NBER WORKING PAPER SERIES

THE VULNERABILITY OF MINORITY HOMEOWNERS IN THE HOUSING BOOM AND BUST

Patrick Bayer
Fernando Ferreira
Stephen L. Ross

Working Paper 19020
http://www.nber.org/papers/w19020

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2013

This paper was conducted with the financial support of the Ford Foundation, Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton, and the Center for Real Estate and Urban Economic Studies at the University of Connecticut. The support was mainly used to purchase a proprietary data set from DataQuick, and also to buy credit scores and credit outcomes information provided by Experian Information Solutions, Inc. Experian is a service mark and registered trademark of Experian Information Solutions, Inc. The substantive content of the paper is the responsibility of the authors and does not reflect the specific views of any credit reporting agencies. Moreover, no conflicts of interest exist with respect to the research completed, the results obtained, and the sources of funding. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w19020.ack

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

The Vulnerability of Minority Homeowners in the Housing Boom and Bust
Patrick Bayer, Fernando Ferreira, and Stephen L. Ross
NBER Working Paper No. 19020
May 2013
JEL No. J15,R2

ABSTRACT

This paper examines mortgage outcomes for a large, representative sample of individual home purchases and refinances linked to credit scores in seven major US markets in the recent housing boom and bust. Among those with similar credit scores, black and Hispanic homeowners had much higher rates of delinquency and default in the downturn. These differences are not readily explained by the likelihood of receiving a subprime loan or by differential exposure to local shocks in the housing and labor market and are especially pronounced for loans originated near the peak of the boom. Our findings suggest that those black and Hispanic homeowners drawn into the market near the peak were especially vulnerable to adverse economic shocks and raise serious concerns about homeownership as a mechanism for reducing racial disparities in wealth.

Patrick Bayer
Department of Economics
Duke University
213 Social Sciences
Durham, NC 27708
and NBER
patrick.bayer@duke.edu

Fernando Ferreira
The Wharton School
University of Pennsylvania
1461 Steinberg - Dietrich Hall
3620 Locust Walk
Philadelphia, PA 19104-6302
and NBER
fferreir@wharton.upenn.edu

Stephen L. Ross
University of Connecticut
Department of Economics
341 Mansfield Road, Unit 1063
Storrs, CT 06269-1063
stephen.l.ross@uconn.edu

1. Introduction

Homeownership has long been viewed as an important mechanism for building wealth and, hence, the substantially lower ownership rates of minority households may be a serious impediment to reducing racial wealth disparities.[1,2] Motivated by this perspective, a number of public policy programs have an explicit goal of encouraging homeownership and many politicians have embraced it as a means of upwards mobility. President George W. Bush famously said in a 2004 speech that "*We're creating... an ownership society in this country, where more Americans than ever will be able to open up their door where they live and say, welcome to my house, welcome to my piece of property*". With this view in mind, the expansion of housing credit in the United States from late 1990s to mid-2000s was largely cheered[3] and homeownership by households of all races and ethnicities reached record high rates in the mid-2000s.[4]

As the subsequent housing and economic crises developed, however, the risks of homeownership became increasingly obvious. Delinquency and foreclosure rates rose sharply, especially in minority and low-income neighborhoods, and many households not only lost substantial housing wealth but also faced the prospect of lower credit scores (higher borrowing costs) for years to come. A comparison of mortgage delinquencies and foreclosures between 2005 and 2009 provides a particularly stark picture of the differential impact of the downturn by race. Figures 1 and 2 shows that while all homeowners had negligible 90-day delinquency and

[1] According to the Census, homeownership rates were 73.4 and 47.4 percent for whites and blacks, respectively, in 2000 and 74.5 and 45.6 in 2010. Moreover, there is evidence that black households historically hold a higher fraction of their wealth in functional assets, such as houses and vehicles. See Terrell (1971) and Oliver and Shapiro (1997).

[2] Purchasing a home often requires a substantial downpayment, and so savings decisions and wealth accumulation often play a substantial role in the decision to rent or own housing (Englehart, 1996 and Brueckner, 1986). Moreover, racial disparities in wealth dwarf those in income. For example, in 1998, the median wealth of black and Hispanic renters was less than $3,000 dollars and the 75th percentile wealth was below $10,000 (Herbert, Haurin, Rosenthal and Duda 2005). In 1988, Oliver and Shapiro (1997) estimate that the financial assets and net worth of black households nearing retirement (aged 50-64) were only 14 and 20 percent of those of white households, respectively, while the analogous figure for income was 67 percent. Also see Deng, Ross and Wachter (2003), Gyourko, Linneman, Wachter (1999), Duca and Rosenthal (1991, 1994), Wachter, S.M., Megbolugbe (1992), Linneman, Wachter, (1989) on the role of borrowing constraints in explaining the black-white homeownership gap.

[3] For more specific arguments in favor of relaxing downpayment constraints see Rohe, Van Zandt and MacCarthy (2002). Di and Liu (2007) document a reduction in the importance of household wealth in the transition to homeownership, and many other papers document efforts to relax downpayment constraints for low income borrowers, such as Belsky, Retsinas and Duda (2005), Herbert, Haurin, Rosenthal and Duda (2005), and Quercia, McCarthy and Wachter (2003).

[4] Homeownership rates peaked in 2004 Q2 (49.7 percent) for blacks and in 2007 Q1 (50.1 percent) for Hispanics.

default rates in 2005 for our sample of seven major markets, large racial differences had emerged by 2009. More than 1 in 10 black and Hispanic homeowners in our sample had a delinquent mortgage by 2009, compared to 1 in 25 for white households, and a similar pattern held for foreclosure rates. The differential impact of the downturn highlights a key concern with homeownership as a means for reducing racial wealth disparities. Namely, if the downside risks associated with owning a home are distributed unequally by race, increased rates of delinquency and default may ultimately exacerbate rather than diminish the racial wealth gap.

In this paper, we examine mortgage outcomes by race during the last housing cycle in a diverse set of U.S. housing markets. The main goal of our analysis is to distinguish among a number of potential explanations for the higher rates of delinquency and default by minority homeowners in the housing market bust, with the ultimate aim of providing a better understanding of the benefits and risks associated with homeownership as a vehicle for building wealth.

While researchers have documented the greater exposure of minority households to income and health shocks,[5] much less is known about the differential impact of credit and financial shocks, especially in housing markets.[6] The literature suggests that subprime lending has been an important factor in explaining rising foreclosure rates in low income and minority neighborhoods.[7] Here we take a different approach by proposing an important explanation for high rates of negative credit market outcomes for minority homeowners during the crisis, i.e., the selection of high-risk households into the housing market close to the peak of the housing cycle.

We explain this mechanism in the context of a simple model of credit markets in which borrowers are heterogeneous in both their risk of having adverse economic events and in their ability to manage adverse events should they occur. We show that an expansion in credit availability selects borrowers into the market that face a higher risk of adverse events, leading to an increase in default rates among borrowers unable to manage adverse events in any subsequent

[5] See, for example, Smith (1995), Altonji and Blank (1999), Shapiro (2004), and Shuey and Wilson (2008).
[6] Hoynes, Miller, and Schaller (2012) document dramatic demographic differences in labor market outcomes following recession shocks.
[7] That the housing market crisis and the preceding growth of subprime lending have disproportionately affected low income and minority neighborhoods has been documented in Gerardi and Willen (2009), Reid and Laderman (2009), Edminston (2009), and Wachter, Russo and Hershaff (2010). More broadly, Mian and Sufi (2009) document large increases in foreclosure rates in neighborhoods that had a large volume of subprime loans and that this growth in subprime loans occurred primarily in neighborhoods with very low income growth.

market downturn. To the extent that wealth and liquidity gaps leave minority households especially vulnerable to negative economic shocks, our model implies that those minority households drawn into homeownership following a major expansion of credit are especially likely to default in a subsequent downturn.[8]

There are significant empirical challenges to studying mortgage outcomes by race and ethnicity. Most directly, data sets linking home purchases and mortgage decisions by race to subsequent loan performance for a representative sample of homeowners have been essentially non-existent. A second complicating factor is that individual credit scores, which are important in explaining the credit decisions of lenders, the loan terms of individual borrowers, and the subsequent performance of loans, are generally unavailable in public data sets.[9]

We overcome these issues by first assembling a unique panel data set that links a representative sample of Home Mortgage Disclosure Act (HMDA) data on home purchase and refinance mortgages originated between May and August in the years between 2004 and 2008, to public records data on housing transactions and liens for approximately 270,000 homeowners in seven distinct metropolitan housing markets.[10] These data contain information on all liens as well as the name and address of the individual purchasing the housing unit or refinancing their mortgage and in many cases the name of the individual's spouse, in addition to information

[8] The evidence presented in Charles and Hurst (2002) suggests that black households may also have fewer family resources to draw on. In particular, they document that, conditional on credit score and income, young black adults are less likely to purchase homes due in part to the ability and willingness of parents to provide down payment assistance. Also, a large literature in economics has analyzed the extent of risk sharing in families and communities – see, for example, Cochrane (1991), Townsend (1994), Hayashi, Altonji, and Kotlikoff (1996), Amin, Rai, and Topa (2003), and Mazzocco (2004).

[9] Other commonly used data sets miss important subsets of the mortgage market. Most research documenting the distribution and performance of high cost or subprime loans have used one of three sources: publically available Home Mortgage Disclosure Act (HMDA) data, proprietary data aggregated from individual lenders, or data obtained directly from individual lenders. Analyses using Home Mortgage Discolusre Act data (Calem, Gillen and Wachter 2004; Avery, Canner and Cooke 2005; Apgar, Calder and Fauth 2004) typically find significant racial and ethnic differences in the allocation of high cost credit, but these studies can only control for the extremely limited borrower and loan attributes in HMDA. This limitation is addressed in part by studies that exploit proprietary data aggregated across many lenders (Mayer and Pence 2007; Mayer, Pence and Sherlund 2009; Reid and Laderman 2009; Fisher, Lambie-Hanson and Willen 2010), but these studies often capture only a select subset of the mortgage market. Further, these studies typically do not contain information on borrower race or ethnicity and have instead documented differences at the neighborhood level. Alternatively, pricing models based on lender specific data typically find at most very small differences the incidence of high cost mortgage credit between white and minority borrowers (Nelson 2005; Courshane 2007).

[10] Chicago IL CMSA, Cleveland OH MSA, Denver CO MSA, Los-Angeles CA CMSA, Miami-Palm Beach Corridor, San Francisco CA CMSA, and Washington DC-Baltimore MD suburban Corridor.

about house prices, housing characteristics, and neighborhood identifiers that allow us to create neighborhood sociodemographic characteristics based on Census data.[11]

We then provided this rich sample to one of the major credit rating agencies. The credit rating agency used the name and address to match borrowers to archival credit reporting data, providing in each year a Vantage credit score plus detailed credit line information from each individual's report. This matched mortgage-home-credit data provides detailed information on the financial circumstances of each homeowner at the time of mortgage origination, the terms of the mortgage including whether it was a high-cost loan, and the performance of the mortgage over the subsequent years.[12]

Our final panel data has two critical advantages relative to others that have been analyzed in the recent literature. First, all of the information is matched at the individual rather than neighborhood level. This allows us to examine how the pricing and performance of loans varies with the race of the homeowner as well as the racial composition of the neighborhood. Second, our data set includes a representative sample of all mortgages that were originated in these markets in the time period.

Our empirical results show that black and Hispanic households are more likely to become delinquent and default on their mortgages than white households with similar credit scores, house type, neighborhood, and loan characteristics, especially for mortgages originated for new home purchases in 2005-06. One prominent, potential explanation for higher delinquency and foreclosure rates is that minority borrowers were concentrated in the subprime sector of the market, where they faced higher interest rates and more onerous loan terms than white borrowers with equivalent credit history and circumstances. Since 2004, the Home Mortgage Disclosure Act data has contained an indicator for high cost (or rate-spread) loans that is often considered a proxy for subprime loans. Yet, while having a high-cost loan is a strong predictor of subsequent delinquency and default, controlling for this variable has only a minor impact on the estimated

[11] Information on subordinate liens is typically not available except in lender provided samples of mortgages because only individual loans are tracked not entire mortgage transactions. See Foote, Gerardi, Goette, and Willen (2010) for another example where HMDA is matched to housing transaction data in order to obtain information on subordinate liens.

