limited number of rates for a given type of loan. To keep exposition simple, suppose that i) the market has only one bank and it is threatened by entry of other banks, ii) regulators impose a minimum quality of borrowers who may obtain loans (e.g., to qualify for Freddie Mac and Fannie Mae guarantees), and iii) the bank can charge only one rate $\bar{R}$.

To model assumption ii), we know that $R^*(y,s)$ can be used as a sufficient statistic for the quality of a borrower. The bank makes a profit on borrowers with (y,s) such that $R^*(y,s) < \bar{R}$ and losses on borrowers with (y,s) such that $R^*(y,s) > \bar{R}$. We will denote the cutoff interest rate R^+ that meets the regulation

[16] Notice that increasing inequality in this manner is not innocuous. If we assumed instead that the variance of income increased, we would generate the opposite dynamic as income would now be a less precise signal of type. Modeling the increase in inequality as an increase in the distance between types of incomes is consistent with the nature of the increase in U.S. inequality. Debaker et al. (2013) decompose the increase in income inequality into permanent and transitory components and find the vast majority of the increase in inequality is due to dispersion in the permanent component of income. We view the spread in mean income between types as analogous to an increased dispersion in the permanent component of income.

[17] Note that the value at which a household does not experience a change in the interest rate is equal to the average income $\bar{y}$. This value is insensitive to the level of inequality because by construction the average income is held constant and at the average income the likelihood ratios are equal to 1 and therefore the posterior probability is equal to 1/2. This value, however, can move in more complex models and alternative parameterizations.

requirements. With this cutoff rate, the threat of entry sets $\bar{R}$ at the level that yields zero profits as implied by assumption i).

$$\bar{R}\frac{\int\int_{(y,s):R^*(y,s)\leq R^+}\{(1-d)\Pr(L|y,s)+\Pr(H|y,s)\}\bar{\phi}(y)\bar{q}(s)dyds}{\int\int_{(y,s):R^*(y,s)\leq R^+}\bar{\phi}(y)\bar{q}(s)dyds}=R_0$$

where $\bar{\phi}(y)\equiv\frac{1}{2}\phi_L(y)+\frac{1}{2}\phi_H(y)$ and $\bar{q}(s)\equiv\frac{1}{2}q_L(s)+\frac{1}{2}q_H(s)$. Using the insight of equation (9), we can find the threshold level of signal s such that a bank will lend to a household with income y:

$$s^+(y)=Q^{-1}\left\{\frac{1}{\Phi(y)}\times\frac{R_0-R^+(1-d)}{R^+-R_0}\right\} \tag{12}$$

As before, we have $\frac{\partial s^+(y)}{\partial y}<0$. The set of households who obtain a loan is:

$$\mathcal{S}^+(R^+)=\left\{(y,s)\colon R_0\frac{\Phi(y)Q(s)+1}{\Phi(y)Q(s)+(1-d)}\geq R^+\right\}$$

The probability that a household with income y is denied a loan is

$$\Pr(denied\ loan|y)=\Pr(s<s^+(y))=\int_{-\infty}^{s^+(y)}\bar{q}(s)ds$$

Since $\frac{\partial s^+(y)}{\partial y}<0$, it follows that $\frac{\partial\Pr(denied\ loan|y)}{\partial y}<0$: the probability of loan denial decreases in income.

Now we repeat the thought experiment with rising inequality. Similar to the perfect competition case, it takes a larger increment in signal s to compensate for a given decrease in income y because income is a more informative signal. As a result, if the quality of lending standard R^+ is held constant, some low-income households may be denied a loan more often (see Panel C of Figure 8). Panel D of Figure 8 shows how the denial probability changes with rising inequality. The probability of denial increases for households with $y<\bar{y}$ and decreases for households with $y>\bar{y}$.

In contrast to the perfect competition case, the monopoly case has two ways to model an increase in the supply of credit. First, one can continue to model it as a reduction in the cost of funds rate R_0. Second, one can model it as an increase in R^+, i.e., relaxing lending standards to cover high-risk borrowers. In the first case, a decrease in R_0 lowers $\bar{R}$ and thus makes credit cheaper for households with $R^*\leq R^+$. However, it does not affect the interest rate for households with $R^*>R^+$ as these continue to receive no loans (they do not meet lending requirements). In the second case, an increase in R^+ *raises* $\bar{R}$ because a wider coverage now includes high risk households and losses made on these high-risk households have to be compensated by larger profit margins on low-risk households. Thus, while credit is now available to a broader spectrum of households, the cost of borrowing increases for relatively high-income borrowers. On the other hand, the probability of obtaining a loan increases for all households as schedule $s^+(y)$ shifts down. Hence, although high-income households pay a higher price for credit, they are denied loans less frequently.

Our model can therefore potentially account for why lower-income households accumulated relatively less debt in high-inequality regions than did similar households in low-inequality regions during the 2000s: if banks in higher-inequality regions placed more weight on applicants' incomes as a signal of their underlying creditworthiness and therefore channeled more funds toward higher-income applicants than did banks in lower-inequality regions. Under perfect competition, this differential access to funds is predicted to happen through higher interest rates being offered to low-income applicants than high-income applicants whereas under monopoly banking, our model predicts that banks will reject low-income applicants more frequently than high-income applicants. Because banking in the U.S. lies in between these two extremes, we expect both margins to be present in the data, a prediction to which we now turn.

5 Results from the Mortgage Application Data

Our model suggests that variation in inequality across regions should be reflected in the lending decisions of -banks if regional inequality can be used to infer applicants' default probabilities. In this section, we use information on mortgage applications from the publicly available Home Mortgage Disclosure Act database (HMDA), 2001-2011, to test these implications.

The HMDA data are compiled from reports filed by mortgage lenders. The HMDA was passed by Congress in 1975 and began requiring lenders to submit data reports in 1989. The initial intention of the act according to the Consumer Financial Protection Bureau (2012) was to monitor the provision of credit in urban neighborhoods. Later requirements to submit data reports were intended to monitor discriminatory lending practices. Dell'Ariccia et al. (2012) find that HMDA covers between 77% and 95% of all mortgage originations from 2000 to 2006. Reporting criteria differ between depository and nondepository institutions and across years. Depository institutions have typically been required to report if they satisfy an asset threshold, make at least one home mortgage, are federally regulated or insured, and have a branch in a metropolitan area. Nondepository institutions were required to report if the share of home mortgages exceeded a threshold of all loan originations, the lender operated in an MSA, and met an asset threshold. In 2004 the share threshold was supplemented with a level of home mortgage originations to increase the coverage of the market. Lenders who file reports include detailed information on every mortgage application received by the lender during a calendar year. All years of the data contain the size of the loan, income on the application, location of the property down to the census tract, demographics of the applicants, a lender identifier, and the action taken on the loan. Since 2004 the data include additional information including a censored picture of interest rates and the loan's lien status. We use a 50% random sample of all HMDA records.

While the data are very detailed in many respects there are some limitations. First, the data do not identify "piggyback" loans, i.e. loans with subordinate liens used to finance a larger first-lien loan. These secondary loans can be used to lower financing costs and to avoid requirements that a loan being sold to Fannie Mae or Freddie Mac be accompanied by private mortgage insurance if a loan does not meet certain standards. The HMDA does not require lenders to report piggyback loans if they are issued as HELOCs and some piggyback loans might be issued by a lender not covered by HMDA. But some piggyback loans are included in the dataset and, given that these loans are not identified as such, a researcher might infer a much lower loan-to-value ratio than the actual loan-to-value on the property. Since we are not able to identify piggyback loans reliably and these loans are relatively small, we drop all applications where the loan-to-income (LTI) ratio is less than one. Second, we conduct the HMDA analysis at the county level rather than the zip code level. Although the data are available at the census tract, we aggregate to the county in order to use measures of inequality consistent with the CCP analysis. Finally, in contrast to the CCP database, the HMDA data set does not track applicants over time and hence we do not have a panel of applicants/borrowers.

We focus on measures of price or borrower cost in line with the theoretical predictions of the model. First, we assess whether the probability of a loan being rejected depends on the applicant's income rank (within the pool of applicants) interacted with regional inequality. Second, we consider whether the probability of the loan being "high-interest" (conditional on a loan application being approved) varies with inequality and the applicant's rank.[18] Both of these can be interpreted as measuring the price of credit incurred by a borrower and so are helpful to combine with the quantity information available in the CCP. If banks use an applicant's position in the income distribution to help make inferences about their underlying default risk, as suggested by the model, then one would expect banks to reject otherwise similar applications by high-income applicants less frequently in high-inequality regions than in low-inequality regions, or equivalently to reject otherwise similar applications by low-income applicants more frequently in high-inequality regions than in low-inequality regions. By the same logic, we should observe low-income applicants being charged higher interest rates more frequently in high-inequality regions than in low-inequality regions.

We test these predictions by OLS using the following regression[19]

$$Outcome_{ict} = \alpha Rank_{ict} + \gamma Rank_{ict} * Inequality_{c,2001} + \beta Z_{ict} + \lambda_c + error, \qquad (6)$$

where $Rank_{ict}$ is the percentile rank of applicant *i's* income within the pool of applicants in area *c* in year *t*.[20] The inequality measure and the income distribution are defined at the county level. The explanatory variables in vector Z_{ict} include indicators for whether or not the loan is for an owner-occupied property, several race categories and gender, as well as the interaction of the applicant's income rank with the share of applicants in the county who are nonwhite.[21] We also control for the loan-to-income ratio in the application. While we estimate these models with county fixed effects λ_c, the results are very similar if we use state fixed effects (Appendix Table A10). We restrict the analysis to loans for home purchases, applications where the loan-to-income ratio is at most eight and not less than one, loans where the reporter was explicitly making the origination decision (i.e. the loan was not purchased), and where the loan did not fail because of incompleteness or because it was not pre-approved. Notice that we retain

[18] The HMDA reporting guidelines require lenders to report the spread between the Treasury yield and the mortgage interest rate if the spread is greater than three percentage points for first-lien loans or five percentage points for subordinate-lien loan.

[19] Our baseline specification includes a county fixed effect because the county-level controls are not as detailed as those we can construct in the CCP data. Specifications with a state-level fixed effect are available in Appendix Table A10.

[20] The results are also robust to measuring an applicant's rank in the distribution of income of all households in the county.

[21] We include this interaction as an additional control because previous studies have suggested that banks may treat differentially areas with predominantly non-white population. See Turner and Skidmore (1996) for a review.

in the sample loans that are not denied but also not originated. Excluding these does not change our results. As before, we are interested in the sign of the interaction term between income rank and inequality, γ. All standard errors are clustered at the county level. The regressions are estimated separately for each year, 2001-2011. We use the log of the 90/10 income ratio derived from the income imputed in the CCP data in 2001 as the measure of inequality, but the results are essentially the same using the Gini coefficient derived from the Census data.

We present the results for the probability of an application being rejected by a bank in Panel A of Table 9 and results for the probability of a loan being high-interest, conditional on origination, in Panel B of Table 9. For the probability of being rejected, the key finding is that estimated γ is consistently negative: applications from high-ranked households in high-inequality regions are less likely to be rejected than those from high-ranked households in low-inequality regions. This result is consistent with the theoretical predictions of the model in which banks use an applicant's position in the local income distribution, along with the dispersion of that distribution, to make inferences about default risk. Using our 2007 estimates, our results suggest that a one standard deviation increase in inequality will decrease the probability of denial of a household in the 80th percentile rank relative to the 20th percentile rank by approximately 2 percentage points. This is comparable in magnitude to the association between rank and the probability of denial. Similar results obtain with the probability of the loan being high-interest (this variable is not available before 2004): high-rank applicants are less likely to face higher rate loans in high-inequality regions than in low-inequality regions. Again, this is precisely the type of price-discrimination predicted by the model. Doing the same calculation as above with the 2007 estimate we find that high-rank households will see the probability that they pay a high interest loan decline by 1.5 percentage points relative to low-rank household.

We also consider whether the size of the mortgage (intensive margin) varies across inequality regions and ranks within the income distribution by using the loan-to-income ratios associated with each *originated* mortgage. We use the same controls as with rejection probabilities (with the exception of LTI ratios) and county fixed effects. The results for each year are presented in Panel C of Table 9. Unlike mortgage rejection rates and interest rate premia, we find little evidence that loan-to-income ratios vary across households in different inequality regions. To the extent that requested loans reflect demand for credit by households, we again find little evidence that demand-side factors related to local inequality levels mattered for the debt-accumulation decisions of households. However, the HMDA dataset does not allow us to establish if households have multiple loans or reliably link piggyback loans to standard loans. Thus, while our results point mainly toward channels operating through credit supply more work needs to be done to better understand the intensive margin.

6 Conclusion

Using household level measures of debt over the course of 2001 - 2012, we document a systematic link between local levels of income inequality and the debt-accumulation decisions of households of different income levels. Specifically, we find that low-income households in low-inequality regions accumulated more debt during the mid-2000s than did low-income households in high-inequality regions, with reverse (albeit smaller) effects operating for high-income households. While these results point to an economic channel linking economic inequality and borrowing by households of different income groups, they are inconsistent with "keeping up" behavior as a significant force behind the increase in household leveraging over this period.

Instead, we argue that inequality can affect household debt accumulation through effects on credit supply. We develop a model where income inequality matters for the information content of income when evaluating a borrower's credit risk. In the model, this channel leads to relatively more credit being allocated to low-income applicants when local inequality is low rather than high, since higher levels of inequality imply that applicant incomes are stronger signals of credit-worthiness. Consistent with this view, we document that lower-income mortgage applicants in high-inequality regions are rejected more frequently and pay higher interest rates than similar applicants in low-inequality regions. While it is possible that income inequality implicitly captures other factors that are not included in the model or data, our findings suggest that the causality between inequality and debt is running through the credit supply channel.

However, to the extent that this expansion in the supply of credit to lower-income households is unlikely to continue (for example if it reflected a one-time securitization of household debt), our results suggest that a continuation of recent trends toward rising inequality is likely to reduce access to credit for lower-income households. Because limited access to credit restricts households' ability to smooth their consumption and to engage in long-term investments (e.g. sending children to college, retraining for different careers), such differential access to credit could ultimately have negative longer term consequences. To the extent that many of these activities likely have positive societal externalities not captured in our model, such a development could have important policy implications.

References

Aguiar, Mark and Mark Bils. 2012. "Has Consumption Inequality Mirrored Income Inequality," mimeo.

Bertrand, Marianne, and Adair Morse. 2013. "Trickle-Down Consumption," NBER WP No. 18883.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. "Consumption Inequality and Partial Insurance," *American Economic Review* 98(5), 1887–1921.

Bordo, Michael D. andChristopher M. Meissner, 2012. "Does Inequality Lead to a Financial Crisis?" *Journal of International Money and Finance* 31(8), 2147-2161.

Brown, Meta, Andrew Haughwout, Donghoon Lee, and Wilbert van der Klaauw. 2011. "Do We Know What We Owe? A Comparison of Borrower- and Lender-Reported Consumer Debt." N.Y. Fed Report no. 523.

Charles, Kerwin K., Erik Hurst, and Nikolai Roussanov, 2009. "Conspicuous Consumption and Race," *Quarterly Journal of Economics* 124(2), 425-467.

Christen, Markus and Ruskin M. Morgan, 2005. "Keeping Up with the Joneses: Analyzing the Effect of Income Inequality on Consumer Borrowing," *Quantitative Marketing and Economics* 3, 145-173.

Consumer Financial Protection Bureau. 2012. "Supervision and Examination Manual"

Daly, Mary C. and Daniel J. Wilson, 2006. "Keeping Up with the Joneses and Staying Ahead of the Smiths: Evidence from Suicide Data," Federal Reserve Bank of San Francisco WP 2006-12.

Debaker, Jason, Brad Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos. 2013. "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of US Tax Returns," *BPEA*, Spring.

Dell'Ariccia , Giovanni, Deniz Igan, and Luc Laeven. 2012. "Credit Booms and Lending Standards: Evidence from the Subprime Mortgage Market," *Journal of Money, Credit and Banking*, Vol. 44 (2-3): 367–384.

Elul, Ronel, Nicholas S. Souleles, Souphala Chomsisengphet, Dennis Glennon, and Robert Hunt, 2010. “What “Triggers” Mortgage Default?” *American Economic Review Papers and Proceedings* 100(2), 490-494.

Fay, Scott, Erik Hurst, and Michelle J. White, 2002. “The Household Bankruptcy Decision,” *American Economic Review* 92(3), 706-718.

Gross, David B., and Nicholas S. Souleles, 2002. “An Empirical Analysis of Personal Bankruptcy and Delinquency,” *Review of Financial Studies* 15(1), 319-347.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst, 2013. “Endogenous Gentrification and Housing Price Dynamics,” *Journal of Public Economics 100, 45-60.*

Guven, Cahit, and Bent E. Sørensen, 2012. “Subjective Well-Being: Keeping Up with the Joneses. Real or Perceived?” *Social Indicators Research* 109(3), 439-469.

Heathcote, Jonathan, Fabrizio Perri, and Gianluca Violante, 2010. “Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States, 1967-2006,” *Review of Economic Dynamics* 13(1), 15-51.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2004. “The Macroeconomic Implications of Rising Wage Inequality in the United States.” *Journal of Political Economy* 118(4), 681-722

Heffetz, Ori, 2011. “A Test of Conspicuous Consumption: Visibility and Income Elasticities,” *Review of Economics and Statistics* 93(4), 1101-1117.

Kumhoff, Michael, Romain Ranciere, and Pablo Winant, 2013. “Inequality, Leverage and Crises,” mimeo.

Iacoviello, Matteo, 2008. “Household Debt and Income Inequality: 1963-2003,” *Journal of Money, Credit and Banking* 40(5), 929-965.

Kennickell, Arthur B., 1998, “Multiple Imputation in the Survey of Consumer Finances,” Working paper, Federal Reserve Board, available at: http://www.federalreserve.gov/pubs/oss/oss2/method.html.

Krueger, Dirk, and Fabrizio Perri. 2006. “Does Income Inequality Lead to Consumption Inequality? Evidence and Theory,” *Review of Economic Studies*, Vol. 73(1): 163-193.

