FURMAN CENTER FOR REAL ESTATE & URBAN POLICY
NEW YORK UNIVERSITY SCHOOL OF LAW • WAGNER SCHOOL OF PUBLIC SERVICE
40 Washington Square South, Suite 314, New York, NY 10012 · Tel: (212) 998 6713 · Fax: (212) 995 4341 · furmancenter.org

The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City

Working Paper 2010

Sewin Chan, Michael Gedal, Vicki Been, Andrew Haughwout

NYU Wagner School and
Furman Center for Real Estate & Urban Policy

The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City

Sewin Chan
Robert F. Wagner Graduate School of Public Service, New York University
Contact: sewin.chan@nyu.edu

Michael Gedal
Robert F. Wagner Graduate School of Public Service, New York University
Contact: michael.gedal@nyu.edu

Vicki Been
New York University School of Law
Contact: vicki.been@nyu.edu

Andrew Haughwout
Federal Reserve Bank of New York
Contact: andrew.haughwout@ny.frb.org

June 29, 2010

The views represented here are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System. We thank Caitlin Coslett, Johanna Lacoe, Mary Weselcouch and Jackie Begley for their assistance on this paper. This paper has benefited from comments by Ingrid Gould Ellen, Joseph Tracy and Josiah Madar, as well as participants at the NYU Law School Faculty Brownbag Workshop, the 2010 Policy Summit at the Federal Reserve Bank of Cleveland, and the annual meetings of the American Law and Economics Association, the Association for Public Policy Analysis and Management, and the Eastern Economic Association. We are also grateful to the staff at the Federal Reserve Bank of New York for their help with the LoanPerformance data and their patience in answering our many questions, as well as to the Public Data Corporation and the New York City Department of Finance for their assistance in providing the additional data needed for this project.

Correspondence should be sent to sewin.chan@nyu.edu.

The Role of Neighborhood Characteristics in Mortgage Default Risk:

Evidence from New York City

Abstract

We construct a database of non-prime hybrid adjustable and fixed rate mortgages from New York City that augments a rich set of loan and borrower risk characteristics with a variety of census tract level neighborhood characteristics. We find that these neighborhood characteristics are important for default behavior, even after an extensive set of controls. First, default rates increase with the rate of foreclosure notices and the number of lender-owned properties (REOs) in the tract. Second, default rates for home purchase mortgages are higher in predominantly black tracts, regardless of the borrower's own race. We explore possible explanations for our findings.

JEL classification: G2, R1

1. Introduction

The wave of delinquencies and foreclosures that began in 2007 and the financial crisis that it engendered have drawn new attention to the different reasons why households may end up in foreclosure. In this paper, we use a newly assembled dataset from New York City to examine the role of borrower, loan and neighborhood characteristics on default rates of non-prime mortgages. The depth of the dataset allows us to provide a more complete set of controls than previous research. In particular we are able to examine the effects of census tract level neighborhood characteristics while controlling for detailed individual borrower and loan characteristics. Our analysis of neighborhood effects casts light on whether public policies to address the foreclosure crisis and to regulate lending practices should be tailored for different types of neighborhoods.

Many researchers have used LoanPerformance from FirstAmerican CoreLogic, a commercial database that is the major source of non-prime mortgage performance information for the mortgage industry. A major limitation of this database is that its most detailed geographic identifier is the zip code of the mortgaged property. A zip code is a good deal bigger than what is generally thought of as a "neighborhood." Researchers studying neighborhoods typically examine census tracts: in New York City there are about 2,100 census tracts, compared to 200 zip codes.

We have matched LoanPerformance records to actual parcels of land with a high level of precision using real property deeds from New York City's Department of Finance. This allows us to merge information about borrowers, their payment histories, and the terms of their loans, with census tract level neighborhood characteristics from a variety of sources, mortgage applicant information from Home Mortgage Disclosure Act (HMDA) records, as well as property characteristics from city tax records. We are not aware of any other mortgage research that has examined such detailed information at so fine a level of geography, and that contains a full set of critical loan and borrower characteristics, especially the borrowers' credit scores.

This paper also advances the literature by using repeat sales house price indices by community district, a political jurisdiction within New York City. Our sample includes 56 of the 59 such districts in the City, averaging just over four square miles each.[1] This is a much finer geography than in the widely-used Case-Shiller or Federal Housing Finance Agency house price indices.