[12] In parallel research efforts, Researchers in the Federal Reserve System have purchased large samples of credit repository data, e.g. Haughwout, Lee, Tracy and van der Klaaw (2011) and Avery, Brevoort, and Canner (2012), but to our knowledge this data has not been used to examine the relative experience of minority borrowers leading up to and during the crisis.

racial and ethnic differences in future credit market outcomes. These differences in delinquency and default are also relatively unaffected by the inclusion of lender and neighborhood fixed effects, and additional controls for the influence of subprime lending. Thus, strikingly, most of the observed differences in credit market outcomes for minority homeowners are not related to differential access to lenders, types of loans, or any observable factor that might have been used to price the mortgages in the first place.

We next consider whether racial and ethnic differences in loan performance might be attributable to differential exposure to the housing market collapse and associated recession. To capture shocks associated with both the labor and housing market, we include a series of controls that measure variation in the severity of the crisis: (i) tract and county by year fixed effects, (ii) individual indicators of a household's equity position in each year, (iii) the interaction of equity position and county by year unemployment rates, and (iv) race-specific measures of unemployment rates by county and year. The addition of these controls has little impact on estimated differences between Hispanic and white homeowners in either the new purchase or refinance sample. And, while the inclusion of these controls does reduce the estimated differences in delinquencies and foreclosures between black and white homeowners to some extent, substantial differences remain, especially in the home purchase sample. Thus, while increased exposure to labor and housing market shocks explains some of the increased delinquency and default rate for minority homeowners, a substantial unexplained gap remains.

As a final test of the predictions of our mortgage market model following a major expansion of credit, we examine whether the timing of the selection into the housing market has an effect *over and above* the other mechanisms proposed in the literature. We find that racial and ethnic differences are largest for home purchase originations in 2006, the peak of the housing boom according to the Case-Shiller price index. Importantly, the larger differences in 2006 remain even after controlling for the subsequent higher rates of negative equity for borrowers who purchased near the peak of the housing market. Along similar lines, we also examine racial and ethnic differences for a subsample of refinance mortgages that were originally purchased between 1998 and 2008 and subsequently refinanced in our sample period. For this subsample, racial and ethnic differences in foreclosure are tiny for homes that were originally purchased from 1998 to 2003, but substantial for homeowners who originally

purchased their homes between 2004 and 2007 – i.e., those drawn into the market at the peak of the credit expansion.

Taken together, our results provide strong evidence that minority households drawn into homeownership late in the recent housing market boom were especially vulnerable in the subsequent downturn in ways that are not explained by (i) exposure to different lenders or loans, (ii) the performance of local labor and housing markets, and (iii) the differences in equity position. These results call into question the idea of encouraging homeownership as a general mechanism for reducing racial disparities in wealth. To the extent that increases in homeownership are driven by the entry of especially vulnerable households into the owner-occupied market, such a push may backfire, leaving vulnerable households in a difficult financial situation and adversely affecting their wealth and credit-worthiness for years.

There are important caveats to our results. First, we do not directly observe household savings or wealth, which might help to explain why some households are able to manage adverse economic shocks more easily than others. Future work could more explicitly measure whether the estimated differences in mortgage outcomes by race reported here are primarily due to differences in wealth. Another limitation of our analysis is that our study period only covers the recent housing crisis. While Hoynes, Miller, and Schaller (2012) shows similar patterns by race for negative labor market shocks in a number of recent recessions, there is a chance that the recent housing market boom and bust may be unique in history.

The remainder of this paper proceeds as follows. The next section presents a simple model of the credit market and defines our notion of financial vulnerability. Section 3 presents the data used in the analysis. Section 4 presents the main results from our analyses of credit outcomes. Section 5 examines the impact of including controls for economic shocks on estimated racial and ethnic differences in credit outcomes. Section 6 presents additional discussions about the behavior of lenders and about the impact of foreclosures on other credit market outcomes. Section 7 concludes.

2. Model

In this section, we introduce a simple theoretical model that distinguishes between the two aspects of the sensitivity of homeowners to housing market and financial shocks highlighted above. We use the expression *exposure* to negative shocks to indicate how likely households are

to experience negative events of a magnitude that could lead to a mortgage default. If, for example, a homeowner works in an industry or occupation that is particularly sensitive to the business cycle and is more likely to be laid off in a recession, we will say that s/he has a greater exposure to negative shocks.

A second aspect of the responses of homeowners to downturns in the local economy or housing market is their ability to manage negative events when they do occur. We use the expression *vulnerability* to negative events to indicate the degree to which a household can absorb a serious negative event without defaulting on her mortgage. Households with greater precautionary savings or with the ability to draw on family or social networks to make ends meet in tough times, for example, will generally be better equipped to make their mortgage payments during a period of unemployment.

To help clarify the roles of the exposure versus vulnerability to negative events in mortgage default behavior, we develop a simple two-period model of home ownership and mortgage default. In the first period, households decide whether to rent or own. In the second period, homeowners decide whether to default or to meet their mortgage obligations conditional on the realization of an employment shock.

We assume that households are heterogeneous in two dimensions that correspond to the exposure and vulnerability to negative shocks, respectively. First, each household i face a probability of job loss between the first and second periods given by p_i. Second, households are distinguished by the cost of borrowing in the second period to meet their mortgage obligations, facing an effective interest rate of r_i. If, on the one hand, a household can draw on its own wealth or borrow easily on the private market, r_i will correspond to the market interest rate that it receives or forgoes. If, on the other hand, a household must use extraordinary means to meet its obligation (e.g., borrowing from several friends and family members) r_i may be very high, indeed approaching infinity in the limit if the household simply has no means to borrow enough money to make its mortgage payment.

For simplicity, we assume that households are identical in all other respects and focus on how the home ownership and default decisions of households are affected by the heterogeneity in the cost of borrowing and the probability of job loss. In particular, we assume that each household:

i. earns income, $\boldsymbol{y}$, in the second period if employed

ii. borrows ***M***, and faces a cost of mortgage borrowing in the first period, ρ

iii. buys a home with value, ***V,*** or rents a home with rent, ***R***

iv. faces a fixed cost of default, ***C***

v. has a utility function ***u(tenure, I)*** defined over tenure (owning versus renting) and net income ***I*** in the second period.[13]

As mentioned above, households make two decisions in the model. In the first period, they decide whether to rent or own taking into account uncertainty about their employment in the second period. In the second period, homeowners decide whether to make their mortgage payment. For simplicity, we assume that the cost of default ***C*** is high enough that all employed households find it in their best interest to avoid default.

Unemployed households face a simple trade-off regarding the default decision in the second period, choosing to default if $u(own, -r_i M) > u(rent, -C)$.[14] Given this simple decision rule, a threshold interest rate, r_i^* exists such that homeowners with $r_i > r_i^*$ default and those who can borrow relatively cheaply meet their mortgage obligations.

Knowing that they face the risk of job loss and possible high-cost borrowing or default in the second period, the decision to rent or buy in the first period is a function of both p_i and r_i. In particular, households with $r_i < r_i^*$ become homebuyers if:

$$(1) \qquad (1-p_i)u(own, y-\rho M) + p_i u(own, -r_i M) > (1-p_i)u(rent, y-R) + p_i u(rent, -R)$$

The decision rule looks similar for households with high borrowing costs, $r_i > r_i^*$. These households account for the fact that they would default following a negative event in the second period and, therefore, buy if:

[13] In characterizing the household's net income ***I*** in the second period below, we abstract from any complications due to unemployment insurance, dual earner households, etc. As a result, the measure of net income that we characterize in a number of situations we present below is negative. This has no real effect our characterization of the solution to the household's problem.

[14] The main purpose of introducing this model is to characterize how a household's exposure and vulnerability to financial shocks affects its decisions to buy a house and meet its mortgage obligations. To keep this analysis as straightforward as possible, we abstract from the role that the household's equity position (i.e., whether it is underwater on the mortgage) has on the default decision. In general, households should default (rather than sell) only when they have a negative equity position and, thus, one can view our model as a characterization of the household's problem in a negative equity environment. We also abstract from other potential differences between owning and rent, such as control of the property, property tax payments, etc.

(2) $$(1-p_i)u(own, y-\rho M)+p_i u(rent, -C) > (1-p_i)u(rent, y-R)+p_i u(rent, -R)$$

Figure 3 illustrates the household decision problem in *(p, r)* space. The vertical line at r_i* divides the space between the region to the left where homeowners meet their mortgage obligations even if they lose their jobs and the default region to the right where homeowners default upon the realization of a job loss.

In the default region, a horizontal line divides those with low probabilities of job loss, who decide to own given the minimal risk, and those with higher probabilities of losing their jobs, who decide to rent. The line is horizontal because the decision to own versus rent is not a function of how high the household's borrowing cost r_i is above r_i*, as can be seen in equation (2).

In the region to the left of r_i*, homeowners borrow (at what might still be pretty high effective interest rates) in the second period to meet their mortgage obligations even if they lose their jobs. Households with lower values of r_i find it less onerous to borrow in the event of a job loss and, therefore, are willing to become a homebuyer for higher values of p_i than those with borrowing rates near r_i.* As a result, the locus of households that are indifferent between buying and renting in this region is downward sloping.

The implications of our model for racial differences in mortgage outcomes depend on the correlation of race with effective borrowing costs, *r*. If wealth and liquidity gaps leave black and Hispanic households especially vulnerable to negative economic shocks, they may be more likely to have $r_i > r_i$* and, therefore, to default upon receiving a negative economic shock.

Marginal Homebuyers and Defaulters

The decision problem depicted in Figure 3 characterizes the household's side of the mortgage market. Of course, banks need to be willing to lend at the rate $\boldsymbol{r}$ assumed above. Their willingness to do so depends on the fraction of homebuyers who will repay their mortgages in the second period and their external cost of capital. Thus, given a distribution of households in *(p, r)* space, the external cost of capital that the banks face pins down the equilibrium mortgage rate ρ.

This simple characterization of the banks' side of the market can be used to examine the set of households that are drawn into home ownership and (following negative events) into default with a decline in mortgage interest rates.[15] In particular, Figure 4 depicts how the boundary locus between buying and renting shifts upward as the result of a decline in the mortgage interest rates $\rho_2 < \rho_1$. This shift is driven by the trade-off in the own versus rent decision of households characterized in equations (1) and (2).

As can be seen in the figure, at every level of r_i, the marginal households drawn into the owner-occupied market have higher probabilities of job loss p_i than those who would own homes at higher mortgage interest rates. Similarly, the marginal defaulters are those households with $r_i > r_i^*$ with relatively high values of p_i that are lured into homeownership by the lower mortgage interest rates. In this way, the marginal households drawn into the market by cheaper access to mortgage credit have the highest chance of losing their jobs and the highest likelihood of defaulting in that case (conditional on r_i).[16] Further, as again can be seen in figure 4, the share of high r_i individuals increases as ρ falls because the flattening of the boundary as r_i increases implies a larger percentage increase for borrowers with higher values of r_i.

The implications of our model for racial differences in mortgage outcomes following a major credit expansion depend on the correlation of race with effective borrowing costs for those households drawn into homeownership by reduced credit costs. Given large racial wealth disparities, it would not be surprising if the marginal black and Hispanic households drawn into the market were more concentrated in the higher effective borrowing cost region, $r_i > r_i^*$ and, as a result, were subsequently more likely to default following an adverse event.

[15] While our model considers a reduction in interest rates as opposed to relaxing downpayment constraints, this model implicitly captures the effect of a falling price of credit at any fixed level of downpayment. Further, shifts towards lower downpayments will only increase the role of financial vulnerability since households who need a lower downpayment to enter the market also likely have less financial assets on which to drawn after experience a negative shock.