Kuhn, Peter, Peter Kooreman, Adriaan R. Soetevent, and Arie Kapteyn, 2011. “The Effects of Lottery Prizes on Winners and their Neighbors: Evidence from the Dutch Postcode Lottery,” *American Economic Review* 101(5), 2226-2247.

Lee, Donghoon, and Wilbert van der Klaauw. 2010. “An Introduction to the FRBNY Consumer Credit Panel.” Federal Reserve Bank of New York, Staff Report no. 4799.

Luttmer, Erzo F. P., 2005. “Neighbors as Negatives: Relative Earnings and Well-Being,” *Quarterly Journal of Economics* 120(3), 963-1002.

Maurer, Jürgen, and André Meier, 2008. “Smooth It Like the Joneses? Estimating Peer-Group Effects in Intertemporal Consumption Choice,” *The Economic Journal* 118(527), 454-476.

Munnell, Alicia H., Geoffrey M. B. Tootell, Lynn E. Browne, and James McEneaney, 1996. “Mortgage Lending in Boston: Interpreting HMDA Data,” *American Economic Review* 86(1), 25-53.

Neumark, David, and Andrew Postlewaite, 1998. “Relative Income Concerns and the Rise in Married Women’s Employment,” *Journal of Public Economics* 70(1), 157-183.

Perugini, Cristiano, Jens Holscher, and Simon Collie, 2013. “Inequality, Credit Expansion and Financial Crises,” Munich Personal RePec Archive Paper 51336.

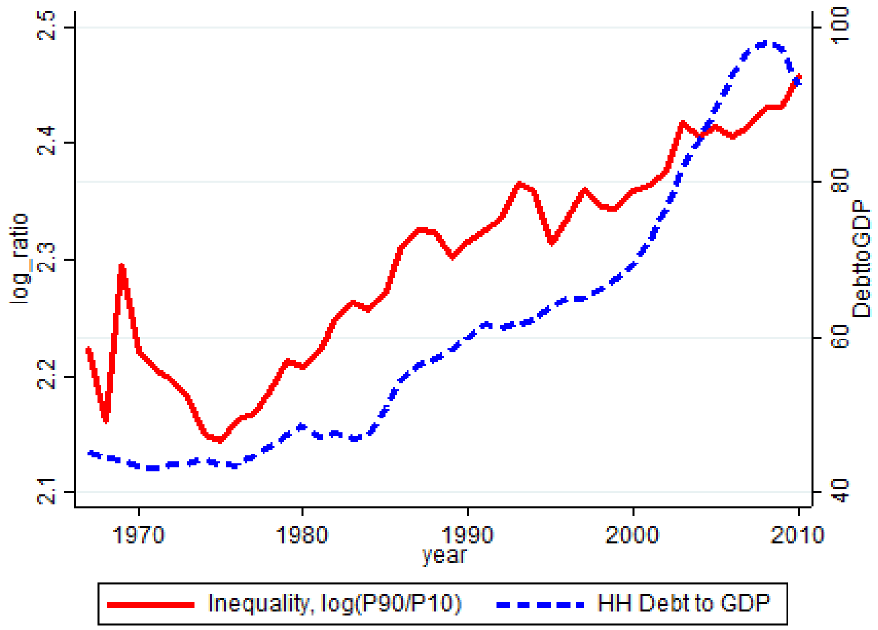
Rajan, Raghuram G., 2010. Fault Lines: How Hidden Fault Lines Still Threaten the World Economy, Princeton University Press, Princeton N.J.

Tootell, Geoffrey M. B., 1996. “Redlining in Boston: Do Mortgage Lenders Discriminate against Neighborhoods?” *Quarterly Journal of Economics* 111(4), 1049-1079.

Turner, Margery Austin, and Felicity Skidmore, 1999. Mortgage Lending Discrimination: A Review of Existing Evidence. The Urban Institute, Washington D.C.

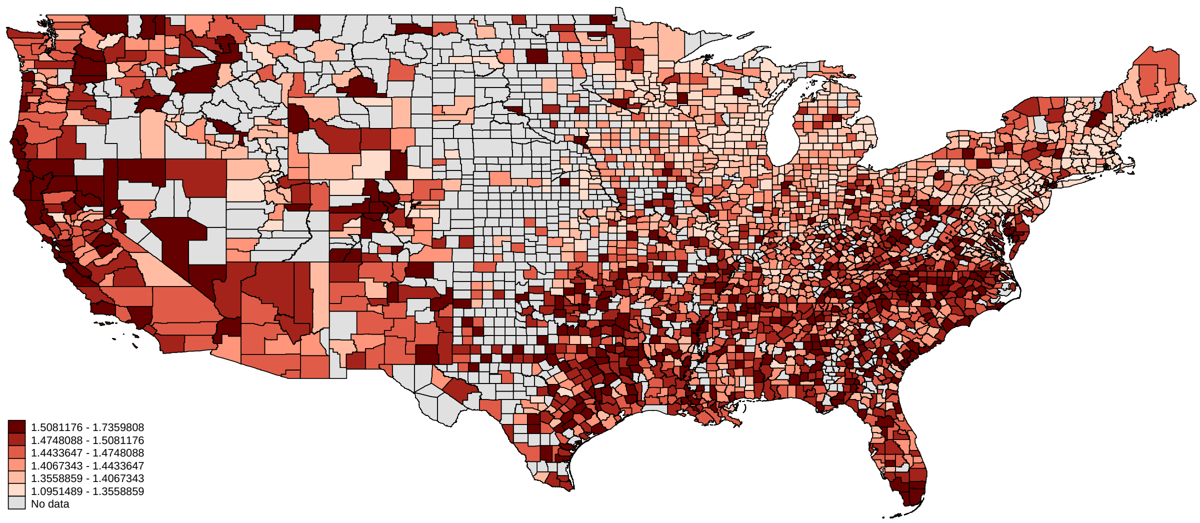
Zizzo, Daniel J. and Andrew J. Oswald, 2001. “Are People Willing to Pay to Reduce Others’ Income,” *Annales d’Economie et de Statistique* 63/64 39-65.

FIGURE 1: INEQUALITY AND DEBT IN THE U.S.



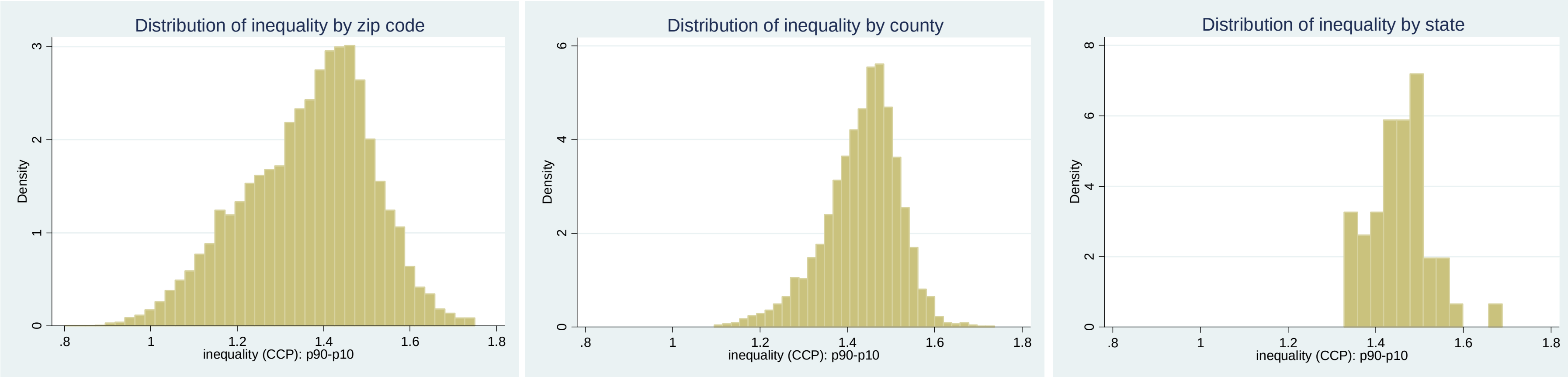
Note: The figure plots the (log) ratio of the 90th percentile to the 10th percentile of incomes of U.S. households (source: U.S. Census Bureau) and the ratio of household (and non-profit) total liabilities relative to GDP (source: Federal Reserve).

FIGURE 2: INEQUALITY ACROSS U.S. COUNTIES



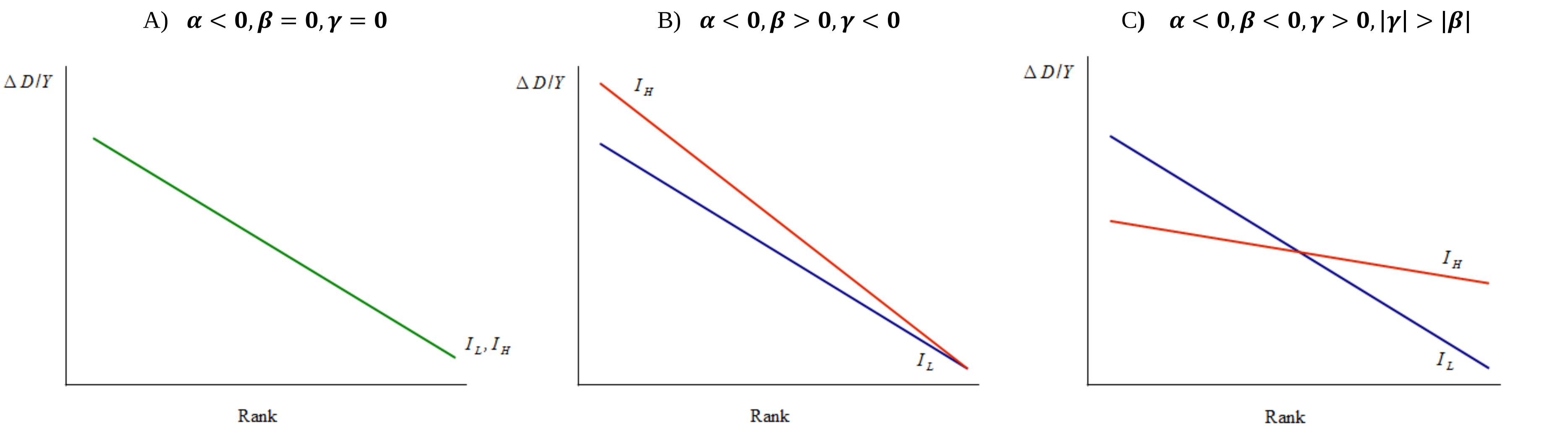
Note: The figure plots inequality in 2001 at the county level. Inequality is measured as the difference in log expected incomes at the 90th and 10th percentiles computed from the CCP. Darker counties are more unequal with each bin representing a quintile of the distribution across counties.

FIGURE 3: CROSS-SECTIONAL INEQUALITY IN THE U.S.



Note: The figures plot the regional distribution of inequality, measured using differences in expected log income between the 90th and 10th percentiles as computed from the CCP, at three levels of aggregation: zip code, county, and state level.

FIGURE 4: DEBT ACCUMULATION, INCOME RANK AND LOCAL INEQUALITY

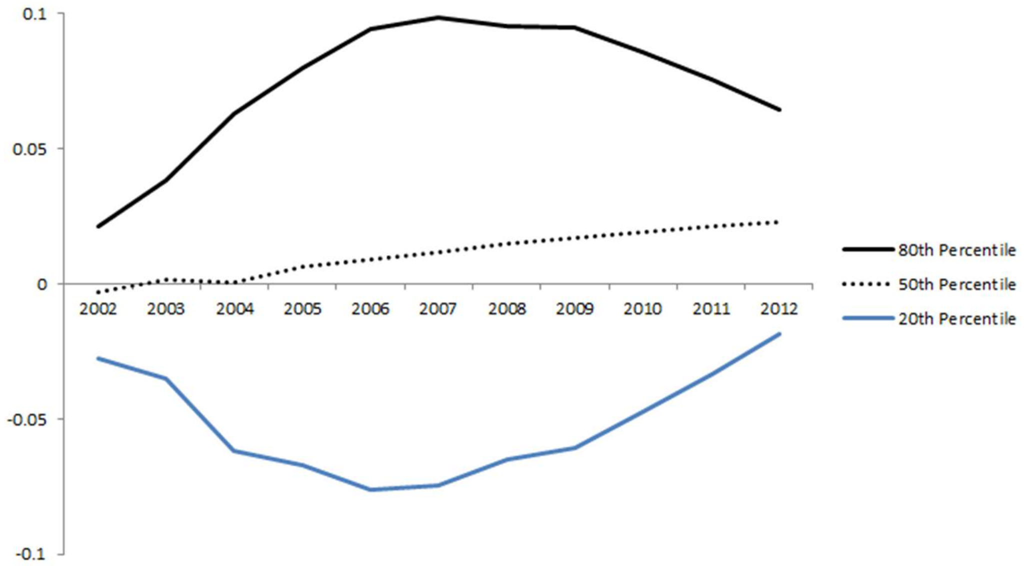


Note: The figure plots qualitative predictions for various theories of how borrowing and inequality interact. Panel A shows a case where the local inequality is irrelevant for borrowing. Panel B demonstrates a special case of "keeping up" when the debt accumulation of the richest household does not depend on the local inequality and inequality increases overall debt accumulation. Panel C shows the case where increased inequality results in high-income households borrowing more and low-income households borrowing less. See section 3.1 in the text for details.

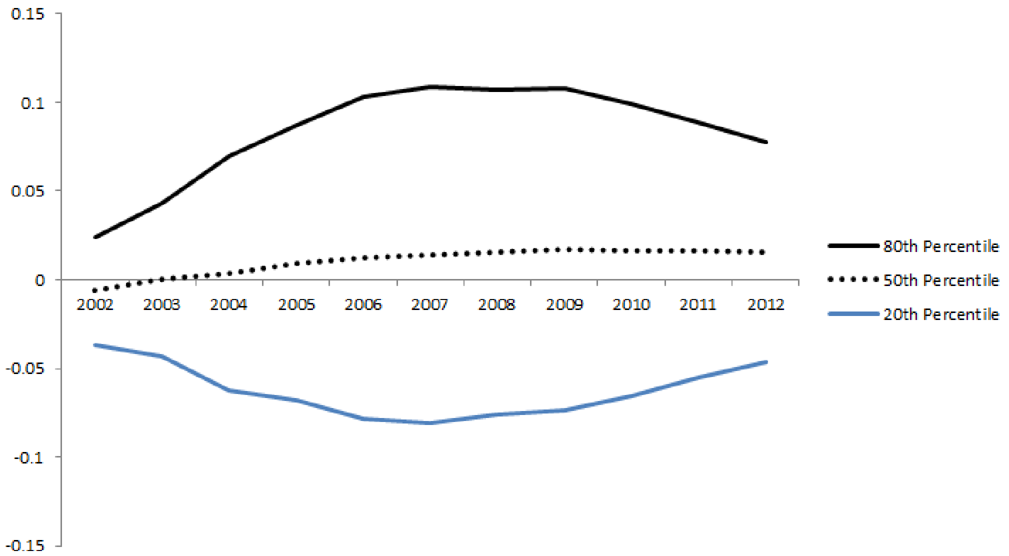
FIGURE 5: THE ESTIMATED EFFECT OF ONE SD INCREASE IN INEQUALITY ON DEBT ACCUMULATION

$$\boldsymbol{\sigma(Inequality) * (\beta + \gamma * Rank)}$$

Panel A: Parsimonious Specification

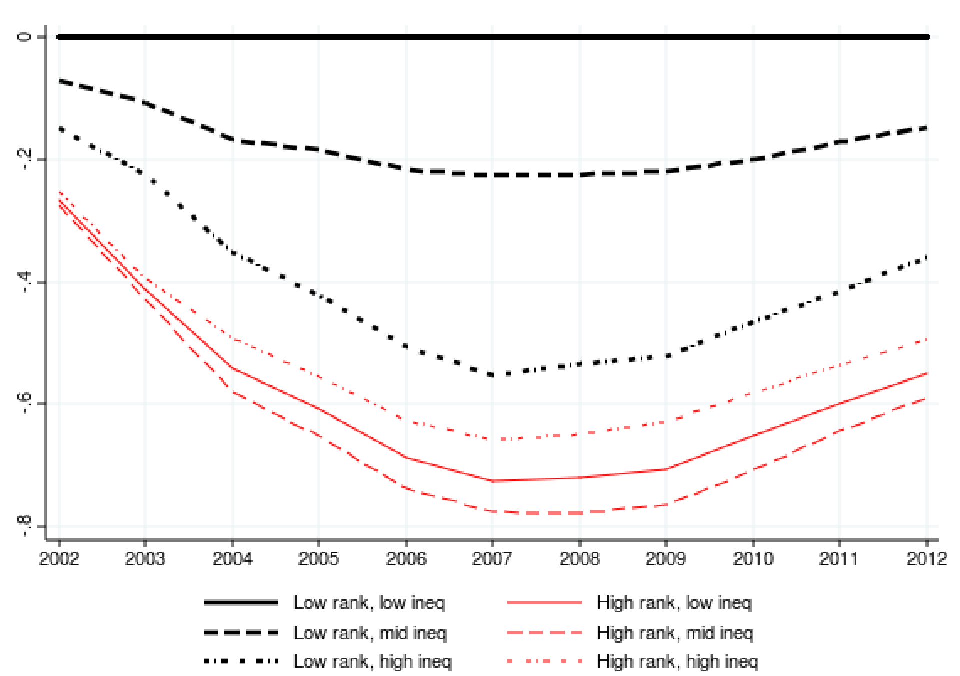


Panel B: Specification with Full Set of Controls



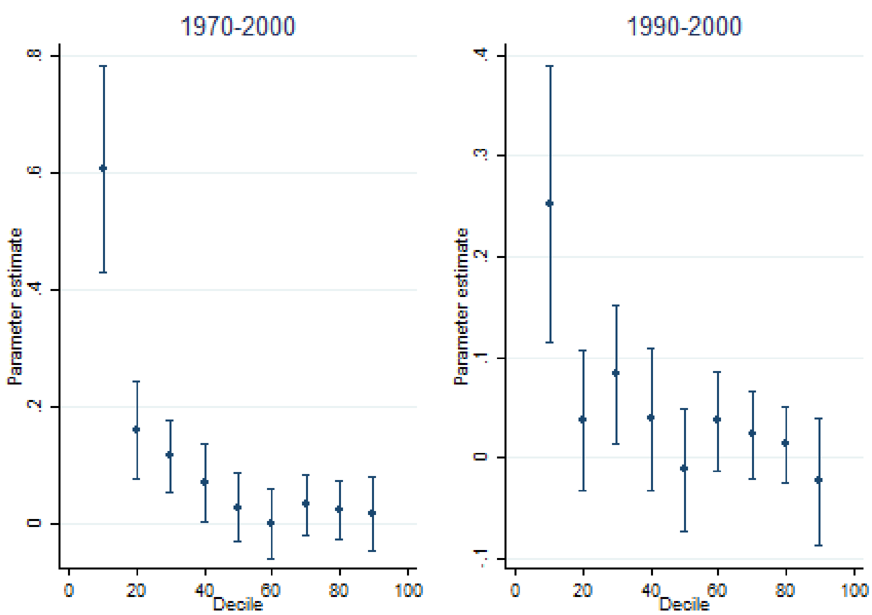
Note: These figures plot the calculated effects of a one standard deviation increase in inequality using estimated coefficients on rank, inequality, and the interaction of rank and inequality from the baseline specification (Table 3: Panel A) and the specification with full controls (Table 3: Panel C).