Importantly for our analysis, we also have precise information at the census tract level on the number of foreclosures and the share of properties that are owned by lenders (banks and other mortgage investors). The latter are the so-called "real estate owned" properties, or REOs, that fail to sell for the lender's reservation price at a foreclosure auction. The foreclosure and REO measures allow us to examine whether foreclosures are contagious, an important policy concern given the geographic concentration of foreclosure across the country.

We report results on the effect of loan and borrower risk characteristics on mortgage default. Not only do these estimates broadly confirm earlier results from the literature while using the additional controls provided by our much richer dataset, the magnitude of these estimates do not change when we add the additional controls. This is an important finding that confirms the validity of existing research on the effect of loan and borrower risk characteristics on non-prime mortgage defaults that do not have the wealth of very local neighborhood level information that we are able to use here.

The primary contribution of the paper, however, is our finding that the neighborhood in which a mortgage holder lives has a powerful impact on the likelihood of default. Most importantly, mortgage holders living in areas with high foreclosure and REO activity have a substantially higher chance of falling behind on their mortgages, as do mortgage holders living in predominantly black neighborhoods. These effects exist even after controlling for loan and borrower characteristics, including the interest rates paid by borrowers, borrower credit score and the borrower's own race.

[1] Our analysis of New York City includes Manhattan, the Bronx, Brooklyn and Queens, but excludes Staten Island.

In the body of the paper, we argue that the foreclosure and REO activity findings reflect a local contagion effect whereby more information about the default process, or reduced stigma surrounding it, may lead to more defaults. In addition, the foreclosure and REO rates may serve as a proxy for very local housing market conditions that are not captured in our community district level house price indices. They also may serve, in part, as proxies for unobserved borrower characteristics or very local economic conditions. Further, we argue that our findings on predominantly black neighborhoods suggest that there may be unobserved loan and borrower characteristics that are both correlated with higher default risk and are more prevalent in black neighborhoods. This is consistent with the non-prime mortgage industry treating black neighborhoods differently from non-black neighborhoods. For all of these interpretations, we consider various measurement error issues, but conclude that measurement alone error cannot be responsible for our findings.

2. Background

There is an extensive and growing body of empirical literature on the determinants of mortgage default.[2] Our goal in this section is to provide some background and context to our analysis, which focuses on the role of borrower, loan and neighborhood characteristics on the default of non-prime mortgages.

Non-prime mortgages have traditionally extended credit to borrowers who could not qualify for prime mortgages, and so by their very nature have higher default risk than do prime mortgages. Okah and Orr (2010) compare non-prime and prime mortgages in New York City and find that on average, non-prime borrowers have, at origination, much lower credit scores, higher debt-to-income

[2] See Mayer et al. (2009) for a recent review.

ratios (DTIs) and loan-to-value ratios (LTVs), as well as substantially higher rates of 90-day default.[3] Data from other locations show similar patterns, and several studies find that non-prime status is one of the strongest predictors of default and foreclosure (Coulton et al., 2008; Ding et al., 2008; Gerardi et al., 2007).

2.1. The Relationship Between Loan and Borrower Characteristics and Default Risk

Both the research on prime and non-prime mortgages has stressed the importance of credit scores, DTIs and LTVs on default behavior. Many earlier studies on mortgage default were not able to observe borrower credit scores and researchers routinely recognize this omission to be problematic. The studies that do have this information find that credit scores play a large role in predicting default (Demyanyk, 2009; Foote et al., 2009; Haughwout et al., 2008). Several studies have found that higher initial DTIs contribute to a higher probability of default, although the effects seem to be less strong than that of LTV, and are somewhat inconsistent over time (Ding et al., 2008; Foote et al., 2009; Haughwout et al., 2008). Recent research on non-prime mortgages has emphasized that it is the current, and not the initial LTV that increases the probability of delinquency and default (for example, Demyanyk, 2009 and Haughwout et al., 2008). Foreclosures happen less frequently in appreciating markets, most likely because financially-distressed borrowers can more easily sell their properties or refinance and prepay the remaining balance on their loans (Danis & Pennington-Cross, 2005; Haughwout et al., 2008; Schloemer et al., 2006).