[16] The model above abstracts from other causes and consequences of lowering borrowing rates, such as its macroeconomic impacts. There is a quite large literature in finance though, suggesting that capital bonanzas that lower loan costs are associated with economic crises, such as debt defaults and banking crises – see Reinhart and Reinhart (2008), for example. But it is beyond the scope of this paper to estimate whether lending availability and low interest rates caused the economic crises.

3. Data

Our data set is based on public Home Mortgage Disclosure Act (HMDA) data from between 2004 and 2008 and proprietary housing transaction/lien and assessor's databases purchased from Dataquick.[17] We begin with a convenience sample of seven major housing markets where Dataquick has information on refinance mortgages going back to at least 2004: Chicago IL CMSA, Cleveland OH MSA, Denver CO MSA, Los-Angeles CA CMSA, Miami-Palm Beach Corridor, San Francisco CA CMSA, and Washington DC-Baltimore MD suburban Corridor. We restrict our HMDA data to home purchase or refinance mortgages on owner-occupied, 1-4 family properties. In the Dataquick sample, we eliminate non-arm's length transactions, transactions where the name field contains the name of a church, trust, or where the first name is missing, and transactions where the address could not be matched to a 2000 Census tract or the zip code was missing.[18] The HMDA and Dataquick data are then merged based on year, loan amount, name of lender, state, county, and census tract. We obtain high quality matches for approximately 50% of our HMDA sample.[19]

Next, we draw a sample of mortgages to provide to a credit-reporting agency. These mortgages were sampled from May through August so that the March 31st archival credit report for the year of the mortgage provides appropriate information on the borrowers' credit quality prior to obtaining the mortgage. We oversample mortgages to minority borrowers, mortgages to white borrowers in minority or low-income neighborhoods, and high cost mortgages as designated in HMDA as high rate spread loans. In order to maximize the number of minority loans given the likelihood of sample saturation, we first draw the following oversamples based on race and ethnicity: 500 in each site, year and group (400 for 2004)[20] selected randomly from mortgages to black borrowers, mortgages to Hispanic borrowers, and mortgages to white

[17] Data provided by DataQuick Information Systems, Inc. www.dataquick.com. The property transaction data is collected by Dataquick or by intermediaries from county assessor's offices and contains a population of all sales and liens of all types including refinance mortgages, home improvement loans, and home equity lines of credit..

[18] This eliminates very few records due to the high quality of the name and address records in the assessor files.

[19] The key factor limiting the match rate is the lender name because the lender of record in the local assessor's data often differs from the respondent in HMDA. Less restrictive match criteria can yield a match rate around 90%, but in order to be conservative we restricted ourselves only to instances where we successfully match on lender name.

[20] The smaller sample in 2004 is driven by budgetary restrictions because costs depend upon the number of records, and earlier years imply more records because we follow every borrower from the origination year until 2009.

borrowers in minority or low-income neighborhoods.[21] We then split the remaining sample into rate spread and non-rate spread loans drawing 1000 borrowers associated with rate spread loans in each year and site (800 for 2004) and 2714 borrowers (2286 for 2004) from the non-rate spread sample in each year and site. Weights are developed based on the probability of selection,[22] and each site receives equal weight in the pooled sample.[23]

This sample is provided to Experian Information Solutions, Inc.[24] who matches the name and address of each borrower and co-borrower to archival credit report data from March 31st preceding the mortgage transaction and March 31st for every year that follows this transaction through 2009. Our match rate for the pre-mortgage archive is 81.4 and 84.5 percent in the home purchase and refinance samples, respectively. For years following the mortgage, the match rate rises by 4 to 5 percentage points. In many cases, these individuals also may not have been found by the credit reporting agencies when the lender requested a report in which case lack of a score matches the information that the lender would have had when approving and pricing the loan, but lenders can enter by hand additional information that is not available to us such as social security number or previous addresses.[25]

Table 1 illustrates the impact of our match process on the sample mean on race and ethnicity of the borrower, whether the primary mortgage is a high cost or rate spread loan,[26] loan amount, family income of the borrower, and census tract variables including median income,

[21] Our budget at the credit reporting agency is based on number of borrowers so whenever a mortgage is sampled which contains the name of the coborrower, typically the borrower's spouse, this was counted as having sampled two borrowers.

[22] The sampling is explicitly based on 8 strata for each site: black borrowers, Hispanic borrowers, white borrowers in minority or low-income neighborhoods, and all other borrowers divided into rate spread and non-rate spread loans. All loans from the same strata and year receive equal weight.

[23] We have a convenience sample of housing markets so it would be inappropriate to weight based on the number of mortgages. In any stratified sampling scheme, Los Angeles, which dominates our sample in terms of total number of HMDA mortgages, would be selected with certainty while housing markets like Denver and Cleveland would be assigned to a stratum with other similarly sized and located metropolitan areas and if chosen would receive a higher weight (offsetting the smaller number of mortgages) based on the probability of being selected from the stratum.

[24] Experian is a service mark and registered trademark of Experian Information Solutions, Inc.

[25] For home purchase mortgages, we only observe the address of the new housing unit, but in practice this does not present a major problem for the credit data match because the archival data can be matched based on current and several past addresses and in practice we observe only a small difference between the home purchase and refinance match rate.

[26] It identifies first liens with a 3 percentage point or more spread between the mortgage Annual Percentage Rate (APR) and the interest rate on 10 year treasury notes. The Annual Percentage Rate or APR includes both the interest or note rate on the loan and the effect of closing costs on the cost of credit. HMDA only reports the actual APR if it exceeds this rate spread, and so our dependent variable is defined as a binary variable capturing whether the APR exceeds the reporting threshold or the rate spread.

percent black, Hispanic and Asian resident, and percent of properties owner-occupied. The first column shows the mean for the entire HMDA sample for our seven sites where each site receives equal weight in the mean. The second column shows the mean for our HMDA-Dataquick match, and the third column restricts our sample to mortgages between May and August. The fourth column shows the weighted mean for the sample of mortgages that was provided to the credit reporting agency. The last column in Table 1 shows the weighted means on these common variables for just the subsample where the name and address was matched to the minimum amount of credit line data in order to generate a record. The sample composition is quite stable except for a moderate decline in share white and moderate increase in loan amount between columns 1 and 2 associated with the difficulty of matching lender names between HMDA and the Dataquick provided assessor files.

Table 2 shows the weighted means for our final samples of post mortgage credit reports for the home purchase and refinance subsamples that were successfully merged to pre-mortgage credit report data.[27] The first three columns show the sample size, the mean and standard deviation for our sample of home purchase mortgages, and the last three columns show these values for refinance mortgages. From the contemporaneous credit history data, we obtain number of 90 to 180 day mortgage delinquencies and foreclosures in the last 12 months. From the pre-mortgage data, we obtain these same outcomes prior to the mortgage and the borrowers' (or if unavailable coborrower's) Vantage score. The first credit report observation following the mortgage is used to obtain monthly mortgage payment, which when combined with HMDA income is used to calculate the mortgage payment to income ratio.[28] The monthly mortgage payment is combined with debt payments from the pre-mortgage credit data and HMDA income to calculate debt payment to income ratio. Finally, age is observed for many borrowers and coborrowers in the credit history files.

The HMDA data contains whether the loan is a high cost or exceeds a standard rate spread above treasury rates, borrower race and ethnicity, gender, loan amount, applicant income,

[27] The samples sizes in Table 2 are substantially larger than the samples sizes in Table 1 because each mortgage in Table 1 will have one post mortgage credit report observation for every year between the year after origination and the final year of our data in 2009.

[28] The mortgage payment for the current mortgage is only observed in the credit line data from the year following the mortgage. However, in most instances, borrowers who are matched by the credit reporting agency prior to the mortgage are also matched in the following year.

and whether a coborrower is present. We use the loan amount to calculate whether the loan is non-conforming or a jumbo loan, i.e. too large to be purchased by the Government Sponsored Enterprises, and based on census tract we observe tract racial and ethnic composition, income, poverty rate, share owner-occupied and the value of median rents to median home price, which is viewed as a proxy for anticipated housing price appreciation. From the match with transaction data, we observe the presence and size of subordinate liens, whether the liens are fixed or variable rate mortgages, the loan to value ratio based on sales price for home purchase mortgages and on an estimated value based on either previous sales price[29] or assessed value for refinance mortgages when a previous sale is unobserved,[30] and detailed property attributes including whether single family home, a condominium, number of units on the property.

4. Delinquencies and Foreclosures by Race

In this section, we present empirical estimates for models of two housing market outcomes, delinquencies and foreclosures. Each mortgage origination contributes one or more observations based on the origination date. Specifically, the 2004 originations contribute housing market outcome observations in 05, 06, 07, 08 and 09, while each 2008 origination contributes only a single post mortgage outcome for 09. We write those outcomes *d* for individual *i*, of race *r*, at time of origination *s*, calendar time *t*, and MSA *m*, as:

$$d_{irstm} = \beta R_{ir} + \gamma X_{is} + \delta_{st} + \theta_m + \varepsilon_{irstm} \quad (3)$$

where *R* includes dummies for blacks, Hispanics and Asians, so all estimates are relative to white households. X_{is} include the set of pre-mortgage origination variables reported in Table 2.[31] Our

[29] We use our extensive housing transaction data to develop both a hedonic and repeat sales quarterly price index for each county. When we observe a previous sale of the property, we simply adjust that earlier sales price to estimate current value based on the hedonic index. However, the repeat sales index yields quite similar estimates.

[30] When a previous sale is not observed, we use the county assessment and adjust that value by the average ratio of sales price to assessed value for that county and quarter, see Clapp, Nanda and Ross (2008). In California, our refinance sample is restricted to mortgages where a previous purchase is observed because property assessments are uninformative as to the value of the underlying property. This restriction is feasible because the Dataquick data in California contains transactions back to the late 1980s.

[31] The loan to value ratio is included as bins below 0.6, 0.6 to 0.8, 0.8 to 0.84, 0.85 to 0.89, 0.90 to 0.94, 0.95 to 1.00, 1.00 to 1.04, and 1.05 and above. In addition to controlling for the combined loan to value ratio, we control for the number of subordinate liens and whether each is a fixed or variable rate loan. The Vantage scores are included as a series of dummy variables based on 20 point intervals. The mortgage payment and debt to income ratios are also

design relies on the fact that we are controlling for almost all characteristics of borrowers, houses, and mortgages that were observed by lenders at time of origination. Especially noteworthy is the information about credit scores, which in principle means that we will compare housing outcomes between blacks and whites, for example, with similar creditworthiness and with similar choices of house type, neighborhoods, and loan characteristics. Finally, the model also includes year of origination by year of the credit profile indicators δ_{st} to deal with common trends, as well as MSA fixed effects. Racial and ethnicity estimates for this baseline model are presented in Table 3, while richer variations of this model – that include tract fixed effects and subprime lending proxies for example – are shown in the subsequent tables.

The first two columns of Table 3 present results for the sample of home purchases and the last two columns for the refinance sample. We find large unexplained differences for both black and Hispanic borrowers in the probability of facing delinquencies and foreclosures. For example, the likelihood of a foreclosure is approximately 4.2 and 3.9 percentage points higher for blacks and Hispanics, respectively, in the home purchase sample. The estimates for the refinance sample also reveal statistically significant differences by race in foreclosure rates, but the point estimates are smaller: 1.7 and 2.4 percentage points for blacks and Hispanics, respectively. Estimates for Asians are also statistically significant but substantially smaller than the differences observed for African-Americans and Hispanics. Similar results obtain for 90-180 day mortgage delinquency rates.[32]

One potential explanation for the differences in housing outcomes by race is that minorities in the last housing cycle had a large fraction of the so-called subprime loans. Although we controlled for many loan characteristics in the baseline model, we add other proxies for subprime lending in Table 4. The first column replicates the main results from Table 3. The second column includes a dummy for rate spread loan, as defined in the data section above. The direct effect of obtaining a rate spread loan on the likelihood of negative credit market outcomes

---

divided into bins. The bins vary in size. For mortgage payment to income ratios, the smallest bins are 0.02 are around the pre-crisis secondary market criteria of 0.28, and for total debt payment to income ratios the smallest bins are 0.03 around the threshold of 0.36.[31] The model also controls for the presence of a coborrower, and for whether the loan amount is non-conforming or the loan is a jumbo. Finally, we include controls for relative lender size and dummy variables for the agency that regulates the lender and in some cases lender fixed effects. Relative lender size varies across site, and agency code sometimes varies over time for the same lender.