FIGURE 6. DEBT ACCUMULATION BY LOW AND HIGH-RANK HOUSEHOLDS AND LOCAL INEQUALITY, NONPARAMETRIC SPECIFICATION



Note: The figure shows the estimated coefficients on the income rank dummies from the nonparametric regressions of the relative household debt accumulation between 2001 and year t. Each regression contains dummies for income ranks and inequality levels (with low-rank households in low-inequality regions being the benchmark), and a full set of controls described in equation (3) and the county-specific fixed effects. Mid-rank households are not shown in Figure. See section 3.4 for details.

FIGURE 7. GROWTH OF AVERAGE INCOME WITHIN DECILE AND INEQUALITY

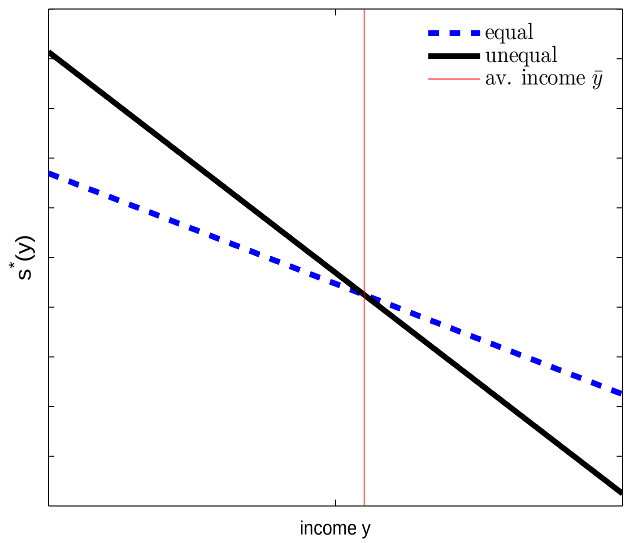


Note: The figure shows the estimated coefficients on inequality in the base year (i.e. 1970 or 1990) from regressing the log difference of average income within a decile across metro areas on measured inequality. Data are from IPUMS. Inequality is measured as the log P90/P10. Confidence intervals are at the 95% level using heteroskedasticity-robust standard errors and each regression contains a constant. See section 3.7 and Appendix C for more details.

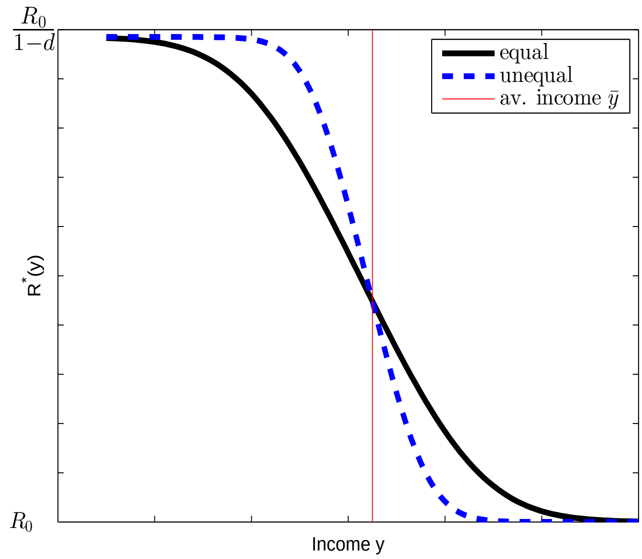
FIGURE 8. THEORETICAL EFFECTS OF A CHANGE IN INEQUALITY ON PROVISION OF CREDIT

Bank Sorting and Inequality under Perfect Competition

Panel A

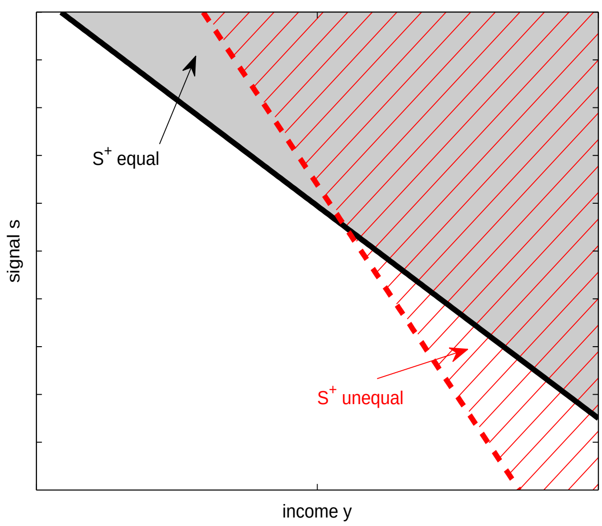


Panel B

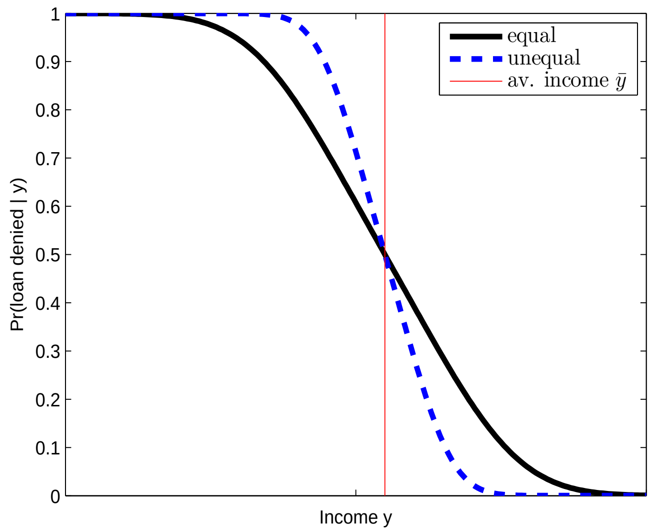


Bank Sorting and Inequality under Monopoly Banking

Panel C



Panel D



Note: Panel A shows the tradeoff $s^*(y)$ for baseline income distribution ("equal") and more unequal income distribution ("unequal"). Panel B plots the interest rate for each income level and for different levels of income inequality. In Panels A and B banks can price discriminate perfectly. Panel C plots sets of households with signals s and y who obtain loans for two "equal" and "unequal" income distributions. Shaded regions indicate combinations of signals that yield an approved loan. Panel D plots loan deny probability as a function of income. In Panels C and D, the bank changes the same rate for all applicants.

TABLE 1: SUMMARY STATISTICS

Category	Mean	St. Dev.	Percentiles 10	25	50	75	90
Panel A: FRBNY Consumer Credit Panel/ Equifax, Q3 2001							
Age of head of household	42.6	11.0	28	34	42	51	58
Household size	3.0	1.7	1	2	3	4	5
Housing debt	56,423	99,938	0	0	12,351	83,255	156,082
Mortgage	54,658	97,202	0	0	8,267	81,163	153,000
HELOC	1,765	12,565	0	0	0	0	0
Auto loans	6,876	11,543	0	0	0	10,805	21,376
Credit card limit	30,459	36,452	1,609	6,127	19,320	42,288	73,009
Credit card balance	8,884	14,812	261	1,120	3,923	10,881	22,893
Student loan	1,639	7,849	0	0	0	0	2,723
Consumer financing	929	5,861	0	0	0	178	2,033
Other debt	4,044	22,158	0	0	0	0	10,410
Total debt	78,794	112,167	1,368	9,437	42,311	111,335	193,395
Bankruptcy rate	0.12	0.32	0.00	0.00	0.00	0.00	1.00
Delinquency rate	0.30	0.46	0.00	0.00	0.00	1.00	1.00
Credit card utilization rate	0.41	0.35	0.02	0.09	0.31	0.71	0.99
Panel B: Survey of Consumer Finances, 2001							
Age of head of household	43.3	11.3	28	35	43	52	59
Household size	2.8	1.4	1	2	2	4	5
Housing debt	60,783	119,310	0	0	29,000	90,000	150,000
Mortgage debt	57,643	90,243	0	0	27,000	88,000	147,000
HELOC	3,140	73,981	0	0	0	0	0
Auto loans	5,182	8,280	0	0	0	8,700	18,000
Credit card limit	19,290	43,636	1,400	4,500	10,000	22,000	42,000
Credit card balance	2,586	5,459	0	0	500	3,000	7,200
Student loan	2,271	9,786	0	0	0	0	5,000
Consumer financing							
Other debt							
Total debt	70,822	121,163	30	6,140	40,000	101,000	164,800
Bankruptcy rate	0.10	0.30	0.00	0.00	0.00	0.00	1.00
Delinquency rate	0.05	0.21	0.00	0.00	0.00	0.00	0.00
Credit card utilization rate	0.27	0.34	0.00	0.00	0.08	0.47	0.93

Note: The sample is restricted to the households with 20-65 year old head of household. The statistics are calculated using sampling weights. Housing debt is the sum of Mortgage and HELOC. The credit card limit is the maximum of the originally recorded credit card limit in the CCP and the credit card balance. The credit card utilization rate is calculated using this credit card limit. The table shows the statistics from the sample restricted to observations with nonzero credit card limit. The delinquency rate is a share of households with at least one member with an account that is 60 day past due or more. The number of observations in Panel A is 7,710,406. The number of observations in Panel B is 14,356.

TABLE 2: INCOME STATISTICS FROM SCF (ACTUAL) AND CCP (IMPUTED)

	Mean	St. dev.	Percentiles				
			10	25	50	75	90
Ln(Y), actual in SCF	10.64	0.97	9.40	10.09	10.69	11.23	11.70
Ln(Y), imputed in CCP	10.91	1.18	9.55	10.15	10.81	11.51	12.36

Note: The sample is restricted to households with the 20-65 y.o. head of household and positive gross income. The sample in the SCF is further restricted to remove outliers. See text for more details.

TABLE 3: BASELINE RESULTS ON HOUSEHOLD DEBT ACCUMULATION

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
						Panel A: Parsimonious Specification					
α	-1.261***	-1.898***	-2.885***	-3.416***	-3.953***	-4.128***	-3.998***	-3.936***	-3.570***	-3.189***	-2.788***
	(0.0237)	(0.0352)	(0.0431)	(0.0521)	(0.0606)	(0.0654)	(0.0649)	(0.0643)	(0.0603)	(0.0562)	(0.0520)
β	-0.294***	-0.398***	-0.689***	-0.776***	-0.889***	-0.883***	-0.791***	-0.753***	-0.610***	-0.466***	-0.309***
	(0.00836)	(0.0121)	(0.0164)	(0.0194)	(0.0227)	(0.0245)	(0.0242)	(0.0240)	(0.0222)	(0.0204)	(0.0188)
γ	0.544***	0.816***	1.387***	1.637***	1.898***	1.925***	1.784***	1.732***	1.477***	1.214***	0.922***
	(0.0158)	(0.0236)	(0.0289)	(0.0351)	(0.0411)	(0.0445)	(0.0442)	(0.0437)	(0.0410)	(0.0382)	(0.0353)
N	5,925,610	5,449,695	4,837,540	4,387,387	4,050,160	3,792,576	3,581,989	3,438,004	3,295,854	3,178,324	3,069,446
R^2	0.018	0.025	0.031	0.038	0.044	0.048	0.052	0.051	0.051	0.053	0.055
						Panel B: Specification with Household Controls					
α	-1.504***	-2.271***	-3.267***	-3.780***	-4.324***	-4.501***	-4.404***	-4.369***	-3.996***	-3.585***	-3.191***
	(0.0219)	(0.0316)	(0.0418)	(0.0516)	(0.0614)	(0.0668)	(0.0666)	(0.0663)	(0.0626)	(0.0581)	(0.0532)
β	-0.376***	-0.478***	-0.708***	-0.800***	-0.924***	-0.959***	-0.916***	-0.897***	-0.802***	-0.690***	-0.586***
	(0.00864)	(0.0119)	(0.0163)	(0.0199)	(0.0238)	(0.0262)	(0.0262)	(0.0261)	(0.0245)	(0.0226)	(0.0207)
γ	0.667***	0.957***	1.465***	1.725***	2.012***	2.102***	2.037***	2.021***	1.826***	1.602***	1.381***
	(0.0147)	(0.0213)	(0.0283)	(0.0351)	(0.0419)	(0.0458)	(0.0456)	(0.0455)	(0.0430)	(0.0398)	(0.0364)
N	5,760,889	5,287,480	4,685,165	4,245,118	3,921,002	3,669,090	3,468,476	3,327,359	3,186,253	3,069,980	2,964,520
R^2	0.050	0.063	0.069	0.076	0.081	0.086	0.095	0.098	0.104	0.114	0.125
						Panel C: Specification with Household and Zip-Level Controls					
α	-1.500***	-2.285***	-3.246***	-3.752***	-4.280***	-4.454***	-4.354***	-4.306***	-3.937***	-3.533***	-3.156***
	(0.0220)	(0.0316)	(0.0419)	(0.0518)	(0.0616)	(0.0670)	(0.0667)	(0.0665)	(0.0627)	(0.0582)	(0.0534)
β	-0.330***	-0.428***	-0.632***	-0.712***	-0.823***	-0.850***	-0.811***	-0.795***	-0.714***	-0.613***	-0.525***
	(0.00803)	(0.0113)	(0.0152)	(0.0186)	(0.0223)	(0.0246)	(0.0246)	(0.0248)	(0.0234)	(0.0218)	(0.0202)
γ	0.673***	0.960***	1.483***	1.750***	2.045***	2.139***	2.078***	2.061***	1.864***	1.636***	1.409***
	(0.0147)	(0.0213)	(0.0283)	(0.0352)	(0.0421)	(0.0459)	(0.0457)	(0.0456)	(0.0430)	(0.0399)	(0.0365)
N	5,760,889	5,287,480	4,685,165	4,245,118	3,921,002	3,669,090	3,468,476	3,327,359	3,186,253	3,069,980	2,964,520
R^2	0.051	0.064	0.070	0.078	0.082	0.088	0.097	0.100	0.105	0.115	0.126
						Panel D: Specification with Zip-Level Fixed Effects					
α	-1.506***	-2.293***	-3.260***	-3.771***	-4.302***	-4.477***	-4.373***	-4.320***	-3.943***	-3.539***	-3.153***
	(0.111)	(0.167)	(0.269)	(0.351)	(0.419)	(0.480)	(0.472)	(0.463)	(0.409)	(0.359)	(0.330)
γ	0.674***	0.962***	1.486***	1.756***	2.052***	2.147***	2.085***	2.066***	1.864***	1.637***	1.404***
	(0.0655)	(0.101)	(0.166)	(0.226)	(0.278)	(0.325)	(0.315)	(0.307)	(0.269)	(0.232)	(0.212)
N	5,760,889	5,287,480	4,685,165	4,245,118	3,921,002	3,669,090	3,468,476	3,327,359	3,186,253	3,069,980	2,964,520
R^2	0.054	0.067	0.074	0.082	0.088	0.094	0.103	0.106	0.111	0.121	0.132

Note: The table presents estimates of specifications (2), (3), (4) and (5) in Panels A through D respectively. Coefficient α corresponds to the partial correlation of household income rank and debt accumulation between 2001 and the year indicated in each column (relative to household's 2001 income). Coefficient β corresponds to the partial correlation of local inequality and household debt accumulation. Coefficient γ is for the interaction of household income and local inequality. Each regression is run at the household level. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. In Panels A-C, the standard errors are clustered by zip code; in Panel D, standard errors are clustered by state. See sections 3.1 and 3.2 in the text for details.