Few models of mortgage default using loan-level data are able to identify and control for borrower race. In studies that do, there is some evidence that black borrowers are more likely to default than white borrowers. Coulton et al. (2008) identify large racial disparities in the probability of receiving a subprime loan, a strong connection between subprime loan status and the probability of foreclosure, and higher rates of foreclosure for black borrowers than for white borrowers, but

[3] Their sample is outstanding loans in January 2009 that were originated during 2004-2007. Data on non-prime loans are from LoanPerformance while data on prime loans are from LP Applied Analytics (formerly McDash).

they are unable to control for borrowers' credit history. Jiang et al. (2009) estimate that delinquency rates are higher for black and Hispanic borrowers, compared to white borrowers and that the difference in the delinquency rates is even higher among broker-originated loans than bank-originated loans.

2.2. The Relationship Between Neighborhood Characteristics and Default Risk

Considerable research documents the negative externalities neighborhoods experience from foreclosures, particularly in terms of the values of nearby housing (for example, Campbell et al., 2009; Rogers & Winter, 2009; Schuetz et al., 2008). Several studies propose mechanisms by which nearby foreclosures could lead to a heightened probability of default or foreclosure. Negative physical externalities caused by foreclosures, including visible deterioration, maintenance deferral or vandalism, may cause declines in neighboring home values (Campbell et al., 2009; Harding et al., 2000; Hartley, 2010; Lee, 2008; Leonard & Murdoch, 2009) and homeowners faced with foreclosure may stop investing in activities that benefit the neighborhood, causing potential homebuyers to view the neighborhood as less attractive (Harding et al., 2009). Foreclosure-induced mobility may increase the number of homes on the market, and thereby drive prices down (Harding et al., 2009; Hartley, 2010; Lee, 2008). Further, the sale of foreclosed properties at discounted prices may lower property appraisals based on comparables for nearby properties (Immergluck & Smith, 2006; Lee, 2008; Lin et al., 2009). There is some evidence that the values individuals place on their homes are influenced by the home valuations of their closest neighbors, so declining property values due to foreclosure may cause neighbors to lower their reservation prices (Harding et al., 2009; Ioannides, 2003).

Research on whether neighborhood racial composition influences default is rather thin, and the results somewhat inconsistent (Berkovec et al., 1994; Firestone et al., 2007). Neighborhood level studies that analyze aggregate default and foreclosure patterns, instead of the behavior of individual borrowers, tend to find a correlation between the share of minority residents or homeowners in a census tract and the tract's share of foreclosures or defaults. These studies concede, however, that minority share in a census tract may merely be a proxy for credit history, wealth, or economic conditions (Berkovec et al., 1994; Pedersen & Delgadillo, 2007), and Van Order & Zorn (2000) find that when credit history is controlled for, the effect of the minority share disappears.[4]

3. Data Description

To investigate the determinants of default, we begin with all first lien hybrid 2/28 and 3/27 adjustable rate mortgages (ARMs)[5] and 30-year fixed rate mortgages (FRMs) originated in New York City from 2004 to 2007 in LoanPerformance, a database of that covers almost 100 percent of all non-prime securitized mortgages in the United States over this time period.[6] We observe monthly updates on these loans until December 2009. Although LoanPerformance provides detailed information on borrower characteristics, loan terms, and payment history, it contains no information on borrower race or gender and provides little in terms of property or neighborhood characteristics. We therefore supplement the loan level data with information from multiple sources.

First, we attach some additional borrower characteristics, including race and ethnicity, from Home Mortgage Disclosure Act (HMDA) data. Second, we incorporate information on whether the borrower took on additional mortgage debt following loan origination, obtained from the New York City Department of Finance (DOF)'s Automated City Register Information System (ACRIS). Third, we merge information from the DOF's Real Property Assessment Database (RPAD) on building characteristics. Fourth, we merge in repeat sales house price indices from The Furman Center for

[4] See also Apgar & Duda (2005) and Grover et al (2008).
[5] 2/28s are 30-year loans with an initial rate that remains in effect for the first 2 years, and then is reset every six months for the remaining 28 years, while on a 3/27, the initial rate is in effect for 3 years and floats for 27 years.
[6] These hybrid ARMs and FRMs represent almost two thirds of all first lien LoanPerformance mortgages in New York City. Because LoanPerformance includes only securitized loans, any inferences should be limited to these types of loans.