[32] The R-squares for these models are not very high in large part because future adverse outcomes are heavily driven by events that happen after the borrower obtains their mortgage, but also because the fit of linear models tends to be lower when events are infrequent.

are sizable and statistically significant for all credit market outcomes in both the home purchase and refinance market. The rate spread loan variable clearly contains information about inherent risk associated with either the borrower and/or the loan terms.[33] However, the inclusion of the rate-spread indicator at most moderately reduces the estimated racial differences in foreclosure and delinquency rates. For example, the delinquency coefficient for blacks in the purchase sample changes from 3.08 to 2.74 percentage points, and that change is not statistically different than zero (declines across all black and Hispanic coefficients range between 2 and 18 percent with an average decline of 9 percent).

Since subprime lending shares are highly concentrated in a subsample of lenders – and such borrowers may exhibit alternative behavior or be affected by the typical attributes of subprime loan products - we include lender fixed effects in column 3. Such controls again lead to only moderate reductions (on average the decline is 7 percent) in the economic magnitude of baseline results and have no impact on their statistical significance. Subprime lending is also potentially concentrated in certain cities and neighborhoods. In column 4 we include a set of tract fixed effects instead of a set of neighborhood level variables (while still controlling for lender fixed effects). The inclusion of tract fixed effects leads to moderate reductions for some coefficients and increases for other, leaving the main results little changed.

Finally, the last column includes additional controls for subprime borrowing recognizing that the impact of key loan terms on borrower outcomes may vary between prime and subprime borrowers. We identify borrowers with Vantage scores below 701 as subprime borrowers[34] and then interact the subprime dummy with dummy variables associated with key thresholds of loan to value ratio, debt to income ratio, mortgage payment to income ratio,[35] as well as with the presence of subordinate debt and whether the primary mortgage is adjustable rate. The estimates of racial and ethnic differences are practically identical to the ones reported in the previous column. Thus, while these variables have considerable explanatory power, they have little effect on estimated racial and ethnic differences in mortgage outcomes.

[33] These effects cannot simply arise from the higher cost of these loans because our model contains detailed controls for the housing and total debt expense to income ratios faced by households.

[34] The credit reporting agencies that developed the Vantage score algorithms describes scores below 701 as non-prime. Further, a Vantage score of 701 is comparable to a FICO score of 660 in that in both cases approximately 30% of individuals have credit scores below these thresholds.

[35] The loan to value thresholds of 0.80, 0.90, 0.95 and 1.00, the debt to income thresholds used are 0.36 and 0.45, and the mortgage payment to income ratio thresholds used are 0.28 and 0.33

Overall, the estimated differences in housing market outcomes by race suggest that most of the racial and ethnic differences in delinquencies and foreclosures are not related to differential access to lenders, types of loans, or any observable factor that might have been used to price the mortgages in the first place. In particular, the racial and ethnic differences decline by only 12 to 29 percent across columns 1 to 5.

5. Mortgage Outcomes and Negative Shocks

We next examine whether the observed racial and ethnic differences in mortgage outcomes might be attributable to local variation in economic shocks. We test this hypothesis in the first column of Table 5 by adding a set of county-year fixed effects to the specification reported in the last column of Table 4 (which includes tract and lender effects). Results are only shown for black and Hispanic homeowners to simplify the exposition. We find that the inclusion of flexible controls for local trends leaves the direct effect of the race and ethnicity indicators on delinquencies and foreclosures virtually unchanged.

In the second column of Table 5 we also include an indicator for negative equity, as homeowners in that situation could strategically default in order to avoid the payment of a loan that is worth more than the value of the asset. We use county-year variation in prices based hedonic models estimated using the Dataquick transaction sample to develop measures of negative equity associated with each housing unit in each credit reporting period.

The literature on mortgage foreclosure and default suggests that the likelihood of foreclosure and the effect of negative equity on foreclosure will vary with the employment outcomes and prospects of homeowners. We do not have individual information about employment. Instead, we create measures of local employment and unemployment rates for prime age males for each county and year using the public use sample of the American Community Survey. Columns 2 and 3 of Table 5 add these local employment measures to the model (also interacting employment with negative equity) for the full population and by race, respectively.

The results reported in Column 2 imply that negative equity has a major direct impact on credit market outcomes. At average local employment levels (93.6 percent), households in our home purchase sample are 1.8 percentage points more likely to be 90 days delinquent on their mortgage payments and 2.7 percentage points more likely to enter foreclosure when their house

is underwater. When local employment rates are one standard deviation higher than the average, the effects of negative equity are near zero.[36] The effect of employment rates on the importance of negative equity can be interpreted as a difference-in-differences estimate because after controlling for county by year FE's and negative equity the employment-negative equity interaction estimates simply indicate delinquency and foreclosure rates rise more sharply with negative equity in the lowest employment county by year cells. Importantly, despite the direct importance of negative equity and local employment rates, the estimated racial and ethnic differences in default and foreclosure rates are unchanged.[37]

The effect of negative equity increases moderately when race-specific employment rates are used in the specification reported in Column 3. Households who are underwater and face the mean employment rate are 2.4 percentage points more likely to 90 days delinquent on their mortgage payments and 3.7 percentage points more likely to enter foreclosure.[38] Importantly, the inclusion of race-specific employment rates affects the race coefficients, reducing, for example, the black coefficient to approximately 1.5 percentage points in the home purchase sample. The effects in the refinance market are even larger reducing black-white differences to about 0.5 percentage points. However, the results for Hispanic borrowers are relatively unaffected by the inclusion of race-specific employment rates. For Hispanics in the home purchase sample, the estimated differences are 1.6 and 2.6 percentage points and for refinance mortgages 1.1 and 1.9 percentage points for delinquency and foreclosure, respectively.[39]

Table 6 investigates whether heterogeneity within minority status could explain conditional differences in delinquencies and foreclosures by race. We interact minority status

[36] The effects of negative equity for the refinance sample are considerably smaller.

[37] Gyourko and Tracy (2013) demonstrate substantial attenuation bias in credit outcome models when county level employment rates are used to proxy for actual future employment outcomes. In our case, however, we find that county employment rates (at least as a proxy for the risk of unemployment) have strong explanatory power for differential rates of mortgage delinquency and foreclosure.

[38] While the estimated coefficient falls from column 2 to 3 and column 5 to 6, the average race specific employment rate is substantially lower across the sample than the average employment race, especially for African-Americans in our sites. As a result, most of the increase in the importance of negative equity is for our minority subsamples. Note that results are quite similar in a model that includes the interaction of negative equity with both the overall and the race specific employment rate variables.

[39] We also ran the results in columns 2, 3, 5 and 6 using a continuous measure of household equity, current loan to rent value. The estimated racial differences and equity effects were very similar and the model fit nearly identical to the results presented in Table 5. Note that in columns 2, 3, 5 and 6 the effect of negative equity (or LTV) is effectively zero when employment rates are 100 percent. This finding is consistent with Foote, Geradi and Willen (2008) who discuss the importance of income shocks for pushing borrowers in negative equity into foreclosure.

with a subprime borrower dummy, based on whether the borrower's initial Vantage score is below 701. The results indicate that the racial and ethnic differences observed for the black home purchase subsample and both Hispanic subsamples persist across the entire sample, not just for high-risk borrowers.[40]

6. Heterogeneity in Mortgage Outcomes by Year of Origination

We close our analysis by examining the pattern of risk based on year of origination. Figures 5 and 6 present unconditional racial differences in delinquencies and foreclosures for White, African-American and Hispanic homeowners in the home purchase and refinance markets by origination year. Delinquencies and foreclosures rise for all groups until 2006 and fall afterwards. However, changes in rates are relatively small for white borrowers, and quite large in both absolute and relative terms for minorities.

Next, we examine conditional differences by year of origination. First, we restrict the sample to 2008 and 2009 delinquencies and foreclosures in order to focus on delinquencies during the crisis period.[41] Table 7 presents the racial and ethnic differences by origination year with 2004 originations as the omitted category for our original OLS specification in Table 3.[42] For black home purchase originations, there is a substantial increase in the likelihood of foreclosure and delinquency for 2006 originations with a decline in later years. For Hispanics, the pattern is less clearly focused on 2006, but nonetheless ethnic differences in foreclosure and delinquency tend to higher in 2005-2007 relative to 2004 and 2008 originations. Overall, mortgages originated for home purchases by black and Hispanic borrowers in 2006 are 5-6 percentage points more likely to experience a foreclosure than 2004 originations.

A natural concern with Table 7 is that origination year may proxy for the equity shocks faced by households during the crisis. Households that buy into a market during the peak of the

[40] We also examined interactions with vantage score, rate spread, credit utilization and bank card delinquencies and found a similar pattern in all cases. The vantage score interactions specifically indicate that unexplained negative outcomes for minorities exist well into the distribution of prime credit scores.

[41] The basic pattern of previous findings are unchanged by the restriction of our sample to 2008 and 2009 delinquencies and foreclosures except that the average incidence of foreclosure rises during this time so that observe racial and ethnic differences are larger. For example, in comparison to Table 3, racial differences (blacks) increased by between 30 and 53 percent while ethnic differences increased by between 62 and 73 percent. Finally, restricting the sample also avoids confounding the direct effect of origination year with differences that arise because earlier origination years allow for foreclosure and delinquency during the pre-crisis period.

[42] Estimates using the subprime specification in the last column of Table 4 are very similar.

crisis likely faced larger declines in prices relative to their purchase price and therefore are more likely to have negative equity levels. Table 8 repeats the analysis for the model with negative equity and race specific employment controls. For blacks in the home purchase subsample and for Hispanics the basic patterns in Table 7 persist with racial differences largest for 2006 home purchase originations and ethnic differences largest for originations between 2005 and 2007. Further, for Hispanics, the largest differences arise for home purchase originations between 2005 and 2007.

We next use the refinance sample in order to examine home purchase originations over a longer timeframe. Specifically, we match refinance mortgages with the original home purchases back to 1997, which is the earliest year for which we observe home purchases in all seven of our markets, and present the racial and ethnic differences in foreclosure by origination year in Table 9. While the sample represents only home purchases that selected into refinance, our estimates demonstrate the significance of purchase year for a representative sample of refinance mortgages. Notably, the origination year coefficients are relatively flat until about 2003 or 2004 when they begin to increase for both the OLS and the negative equity specifications. For both specifications, we see the largest racial and ethnic differences in foreclosures for homes that were originally purchased in 2004 to 2007 late in the boom.[43]

Finally, in order to assure that the estimated racial differences are not entirely driven by lending in the subprime sector, we repeat our analysis for a sample of prime borrowers with vantages scores of 701 or higher. Table 10 shows origination year results for the negative equity model for both our sample of home purchase and refinance models. The results are very similar to the estimates for the home purchase sample shown in Table 8 and for the refinance sample reported in Table 9. Racial and ethnic differences for home purchase loans are substantially larger for 2006 originations, and for refinance loans the differences are substantially larger for 2005 and 2006 for blacks and between 2004 and 2007 for Hispanics.[44]

Overall, these results corroborate the intuition presented in the credit market model that we developed in Section 2. As credit expanded during the housing boom, households with

[43] Results for delinquencies are similar except that the increase in racial differences begins a little earlier, which makes sense given the lag between delinquencies and foreclosure.

[44] The results for the OLS and Subprime model specifications are similar except as in Tables 8 and 9 that the differences for blacks in the subprime model are larger and have greater statistical significance than the results for the negative equity model.

higher unobserved risk of defaults and delinquencies entered the housing market and, empirically, black and Hispanic borrowers were much more likely to be in this group. The selection of these especially high-risk households in to the market translated into much higher rates of delinquencies and defaults once the recession started, especially among minority borrowers.