TABLE 4: INTERACTIONS OF RANK WITH CREDIT SCORES AND INITIAL DEBT LEVELS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Panel A: Include Interaction of Household Credit Score and Local Inequality											
α	-1.361***	-2.046***	-2.876***	-3.340***	-3.827***	-4.036***	-4.003***	-3.962***	-3.625***	-3.244***	-2.914***
	(0.0235)	(0.0330)	(0.0440)	(0.0537)	(0.0628)	(0.0686)	(0.0679)	(0.0676)	(0.0631)	(0.0581)	(0.0538)
β	-0.708***	-1.076***	-1.631***	-1.861***	-2.133***	-2.106***	-1.905***	-1.890***	-1.729***	-1.583***	-1.354***
	(0.0198)	(0.0309)	(0.0412)	(0.0519)	(0.0648)	(0.0749)	(0.0783)	(0.0790)	(0.0744)	(0.0695)	(0.0655)
γ	0.577***	0.795***	1.227***	1.465***	1.731***	1.849***	1.835***	1.823***	1.647***	1.436***	1.241***
	(0.0158)	(0.0223)	(0.0298)	(0.0366)	(0.0430)	(0.0471)	(0.0465)	(0.0464)	(0.0433)	(0.0399)	(0.0368)
φ	-0.307***	-0.690***	-1.386***	-1.727***	-2.128***	-2.007***	-1.553***	-1.359***	-1.269***	-1.281***	-1.113***
	(0.0383)	(0.0584)	(0.0769)	(0.0958)	(0.117)	(0.136)	(0.142)	(0.142)	(0.132)	(0.123)	(0.116)
σ	0.512***	0.879***	1.353***	1.545***	1.751***	1.668***	1.445***	1.441***	1.333***	1.268***	1.082***
	(0.0255)	(0.0392)	(0.0520)	(0.0650)	(0.0799)	(0.0925)	(0.0966)	(0.0970)	(0.0901)	(0.0837)	(0.0789)
N	5,760,889	5,287,480	4,685,165	4,245,118	3,921,002	3,669,090	3,468,476	3,327,359	3,186,253	3,069,980	2,964,520
R^2	0.051	0.064	0.070	0.078	0.083	0.088	0.097	0.100	0.106	0.115	0.126
Panel B: Include Interaction of Initial Household Debt Level and Local Inequality											
α	-0.516***	-1.171***	-2.017***	-2.422***	-2.970***	-3.069***	-2.916***	-2.814***	-2.316***	-1.848***	-1.309***
	(0.0275)	(0.0387)	(0.0489)	(0.0605)	(0.0732)	(0.0814)	(0.0849)	(0.0857)	(0.0802)	(0.0769)	(0.0710)
β	-0.312***	-0.452***	-0.670***	-0.758***	-0.878***	-0.910***	-0.881***	-0.857***	-0.770***	-0.659***	-0.556***
	(0.0118)	(0.0170)	(0.0224)	(0.0273)	(0.0329)	(0.0357)	(0.0370)	(0.0374)	(0.0365)	(0.0348)	(0.0328)
γ	0.233***	0.530***	0.987***	1.203***	1.481***	1.529***	1.460***	1.433***	1.221***	1.014***	0.744***
	(0.0200)	(0.0282)	(0.0359)	(0.0443)	(0.0540)	(0.0600)	(0.0627)	(0.0631)	(0.0591)	(0.0564)	(0.0520)
φ	-2.97***	-3.79***	-4.09***	-4.47***	-4.59***	-5.00***	-5.37***	-5.49***	-6.05***	-6.21***	-6.876***
	(0.089)	(0.115)	(0.125)	(0.147)	(0.167)	(0.200)	(0.214)	(0.213)	(0.199)	(0.214)	(0.195)
σ	1.67***	2.15***	2.49***	2.81***	3.05***	3.38***	3.54***	3.55***	3.67***	3.53***	3.71***
	(0.063)	(0.0824)	(0.891)	(0.105)	(0.122)	(0.147)	(0.158)	(0.153)	(0.144)	(0.152)	(0.140)
N	3,989,837	3,643,849	3,203,783	2,882,349	2,650,275	2,470,570	2,329,399	2,228,828	2,128,927	2,047,809	1,974,388
R^2	0.053	0.061	0.064	0.070	0.074	0.079	0.088	0.091	0.098	0.109	0.124

Note: The table presents estimates of specification (3') and (3'') in section 3.2. Coefficient α corresponds to the partial correlation of household income rank and debt accumulation between 2001 and the year indicated in each column (relative to household's 2001 income). Coefficient β corresponds to the partial correlation of local inequality and household debt accumulation. Coefficient γ is for the interaction of household income and local inequality. Coefficient φ represent the effects of each additional variable (household credit score in Panel A and initial household debt level in Panel B) while σ captures the interaction of this household variable with local inequality. Each regression is run at the household level. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. The standard errors are clustered by zip code. In Panel B, coefficients φ and σ and the respective standard errors are multiplied by 10^6.

TABLE 5: HOUSEHOLD DEBT ACCUMULATION ALONG SUBSETS OF DATA

		α	β	γ	N	R^2
Grouping Zip Codes by Census Region	Midwest	-3.352***	-0.434***	1.376***	872,335	0.107
		(0.135)	(0.0526)	(0.0965)		
	Northeast	-4.440***	-0.908***	2.316***	739,940	0.076
		(0.130)	(0.0494)	(0.0945)		
	South	-4.619***	-0.802***	2.157***	1,328,024	0.101
		(0.126)	(0.0443)	(0.0848)		
	West	-6.233***	-1.369***	3.101***	728,791	0.061
		(0.187)	(0.0638)	(0.121)		
Grouping Zip Codes by Average Credit Ratings	Low	-6.205***	-1.476***	3.375***	999,984	0.093
		(0.146)	(0.0418)	(0.0994)		
	Middle	-5.130***	-1.052***	2.548***	1,185,568	0.102
		(0.106)	(0.0404)	(0.0731)		
	High	-2.515***	-0.218***	1.214***	1,483,538	0.101
		(0.0705)	(0.0286)	(0.056)		
Grouping Zip Codes by Initial Average Debt-to-Income Ratios	Low	-3.253***	-0.631***	1.512***	951,154	0.072
		(0.166)	(0.0599)	(0.111)		
	Middle	-4.175***	-0.772***	1.933***	1,244,905	0.088
		(0.120)	(0.0443)	(0.0819)		
	High	-4.468***	-0.834***	2.083***	1,473,031	0.100
		(0.0893)	(0.0342)	(0.0621)		
Grouping Zip Codes by House Price Growth (2001-2005)	Low	-3.872***	-0.577***	1.677***	836,451	0.114
		(0.135)	(0.0510)	(0.0941)		
	Middle	-5.136***	-1.024***	2.603***	820,675	0.083
		(0.134)	(0.0501)	(0.0919)		
	High	-5.650***	-1.206***	2.828***	799,557	0.061
		(0.179)	(0.0614)	(0.119)		
Grouping Zip Codes by 2001 Average House Price to Median Income Ratio	Low	-4.707***	-0.915***	2.232***	795,208	0.051
		(0.144)	(0.050)	(0.093)		
	Middle	-4.256***	-0.728***	1.847***	830,645	0.103
		(0.150)	(0.057)	(0.103)		
	High	-3.702***	-0.566***	1.585***	834,311	0.115
		(0.151)	(0.059)	(0.106)		

Note: The table presents estimates of specification (4) in the text using household debt accumulation from 2001 to 2007. Panel A presents separate estimates for households located in each of four Census regions. Panel B presents estimates for households in zip codes with low, medium, or high initial average credit ratings. Panel C presents estimates for households in zip codes with low, medium, or high initial average debt-to-income ratios. Panel D decomposes zip codes by growth of house prices between 2001 and 2005. See section 3.3 in the text for details. Coefficient α corresponds to the partial correlation of household income rank and debt accumulation between 2001 and the year indicated in each column (relative to household's 2001 income). Coefficient β corresponds to the partial correlation of local inequality and household debt accumulation. Coefficient γ is for the interaction of household income and local inequality. Each regression is run at the household level. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. The standard errors are clustered by zip code.

TABLE 6: MEASURING INEQUALITY AT DIFFERENT LEVELS OF AGGREGATION

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
	Panel A: Inequality at the County Level										
α	-1.174***	-2.073***	-3.108***	-3.949***	-4.756***	-5.179***	-5.055***	-4.996***	-4.560***	-4.176***	-3.631***
	(0.0865)	(0.134)	(0.252)	(0.321)	(0.417)	(0.475)	(0.493)	(0.475)	(0.452)	(0.445)	(0.382)
β	-0.241***	-0.310***	-0.456***	-0.548***	-0.570***	-0.578**	-0.519**	-0.501**	-0.475**	-0.467**	-0.426**
	(0.0423)	(0.0671)	(0.118)	(0.156)	(0.202)	(0.232)	(0.237)	(0.227)	(0.209)	(0.200)	(0.174)
γ	0.583***	0.986***	1.531***	1.993***	2.413***	2.626***	2.545***	2.534***	2.343***	2.170***	1.861***
	(0.0606)	(0.0943)	(0.175)	(0.224)	(0.293)	(0.334)	(0.344)	(0.330)	(0.314)	(0.309)	(0.264)
N	6,640,570	6,257,495	5,782,494	5,435,548	5,172,907	4,966,746	4,793,457	4,661,838	4,531,493	4,421,495	4,319,303
R^2	0.048	0.060	0.070	0.079	0.086	0.091	0.098	0.100	0.105	0.115	0.125
	Panel B: Inequality at the State Level										
α	-0.926**	-1.710***	-2.852**	-4.036***	-5.283***	-5.651***	-5.592***	-5.545***	-4.969***	-4.482***	-3.795***
	(0.359)	(0.543)	(1.114)	(1.412)	(1.667)	(1.697)	(1.612)	(1.525)	(1.476)	(1.391)	(1.224)
β	0.0490	0.0832	0.254	0.478	0.839**	1.317***	1.472***	1.386***	1.193**	1.001**	0.863*
	(0.114)	(0.163)	(0.259)	(0.324)	(0.394)	(0.458)	(0.469)	(0.483)	(0.479)	(0.468)	(0.447)
γ	0.393	0.695*	1.280*	1.937**	2.616**	2.765**	2.711**	2.708**	2.409**	2.170**	1.770**
	(0.242)	(0.367)	(0.754)	(0.954)	(1.125)	(1.144)	(1.080)	(1.019)	(0.988)	(0.929)	(0.815)
N	7,015,125	6,704,094	6,344,116	6,088,596	5,893,406	5,737,576	5,600,035	5,490,380	5,383,103	5,293,822	5,209,929
R^2	0.049	0.062	0.071	0.082	0.088	0.092	0.099	0.100	0.108	0.119	0.130

Note: The table presents estimates of specification (4) while measuring inequality at different levels of aggregation: county level in Panel A and state level in Panel B. Coefficient α corresponds to the partial correlation of household income rank and debt accumulation between 2001 and the year indicated in each column (relative to household's 2001 income). Coefficient β corresponds to the partial correlation of local inequality and household debt accumulation. Coefficient γ is for the interaction of household income and local inequality. Each regression is run at the household level. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. See section 3.4 in the text for details.

TABLE 7: RESULTS BY FORM OF DEBT

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
	Panel A: Mortgage Debt Accumulation										
α	-1.280***	-1.991***	-2.840***	-3.243***	-3.727***	-3.981***	-3.873***	-3.779***	-3.504***	-3.192***	-2.868***
	(0.0184)	(0.0274)	(0.0375)	(0.0452)	(0.0538)	(0.0594)	(0.0587)	(0.0574)	(0.0558)	(0.0526)	(0.0479)
β	-0.320***	-0.444***	-0.631***	-0.699***	-0.798***	-0.846***	-0.805***	-0.778***	-0.707***	-0.617***	-0.539***
	(0.00659)	(0.00962)	(0.0134)	(0.0161)	(0.0193)	(0.0215)	(0.0213)	(0.0211)	(0.0204)	(0.0193)	(0.0177)
γ	0.660***	0.985***	1.452***	1.673***	1.938***	2.078***	1.993***	1.932***	1.757***	1.555***	1.358***
	(0.0123)	(0.0184)	(0.0254)	(0.0309)	(0.0370)	(0.0410)	(0.0404)	(0.0395)	(0.0384)	(0.0362)	(0.0329)
N	5,759,852	5,286,511	4,684,155	4,244,067	3,919,926	3,667,964	3,467,395	3,326,197	3,185,052	3,068,773	2,963,305
R^2	0.052	0.063	0.068	0.078	0.082	0.087	0.096	0.099	0.109	0.122	0.138
	Panel B: Auto Debt Accumulation										
α	-0.084***	-0.162***	-0.210***	-0.231***	-0.228***	-0.215***	-0.187***	-0.155***	-0.132***	-0.133***	-0.142***
	(0.00311)	(0.00427)	(0.00507)	(0.00543)	(0.00577)	(0.00581)	(0.00584)	(0.00567)	(0.00547)	(0.00556)	(0.00553)
β	-0.021***	-0.032***	-0.038***	-0.039***	-0.037***	-0.037***	-0.030***	-0.024***	-0.019***	-0.020***	-0.022***
	(0.0012)	(0.0017)	(0.0021)	(0.0022)	(0.0023)	(0.0023)	(0.0023)	(0.0022)	(0.0021)	(0.0021)	(0.0021)
γ	0.0185***	0.0302***	0.0426***	0.0494***	0.0486***	0.0450***	0.0365***	0.0249***	0.0209***	0.0268***	0.0325***
	(0.0020)	(0.0028)	(0.0034)	(0.0036)	(0.0038)	(0.0039)	(0.0039)	(0.0038)	(0.0037)	(0.0038)	(0.0037)
N	5,761,635	5,287,863	4,684,952	4,244,817	3,920,756	3,669,005	3,468,554	3,327,421	3,186,260	3,069,941	2,964,809
R^2	0.083	0.110	0.123	0.134	0.144	0.157	0.181	0.199	0.218	0.225	0.223
	Panel C: Credit Card Balance Accumulation										
α	-0.025***	-0.010***	0.001	0.009**	0.0161***	0.006	0.011**	0.014**	0.030***	0.035***	0.042***
	(0.0024)	(0.0033)	(0.0039)	(0.0043)	(0.0048)	(0.0051)	(0.0054)	(0.0056)	(0.0053)	(0.0050)	(0.0049)
β	-0.001	0.001	0.000	0.004***	0.004**	-0.0006	-0.003	-0.003	-0.002	-0.004**	-0.002
	(0.0009)	(0.0012)	(0.0015)	(0.0016)	(0.0018)	(0.0019)	(0.0021)	(0.0022)	(0.0020)	(0.0019)	(0.0019)
γ	0.002	0.001	0.004	0.000	0.003	0.009***	0.0109***	0.011***	0.018***	0.026***	0.025***
	(0.0016)	(0.0022)	(0.0026)	(0.0029)	(0.0032)	(0.0034)	(0.00361)	(0.0038)	(0.0035)	(0.0034)	(0.0034)
N	5,237,881	4,732,993	4,180,223	3,803,376	3,512,256	3,293,489	3,111,432	2,946,655	2,798,244	2,699,678	2,602,128
R^2	0.085	0.119	0.144	0.155	0.168	0.162	0.161	0.166	0.204	0.234	0.252
	Panel D: Credit Card Limits										
α	-0.171***	-0.231***	-0.282***	-0.405***	-0.409***	-0.476***	-0.473***	-0.404***	-0.337***	-0.315***	-0.303***
	(0.00650)	(0.00881)	(0.0103)	(0.0136)	(0.0140)	(0.0165)	(0.0175)	(0.0163)	(0.0147)	(0.0143)	(0.0150)
	-	-		-	-	-	-	-	-	-	-
β	0.0177***	0.0256***	-0.0441***	0.0441***	0.0491***	0.0599***	0.0478***	0.0787***	0.0898***	0.0768***	0.0604***
	(0.00219)	(0.00295)	(0.00354)	(0.00460)	(0.00482)	(0.00577)	(0.00630)	(0.00584)	(0.00532)	(0.00545)	(0.00577)
γ	0.00691	0.0268***	0.0634***	0.0378***	0.0637***	0.0618***	0.0403***	0.138***	0.171***	0.183***	0.171***
	(0.00431)	(0.00589)	(0.00693)	(0.00913)	(0.00946)	(0.0112)	(0.0119)	(0.0111)	(0.00992)	(0.00972)	(0.0102)
N	5,761,303	5,287,941	4,685,242	4,245,256	3,920,953	3,669,293	3,468,772	3,327,343	3,186,164	3,069,851	2,964,562
R^2	0.043	0.070	0.103	0.128	0.131	0.139	0.143	0.164	0.203	0.226	0.236

Note: The table presents estimates of specification (4) for different forms of household debt: mortgage debt in Panel A, auto debt in Panel B, credit card balances in Panel C and credit card limits in Panel D. Coefficient α corresponds to the partial correlation of household income rank and debt accumulation between 2001 and the year indicated in each column (relative to household's 2001 income). Coefficient β corresponds to the partial correlation of local inequality and household debt accumulation. Coefficient γ is for the interaction of household income and local inequality. Each regression is run at the household level. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. See section 3.6 in the text for details.

TABLE 8: THE LOG P90/P10 RATIO OF INCOME IN 2000 AND EARLIER YEARS ACROSS METRO AREAS

	1970	1980	1990
β	0.328***	0.697***	0.734***
	(0.0615)	(0.0843)	(0.0643)
N	117	117	117
R^2	0.204	0.379	0.526

Note: The table presents estimates of the extent to which lagged measured inequality predicts current measured inequality. For example, the column labeled 1970 regresses the log p90/p10 ratio for metro areas in 2000 on the same measure from 1970. The same metro areas are used in every year. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. See section 3.7 in the text for more details.

TABLE 9: MORTGAGE APPLICATIONS AND LOCAL INEQUALITY

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
					Panel A: Probability of Mortgage Application Being Rejected						
α	-0.084***	-0.066***	-0.055***	-0.055***	-0.057***	-0.045***	-0.038***	-0.041***	-0.039***	-0.055***	-0.060***
	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
γ	-0.394***	-0.336***	-0.279***	-0.343***	-0.305***	-0.309***	-0.239***	-0.179***	-0.201***	-0.273***	-0.375***
	(0.083)	(0.069)	(0.051)	(0.035)	(0.033)	(0.036)	(0.035)	(0.038)	(0.030)	(0.035)	(0.042)
N	2,244,576	2,264,842	2,520,425	2,635,465	2,970,262	2,663,236	1,921,810	1,319,589	1,240,372	1,275,372	1,196,404
R^2	0.124	0.095	0.068	0.061	0.056	0.056	0.059	0.048	0.042	0.055	0.071
					Panel B: Probability of Mortgage Being High-Interest (conditional on origination)						
α				-0.038***	-0.062***	-0.051***	-0.037***	-0.045***	-0.024***	-0.023***	-0.029***
				(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
γ				-0.183***	-0.220***	-0.202***	-0.171***	-0.126***	-0.073***	-0.106***	-0.132***
				(0.026)	(0.038)	(0.038)	(0.030)	(0.027)	(0.014)	(0.015)	(0.021)
N				1,995,005	2,148,955	1,892,164	1,384,324	959,930	944,620	955,348	894,997
R^2				0.110	0.173	0.139	0.080	0.065	0.047	0.082	0.082
					Panel C: Loan-to-Income Ratios of Mortgage Applications (conditional on origination)						
α	-0.164***	-0.174***	-0.184***	-0.174***	-0.165***	-0.167***	-0.182***	-0.183***	-0.189***	-0.192***	-0.186***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
γ	0.094	0.071	0.129**	0.145***	0.085*	0.089*	0.174***	0.107**	0.029	0.110**	0.084
	(0.067)	(0.065)	(0.065)	(0.050)	(0.044)	(0.046)	(0.045)	(0.054)	(0.054)	(0.054)	(0.057)
N	1,746,160	1,794,892	1,971,148	1,995,005	2,148,955	1,892,164	1,384,324	959,930	944,620	955,348	894,997
R^2	0.327	0.352	0.376	0.355	0.340	0.351	0.369	0.379	0.404	0.409	0.391

Note: The table presents estimates of specification (6) for different dependent variables as indicated in each panel. Coefficient α corresponds to the partial correlation of applicant's income rank and the dependent variable in the year indicated by each column. Coefficient γ corresponds to the interaction of local inequality and applicant's income rank. Standard errors are clustered at the county level and each regression includes a county fixed effects as well as controls for race, sex, occupancy, the LTI, and an interaction of rank with the fraction of non-white applicants. The sample is restricted to home purchase loans with an LTI between 1 and 8 and where the application was not rejected by the borrower or failed for a reason other than denial. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively. See section 5 in the text for more details.