Real Estate and Urban Policy that track appreciation in 56 different community districts of New York City.[7] Fifth, we link information on the demographic characteristics of census tracts using the 2000 Census. Sixth, we add the share of all originations that are non-prime, and loan application denial rates at the census tract level using HMDA data. Seventh, we add the rate of mortgage foreclosure notices (*lis pendens*) and the share of properties that are owned by lenders (REOs) at the census tract level.[8] Finally, we attach monthly crime rates at the census tract level using a database obtained from the New York City Police Department.

To match loan level information from the LoanPerformance database to these diverse sources, we relied on mortgage deeds contained within ACRIS. Using a hierarchical matching algorithm, we were able to match 93 percent of the loans in the LoanPerformance database back to the deeds records, which thus gave us the exact location of the mortgaged property and allowed us to merge on the additional information noted above.[9]

---

[7] See Armstrong et al. (2009) for a description. We transform quarterly indices into monthly series by linear interpolation.

[8] The *lis pendens* are from Public Data Corporation and the REOs are from DOF property sales data.

[9] Our procedure for matching LoanPerformance to ACRIS is similar to the method used by Haughwout et al. (2009) to match LoanPerformance to HMDA. Our data from ACRIS do not include Staten Island and thus we had to drop this borough from our analysis. We merged LoanPerformance loans to ACRIS mortgage deeds using three common fields: origination or deed date, loan amount and zip code, using six stages of hierarchical matching. At the end of each stage, loans and deeds that uniquely matched each other were set aside and considered matched, while all other loans and deeds enter the next stage. Stage 1 matched loans and deeds on the raw values of date, loan amount and zip code. Stage 2 matched the remaining loans and deeds on the raw values of date and zip code, and the loan amount rounded to $1,000. Stage 3 matched on the raw values of date and zip code, and the loan amount rounded to $10,000. Stage 4 matched on the raw values of zip code and loan amount, and allowed dates to differ by up to 90 days. Stage 5 matched on the raw value of zip code, loan amount rounded to $1,000, and allowed dates to differ by up to 90 days. Stage 6 matched on the raw value of zip code, loan amount rounded to $10,000, and allowed dates to differ by up to 90 days. We believe it is valid to introduce a 90-day window because for a good fraction of LoanPerformance loans, the origination date is imputed by backdating the first payment date by one month, and in ACRIS, there may be administrative lags in the recording of the deeds data. The chance of false positive matching is low because we are matching loans to the full universe of deed records, and only considering unique matches.

In the analysis below, we use the 78 percent of LoanPerformance hybrid ARMs and FRMs that matched both the deeds records and a unique loan in the HMDA database. [10] This 78 percent sample is not significantly different from the full universe in terms of the loan, borrower, and neighborhood characteristics that we use in the analyses below. Figure 1 plots the number of originations by quarter in our final sample. The majority of originations occurred in 2004 and 2005 and the sample is split roughly equally between FRMs and hybrid ARMs, with approximately three-quarters of the ARMs being 2/28s. Home purchases make up 39 percent of ARMs and 28 percent of FRMs.

3.1. Loan and Property Characteristics

Table 1 displays summary statistics for loan and borrower characteristics. Average loan amounts are similar across the two loan types. About 90 percent of borrowers claim to be owner-occupants, although these self reports may be unreliable as owner-occupiers tend to get lower cost mortgages than investors, all else equal. The relative interest rate at origination is calculated as the interest rate minus the six-month London Interbank Offered Rate (LIBOR) at origination for ARMs, and the interest rate minus the Freddie Mac average interest rate for prime 30-year fixed rate mortgages during the month of origination for FRMs. The ARM margin is the amount added to the six-month LIBOR to determine the interest rate at future rate adjustments; the margin is between 6 and 7 percent for almost half of the ARMs in our sample. Payment shock at the time of the first ARM readjustment represents the jump in monthly payments that 2/28 borrowers experience at month 25 and 3/27 borrowers experience at month 37. The vast majority of loans experience a payment shock of less than 20 percent at this time. While payments will also adjust every six

[10] We merged HMDA records to ACRIS deeds based on date, loan amount and census tract, using the same six stage hierarchical matching technique as for the LoanPerformance-ACRIS match. We then uniquely paired the LoanPerformance records with HMDA records based on the unique deed identification number from ACRIS. While other researchers have matched loan level data (such as LoanPerformance) directly to HMDA by using the zip code as a common geographic identifier, our matching strategy is likely more reliable as it uses a more precise common geographical identifier (census tract).

months following the first adjustment, these subsequent adjustments are generally tiny in comparison because of the wide-spread use of teaser rates for the initial fixed period.