7. Summary and Conclusion

In this paper, we identify robust racial and ethnic differences in the likelihood of mortgage delinquencies and foreclosures in the recent housing market bust and associated recession. Substantial differences remain after controlling for (i) the borrower, home, and market attributes, including the individual's credit score, that would have been observed by lenders *ex ante*, (ii) the characteristics of the lender and loan itself and (iii) *ex post* measures of exposure to local housing and labor market shocks. Collectively, these results imply that the relatively poor mortgage outcomes for minority borrowers are not simply a function of greater participation in the subprime sector or greater exposure to neighborhood housing price declines or unemployment rates.

A further decomposition of our main findings by origination year reveals that the large estimated racial and ethnic differences in mortgage outcomes are concentrated in mortgages that were originated in 2005-06. That is, black and Hispanic households drawn into homeownership at the peak of the credit expansion and housing boom were especially likely to subsequently become delinquent and default on their mortgages. This finding is again robust to the inclusion of a broad set of controls for borrower, lender, loan, and neighborhood attributes.

The simple credit market model that we develop early in the paper highlights a potential unified explanation for the full set of empirical results. Drawing a distinction between the exposure to adverse economic shocks and the ability to deal with these events when they occur (through, for example, personal savings, access to credit, financial help from friends or family members), the model implies that (i) those households drawn into homeownership late in a major credit expansion are especially at risk for adverse economic outcomes and (ii) the financially vulnerable among these households are especially likely to default when such shocks occur. This model provides a coherent explanation for our empirical findings that black and Hispanic

households drawn into homeownership late in the housing boom were especially vulnerable to the adverse economic shocks that occurred during the recent financial crisis.

Our results complement a recent literature that aims to understand the causes and consequences of the last housing cycle. For example, while Mian and Sufi (2009) demonstrate a significant role for subprime lending in explaining overall neighborhood level foreclosure rates, our analysis implies that subprime lending can explain at most a modest fraction of observed racial and ethnic differences in credit market outcomes. In addition, tighter underwriting standards and increased financial oversight arising from recent financial reforms are unlikely to address these concerns because the observed differences arise after controlling for all traditional underwriting variables, are based on comparisons within lenders and neighborhoods, and occur across a broad spectrum of minority borrowers.

Finally, regardless of the ultimate explanation for the observed higher rates of negative mortgage outcomes, our study raises serious concerns about homeownership as a vehicle for reducing racial wealth disparities. Our findings suggest instead that homeownership may be especially risky for households with a low initial level of wealth (savings) or fewer family resources on which to draw when hit with an adverse economic shock. Because delinquencies and default have consequences that go beyond the direct loss of housing equity/wealth, they can contribute substantially to perpetuating the wealth gap across generations. The increased cost of all subsequent borrowing (through the lower credit scores), in particular, makes future wealth accumulation much more difficult.

References

Altonji, Joseph G. and Rebecca M. Blank. 1999. Chapter 48, Race and gender in the labor market, In: Orley C. Ashenfelter and David Card, Editor(s), *Handbook of Labor Economics*, Elsevier, Volume 3, Part C, p. 3143-3259.

Amin, Sajeda, Ashok S. Rai, and Giorgio Topa. 2003. "Does Microcredit Reach the Poor and Vulnerable? Evidence from Northern Bangladesh." *Journal of Development Economics* 70 (1): p. 59–82.

Apgar, William, Alegra Calder, and Gary Fauth. 2004. *Credit, Capital and Communities: The Implications of the Changing Mortgage Banking Industry for Community-Based Organizations*. Cambridge, MA: Joint Center for Housing Studies, Harvard University.

Avery, Robert B., Kenneth P. Brevoort, and Glenn B. Canner. 2012. "Does Credit Scoring Produce a Disparate Impact?" *Real Estate Economics*, Vol. 40, p. S65-S114.

Avery, Robert B., Glen B. Canner, and Robert E. Cooke. 2005. "New Information Reported under HMDA and Its Implications in Fair Lending Enforcement." *Federal Reserve Bulletin* (Summer 2005).

Belsky, Eric S., Nicolas P. Retsinas and Mark Duda. 2007. Chapter 9, The Financial Returns to Low-Income Homeownership, In: *Chasing the American Dream: new perspectives on affordable homeownership,* ed. William M. Rohe and Harry L. Watson, p. 191-212. Ithaca and London: Cornell University Press.

Brueckner, Jan K. 1986. 'The Downpayment Constraint and Housing Tenure Choice." *Regional Science and Urban Economics* 16, p. 519–525.

Calem, Paul S., Kevin Gillen, and Susan Wachter. 2004. "The Neighborhood Distribution of Subprime Mortgage Lending." *Journal of Real Estate Finance and Economics*, 29, p. 393-410.

Clapp, John, Anupam Nanda and Stephen L. Ross. 2008. "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values." *Journal of Urban Economics,* Volume 63, Issue 2, p. 451-466.

Charles, Kerwin, and Erik Hurst. 2002. "The Transition to Home Ownership and the Black-White Wealth Gap." *Review of Economics and Statistics*, 84 (2), p. 281–97.

Cochrane, John. 1991. "A Simple Test of Consumption Insurance." *The Journal of Political Economy*, 99 (5), p. 957–976.

Courchane, Marsha J. 2007. "The Pricing of Home Mortgage Loans to Minority Borrowers: How Much of the APR Differential Can We Explain?" *Journal of Real Estate Research*, 29, p. 399-439.

Deng, Yongheng, Stephen L. Ross, and Susan M. Wachter. 2003. "Racial Differences in Homeownership: The Effect of Residential Location," *Regional Science and Urban Economics* 33 (5), p. 517–556.

Duca, John V., and Stuart S. Rosenthal. 1994. "Borrowing Constraints and Access to Owner-Occupied Housing," *Regional Science and Urban Economics* 24, p. 301–322.

Edminston, Kelly D. 2009. "Characteristics of High-Foreclosure Neighborhoods in the Tenth District." *Economic Review (Quarter 2)*, p. 51-75.

Engelhart, Gary V. 1996. "Consumption, Down Payments, and Liquidity Constraints," *Journal of Money, Credit, and Banking*, 28:2 (May), p. 255-271.

Fisher, Lynn M., Lauren Lambie-Hanson and Paul Willen. 2010. “A Profile of the Mortgage Crisis in a Low-and-Moderate-Income Community.” FRB of Boston Public Policy Discussion Paper No. 10-6.

Foote, Christopher, Kristopher Gerardi, Lorenz Goette, and Paul S. Willen. 2010. “Reducing Foreclosures: No Easy Answers.” In *NBER Macroeconomics Annual* Volume 24, ed. Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, p. 89–138. Chicago: University of Chicago Press.

Gerardi, Kristopher S. and Paul Willen. 2009. “Subprime Mortgages, Foreclosures, and Urban Neighborhoods.” *BE Journal of Economic Analysis and Policy*, 9 (3), Symposium Article 12.

Foote, Christopher L., Kristopher Gerardi, Paul S. Willen. 2008. “Negative Equity and Foreclosure: Theory and Evidence.” *Journal of Urban Economics*, Volume 64, Issue 2, p. 234-245.

Gerardi, Kristopher S., Adam H. Shapiro and Paul Willen. 2007. “Subprime Outcomes: Risky Mortgages, Homeownership Experiences, and Foreclosures.” FRB of Boston Working Paper No. 07-15.

Gittleman, Maury and Edward Wolff. 2004. “Racial Differences in Patterns of Wealth Accumulation.” *Journal of Human Resources*, 39(1), p. 193-227.

Gyourko, Joseph, Peter Linneman and Susan Wachter. 1999. “Analyzing the Relation among Race, Wealth, and Homeownership in America.” *Journal of Housing Economics* 8, p. 63–89.

Gyourko, Joseph and Joseph Tracy. 2013. “Unemloyment and Unobserved Credit Risk in the FHA Single Family Mortgage Insurance Fund.” NBER Working Paper No. 18880.

Haughwout, Andrew, Donghoon Lee, Joseph Tracy and Wilbert van der Klaaw. 2011. “Real Estate Investors, the Leverage Cycle and the Housing Market Crisis.” New York Federal Reserve Working Paper.

Hayashi, Fumio, Joseph Altonji, and Laurence Kotlikoff. 1996. “Risk-Sharing between and within Families.” *Econometrica,* 64 (2), p. 261–294.

Herbert, Christopher E., Donald R. Haurin, Stuart S. Rosenthal, and Mark Duda. 2005. “Homeownership Gaps Among Low Income and Minority Borrowers and Neighborhoods.” Washington, DC: U.S. Department of Housing and Urban Development. http://www.huduser.org/publications/HOMEOWN/HGapsAmongLInMBnN.html.

Hoynes, Hilary, Douglas L. Miller and Jessamyn Schaller. 2012. "Who Suffers during Recessions?" *Journal of Economic Perspectives*, American Economic Association, vol. 26(3), p. 27-48, Summer.

Linneman, Peter, Susan Wachter, 1989. “The Impacts of Borrowing Constraints on Homeownership.” *AREUEA Journal* 17, p. 389–402.

Mayer, Chris and Karen Pence. 2008. “Subprime Mortgages: What, Where and To Whom?” NBER Working Paper No. 14083.

Mayer, Chris, Karen Pence and Shane M. Sherlund. 2009. “The Rise in Mortgage Defaults.” *Journal of Economic Perspectives*, volume 23(1), p. 27-50.

Mazzocco, Maurizio. 2004. “Savings, Risk Sharing and Preferences for Risk,” *American Economic Review* 94 (4), p. 1169–1182

Mian, Atif and Amir Sufi . 2009. “The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis.” *Quarterly Journal of Economics*, 124(4), p. 1449–96.

Pendall, Rolf, Arthur C. Nelson, Casey J. Dawkins, and Gerrit J. Knaap. 2005. Chapter 10, Connecting Smart Growth, Housing Affordability, and Racial Equity, In *The geography of opportunity: Race and housing choice in Metropolitan America*, ed. William J. Wilson, p. 219-46. Brookings Institution Press.

Oliver, Melvin and Thomas M. Shapiro. 1997. *Black Wealth, White Wealth: A New Perspective on Racial Inequality*, Routledge.

Quercia, Roberto G., George W. McCarthy, and Susan M. Wacther. 2003. “The Impacts of Affordable Lending Efforts on Homeownership rates.” *Journal of Housing Economics* 12(1), p. 29-59.

Reid, Carolina and Elizabeth Laderman. 2009. “The Untold Costs of Subprime Lending: Examining the Links among Higher-Priced Lending, Foreclosures and Race in California.” Federal Reserve Bank of San Francisco Working Paper.

Nicholas P. Retsinas and Eric Belsky. 2006. *Building Assets, Building Credit: Creating Wealth in Low-Income Communities*. Washington, DC: JCHS Brookings.

Reinhart, Carmen M. and Kenneth S. Rogoff, 2008. "Is the 2007 US Sub-prime Financial Crisis So Different? An International Historical Comparison." *American Economic Review*, American Economic Association, vol. 98(2), p. 339-44.

Rohe, William M., Shannon Van Zandt, and George McCarthy. 2002. “Social Benefits and Costs of Homeownership.” In *Low-Income Homeownership*, edited by Nicolas P. Retsinas and Eric S. Belsky, p. 381–406. Cambridge, MA: Joint Center for Housing Studies; Washington DC: Brookings Institution Press.

Shapiro, Thomas. 2004. *The Hidden Cost of Being African American: How Wealth Perpetuates Inequality*, Oxford University Press

Shuey, Kim and Andrea Wilson. 2008. “Cumulative Disadvantage and Black-White Disparities in Life-Course Health Trajectories.” *Research on Aging* vol. 30 no. 2, p. 200-225.

Smith, Jeffrey. 1995. “Racial and Ethnic Differences in Wealth in the Health and Retirement Study.” *Journal of Human Resources*, 30: S, p. 158-83.