APPENDIX

NOT FOR PUBLICATION

APPENDIX A: ADDITIONAL TABLES AND FIGURES

APPENDIX TABLE A1: ROBUSTNESS TO USING IRS MEASURE OF INEQUALITY

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: Parsimonious Specification						
α	-1.253***	-1.979***	-2.583***	-3.012***	-3.382***	-3.515***	-3.494***	-3.496***	-3.397***	-3.246***	-3.066***
	(0.0226)	(0.0339)	(0.0450)	(0.0540)	(0.0643)	(0.0698)	(0.0701)	(0.0686)	(0.0645)	(0.0588)	(0.0538)
β	-0.989***	-1.443***	-2.071***	-2.328***	-2.574***	-2.579***	-2.375***	-2.271***	-2.024***	-1.776***	-1.465***
	(0.0273)	(0.0400)	(0.0569)	(0.0678)	(0.0824)	(0.0884)	(0.0896)	(0.0879)	(0.0814)	(0.0731)	(0.0665)
γ	1.840***	2.972***	4.036***	4.646***	5.141***	5.133***	4.901***	4.872***	4.620***	4.256***	3.772***
	(0.0507)	(0.0761)	(0.101)	(0.121)	(0.144)	(0.156)	(0.157)	(0.154)	(0.146)	(0.133)	(0.122)
N	5,924,528	5,448,827	4,837,107	4,387,141	4,049,986	3,792,441	3,581,901	3,437,924	3,295,791	3,178,262	3,069,405
R^2	0.019	0.025	0.031	0.037	0.044	0.048	0.052	0.051	0.051	0.053	0.055
					Panel B: Specification with Household and Regional Controls						
α	-1.111***	-1.864***	-2.504***	-2.903***	-3.294***	-3.398***	-3.348***	-3.350***	-3.131***	-2.861***	-2.602***
	(0.0239)	(0.0347)	(0.0481)	(0.0582)	(0.0697)	(0.0756)	(0.0760)	(0.0749)	(0.0714)	(0.0656)	(0.0596)
β	-0.735***	-1.066***	-1.482***	-1.690***	-1.918***	-1.941***	-1.828***	-1.802***	-1.662***	-1.475***	-1.280***
	(0.0285)	(0.0406)	(0.0571)	(0.0690)	(0.0848)	(0.0923)	(0.0940)	(0.0937)	(0.0891)	(0.0822)	(0.0767)
γ	1.399***	2.309***	3.349***	4.014***	4.702***	4.856***	4.764***	4.822***	4.498***	4.033***	3.527***
	(0.0535)	(0.0782)	(0.109)	(0.132)	(0.159)	(0.172)	(0.173)	(0.171)	(0.164)	(0.151)	(0.137)
N	5,759,823	5,286,632	4,684,753	4,244,903	3,920,861	3,668,986	3,468,411	3,327,299	3,186,211	3,069,940	2,964,489
R^2	0.051	0.063	0.069	0.077	0.082	0.087	0.096	0.099	0.105	0.115	0.126

Note: The table reproduces the results in Table 3 of the text using the IRS measure of inequality rather than the CCP measure. See section 3.2 in the text for details.

APPENDIX TABLE A2: ALTERNATIVE SPECIFICATIONS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
	Panel A: Inverse of Expected Income Replaces Rank										
α	12,256***	20,148***	31,725***	41,280***	51,544***	57,399***	57,878***	57,950***	54,275***	49,893***	45,220***
	(322.6)	(532.1)	(709.6)	(888.7)	(1,092)	(1,236)	(1,285)	(1,280)	(1,226)	(1,162)	(1,104)
β	0.0232***	0.0949***	0.184***	0.285***	0.373***	0.417***	0.413***	0.418***	0.384***	0.340***	0.285***
	(0.00501)	(0.00775)	(0.0104)	(0.0125)	(0.0154)	(0.0171)	(0.0176)	(0.0174)	(0.0168)	(0.0160)	(0.0151)
γ	-5,710***	-9,588***	-16,741***	-21,889***	-27,505***	-30,109***	-29,449***	-29,231***	-26,394***	-23,090***	-19,328***
	(210.5)	(347.9)	(462.3)	(580.3)	(716.5)	(812.8)	(845.9)	(842.0)	(806.7)	(766.2)	(728.2)
N	5,925,610	5,449,695	4,837,540	4,387,387	4,050,160	3,792,576	3,581,989	3,438,004	3,295,854	3,178,324	3,069,446
R^2	0.009	0.013	0.017	0.023	0.030	0.035	0.038	0.037	0.037	0.038	0.040
	Panel B: Outcome is the Log Difference of Debt										
α	-0.968***	-1.052***	-1.138***	-1.087***	-1.072***	-1.052***	-1.003***	-1.032***	-0.979***	-0.688***	-0.497***
	(0.0468)	(0.0533)	(0.0606)	(0.0655)	(0.0704)	(0.0756)	(0.0789)	(0.0830)	(0.0865)	(0.0878)	(0.0888)
β	-0.224***	-0.220***	-0.271***	-0.190***	-0.131***	-0.143***	-0.0965***	-0.0860**	-0.0696*	0.0652	0.157***
	(0.0180)	(0.0245)	(0.0280)	(0.0304)	(0.0328)	(0.0358)	(0.0372)	(0.0391)	(0.0407)	(0.0408)	(0.0411)
γ	0.305***	0.317***	0.375***	0.305***	0.284***	0.275***	0.252***	0.280***	0.258***	0.0548	-0.0890
	(0.0317)	(0.0392)	(0.0445)	(0.0482)	(0.0519)	(0.0559)	(0.0584)	(0.0615)	(0.0641)	(0.0652)	(0.0659)
N	5,902,373	5,415,846	4,799,396	4,348,711	4,016,151	3,758,688	3,552,808	3,407,838	3,263,343	3,144,516	3,036,915
R^2	0.062	0.074	0.078	0.082	0.083	0.085	0.085	0.080	0.078	0.085	0.091

Note: This table estimates two alternative specifications to check if the imputation is inducing a spurious correlation. Panel A replaces rank with the inverse of expected income while Panel B uses the log difference of debt as the outcome instead of the change in debt normalized by initial income. See section 3.2 in the text for details.

Appendix Table A3: Robustness to Geographic Region

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
						Panel A: Midwest					
α	-1.424***	-2.168***	-2.911***	-3.107***	-3.431***	-3.352***	-3.212***	-3.219***	-2.867***	-2.581***	-2.289***
	(0.0492)	(0.0655)	(0.0914)	(0.108)	(0.129)	(0.135)	(0.134)	(0.133)	(0.125)	(0.121)	(0.111)
β	-0.316***	-0.388***	-0.512***	-0.482***	-0.486***	-0.434***	-0.365***	-0.360***	-0.312***	-0.241***	-0.186***
	(0.0196)	(0.0254)	(0.0350)	(0.0407)	(0.0496)	(0.0526)	(0.0533)	(0.0524)	(0.0494)	(0.0473)	(0.0439)
γ	0.633***	0.898***	1.282***	1.329***	1.477***	1.376***	1.298***	1.305***	1.121***	0.977***	0.796***
	(0.0346)	(0.0463)	(0.0653)	(0.0770)	(0.0918)	(0.0965)	(0.0964)	(0.0951)	(0.0900)	(0.0866)	(0.0802)
N	1,308,806	1,212,818	1,087,589	992,805	925,225	872,335	828,437	798,196	766,619	741,063	716,769
R^2	0.058	0.071	0.080	0.091	0.099	0.107	0.118	0.122	0.132	0.146	0.160
						Panel B: Northeast					
α	-1.340***	-2.191***	-3.168***	-3.593***	-4.230***	-4.440***	-4.409***	-4.348***	-4.278***	-3.908***	-3.546***
	(0.0420)	(0.0597)	(0.0845)	(0.101)	(0.118)	(0.130)	(0.140)	(0.141)	(0.131)	(0.123)	(0.113)
β	-0.288***	-0.432***	-0.677***	-0.721***	-0.860***	-0.908***	-0.891***	-0.880***	-0.901***	-0.795***	-0.724***
	(0.0157)	(0.0227)	(0.0313)	(0.0377)	(0.0445)	(0.0494)	(0.0526)	(0.0539)	(0.0503)	(0.0479)	(0.0439)
γ	0.649***	1.016***	1.609***	1.821***	2.190***	2.316***	2.284***	2.236***	2.224***	1.998***	1.769***
	(0.0300)	(0.0431)	(0.0615)	(0.0734)	(0.0858)	(0.0945)	(0.102)	(0.103)	(0.0960)	(0.0907)	(0.0830)
N	1,106,735	1,026,724	920,777	844,493	786,659	739,940	702,595	674,926	646,314	624,174	603,615
R^2	0.046	0.056	0.060	0.068	0.072	0.076	0.083	0.086	0.091	0.099	0.108
						Panel C: South					
α	-1.644***	-2.445***	-3.515***	-4.054***	-4.570***	-4.619***	-4.487***	-4.376***	-3.897***	-3.449***	-3.000***
	(0.0428)	(0.0647)	(0.0825)	(0.0995)	(0.118)	(0.126)	(0.126)	(0.128)	(0.126)	(0.117)	(0.110)
β	-0.370***	-0.453***	-0.677***	-0.755***	-0.859***	-0.802***	-0.740***	-0.721***	-0.607***	-0.511***	-0.401***
	(0.0149)	(0.0218)	(0.0283)	(0.0345)	(0.0407)	(0.0443)	(0.0447)	(0.0457)	(0.0448)	(0.0423)	(0.0404)
γ	0.738***	1.026***	1.608***	1.886***	2.161***	2.157***	2.090***	2.059***	1.811***	1.576***	1.314***
	(0.0281)	(0.0428)	(0.0548)	(0.0662)	(0.0791)	(0.0848)	(0.0844)	(0.0860)	(0.0844)	(0.0784)	(0.0736)
N	2,102,122	1,929,243	1,706,947	1,545,476	1,423,138	1,328,024	1,251,862	1,200,950	1,150,984	1,107,236	1,069,051
R^2	0.058	0.073	0.082	0.091	0.096	0.101	0.110	0.114	0.121	0.133	0.145
						Panel D: West					
α	-2.053***	-3.262***	-4.642***	-5.396***	-5.951***	-6.233***	-6.116***	-6.141***	-5.745***	-5.119***	-4.680***
	(0.0603)	(0.0884)	(0.111)	(0.146)	(0.171)	(0.187)	(0.183)	(0.184)	(0.168)	(0.154)	(0.134)
β	-0.482***	-0.707***	-1.009***	-1.178***	-1.307***	-1.369***	-1.334***	-1.333***	-1.234***	-1.079***	-0.969***
	(0.0206)	(0.0290)	(0.0377)	(0.0485)	(0.0569)	(0.0638)	(0.0607)	(0.0618)	(0.0565)	(0.0518)	(0.0458)
γ	0.970***	1.500***	2.221***	2.630***	2.933***	3.101***	3.015***	3.034***	2.827***	2.462***	2.214***
	(0.0381)	(0.0563)	(0.0707)	(0.0939)	(0.110)	(0.121)	(0.118)	(0.118)	(0.108)	(0.0991)	(0.0857)
N	1,243,226	1,118,695	969,852	862,344	785,980	728,791	685,582	653,287	622,336	597,507	575,085
R^2	0.042	0.053	0.055	0.058	0.059	0.061	0.067	0.068	0.071	0.078	0.089

Note: The table replicates the results in Panel A of Table 5 in the main text for each year in our sample.

APPENDIX TABLE A4: ROBUSTNESS TO AVERAGE LOCAL CREDIT RATINGS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: Low Average Credit Ratings						
α	-1.156***	-2.037***	-3.231***	-4.323***	-5.510***	-6.205***	-6.321***	-6.186***	-5.658***	-5.038***	-4.503***
	(0.0397)	(0.0576)	(0.0795)	(0.102)	(0.129)	(0.146)	(0.149)	(0.149)	(0.143)	(0.134)	(0.128)
β	-0.301***	-0.480***	-0.778***	-1.018***	-1.317***	-1.476***	-1.467***	-1.439***	-1.326***	-1.163***	-1.019***
	(0.0109)	(0.0160)	(0.0222)	(0.0289)	(0.0366)	(0.0418)	(0.0431)	(0.0439)	(0.0428)	(0.0406)	(0.0390)
γ	0.527***	0.930***	1.600***	2.241***	2.940***	3.375***	3.445***	3.383***	3.109***	2.746***	2.415***
	(0.0265)	(0.0386)	(0.0533)	(0.0691)	(0.0876)	(0.0994)	(0.101)	(0.102)	(0.0974)	(0.0910)	(0.0868)
N	1,811,119	1,646,108	1,417,541	1,237,579	1,104,956	999,984	917,093	864,212	812,178	763,809	724,970
R^2	0.056	0.074	0.078	0.088	0.091	0.093	0.099	0.101	0.111	0.126	0.140
					Panel B: Medium Average Local Credit Ratings						
α	-1.823***	-2.782***	-3.850***	-4.408***	-4.945***	-5.130***	-5.130***	-5.097***	-4.605***	-4.210***	-3.735***
	(0.0350)	(0.0501)	(0.0672)	(0.0821)	(0.0964)	(0.106)	(0.107)	(0.109)	(0.103)	(0.0980)	(0.0929)
β	-0.456***	-0.590***	-0.836***	-0.909***	-1.016***	-1.052***	-1.035***	-1.016***	-0.891***	-0.793***	-0.675***
	(0.0131)	(0.0187)	(0.0252)	(0.0306)	(0.0364)	(0.0404)	(0.0410)	(0.0422)	(0.0399)	(0.0384)	(0.0361)
γ	0.858***	1.248***	1.845***	2.139***	2.446***	2.548***	2.557***	2.543***	2.269***	2.059***	1.784***
	(0.0235)	(0.0338)	(0.0456)	(0.0560)	(0.0662)	(0.0731)	(0.0734)	(0.0749)	(0.0706)	(0.0673)	(0.0636)
N	1,909,729	1,731,649	1,518,184	1,372,935	1,266,001	1,185,568	1,121,637	1,075,671	1,029,356	992,664	958,771
R^2	0.056	0.070	0.082	0.092	0.098	0.102	0.111	0.113	0.118	0.128	0.137
					Panel C: High Average Local Credit Ratings						
α	-1.209***	-1.654***	-2.103***	-2.243***	-2.415***	-2.515***	-2.449***	-2.459***	-2.381***	-2.170***	-2.063***
	(0.0312)	(0.0417)	(0.0523)	(0.0590)	(0.0654)	(0.0705)	(0.0698)	(0.0721)	(0.0699)	(0.0656)	(0.0610)
β	-0.208***	-0.195***	-0.238***	-0.222***	-0.228***	-0.218***	-0.199***	-0.199***	-0.191***	-0.140***	-0.120***
	(0.0120)	(0.0165)	(0.0210)	(0.0234)	(0.0260)	(0.0286)	(0.0285)	(0.0286)	(0.0278)	(0.0263)	(0.0243)
γ	0.503***	0.577***	0.831***	0.888***	0.981***	1.016***	0.965***	0.960***	0.890***	0.740***	0.634***
	(0.0212)	(0.0285)	(0.0358)	(0.0404)	(0.0451)	(0.0486)	(0.0481)	(0.0497)	(0.0483)	(0.0452)	(0.0419)
N	2,040,041	1,909,723	1,749,440	1,634,604	1,550,045	1,483,538	1,429,746	1,387,476	1,344,719	1,313,507	1,280,779
R^2	0.063	0.075	0.089	0.094	0.097	0.101	0.111	0.113	0.117	0.125	0.134

Note: The table replicates the results in Panel B of Table 5 in the main text for each year in our sample.

APPENDIX TABLE A5: ROBUSTNESS TO AVERAGE INITIAL DEBT-TO-INCOME RATIOS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: Low Average Initial Debt-to-Income Ratio						
α	-0.995***	-1.453***	-2.202***	-2.675***	-3.178***	-3.253***	-3.117***	-3.070***	-2.738***	-2.453***	-2.235***
	(0.0410)	(0.0668)	(0.0934)	(0.122)	(0.148)	(0.166)	(0.165)	(0.163)	(0.165)	(0.152)	(0.139)
β	-0.234***	-0.262***	-0.410***	-0.505***	-0.619***	-0.631***	-0.592***	-0.565***	-0.503***	-0.431***	-0.378***
	(0.0147)	(0.0232)	(0.0326)	(0.0420)	(0.0522)	(0.0599)	(0.0605)	(0.0602)	(0.0600)	(0.0560)	(0.0523)
γ	0.442***	0.560***	0.968***	1.227***	1.487***	1.512***	1.433***	1.421***	1.268***	1.120***	0.994***
	(0.0272)	(0.0448)	(0.0622)	(0.0816)	(0.0985)	(0.111)	(0.110)	(0.109)	(0.111)	(0.102)	(0.0936)
N	1,536,549	1,405,965	1,234,921	1,113,369	1,023,921	951,154	892,311	853,127	813,229	779,065	749,549
R^2	0.045	0.056	0.059	0.066	0.068	0.072	0.080	0.086	0.096	0.110	0.125
					Panel B: Medium Average Initial Debt-to-Income Ratio						
α	-1.292***	-1.913***	-2.915***	-3.489***	-3.990***	-4.175***	-4.083***	-4.005***	-3.599***	-3.290***	-2.833***
	(0.0345)	(0.0502)	(0.0707)	(0.0862)	(0.107)	(0.120)	(0.122)	(0.124)	(0.115)	(0.109)	(0.101)
β	-0.259***	-0.310***	-0.532***	-0.632***	-0.738***	-0.772***	-0.730***	-0.716***	-0.629***	-0.556***	-0.437***
	(0.0129)	(0.0183)	(0.0261)	(0.0320)	(0.0399)	(0.0443)	(0.0449)	(0.0466)	(0.0433)	(0.0411)	(0.0384)
γ	0.546***	0.721***	1.267***	1.564***	1.841***	1.933***	1.884***	1.849***	1.638***	1.485***	1.209***
	(0.0230)	(0.0339)	(0.0476)	(0.0581)	(0.0732)	(0.0819)	(0.0828)	(0.0844)	(0.0782)	(0.0741)	(0.0686)
N	1,945,720	1,788,142	1,583,443	1,438,108	1,328,280	1,244,905	1,177,341	1,130,314	1,083,891	1,044,828	1,009,820
R^2	0.050	0.063	0.067	0.076	0.081	0.088	0.098	0.101	0.109	0.121	0.133
					Panel C: High Average Initial Debt-to-Income Ratio						
α	-1.654***	-2.489***	-3.413***	-3.833***	-4.313***	-4.468***	-4.367***	-4.356***	-4.026***	-3.591***	-3.249***
	(0.0324)	(0.0442)	(0.0573)	(0.0711)	(0.0838)	(0.0893)	(0.0889)	(0.0884)	(0.0825)	(0.0757)	(0.0705)
β	-0.356***	-0.470***	-0.647***	-0.705***	-0.803***	-0.834***	-0.802***	-0.790***	-0.709***	-0.605***	-0.537***
	(0.0123)	(0.0168)	(0.0215)	(0.0265)	(0.0309)	(0.0342)	(0.0341)	(0.0341)	(0.0323)	(0.0300)	(0.0280)
γ	0.730***	1.030***	1.517***	1.728***	1.995***	2.083***	2.012***	2.016***	1.829***	1.574***	1.374***
	(0.0222)	(0.0304)	(0.0393)	(0.0492)	(0.0581)	(0.0621)	(0.0618)	(0.0615)	(0.0573)	(0.0526)	(0.0488)
N	2,278,620	2,093,373	1,866,801	1,693,641	1,568,801	1,473,031	1,398,824	1,343,918	1,289,133	1,246,087	1,205,151
R^2	0.058	0.071	0.079	0.086	0.092	0.100	0.109	0.112	0.115	0.122	0.131

Note: The table replicates the results in Panel C of Table 5 in the main text for each year in our sample.

APPENDIX TABLE A6: ROBUSTNESS TO AVERAGE HOUSE PRICE GROWTH (2001-2005)

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: Low Average House Price Growth						
α	-1.703***	-2.689***	-3.509***	-3.745***	-3.965***	-3.872***	-4.611***	-5.124***	-4.311***	-3.800***	-3.184***
	(0.0495)	(0.0688)	(0.0940)	(0.108)	(0.129)	(0.135)	(0.147)	(0.149)	(0.138)	(0.127)	(0.118)
β	-0.388***	-0.527***	-0.668***	-0.640***	-0.633***	-0.577***	-0.788***	-0.975***	-0.746***	-0.613***	-0.460***
	(0.0181)	(0.0244)	(0.0347)	(0.0399)	(0.0471)	(0.0510)	(0.0553)	(0.0577)	(0.0523)	(0.0474)	(0.0461)
γ	0.778***	1.195***	1.608***	1.690***	1.773***	1.677***	2.215***	2.552***	2.055***	1.763***	1.379***
	(0.0331)	(0.0463)	(0.0639)	(0.0743)	(0.0889)	(0.0941)	(0.103)	(0.103)	(0.0955)	(0.0879)	(0.0818)
N	1,291,537	1,189,220	1,049,983	956,487	888,735	836,451	782,371	733,143	697,338	672,647	658,245
R^2	0.059	0.074	0.090	0.103	0.108	0.114	0.119	0.117	0.125	0.134	0.148
					Panel B: Medium Average House Price Growth						
α	-1.748***	-2.605***	-3.532***	-3.894***	-4.612***	-5.136***	-4.832***	-4.470***	-4.317***	-3.855***	-3.553***
	(0.0448)	(0.0666)	(0.0826)	(0.0983)	(0.121)	(0.134)	(0.145)	(0.142)	(0.136)	(0.127)	(0.116)
β	-0.416***	-0.527***	-0.686***	-0.718***	-0.865***	-1.024***	-0.915***	-0.778***	-0.778***	-0.652***	-0.613***
	(0.0174)	(0.0254)	(0.0313)	(0.0368)	(0.0457)	(0.0501)	(0.0554)	(0.0531)	(0.0508)	(0.0485)	(0.0445)
γ	0.851***	1.191***	1.688***	1.867***	2.281***	2.603***	2.368***	2.132***	2.070***	1.795***	1.643***
	(0.0300)	(0.0454)	(0.0564)	(0.0682)	(0.0839)	(0.0919)	(0.0987)	(0.0964)	(0.0923)	(0.0863)	(0.0787)
N	1,314,237	1,194,454	1,059,984	971,383	899,143	820,675	755,509	730,221	702,186	674,141	655,088
R^2	0.054	0.067	0.069	0.073	0.077	0.083	0.099	0.104	0.109	0.119	0.127
					Panel C: High Average House Price Growth						
α	-1.643***	-2.504***	-3.838***	-5.022***	-5.690***	-5.650***	-5.236***	-5.035***	-4.649***	-4.289***	-3.810***
	(0.0450)	(0.0663)	(0.0947)	(0.136)	(0.164)	(0.179)	(0.155)	(0.143)	(0.139)	(0.126)	(0.116)
β	-0.357***	-0.484***	-0.797***	-1.077***	-1.259***	-1.206***	-1.107***	-1.038***	-0.959***	-0.864***	-0.704***
	(0.0161)	(0.0235)	(0.0333)	(0.0466)	(0.0559)	(0.0614)	(0.0534)	(0.0508)	(0.0489)	(0.0450)	(0.0417)
γ	0.745***	1.065***	1.810***	2.480***	2.864***	2.828***	2.607***	2.522***	2.314***	2.130***	1.803***
	(0.0295)	(0.0436)	(0.0621)	(0.0890)	(0.108)	(0.119)	(0.103)	(0.0964)	(0.0940)	(0.0850)	(0.0777)
N	1,368,563	1,240,625	1,075,547	937,809	846,694	799,557	779,330	754,477	719,891	692,720	653,636
R^2	0.046	0.054	0.054	0.057	0.056	0.061	0.070	0.077	0.080	0.089	0.098

Note: The table replicates the results in Panel D of Table 5 in the main text for each year in our sample.

APPENDIX TABLE A7: ROBUSTNESS TO INITIAL LEVELS OF HOUSE PRICES RELATIVE TO INCOME

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
						Panel A: Low Initial Relative House Prices					
α	-1.417***	-2.150***	-3.125***	-3.728***	-4.367***	-4.707***	-4.714***	-4.722***	-4.351***	-3.949***	-3.569***
	(0.042)	(0.063)	(0.084)	(0.104)	(0.124)	(0.144)	(0.143)	(0.140)	(0.133)	(0.125)	(0.113)
β	-0.303***	-0.399***	-0.572***	-0.697***	-0.829***	-0.915***	-0.914***	-0.893***	-0.811***	-0.728***	-0.632***
	(0.014)	(0.021)	(0.029)	(0.036)	(0.043)	(0.050)	(0.049)	(0.049)	(0.046)	(0.043)	(0.040)
γ	0.624***	0.872***	1.363***	1.682***	2.037***	2.232***	2.231***	2.224***	2.022***	1.794***	1.560***
	(0.026)	(0.040)	(0.053)	(0.066)	(0.080)	(0.093)	(0.092)	(0.091)	(0.086)	(0.080)	(0.072)
N	1,346,793	1,210,187	1,047,956	935,253	855,929	795,208	748,478	712,722	677,495	650,400	624,841
R^2	0.036	0.043	0.043	0.045	0.047	0.051	0.058	0.061	0.064	0.071	0.081
						Panel B: Medium Initial Relative House Prices					
α	-1.595***	-2.489***	-3.304***	-3.689***	-4.152***	-4.256***	-4.190***	-4.054***	-3.723***	-3.283***	-2.991***
	(0.051)	(0.073)	(0.099)	(0.120)	(0.139)	(0.150)	(0.149)	(0.153)	(0.142)	(0.132)	(0.124)
β	-0.330***	-0.441***	-0.607***	-0.627***	-0.724***	-0.728***	-0.676***	-0.613***	-0.548***	-0.451***	-0.406***
	(0.020)	(0.028)	(0.038)	(0.045)	(0.053)	(0.057)	(0.057)	(0.059)	(0.056)	(0.052)	(0.049)
γ	0.670***	0.999***	1.402***	1.557***	1.802***	1.847***	1.811***	1.737***	1.571***	1.327***	1.176***
	(0.035)	(0.050)	(0.068)	(0.082)	(0.095)	(0.103)	(0.102)	(0.106)	(0.098)	(0.092)	(0.086)
N	1,333,467	1,220,350	1,076,042	968,303	890,466	830,645	783,737	751,365	719,215	692,286	668,525
R^2	0.062	0.076	0.084	0.092	0.096	0.103	0.113	0.116	0.122	0.132	0.142
						Panel C: High Initial Relative House Prices					
α	-1.419***	-2.161***	-3.015***	-3.381***	-3.641***	-3.702***	-3.485***	-3.538***	-3.291***	-2.890***	-2.515***
	(0.056)	(0.076)	(0.099)	(0.120)	(0.146)	(0.151)	(0.157)	(0.152)	(0.146)	(0.130)	(0.119)
β	-0.293***	-0.376***	-0.544***	-0.585***	-0.591***	-0.566***	-0.481***	-0.524***	-0.506***	-0.417***	-0.310***
	(0.021)	(0.028)	(0.037)	(0.045)	(0.056)	(0.059)	(0.061)	(0.060)	(0.057)	(0.052)	(0.047)
γ	0.596***	0.858***	1.308***	1.480***	1.577***	1.585***	1.445***	1.509***	1.416***	1.208***	0.993***
	(0.039)	(0.053)	(0.069)	(0.084)	(0.102)	(0.106)	(0.110)	(0.107)	(0.103)	(0.091)	(0.084)
N	1,299,320	1,198,652	1,065,879	966,058	891,869	834,311	788,325	756,972	725,798	699,816	676,498
R^2	0.065	0.082	0.091	0.104	0.109	0.115	0.126	0.129	0.136	0.149	0.162

Note: The table replicates the results in Panel E of Table 5 in the main text for each year in our sample.

APPENDIX TABLE A8

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
	Panel A: Households with No Mortgage Debt in 2001										
α	-0.934***	-1.897***	-3.120***	-3.780***	-4.622***	-4.915***	-4.832***	-4.786***	-4.314***	-3.987***	-3.562***
	(0.032)	(0.041)	(0.055)	(0.069)	(0.083)	(0.091)	(0.091)	(0.093)	(0.088)	(0.084)	(0.079)
β	-0.297***	-0.449***	-0.713***	-0.818***	-0.972***	-1.009***	-0.971***	-0.969***	-0.897***	-0.814***	-0.725***
	(0.009)	(0.013)	(0.017)	(0.022)	(0.026)	(0.029)	(0.030)	(0.030)	(0.028)	(0.027)	(0.026)
γ	0.431***	0.835***	1.472***	1.815***	2.258***	2.409***	2.372***	2.393***	2.169***	2.033***	1.813***
	(0.022)	(0.028)	(0.038)	(0.047)	(0.056)	(0.062)	(0.062)	(0.064)	(0.060)	(0.057)	(0.054)
N	2,748,810	2,482,153	2,149,720	1,912,682	1,743,540	1,609,502	1,500,510	1,425,800	1,351,290	1,289,411	1,236,456
R^2	0.035	0.048	0.062	0.068	0.074	0.077	0.082	0.083	0.084	0.085	0.085
	Panel B: Households with Positive Mortgage Debt in 2001										
α	-0.994***	-1.422***	-1.758***	-1.951***	-2.144***	-2.215***	-2.223***	-2.264***	-2.117***	-1.853***	-1.696***
	(0.031)	(0.043)	(0.053)	(0.062)	(0.070)	(0.076)	(0.078)	(0.080)	(0.078)	(0.073)	(0.068)
β	-0.088***	-0.074***	-0.037*	-0.030	-0.046	-0.062*	-0.083**	-0.100***	-0.109***	-0.088***	-0.104***
	(0.013)	(0.018)	(0.022)	(0.025)	(0.029)	(0.032)	(0.032)	(0.033)	(0.032)	(0.031)	(0.028)
γ	0.288***	0.360***	0.438***	0.516***	0.594***	0.643***	0.690***	0.744***	0.759***	0.680***	0.662***
	(0.022)	(0.030)	(0.036)	(0.043)	(0.049)	(0.053)	(0.054)	(0.055)	(0.054)	(0.051)	(0.047)
N	3,012,079	2,805,327	2,535,445	2,332,436	2,177,462	2,059,588	1,967,966	1,901,559	1,834,963	1,780,569	1,728,064
R^2	0.040	0.046	0.061	0.066	0.072	0.077	0.081	0.081	0.076	0.076	0.075

Note: This table presents results from estimating the same specification as in Panel C of Table 3 for two subsets of the data: households with no mortgage debt in 2001 (Panel A) and households with positive mortgage debt in 2001 (Panel B).

APPENDIX TABLE A9-1: ROBUSTNESS TO ADDITIONAL INTERACTIONS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: Includes Interaction of Rank with Rate of Homeownership						
α	-0.980***	-1.368***	-1.767***	-1.951***	-2.115***	-2.107***	-2.005***	-2.095***	-1.885***	-1.692***	-1.552***
	(0.0249)	(0.0356)	(0.0462)	(0.0578)	(0.0694)	(0.0762)	(0.0771)	(0.0769)	(0.0733)	(0.0682)	(0.0643)
β	-0.259***	-0.311***	-0.406***	-0.434***	-0.487***	-0.486***	-0.442***	-0.444***	-0.385***	-0.317***	-0.272***
	(0.00819)	(0.0118)	(0.0155)	(0.0190)	(0.0228)	(0.0255)	(0.0256)	(0.0258)	(0.0243)	(0.0226)	(0.0211)
γ	0.516***	0.683***	1.022***	1.186***	1.364***	1.403***	1.337***	1.360***	1.214***	1.056***	0.906***
	(0.0143)	(0.0205)	(0.0264)	(0.0330)	(0.0396)	(0.0435)	(0.0435)	(0.0439)	(0.0416)	(0.0383)	(0.0359)
N	5,727,356	5,257,066	4,658,759	4,221,379	3,899,085	3,648,535	3,449,008	3,308,587	3,168,380	3,052,691	2,947,893
R^2	0.051	0.063	0.070	0.078	0.083	0.088	0.097	0.100	0.106	0.116	0.126
					Panel B: Includes Interaction of Rank with Fraction of Black Residents						
α	-1.514***	-2.294***	-3.284***	-3.795***	-4.335***	-4.514***	-4.405***	-4.366***	-3.995***	-3.586***	-3.197***
	(0.0220)	(0.0316)	(0.0418)	(0.0518)	(0.0615)	(0.0670)	(0.0668)	(0.0666)	(0.0630)	(0.0584)	(0.0538)
β	-0.374***	-0.474***	-0.704***	-0.794***	-0.915***	-0.948***	-0.901***	-0.881***	-0.786***	-0.677***	-0.578***
	(0.00863)	(0.0119)	(0.0164)	(0.0201)	(0.0239)	(0.0264)	(0.0264)	(0.0263)	(0.0247)	(0.0229)	(0.0210)
γ	0.660***	0.943***	1.448***	1.709***	1.992***	2.081***	2.011***	1.994***	1.801***	1.582***	1.363***
	(0.0147)	(0.0213)	(0.0283)	(0.0353)	(0.0421)	(0.0460)	(0.0457)	(0.0456)	(0.0432)	(0.0400)	(0.0367)
N	5,727,471	5,257,165	4,658,826	4,221,433	3,899,132	3,648,580	3,449,048	3,308,627	3,168,414	3,052,725	2,947,921
R^2	0.050	0.063	0.069	0.076	0.081	0.086	0.095	0.098	0.104	0.114	0.125

Note: This table augments the specification in Panel C of Table 3 of the main text by adding the level of the listed variable and its interaction with rank. Panel A includes the fraction of residents in a zipcode who own their home calculated from the Census. Panel B includes the fraction of residents who identify as black calculated from the Census.