Terrell, Henry. 1971. “Wealth Accumulation of Black and White Families: The Empirical Evidence.” *Journal of Finance* 26, p. 363-77.

Townsend, Robert M. 1994. “Risk and Insurance in Village India.” *Econometrica,* 62 (3), p. 539–591.

Wachter, Susan M. and Isaac F. Megbolugbe. 1992. “Racial and Ethnic Disparities in Homeownership.” *Housing Policy Debate* 3 (2), p. 353–370.

Wachter, Susan M., Karl Russo and Jonathan E. Hershaff. 2010. “Subprime Lending: Neighborhood Patterns Over Time in US Cities.” Working Paper.

Zhu, Xiao Di and Xiaodong Liu. 2007. “The Importance of Wealth and Income in the Transition to Homeownership.” http://www.huduser.org/periodicals/cityscpe/vol9num2/

Table 1: Sample Selection of HMDA Variables

	HMDA full sample		DQ High quality match		DQ High quality match--May-Aug		Sample, weighted		Experian Matched Sample, weighted	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Black	11.1%	33.9%	11.6%	34.5%	11.4%	34.3%	11.4%	34.3%	11.2%	31.5%
Hispanic	17.4%	41.0%	19.4%	42.7%	19.3%	42.6%	19.2%	42.6%	18.5%	38.8%
Asian	7.5%	28.4%	8.6%	30.3%	8.5%	30.1%	8.6%	30.3%	8.9%	28.4%
White	67.8%	50.5%	60.1%	52.9%	60.5%	52.8%	60.5%	52.8%	61.1%	48.7%
Loan Amount	$247	$243	$271	$221	$274	$224	$274	$227	$278	$211
Applicant Income	$107	$142	$105	$128	$105	$127	$106	$132	$106	$115
Tract Median Inc	$59.1	$25.6	$59.6	$25.2	$59.7	$25.2	$59.7	$25.2	$60.4	$23.5
Tract Pct Black	12.6%	23.8%	11.6%	22.5%	11.5%	22.4%	11.5%	22.4%	11.3%	20.4%
Tract Pct Hispanic	16.9%	22.7%	16.5%	22.1%	16.4%	22.0%	16.5%	22.1%	16.3%	20.2%
Tract Pct Asian	6.3%	10.9%	6.5%	11.2%	6.5%	11.1%	6.5%	11.1%	6.6%	10.4%
Number of Observations	9,345,709		4,002,996		1,459,468		273,589		238,785	

Notes: The first column presents the means and standard deviations for all Home Mortgage Disclosure Act (HMDA) home purchase and refinance mortgages in our seven market areas between 2004 and 2008 where each market is given equal weight in the means. The second column presents the means for the subsample where we have a high quality match between HMDA and the housing transaction file based on loan amount, type of loan, census tract and lender name, and the third column is based on further restricting the sample to mortgages originated between May and August. The fourth column presents means for the stratified sample that was merged to the credit history data weighted by the inverse of the sampling probabilities again with equal weights for each market, and the fifth column presents weighted means for the subsample that was successfully merged to an established credit history prior to mortgage origination.

Table 2: Descriptive Statistics

	purchase		refinance	
Credit Data	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Mortgage Delinquency 90-180 Days	0.035	0.214	0.030	0.196
Foreclosure	0.053	0.319	0.037	0.250
Delinquency prior to Mortgage	0.001	0.025	0.002	0.046
Foreclosure prior to Mortgage	0.001	0.037	0.006	0.082
Vantage Score	781	104	775	110
Mortgage Payment to Income Ratio	0.256	0.280	0.247	1.064
Debt Payment to Income Ratio	0.321	0.336	0.348	1.191
Borrower Age	27.1	23.1	34.3	24.8
HMDA Data				
Rate Spread	0.151	0.358	0.168	0.374
American Indian	0.003	0.054	0.003	0.056
Asian	0.097	0.296	0.086	0.280
Black	0.089	0.284	0.126	0.331
Hispanic	0.195	0.397	0.182	0.386
White	0.616	0.486	0.604	0.489
Male	0.643	0.479	0.643	0.479
Female	0.355	0.478	0.355	0.478
Loan Amount	106	106	101	109
Applicant Income	286	210	262	199
Coborrower Present	0.354	0.478	0.476	0.499
Jumbo Loan Amount	0.280	0.449	0.189	0.392
Tract Median Income	60,041	23,427	60,436	23,543
Tract Pct Black	10.0	18.3	12.3	21.7
Tract Pct Hispanic	16.2	20.2	17.1	20.6
Tract Pct Asian	6.5	10.3	7.1	10.8
Tract Pct Owner Occ	68.3	23.9	69.2	22.7
Tract Pct in Poverty	7.71	7.27	7.96	7.49
Tract Price/Rent	0.004	0.002	0.004	0.002
Dataquick Data				
Loan to Value Ratio	0.885	0.242	0.576	0.416
Subordinate Lien	0.428	0.495	0.016	0.127
First Lien Adjustable Rate	0.515	0.500	0.478	0.500
condo	0.218	0.413	0.139	0.346
mobile	0.001	0.035	0.001	0.034
Single Family	0.774	0.418	0.837	0.369
lot size	15,675	575,974	15,051	397,858
Unit square feet	1,798	26,193	1,766	20,119
number of bathroom	1.99	6.60	2.00	1.15
number of bedroom	2.17	6.85	2.04	1.77
stories	1.17	1.58	1.22	1.33
units	1.41	14.83	1.49	19.88
Number of observations	327124		306213	

Notes: The first two columns contain means and standard deviations for our sample of post-mortgage credit reports for each home purchase mortgage borrower for which a credit score was observed prior to mortgage origination. The last two columns contain the same information for the post-mortgage reports of mortgage refinancers.

Table 3: Racial Differences in Credit Outcomes

	Home Purchase Sample		Refinance Sample	
Race	**Delinquencies**	**Foreclosures**	**Delinquencies**	**Foreclosures**
Black	0.030816***	0.041926***	0.015711***	0.016984***
	(0.002115)	(0.003613)	(0.001804)	(0.002377)
Hispanic	0.023411***	0.038697***	0.015130***	0.023853***
	(0.001558)	(0.002504)	(0.001359)	(0.001980)
Asian	0.003359**	0.005623**	0.007799***	0.011862***
	(0.001673)	(0.002466)	(0.002185)	(0.002760)
R-square	0.057	0.074	0.046	0.046

Notes: This table presents the OLS estimates for number of 90 to 180 day mortgage delinquencies and number of foreclosures within the last 12 months for the samples described in Table 2. The models include detailed controls for pre-origination vantage credit score, loan to value ratio, mortgage and total debt payment to income ratios, whether interest rate is adjustable, presence of subordinate debt, whether a jumbo loan, borrower income, race, ethnicity, gender, age, presence of coborrower, census tract demographics and detailed unit attributes, as well as origination year by credit year fixed effects. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 4: Credit Outcomes and Subprime Controls

Home Purchase Sample						
Model	**Race**	**Baseline**	**Rate Spread**	**Lender FE's**	**Lender and Tract FE**	**Add'l Subprime Controls**
Delinquencies	Black	0.030816***	0.027439***	0.026969***	0.023542***	0.023164***
		(0.002115)	(0.002106)	(0.002125)	(0.002315)	(0.002314)
	Hispanic	0.023411***	0.020849***	0.018841***	0.018888***	0.018885***
		(0.001558)	(0.001561)	(0.001574)	(0.001652)	(0.001646)
	Rate		0.036033***	0.034187***	0.033659***	0.033330***
	Spread		(0.002016)	(0.002464)	(0.002441)	(0.002461)
R-square		0.057	0.060	0.064	0.098	0.099
Foreclosures	Black	0.041926***	0.034397***	0.033230***	0.030188***	0.029918***
		(0.003613)	(0.003558)	(0.003587)	(0.003921)	(0.003910)
	Hispanic	0.038697***	0.032979***	0.029501***	0.030629***	0.030823***
		(0.002504)	(0.002521)	(0.002521)	(0.002749)	(0.002754)
	Rate		0.080369***	0.061783***	0.058779***	0.056879***
	Spread		(0.003226)	(0.004032)	(0.004001)	(0.004051)
R-square		0.074	0.080	0.086	0.123	0.123

Refinance Sample						
Model	**Race**	**Baseline**	**Rate Spread**	**Lender FE's**	**Lender and Tract FE**	**Add'l Subprime Controls**
Delinquencies	Black	0.015711***	0.014743***	0.013600***	0.012385***	0.012413***
		(0.001804)	(0.001802)	(0.001830)	(0.001904)	(0.001901)
	Hispanic	0.015130***	0.014763***	0.012914***	0.012425***	0.012668***
		(0.001359)	(0.001358)	(0.001366)	(0.001455)	(0.001456)
	Rate		0.018598***	0.017965***	0.018793***	0.018268***
	Spread		(0.001572)	(0.001941)	(0.001891)	(0.001908)
R-square		0.046	0.047	0.052	0.090	0.091
Foreclosures	Black	0.016984***	0.015506***	0.014384***	0.014745***	0.014787***
		(0.002377)	(0.002379)	(0.002399)	(0.002547)	(0.002549)
	Hispanic	0.023853***	0.023282***	0.021612***	0.020747***	0.020975***
		(0.001980)	(0.001978)	(0.001994)	(0.002116)	(0.002121)
	Rate		0.028534***	0.024762***	0.023953***	0.023098***
	Spread		(0.002178)	(0.002592)	(0.002523)	(0.002588)
R-square		0.046	0.047	0.052	0.096	0.096

Notes: The first column presents the estimates from Table 3. Panel 1 contains the estimates for the home purchase sample, and Panel 2 for the refinance sample. The second column adds a dummy variable to control for whether the interest rate was high cost or above the HMDA rate spread threshold, the third column presents estimates after also adding lender fixed effects, and the fourth column adds census tract fixed effects. The final column estimates are based on the same specification in the fourth column except that borrowers are divided between those with vantage scores less than or greater than or equal to 701. The resulting subprime dummy variable is interacted with key loan term variables based on loan to value ratio, mortgage and total debt payment to income ratios, whether interest rate is adjustable, and presence of subordinate debt. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 5: Credit Outcomes and Contemporaneous Controls

	Home Purchase Sample					
	Delinquencies			Foreclosures		
Race	Cnty-Year FE	Negative Equity and Employment	Race Specific Employoment	Cnty-Year FE	Negative Equity and Employment	Race Specific Employoment
Black	0.022986***	0.022892***	0.013658***	0.029607***	0.029410***	0.015540***
	(0.002316)	(0.002317)	(0.002870)	(0.003916)	(0.003918)	(0.004372)
Hispanic	0.018459***	0.018263***	0.015547***	0.030120***	0.029758***	0.025568***
	(0.001644)	(0.001646)	(0.001686)	(0.002747)	(0.002758)	(0.002829)
Rate Spread	0.032589***	0.031392***	0.031083***	0.055495***	0.053252***	0.052724***
	(0.002454)	(0.002454)	(0.002462)	(0.004044)	(0.004028)	(0.004005)
Negative Equity		0.519101***	0.443925***		1.059106***	0.752073***
		(0.058957)	(0.031764)		(0.094168)	(0.051406)
Emp*Neg Equity		-0.535920***			-1.102721***	
		(0.062170)			(0.099259)	
Race Emp*Neg Eq			-0.454918***			-0.774751***
			(0.033350)			(0.054018)
R-squared	0.105	0.106	0.107	0.133	0.134	0.135