APPENDIX TABLE A9-2: ROBUSTNESS TO ADDITIONAL INTERACTIONS

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
						Panel C: Includes Interaction of Rank with House Quality Dispersion					
α	-1.617***	-2.488***	-3.554***	-4.125***	-4.777***	-4.982***	-4.888***	-4.828***	-4.429***	-3.972***	-3.544***
	(0.0285)	(0.0420)	(0.0564)	(0.0697)	(0.0831)	(0.0905)	(0.0909)	(0.0910)	(0.0871)	(0.0803)	(0.0742)
β	-0.395***	-0.485***	-0.762***	-0.861***	-1.004***	-1.053***	-1.021***	-0.978***	-0.864***	-0.748***	-0.645***
	(0.0117)	(0.0167)	(0.0230)	(0.0283)	(0.0347)	(0.0382)	(0.0381)	(0.0383)	(0.0359)	(0.0334)	(0.0306)
γ	0.727***	1.016***	1.570***	1.843***	2.162***	2.278***	2.219***	2.155***	1.943***	1.708***	1.493***
	(0.0196)	(0.0294)	(0.0398)	(0.0499)	(0.0600)	(0.0662)	(0.0668)	(0.0667)	(0.0633)	(0.0585)	(0.0539)
N	3,134,287	2,866,480	2,531,193	2,286,429	2,109,396	1,974,580	1,867,883	1,791,116	1,715,264	1,653,681	1,597,314
R^2	0.052	0.064	0.070	0.078	0.082	0.088	0.098	0.100	0.106	0.115	0.125
						Panel D: Includes Interaction of Rank with County-Level Crime Rate					
α	-1.506***	-2.269***	-3.264***	-3.774***	-4.321***	-4.497***	-4.402***	-4.363***	-3.992***	-3.580***	-3.186***
	(0.0221)	(0.0317)	(0.0419)	(0.0517)	(0.0615)	(0.0668)	(0.0668)	(0.0665)	(0.0629)	(0.0582)	(0.0535)
β	-0.373***	-0.472***	-0.701***	-0.792***	-0.915***	-0.946***	-0.905***	-0.883***	-0.794***	-0.685***	-0.581***
	(0.00870)	(0.0120)	(0.0164)	(0.0201)	(0.0240)	(0.0264)	(0.0265)	(0.0264)	(0.0247)	(0.0229)	(0.0209)
γ	0.661***	0.945***	1.451***	1.707***	1.993***	2.076***	2.014***	1.995***	1.810***	1.592***	1.373***
	(0.0148)	(0.0215)	(0.0285)	(0.0353)	(0.0423)	(0.0462)	(0.0462)	(0.0461)	(0.0435)	(0.0403)	(0.0368)
N	5,712,121	5,243,998	4,648,163	4,212,602	3,892,093	3,642,926	3,444,118	3,304,200	3,164,169	3,048,826	2,944,256
R^2	0.050	0.063	0.069	0.076	0.081	0.087	0.095	0.098	0.105	0.115	0.126

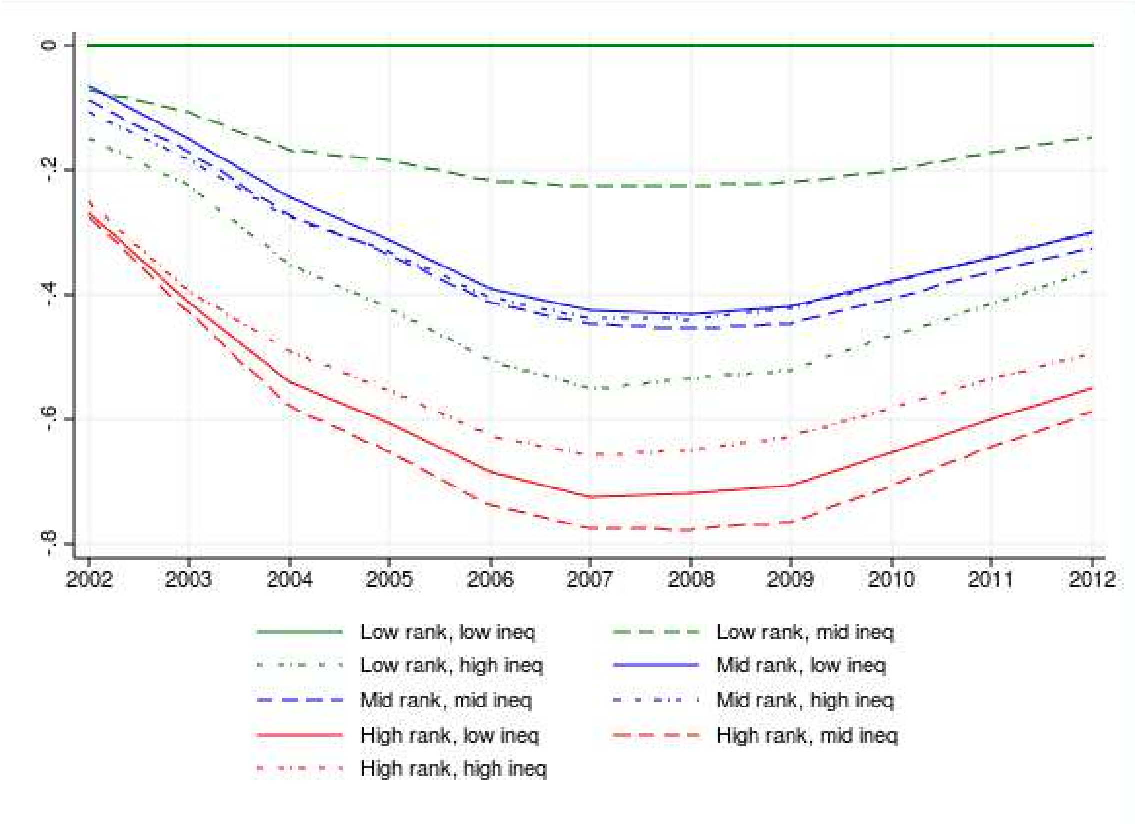
Note: This table augments the specification in Panel C of Table 3 of the main text by adding the level of the listed variable and its interaction with rank. Panel C includes the log of the ratio of average house prices in the top and bottom third of the price distribution as calculated by Zillow. Panel B includes the crime rate (reported crimes) as reported in the Uniform Crime Reporting Statistics at the county level.

Appendix Table A10: Mortgage Applications and Local Inequality with State FE

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
					Panel A: Probability of Mortgage Application Being Rejected						
α	-0.087***	-0.068***	-0.057***	-0.056***	-0.058***	-0.046***	-0.039***	-0.043***	-0.041***	-0.057***	-0.062***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
β	0.412***	0.354***	0.324***	0.376***	0.351***	0.356***	0.300***	0.244***	0.218***	0.250***	0.300***
	(0.048)	(0.037)	(0.027)	(0.026)	(0.024)	(0.024)	(0.024)	(0.023)	(0.021)	(0.028)	(0.030)
γ	-0.415***	-0.343***	-0.290***	-0.368***	-0.331***	-0.330***	-0.255***	-0.184***	-0.204***	-0.278***	-0.389***
	(0.083)	(0.070)	(0.053)	(0.035)	(0.033)	(0.036)	(0.036)	(0.039)	(0.031)	(0.036)	(0.042)
N	2,244,576	2,264,842	2,520,425	2,635,465	2,970,262	2,663,236	1,921,810	1,319,589	1,240,372	1,275,372	1,196,404
R^2	0.088	0.071	0.051	0.047	0.045	0.046	0.047	0.035	0.029	0.037	0.046
					Panel B: Probability of Mortgage Being High-Interest (conditional on origination)						
α				-0.039***	-0.063***	-0.052***	-0.038***	-0.046***	-0.025***	-0.023***	-0.029***
				(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
β				0.238***	0.284***	0.274***	0.196***	0.143***	0.079***	0.108***	0.119***
				(0.025)	(0.036)	(0.035)	(0.024)	(0.024)	(0.014)	(0.013)	(0.017)
γ				-0.201***	-0.264***	-0.247***	-0.189***	-0.130***	-0.072***	-0.101***	-0.124***
				(0.026)	(0.038)	(0.039)	(0.030)	(0.027)	(0.014)	(0.015)	(0.021)
N				1,995,005	2,148,955	1,892,164	1,384,324	959,930	944,620	955,348	894,997
R^2				0.098	0.159	0.123	0.063	0.046	0.027	0.042	0.043
					Panel C: Loan-to-Income Ratios of Mortgage Applications (conditional on origination)						
α	-0.165***	-0.174***	-0.185***	-0.175***	-0.166***	-0.168***	-0.183***	-0.184***	-0.190***	-0.193***	-0.187***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
β	-0.247***	-0.237***	-0.265***	-0.290***	-0.250***	-0.254***	-0.294***	-0.238***	-0.133***	-0.151***	-0.126***
	(0.046)	(0.046)	(0.047)	(0.045)	(0.040)	(0.037)	(0.040)	(0.047)	(0.047)	(0.044)	(0.046)
γ	0.114*	0.090	0.147**	0.173***	0.112**	0.111**	0.181***	0.108**	0.030	0.109**	0.074
	(0.066)	(0.066)	(0.066)	(0.052)	(0.046)	(0.047)	(0.046)	(0.053)	(0.054)	(0.056)	(0.059)
N	1,746,160	1,794,892	1,971,148	1,995,005	2,148,955	1,892,164	1,384,324	959,930	944,620	955,348	894,997
R^2	0.299	0.323	0.346	0.328	0.317	0.329	0.345	0.350	0.371	0.380	0.364

Note: The table replicates the results in Table 9 using state fixed effects rather than county fixed effects. Coefficient α corresponds to the partial correlation of applicant's income rank and the dependent variable in the year indicated by each column. Coefficient β corresponds to local inequality. Coefficient γ corresponds to the interaction of local inequality and applicant's income rank. Standard errors are clustered at the county level and each regression includes a state fixed effect as well as controls for race, sex, occupancy, the LTI, and an interaction of rank with the fraction of non-white applicants. The sample is restricted to home purchase loans with an LTI between 1 and 8 and where the application was not rejected by the borrower or failed for a reason other than denial. Statistical significance at the 1%, 5%, and 10% levels are indicated by ***, **, and * respectively.

APPENDIX FIGURE 1 DEBT ACCUMULATION BY LOW, MEDIUM AND HIGH-RANK HOUSEHOLDS AND LOCAL INEQUALITY, NONPARAMETRIC SPECIFICATION



Note: The figure shows the full set of estimated coefficients on the income rank dummies from the nonparametric regressions of the relative household debt accumulation between 2001 and year t. Each regression contains dummies for income ranks and inequality levels (with low-rank households in low-inequality regions being the benchmark), and a full set of controls described in equation (3) and the county-specific fixed effects. See section 3.4 for details.

APPENDIX B: ADDITIONAL INFORMATION ON CCP DATA

The Equifax FRBNY Consumer Credit Panel is a longitudinal database with detailed information on consumer debt and credit. The core of the database constitutes a 5% random sample of all U.S. individuals with credit (i.e., the primary sample). The database also contains information on all individuals with credit files residing in the same household as the individuals in the primary sample. The household members are added to the sample based on the mailing address in the existing credit files. Thus, the resulting sample is a sample of U.S. households in which at least one member has a credit file.

The individual records in the CCP contain information on the mortgage debt, credit card debt and credit card limits, home equity lines of credit, student loans, auto loans, bankruptcy and delinquencies. The data include residential location on the census block level and the birth year of individuals. The data in the CCP are updated quarterly. We use 100% of the CCP sample.

The unit of the analysis in the paper is a household. The CCP is primarily an individual-level dataset; however, it contains two identifiers that allow us to construct the household records in each period and then link the household records from period to period. In each quarter, a unique (household) identifier is given for all individuals who reside in the same household as an individual in the primary sample. We use this identifier to aggregate the individual level information to construct the household level credit variables. We restrict the analysis to households with at most 10 members.

The household identifier identifies household members only in one period. We then use the second identifier in the CCP data, an individual identifier that remains constant from period to period, to link household records from one quarter to another. To construct the longitudinal household record, we proceed as follows. Let i denote the identification number of a household in 2001. To identify the continuation of household i in year t, t > 2001, we first determine what members of household i are present in year t using individual identifiers. We then determine the identification number of the household to which each member of household i belongs to in year t. If there is more than one such household, we flag the modal household, if one exists. Let j denote this modal household. We then repeat the procedure in reverse: consider all members of household j who are present in year t and determine what members of household j are present in year 2001 using individual identifiers, determine the identification number of the household to which each member of household j belongs to in year 2001. If there is more than one such household, we flag the modal household, if one exists. Let i' denote this modal household. If i' equals i', we identify j as a continuation record for household i. While the primary sample of individuals in the CCP is a random sample of all U.S. households with credit reports; the resulting sample of the households is not random. Following, Lee and van der Klaauw (2010) we define the sampling weights as the inverse of the probability to be included in the sample, $w_h = \frac{1}{1-.05^N}$, where N is the number of individuals in the household who are in the primary sample.

For each individual, the data contain a record of her debt by detailed category as well as a record of the balances on the joint or cosigned accounts. In aggregating the debt on the household level, we use a correction to avoid double counting of the balances on joint accounts. This choice follows Brown, Haughwout, Lee and van der Klaauw (2011). In particular, while aggregating, we discount the total debt of the household members by 50% of the total debt on joint accounts of the household members. The exact formula that we use is

$$d_{h,j} = \max\{ \textstyle\sum_i (d^i_{h,j} - .5d^{i,c}_{h,j}), .5d^{i,c}_{h,j}\}.$$

Where $d^i_{h,j}$ is the total debt in category j of member i in household h and $d^{i,c}_{h,j}$ is the debt in joint accounts. The second input to the maximum function addresses the situation that arises with so-called "thin" credit records, or records with at most two credit report-worthy debts. The individuals with thin records are not included in the primary sample, but they are included in the additional sample. These individuals might have records on joint accounts that are missed on individual accounts. We thank Donghoon Lee for this suggestion.

Variable Descriptions

Here we provide a short description of the variables used in the CCP analysis. For a detailed description of the CCP dataset please see Lee and van der Klaauw (2010).

Age: We follow Brown, Haughwout, Lee, and van der Klaauw (2011) and define age as the median age of adult members of the house.

Auto debts: These are any loans taken out explicitly for the purchase of a car including loans from banks and those from automobile financing institutions.

Bankruptcy: An indicator in the CCP taken from public records that detail whether or not an individual has filed for bankruptcy.

Credit Card Balance: The sum of reported balances across bank cards as well as retail cards. These cards reflect revolving accounts at banks, credit unions, credit card companies, and others. Importantly, the CCP does not distinguish between balances rolled over billing periods (and so potentially subject to interest charges) and cards where the balance is paid every month.

Credit Card Limits: We take the maximum of reported limits and balances across all bank and retail cards to ensure that reported utilization is not greater than one.

Credit Card Utilization Rate: This is the ratio of the credit card balance and credit card limit.

Delinquency: Indicator for whether or not a household is at least 60 days delinquent on any of its accounts in the current quarter.

HELOC Debt: The sum of home equity lines of credit, or home equity revolving accounts. We use the classification of HELOCs vs. installment loans provided by the CCP data.

Mortgage Debt: The sum of all mortgage installment loans.

Riskscore: A variable constructed by Equifax and similar to FICO. A higher number is interpreted as a lower default risk. We construct the household riskscore by taking the average of individual riskscores within the household.

Size: Household size sums the number of distinct social security numbers that can be linked by household identifiers in a specific time period. We restrict the household size to at most 10.

Student Loans: These include loans financing education from private and public institutions.

Total debt: Constructed as the sum of mortgage debt balance, credit card balances, auto debts, balance on home equity lines of credit, and student loans.

Appendix C: Decomposing U.S. Inequality since 1970

The decomposition is constructed using the following IPUMS samples: 1970, 1980, 1% metro samples and the 1990 and 2000 1% unweighted sample. Within each of these samples we use the metro area geographies defined by IPUMS in the following way:

> "Metropolitan areas are counties or combinations of counties centering on a substantial urban area. METAREA identifies the metropolitan area where the household was enumerated, if that metropolitan area was large enough to meet confidentiality requirements."

We restrict the sample to the set of metro areas that can be identified in each year to get 117 metro areas containing roughly 60% of the entire sample within each year. We also restrict the sample to households where the respondent's age is between 25 and 65 and the respondent is the head of the household or the spouse of the head of the household. These restrictions are not important for the results.

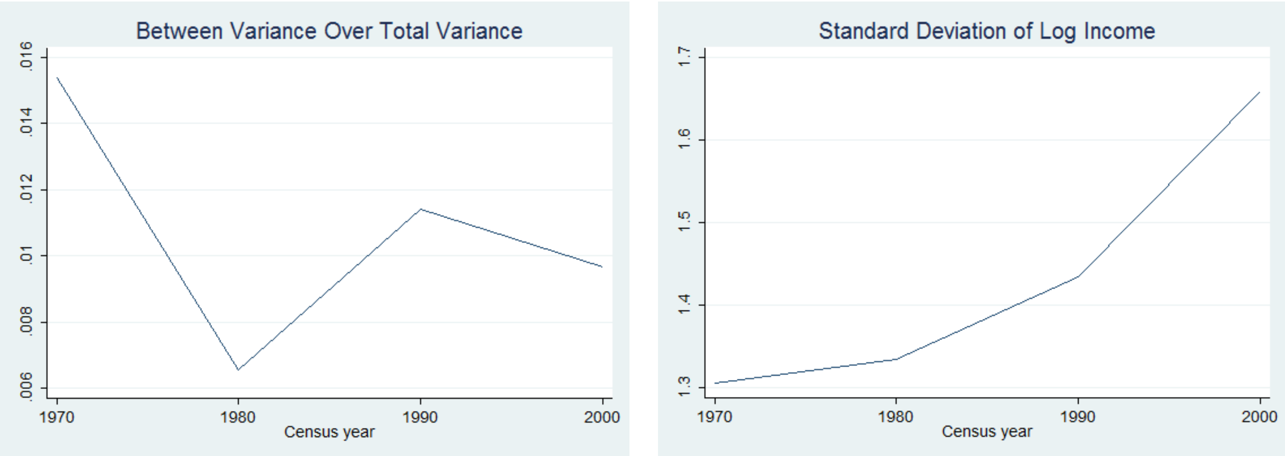
To calculate income we use family total income. While not exactly the same as household income it is available for all years whereas household income is not available in 1970. We estimate the following model of log family income on each year of the sample:

$$\log(y_{ia}) = \alpha_a + \epsilon_i$$

Estimating this function gives estimates of the variance of the fixed effects and the variance of the residuals for each year. We then calculate the share of variance explained by variance of the fixed effects as:

$$Share = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_i^2}$$

Appendix Figure C1: Decomposing Aggregate U.S. Inequality



Note: The left-hand figure plots the ratio of "between" variance of mean incomes to the total variance of incomes. The right-hand figure plots the standard deviation of log income across all households.

APPENDIX D: TIME VARIATION IN LOCAL INEQUALITY RATES

To get a sense of how inequality within counties has varied across time we computed Gini coefficients at the county level using 1970 and 2000 Census aggregates available from ICPSR. To compute the Gini coefficient we follow the same procedure outlined in the Appendix and reproduced below. Because the number of bins used to compute the coefficient is not the same in both years (1970 has fewer bins) the levels of the Gini coefficients are not directly comparable. Using the Census data we match 3,122 counties.

Let $f(y_i)$ be a discrete probability function where $i = 1, \dots, n$ and $y_i < y_{i+1}$. Then the Gini coefficient G is defined as

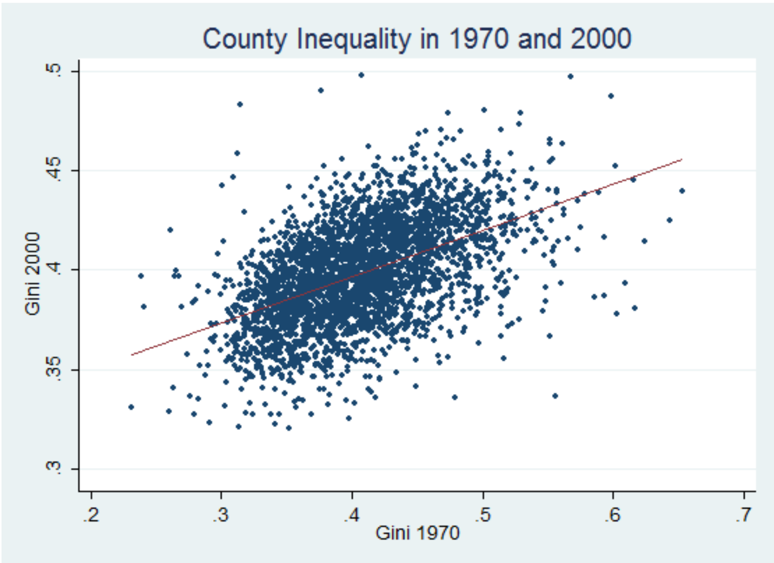
$$G = 1 - \frac{\sum_{i=1}^{n} f(y_i)(S_{i-1} + S_i)}{S_n}$$

where $S_i = \sum_{j=1}^{i} f(y_j) y_j$ and $S_0 = 0$.

We approximate the discrete probability function with the share of a location's population within each bin reported by the Census. For all bins but the last we assume all the mass is distributed at the midpoint of the bin. For the very last bin we add the last increment to the lower boundary. For example, if the last bin is incomes of $200,000 and up and the bin before was $150,000 to $199,999 we assign the last bin to have the value $250,000. This assumption limits the impact the very top bin will have on the coefficient, but should provide a reasonable approximation of inequality at low levels of aggregation.

The figure reported below shows a high degree of correlation between inequality in 1970 and inequality in 2000. The R-squared is 0.26 and the Spearman correlation is 0.52, suggesting inequality is quite persistent.

APPENDIX FIGURE D1: PERSISTENCE OF LOCAL INEQUALITY



Note: The figure plots Gini coefficients for income inequality in U.S counties in 1970 versus 2000.

APPENDIX E: SUMMARY STATISTICS FROM HMDA DATA

Table 1 in this appendix provides summary statistics from the 15% HMDA samples. We report the fraction of applications denied, originated, for owner-occupied properties, high interest, the race of the primary applicant, and the regulator of the lender. When using the HMDA data it is important to recognize that changes in reporting requirements from 2003 to 2004 had significant effects on the coverage of the mortgage market and so statistics we calculate. This can be seen clearly when comparing the change in racial composition of applicants from 2003 to 2004. While some of this might reflect real shifts in the provision of credit to non-white groups it also reflects the increased coverage of rural areas and smaller, non-bank lenders. This can also be seen by the large increase in applications filed at lenders regulated by HUD. While mortgage company activity was almost certainly increasing over this period many lenders were simply not reporting in the HMDA data.

The health of the mortgage market can be traced out by changes in the sample size. The number of applications reported peaked in 2007 and then declined steadily until 2011. Interestingly, the fraction of loans with high interest rates has also declined sharply, probably reflecting fewer loans with junior liens.

Notice that the mean applicant income reported in the HMDA data is substantially higher than the average household income reported in the SCF data and the imputed CCP data. However, average income is comparable to the average income of homeowners as reported in the 2007 SCF, which is about $99,500.

Table 2 provides some sample correlations from 2007, most of which are qualitatively similar to other years. Owner-occupied applications are less likely to be denied while applications with high LTI ratios are more likely to be denied. Applicants applying to HUD-regulated lenders are more likely to be denied, which could reflect the stress of mortgage companies in this period or an increased likelihood that the applicant is subprime. Applicants to HUD lenders tend to have smaller incomes and higher LTI ratios.

APPENDIX TABLE E1: SUMMARY STATISTICS FROM HMDA

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011
Denied	0.15	0.13	0.13	0.15	0.16	0.18	0.18	0.17	0.14	0.15	0.15
Originated	0.78	0.79	0.78	0.76	0.72	0.71	0.72	0.73	0.76	0.75	0.75
OOC	0.94	0.93	0.92	0.90	0.88	0.90	0.91	0.92	0.94	0.94	0.93
LTI	2.31	2.43	2.58	2.65	2.67	2.63	2.72	2.72	2.81	2.79	2.70
sd	0.88	0.94	1.03	1.08	1.08	1.04	1.11	1.10	1.12	1.12	1.10
Loan	140.16	154.40	168.24	193.11	212.85	223.00	226.41	207.03	198.34	203.31	200.69
sd	96.03	104.30	111.90	147.30	165.15	173.16	180.86	155.68	141.21	148.88	151.88
Income	64.84	68.46	70.72	78.13	85.41	91.21	91.01	84.15	78.02	80.84	82.38
sd	47.46	49.75	50.95	63.29	70.48	76.46	81.55	73.44	65.42	68.73	71.28
High Int				0.08	0.16	0.16	0.08	0.06	0.04	0.02	0.03
White	0.89	0.88	0.88	0.74	0.71	0.69	0.73	0.76	0.76	0.76	0.77
Black	0.08	0.07	0.08	0.08	0.10	0.11	0.10	0.08	0.07	0.07	0.07
OCC	0.28	0.27	0.26	0.23	0.20	0.23	0.32	0.32	0.29	0.31	0.06
FRS	0.11	0.18	0.18	0.15	0.16	0.16	0.17	0.09	0.09	0.08	0.04
FDIC	0.09	0.08	0.07	0.07	0.06	0.06	0.06	0.09	0.11	0.11	0.09
OTS	0.11	0.10	0.09	0.08	0.09	0.08	0.10	0.08	0.06	0.05	0.00
NCUA	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.05	0.04	0.04	0.04
HUD	0.39	0.36	0.38	0.45	0.47	0.45	0.33	0.36	0.40	0.41	0.43
N	644680	647685	722326	790699	890889	798332	577110	395574	371967	382851	359100

Note: The table provides sample means for all variables and standard deviations for continuous variables for all years of the HMDA data under the sample restrictions identified in the text. Denied gives the probability that an application was formally denied while originated gives the probability a loan was approved and the funds disbursed to the borrower. OOC indicates that the application is for an owner-occupied home. LTI is the loan-to-income ratio on the application constructed from the application's stated loan and income. High Int indicates if a loan was ultimately originated as a high interest loan. While and black both refer to the race of the primary applicant. OCC indicates a loan filed at a lender regulated by the Office of the Comptroller of the Currency. Similarly, FRS indicates a lender regulated by the Federal Reserve System, OTS regulated by the Office of Thrift Supervision, NCUA the National Credit Union Administration, and HUD the Department of Housing and Urban Development.

APPENDIX TABLE E2: SAMPLE CORRELATIONS FROM 2007 HMDA

	Denied	Originated	OOC	LTI	Loan	Inc	White	Black
Denied	1.000							
Originated	-0.762***	1.000						
OOC	-0.0192***	0.021***	1.000					
LTI	0.053***	-0.060***	0.200***	1.000				
Loan	0.001	-0.020***	-0.0308***	0.208***	1.000			
Income	-0.028***	0.014***	-0.169***	-0.238***	0.815***	1.000		
White	-0.145***	0.146***	-0.0105***	-0.116***	-0.033***	0.034***	1.000	
Black	0.116***	-0.113***	0.007***	0.050***	-0.053***	-0.074***	-0.545***	1.000
OCC	-0.066***	0.120***	-0.005***	-0.012***	0.056***	0.063***	0.006***	-0.025***
FRS	0.051***	-0.070***	-0.002	-0.022***	-0.023***	-0.011***	0.001	0.004**
FDIC	-0.044***	0.045***	-0.031***	-0.031***	-0.060***	-0.041***	0.078***	-0.037***
OTS	0.0547***	-0.009***	-0.022***	-0.003*	0.081***	0.070***	-0.027***	0.006***
NCUA	-0.025***	0.008***	0.029***	-0.004**	-0.042***	-0.040***	0.039***	-0.020***
HUD	0.022***	-0.084***	0.026***	0.048***	-0.042***	-0.062***	-0.044***	0.044***
N	577110							

Note: The table provides correlations for all years of the HMDA data under the sample restrictions identified in the text. Denied gives the probability that an application was formally denied while originated gives the probability a loan was approved and the funds disbursed to the borrower. OOC indicates that the application is for an owner-occupied home. LTI is the loan-to-income ratio on the application constructed from the application's stated loan and income. High Int indicates if a loan was ultimately originated as a high interest loan. White and black both refer to the race of the primary applicant. OCC indicates a loan filed at a lender regulated by the Office of the Comptroller of the Currency. Similarly, FRS indicates a lender regulated by the Federal Reserve System, OTS regulated by the Office of Thrift Supervision, NCUA the National Credit Union Administration, and HUD the Department of Housing and Urban Development.

APPENDIX F: INCOME AND DEFAULT

We use the CCP data to verify our assumption about probability of default conditional on income. In particular, we estimate a linear probability model of the probability of default as a function of household income.

The dependent variable takes value 1 if any member of the household in year *t* is 60-day past due or longer on any account (mortgage, auto loan, credit card, etc.). The explanatory variable of interest is the (log of the) household income in year 2001 (using the expected imputed income). We first estimate a parsimonious specification with only the income measure. We then estimate a specification with the measure of income and the full set of household and regional controls. These household-level controls are the following variables measured at 2001: dummies for age of the head of household and for the size of the household; amount of mortgage, auto loan, credit card balance, credit card limit, HELOC, student loan; dummies for bankruptcy and 60 DPD or longer, and risk score. The regional-level controls are the following zip code-level variables measured in 2001: income inequality, median of total household debt, median of household mortgage, house price growth between 2001 and year *t*, the ratio of the median house price to the median income, and the county level fixed effects. In the estimation, the standard errors are clustered by zip code. We use a linear probability model since the mean of the dependent variable is in the range 0.25-0.30. The equation is estimated for each year from 2002 to 2012 for the sample of the households use in the benchmark regression of our analysis (i.e., the households that do not change location between year 2001 and year *t*).

We report results in Appendix Table E1. We find that higher-income households and households with higher income ranks have lower probability of default.

Appendix Table F1. Income and default.

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
					Panel A: No Controls						
rank	-0.387***	-0.337***	-0.347***	-0.314***	-0.294***	-0.264***	-0.244***	-0.219***	-0.206***	-0.199***	-0.205***
	(0.00181)	(0.00184)	(0.00186)	(0.00182)	(0.00179)	(0.00179)	(0.00180)	(0.00183)	(0.00185)	(0.00188)	(0.00191)
N	6,172,512	5,676,766	5,039,109	4,570,211	4,218,948	3,950,618	3,731,267	3,581,280	3,433,201	3,310,773	3,197,351
R^2	0.029	0.022	0.023	0.019	0.017	0.014	0.012	0.010	0.008	0.008	0.008
					Panel B: County Fixed Effects						
rank	-0.385***	-0.335***	-0.345***	-0.312***	-0.293***	-0.263***	-0.245***	-0.220***	-0.208***	-0.201***	-0.208***
	(0.00184)	(0.00186)	(0.00189)	(0.00184)	(0.00179)	(0.00178)	(0.00177)	(0.00180)	(0.00180)	(0.00183)	(0.00186)
N	6,172,512	5,676,766	5,039,109	4,570,211	4,218,948	3,950,618	3,731,267	3,581,280	3,433,201	3,310,773	3,197,351
R^2	0.058	0.051	0.055	0.051	0.048	0.043	0.039	0.035	0.034	0.033	0.033
					Panel C: Household-specific Characteristics and County Fixed Effects						
rank	-0.0381***	-0.0422***	-0.0443***	-0.0489***	-0.0521***	-0.0458***	-0.0288***	-0.0083***	0.00125	0.00724***	0.0146***
	(0.00168)	(0.00189)	(0.00209)	(0.00221)	(0.00230)	(0.00245)	(0.00251)	(0.00260)	(0.00268)	(0.00270)	(0.00276)
N	4,195,007	3,836,566	3,380,052	3,047,381	2,803,886	2,619,591	2,470,908	2,367,350	2,265,545	2,182,951	2,105,700
R^2	0.460	0.359	0.326	0.274	0.244	0.213	0.187	0.177	0.171	0.161	0.159
					Panel D: No Controls						
ln(y)	-0.163***	-0.149***	-0.157***	-0.147***	-0.142***	-0.131***	-0.122***	-0.111***	-0.105***	-0.102***	-0.105***
	(0.000620)	(0.000600)	(0.000621)	(0.000627)	(0.000634)	(0.000632)	(0.000649)	(0.000675)	(0.000697)	(0.000709)	(0.000730)
N	6,172,512	5,676,766	5,039,109	4,570,211	4,218,948	3,950,618	3,731,267	3,581,280	3,433,201	3,310,773	3,197,351
R^2	0.049	0.041	0.045	0.041	0.038	0.033	0.029	0.023	0.021	0.020	0.021
					Panel E: County Fixed Effects						
ln(y)	-0.152***	-0.136***	-0.143***	-0.133***	-0.127***	-0.117***	-0.111***	-0.102***	-0.0972***	-0.0943***	-0.0977***
	(0.000625)	(0.000616)	(0.000633)	(0.000626)	(0.000619)	(0.000611)	(0.000615)	(0.000632)	(0.000635)	(0.000640)	(0.000654)
N	6,172,512	5,676,766	5,039,109	4,570,211	4,218,948	3,950,618	3,731,267	3,581,280	3,433,201	3,310,773	3,197,351
R^2	0.070	0.062	0.067	0.062	0.059	0.053	0.049	0.043	0.042	0.040	0.041
					Panel F: Household-specific Characteristics and County Fixed Effects						
ln(y)	-0.0107***	-0.0115***	-0.0128***	-0.0147***	-0.0161***	-0.0138***	-0.0081***	-0.00102	0.00211**	0.00425***	0.00649***
	(0.000599)	(0.000676)	(0.000742)	(0.000789)	(0.000820)	(0.000873)	(0.000895)	(0.000936)	(0.000966)	(0.000974)	(0.00100)
N	4,195,007	3,836,566	3,380,052	3,047,381	2,803,886	2,619,591	2,470,908	2,367,350	2,265,545	2,182,951	2,105,700
R^2	0.460	0.359	0.326	0.274	0.244	0.213	0.187	0.177	0.171	0.161	0.159

Note: The table reports estimated coefficients on income rank (Panels A-C) and log income (Panels D-F) in the linear regression where the dependent variable is a dummy variable equal to one if a household defaults in a given year and zero otherwise. Standard errors (clustered by zip code) are reported in parentheses. ***,**,*denote statistical significance at 1%, 5% and 10%.

Appendix G: Imputation of Income

In the first step of our work, we estimate the relationship between income and observables in the SCF and then use this relationship to impute income in the CCP. In this appendix, we describe how variables are constructed and what specification is estimated.

In the table below, we describe how variables are constructed in CCP and SCF. We use only variables which are available in both CCP and SCF. While there are some differences in the definitions across datasets, we made every effort to make it as comparable as possible.

Variable	SCF	Counterpart in CCP
Auto loans	X2218 + X2318 + X2418 + X7169 + X2424 + X2507 + X2607	Auto loan bank and auto loan finance balance
Bankruptcy flag	X6772	Chapter 7 or Chapter 13 bankruptcy flag
Credit Card Limit[22]	X414	Bank card + retail card high credit
Credit Card Balance	X413 + X427+ X421 + X424 + X430	Bank card + retail card balance
Delinquency flag	X3005	A flag if any account is 60 DPD or more
HELOC Balance	X1108 + X1119 + X1130 + X1136	Home equity revolving balance
Income	X5729	None
Mortgage Debt	X805 + X905 + X1005	First mortgage balance + home equity installment balance
Student Loans	X7824 + X7847 + X7870 + X7924 + X7947 + X7970	Student loans balance

We also use household size and head of household age. The CCP does not include racial identifiers so we do not use these. In our imputation, we use all of the SCF replicates, which are discussed in detail by Kennickell (1998). Because the SCF intentionally oversamples wealthy households, we apply the SCF-computed weights X42001. Note that we take the natural log of one plus the level for all continuous variables to make the distribution of these variables more well-behaved and to avoid dropping observations with zero values. We also restrict the sample to households where the head's age was between 20 and 65. We dropped outliers using Cook's distance.

As discussed in the text, our regression has the general form

$$\log(Y_{i,SCF}) = \beta f(\, X_{i,SCF}) + \; \epsilon_{i,SCF}.$$

In choosing the specific form of *f*, we aimed to capture as much of joint distribution of the observables and income as we could with a flexible assumption. Terms were added if it was found that they were meaningful

[22] We code responses of "no limit" in the SCF as 1,000,000.

predictors of log income. Households with missing values are dropped, although results are essentially the same if we keep them and add one before taking logs. The function *f* was composed of

1. Third-order Chebychev polynomials of mortgage, auto, and credit card limits,
2. Credit card, HELOC, and student loan balances,
3. Nine age bins in five year intervals,
4. Interactions of all age bins with each type of debt balance,
5. Household size and interactions of household size with debt balances and age bins,
6. Indicators for bankruptcy and delinquency and interactions of these indicators with other indicators,
7. Indicators for positive credit card limit and interactions of this variable with various variables,
8. Interactions of household size, age, and debt levels.

Table 2 shows that using data from 2001 the aggregate income statistics computed directly from the SCF match those we impute in the CCP very closely.