	Refinance Sample					
	Delinquencies			Foreclosures		
Race	Cnty-Year FE	Negative Equity and Employment	Race Specific Employoment	Cnty-Year FE	Negative Equity and Employment	Race Specific Employoment
Black	0.012350***	0.012364***	.0047619 *	0.014788***	0.014908***	0.005883*
	(0.001902)	(0.001905)	(.0026721)	(0.002547)	(0.002548)	(0.003024)
Hispanic	0.012422***	0.012419***	.0106979***	0.020593***	0.020679***	0.018589***
	(0.001455)	(0.001460)	(.0015142)	(0.002119)	(0.002109)	(0.002144)
Rate	0.017991***	0.018173***	.0179044***	0.022746***	0.022834***	0.022708***
Spread	(0.001910)	(0.001915)	(.0019146)	(0.002591)	(0.002602)	(0.002609)
Negative Equity		0.184635***	.1647676***		0.230915***	0.142201***
		(0.049956)	(.0347953)		(0.057935)	(0.044460)
Emp*Neg Equity		-0.197139***			-0.243035***	
		(0.052326)			(0.060643)	
Race Emp*Neg Eq			-.1759232***			-0.149347***
			(.0363514)			(0.046510)
R-squared	0.096	0.097	0.097	0.102	0.103	0.103

Notes: The models in the first and fourth columns add county by credit profile year fixed effects to the last specification shown in Table 4. The models in the second and fifth columns also add a control for whether the borrower has negative equity based on the mortgage amount and a county specific hedonic housing price index, and the interaction between the negative equity dummy and a county specific employment rate for prime age males for each credit year. The models in the third and sixth columns replace the overall employment rate for each county and credit year with employment rate for prime age adults of each race or ethnicity. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 6: Race Effects by Subprime Credit Score

	Home Purchase Sample					
	Mortgage Trades			Foreclosures		
Race	OLS	Subprime Controls	Negative Equity and Employment	OLS	Subprime Controls	Negative Equity and Employment
Black	0.022166***	0.017827***	0.009136***	0.036955***	0.030976***	0.017679***
	(0.002405)	(0.002591)	(0.003101)	(0.004138)	(0.004580)	(0.005041)
Hispanic	0.025168***	0.021376***	0.017895***	0.036867***	0.030556***	0.024973***
	(0.001712)	(0.001807)	(0.001851)	(0.002886)	(0.003102)	(0.003167)
Black*Subprime	0.018113***	0.011277***	0.009912**	0.013229**	-0.000895	-0.003200
	(0.003747)	(0.003855)	(0.003863)	(0.006108)	(0.006041)	(0.006079)
Hispanic*Subprime	-0.004326	-0.006734**	-0.006278*	0.007685	0.001966	0.002898
	(0.003203)	(0.003323)	(0.003329)	(0.004759)	(0.004868)	(0.004864)
R-squared	0.058	0.099	0.107	0.075	0.123	0.135

	Refinance Sample					
	Mortgage Trades			Foreclosures		
Race	OLS	Subprime Controls	Negative Equity and Employment	OLS	Subprime Controls	Negative Equity and Employment
Black	0.009922***	0.006531***	0.000113	0.016697***	0.015956***	0.008201**
	(0.001922)	(0.002044)	(0.002795)	(0.002734)	(0.002841)	(0.003311)
Hispanic	0.017567***	0.014997***	0.013044***	0.023788***	0.020817***	0.018434***
	(0.001512)	(0.001621)	(0.001675)	(0.002222)	(0.002364)	(0.002400)
Black*Subprime	0.012236***	0.012533***	0.012980***	0.001079	-0.002246	-0.002695
	(0.002954)	(0.003004)	(0.003022)	(0.003771)	(0.003786)	(0.003811)
Hispanic*Subprime	-0.007245**	-0.007015**	-0.006310**	0.000546	0.000664	0.001243
	(0.002863)	(0.002870)	(0.002890)	(0.003782)	(0.003985)	(0.003996)
R-squared	0.046	0.091	0.097	0.046	0.096	0.103

Notes: The models in Table 6 are based on the models in columns 3 and 6 of Table 5 plus the inclusion of interactions between a dummy for whether the vantage score is 701 or higher with borrower race and ethnicity. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 7: Credit Outcomes and Origination Year (OLS)

Black Estimates				
	Delinquencies		Foreclosures	
Race Interactions	**Home Purchase**	**Refinance**	**Home Purchase**	**Refinance**
Level	0.028680***	0.020363***	0.042366***	0.038888***
	(0.005630)	(0.005529)	(0.010005)	(0.007163)
2008	-0.030374***	-0.023977***	-0.054060***	-0.050425***
	(0.006540)	(0.006136)	(0.011023)	(0.007394)
2007	0.003047	-0.008192	-0.024935**	-0.031753***
	(0.006785)	(0.006155)	(0.011110)	(0.007693)
2006	0.034841***	0.014496**	0.054604***	-0.015939*
	(0.007645)	(0.007027)	(0.013140)	(0.008706)
2005	0.018363**	0.007660	0.013770	-0.008627
	(0.007738)	(0.006634)	(0.012908)	(0.008289)
R-squared	0.072	0.053	0.103	0.06
Observations	204844	190844	204844	190844

Hispanic Estimates				
	Delinquencies		Foreclosures	
Ethnicity Interactions	**Home Purchase**	**Refinance**	**Home Purchase**	**Refinance**
Level	0.021869***	0.022834***	0.040405***	0.035913***
	(0.004992)	(0.004902)	(0.009293)	(0.006843)
2008	-0.018115***	-0.020398***	-0.047709***	-0.044322***
	(0.006113)	(0.005647)	(0.009911)	(0.007272)
2007	0.024925***	0.007498	0.008693	-0.014604*
	(0.006201)	(0.006075)	(0.011466)	(0.007464)
2006	0.029054***	0.006854	0.065120***	0.016418**
	(0.006722)	(0.006265)	(0.011730)	(0.008267)
2005	0.021854***	0.003149	0.032848***	0.016267*
	(0.006776)	(0.006238)	(0.010851)	(0.008330)
R-squared	0.072	0.053	0.103	0.06
Observations	204844	190844	204844	190844

Notes: The models in Table 7 are based on the OLS models in Table 3 plus interactions of origination year dummy variables with borrower race and ethnicity. The first panel presents the estimates of origination year interacted with borrower black, and the second panel presents the estimates of origination year interacted with borrower Hispanic. The first row in each panel labeled Level represents the effect for the omitted category, 2004 originations. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 8: Negative Equity and Employment Controls

Black Estimates				
	Delinquencies		Foreclosures	
Race Interactions	**Home Purchase**	**Refinance**	**Home Purchase**	**Refinance**
Level	0.018953***	0.010551*	0.034470***	0.033437***
	(0.006482)	(0.005951)	(0.011669)	(0.007486)
2008	-0.031404***	-0.024726***	-0.050213***	-0.049291***
	(0.007141)	(0.006707)	(0.012362)	(0.007628)
2007	0.003604	-0.009586	-0.031485***	-0.034685***
	(0.007193)	(0.006417)	(0.012078)	(0.007723)
2006	0.026359***	0.011544	0.027268*	-0.021409**
	(0.007739)	(0.007079)	(0.014066)	(0.008405)
2005	0.012104	0.005422	-0.009892	-0.016287**
	(0.007943)	(0.006742)	(0.013599)	(0.007984)
R-squared	0.139	0.131	0.181	0.145
Observations	202298	186,066	202298	186,066

Hispanic Estimates				
	Delinquencies		Foreclosures	
Ethnicity Interactions	**Home Purchase**	**Refinance**	**Home Purchase**	**Refinance**
Level	0.016247***	0.014552***	0.036669***	0.028059***
	(0.005184)	(0.004923)	(0.010347)	(0.007381)
2008	-0.026338***	-0.022186***	-0.060941***	-0.051474***
	(0.006581)	(0.005932)	(0.010699)	(0.008012)
2007	0.024022***	0.010803*	0.009659	-0.008811
	(0.006376)	(0.005878)	(0.012251)	(0.008092)
2006	0.023450***	0.008561	0.042193***	0.014595*
	(0.006596)	(0.006060)	(0.012263)	(0.008490)
2005	0.015738**	0.005834	0.013017	0.020627**
	(0.006665)	(0.006003)	(0.011217)	(0.008417)
R-squared	0.139	0.131	0.181	0.145
Observations	202298	186,066	202298	186,066

Notes: The models in Table 8 are based on the models in Table 5 columns 3 and 6 plus interactions of origination year dummy variables with borrower race and ethnicity. The first panel presents the estimates of origination year interacted with borrower black, and the second panel presents the estimates of origination year interacted with borrower Hispanic. The first row in each panel labeled Level represents the effect for the omitted category, 2004 originations. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 9: Foreclosure Models with Home Purchase Origination for Refinance Subsample

	OLS Model		Negative Equity	
Race Interactions	**Black**	**Hispanic**	**Black**	**Hispanic**
Level	0.017925***	0.008026*	0.012620**	0.008115*
	(0.004444)	(0.004214)	(0.005051)	(0.004617)
2008	-0.017951	0.002973	-0.006973	-0.046481
	(0.012054)	(0.019965)	(0.020435)	(0.037424)
2007	0.053288	0.107388***	0.046266	0.102502**
	(0.035553)	(0.040368)	(0.036444)	(0.040059)
2006	0.028865	0.017662	0.025089	0.009102
	(0.024466)	(0.014825)	(0.024910)	(0.015319)
2005	0.028824**	0.034287***	0.028214**	0.026152***
	(0.011744)	(0.008704)	(0.011846)	(0.008737)
2004	0.010658	0.026000***	0.009023	0.024183***
	(0.006974)	(0.006249)	(0.007089)	(0.006423)
2003	0.008739	0.011526**	0.006163	0.008807
	(0.006990)	(0.005739)	(0.006794)	(0.005865)
2002	0.008027	0.008392	0.010266	0.004589
	(0.006469)	(0.005924)	(0.006913)	(0.006466)
2001	0.000570	0.007005	0.003163	0.004757
	(0.006268)	(0.005923)	(0.006668)	(0.006447)
2000	-0.004354	-0.002058	-0.005725	-0.004767
	(0.005817)	(0.006254)	(0.006228)	(0.006255)
1999	0.004230	-0.005664	0.002234	-0.008100
	(0.006568)	(0.005606)	(0.006709)	(0.006280)
1998	-0.007833	0.004234	-0.007542	0.005084
	(0.006442)	(0.006132)	(0.006813)	(0.006954)
R-squared	0.048	0.048	0.105	0.105
Observations	306,213	306,213	300,082	300,082

Notes: The first two columns present the OLS estimates based on the model in Table 3 plus the interactions of the original home purchase year with race and ethnicity for a sample of refinance mortgages where the original home purchase occurred between 1997 and 2008. The dependent variable is number of foreclosures in the last 12 months for the 2008 and 2009 credit years only. The first panel presents the estimates of home purchase year interacted with borrower black, and the second panel presents the estimates of home purchase year interacted with borrower Hispanic. The first row in each panel labeled Level represents the effect for the omitted category, 1997 originations. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Table 10: Foreclosure Negative Equity Models for Prime Borrowers

	Home Purchase		Refinance	
Race Interactions	**Black**	**Hispanic**	**Black**	**Hispanic**
Level	0.020717	0.041393***	0.015688	0.031470**
	(0.017068)	(0.014579)	(0.009936)	(0.013622)
2008	-0.031124*	-0.067384***	-0.031704	-0.014629
	(0.018209)	(0.014491)	(0.037829)	(0.031693)
2007	-0.012031	0.005029	-0.011940	0.114720**
	(0.016941)	(0.016864)	(0.022342)	(0.047809)
2006	0.031988	0.046577***	0.045195	0.002132
	(0.019512)	(0.017218)	(0.030263)	(0.016301)
2005	0.008042	0.001419	0.055406***	0.021339*
	(0.019242)	(0.015471)	(0.019379)	(0.012734)
2004			0.015458	0.036888***
			(0.014035)	(0.012495)
2003			0.021263	0.003455
			(0.014147)	(0.011321)
2002			0.023861	0.010256
			(0.017755)	(0.012638)
2001			0.023333	-0.004721
			(0.015028)	(0.013581)
2000			-0.000736	-0.014842
			(0.012834)	(0.012840)
1999			0.027290*	-0.020838*
			(0.016373)	(0.012519)
1998			-0.010367	-0.007253
			(0.011529)	(0.014541)
R-squared	0.195		0.164	
Observations	146,693		133,441	

Notes: This table presents the results of number of foreclosures regressed on the interactions of home purchase year with race and ethnicity for 2008 and 2009 credit years and for a subsample of borrowers with credit scores above 701. The first two columns present the estimates for the home purchase mortgage sample based on the Negative Equity and Race Specific Employment models presented in Table 8 for number of foreclosures. The last two columns present the estimates for the refinance mortgage sample based on the Negative Equity and Race Specific Employment models presented in columns 3 and 4 of Table 9. Standard errors (in parentheses) are robust to heteroscedasticity, and clustered at tract-credit year level. *** indicates significant at 1% level, ** means significant at 5% level, and * is significant at 10% level.

Figure 1. Unconditional rates of mortgage delinquency by race and year

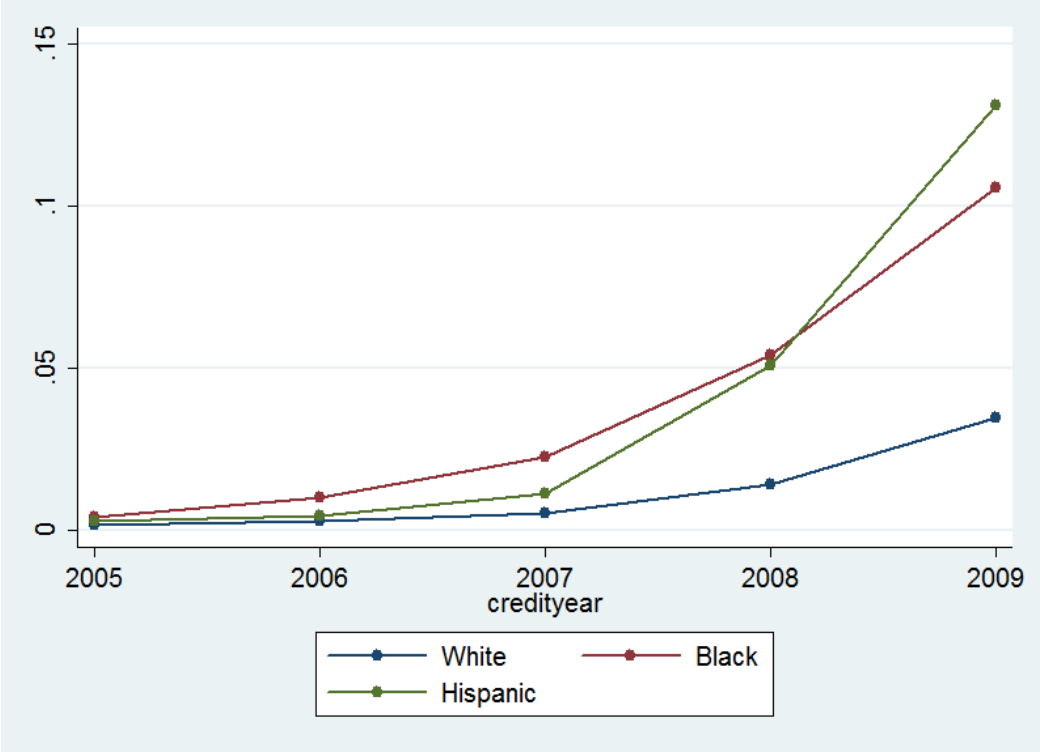

Figure 2. Unconditional rates of mortgage foreclosures by race and year

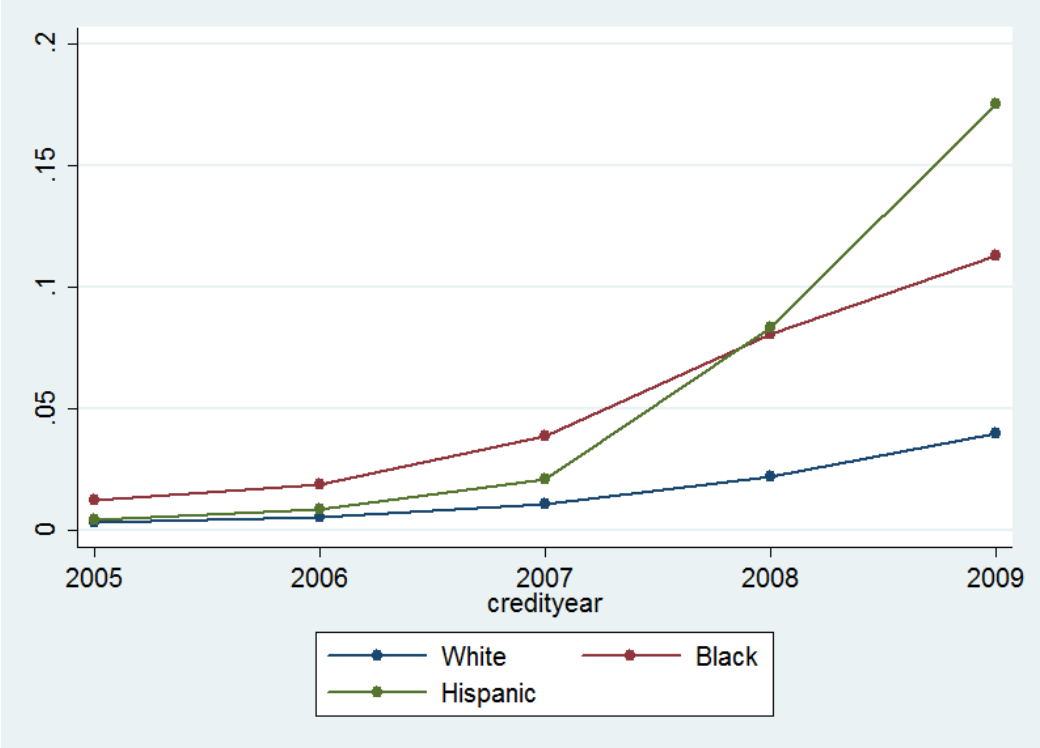

Figure 3: Homeownership and Default Decisions in (p, r) Space

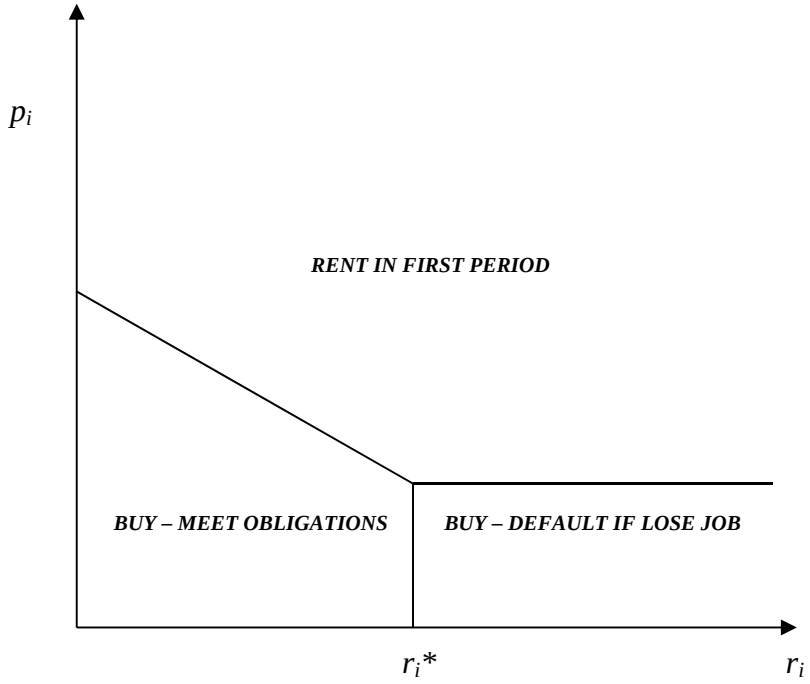

Figure 4: Effect of Cheaper Mortgage Credit on Homeownership and Default

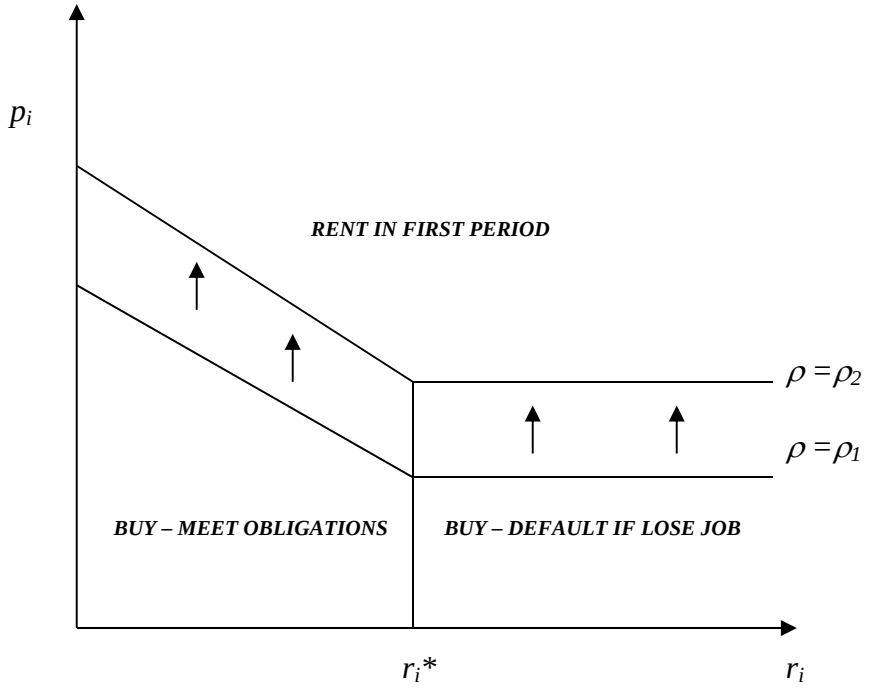

Figure 5. Unconditional rates of mortgage delinquency by race and origination year

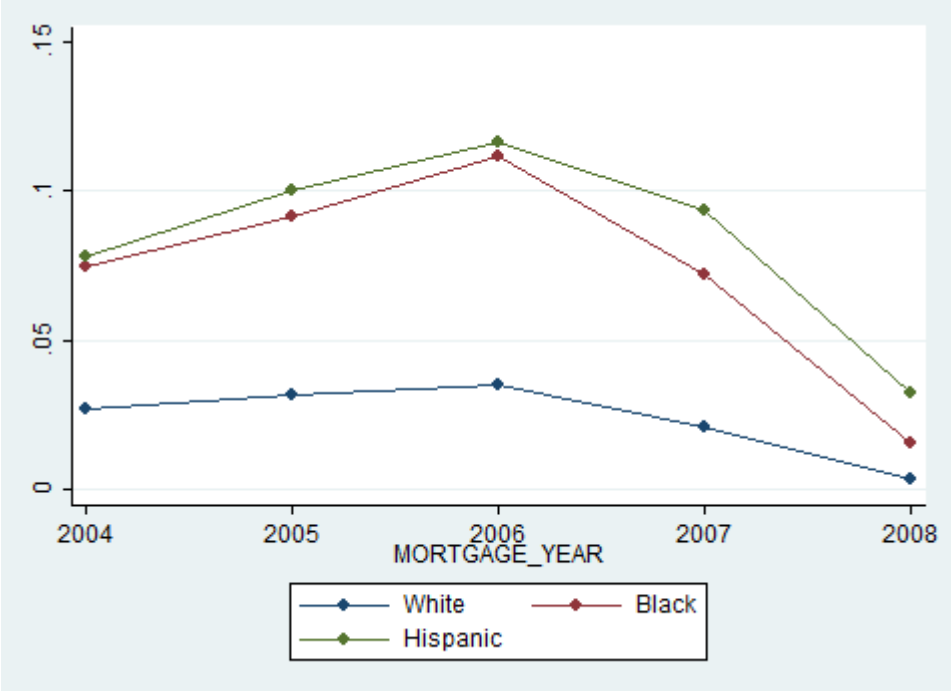

Figure 6. Unconditional rates of mortgage foreclosures by race and origination year