In terms of borrower risk, ARM borrowers tend to have lower FICO scores,[11] higher debt-to-income ratios (DTIs) and higher combined loan-to-value ratios (LTVs) at origination than FRM borrowers. The combined LTV measure is reported in LoanPerformance and it combines the loan amounts for the first lien mortgage (the focus of our default analysis) as well as any other liens in existence at the time of origination. Any new liens taken out afterwards will not be reflected in this measure; however, we have information on these liens from ACRIS: 5 percent of ARM borrowers and 11 percent of FRM borrowers took on additional debt against the same property (in any form, including second mortgages, home equity loans and lines of credit) that totaled at least 5 percent of the first lien's original loan amount. About 40 percent of borrowers provided full loan documentation.

We rely on HMDA for some additional borrower characteristics: a majority of primary borrowers in our sample are male, and 21 percent of ARMs and 28 percent of the FRMs included a coborrower. Almost half of ARM borrowers and over one third of FRM borrowers are black.[12] By contrast, blacks made up just 20 percent of borrowers who originated a loan in New York City during 2004-2007, according to HMDA.

Finally, our analysis also relies on measures of building characteristics (these statistics are not displayed in table 1). The majority of loans are collateralized by 2-4 family buildings (60 percent), followed by single-family (35 percent), condominiums (5 percent), and 5+ family buildings (1 percent). The distribution across building types is similar for ARMs and FRMs. The median

---

[11] The Fair Isaac Corporation (FICO) credit score is the most widely used credit score model in the United States. The exact model for calculating the score is a trade secret, but it depends on payment history, credit utilization, length of credit history, types of credit used and recent searches for credit.

[12] For the remainder of the paper, we refer to these non-missing, mutually exclusive race/ethnicity categories - white non-Hispanic, black non-Hispanic, Hispanic white, Asian non-Hispanic, and all others - as "white," black," "Hispanic white," "Asian," and "other," respectively. The "other" category can be further disaggregated to Hispanic black (about 1 percent of the entire sample), Hispanic other (about 3 percent), and other non-Hispanic (about 2 percent).

building age is 80 years for ARMs and 77 years for FRMs - comparable to a median age of 76 years for all residential properties in New York City. A relatively small share (about 5 percent) of structures was built in the 10 years preceding loan origination.

3.2. Neighborhood Characteristics

Figure 2 provides a snapshot of the distribution of foreclosures across New York City in 2009. Each dot represents one notice of foreclosure for a 1-4 family building. The map shows that foreclosures are concentrated, especially in neighborhoods where the majority of residents are black, according to the 2000 Census.[13] To investigate the link between neighborhood racial composition and foreclosures, we constructed a neighborhood foreclosure rate measure defined as the number of foreclosure notices (*lis pendens*) issued on 1-4 family buildings in a census tract during a six-month period, divided by the stock of 1-4 family buildings in that tract.[14] Figure 3 shows average neighborhood foreclosure rates on a semiannual basis, broken down by neighborhood racial composition. Overall, neighborhood foreclosure rates in New York City have trended upwards since 2004. In tracts where more than 60 percent of residents are black, the average foreclosure rate in the first half of 2004 was 0.4 percent, almost double the rate in tracts that are less than 40 percent black. By the last half of 2009, foreclosure rates in black neighborhoods had doubled to about 0.8 percent, compared to 0.4 percent for non-black neighborhoods.

Table 2 summarizes the distribution of loans in our sample across neighborhoods in terms of demographics in 2000. Just over half of loans were originated in census tracts where the median household income was greater than $40,000. The share of adults with a high school diploma was somewhat higher on average for FRMs than for ARMs. The average share of non-native born

---

[13] Throughout the paper, we classify neighborhoods based on the distribution of residents by race/ethnicity as follows: "black" residents include non-Hispanic blacks, "white" residents include non-Hispanic whites, "Asian" residents include non-Hispanic Asians, and "Hispanic" residents include all individuals reporting Hispanic origin regardless of race.

[14] Our neighborhood foreclosure measure is based on a sample that differs in three important ways from the sample of loans we analyze: it includes only 1-4 family homes, it is drawn from the full universe of active mortgages for 1-4 family homes, and it may include loans that were originated before 2004.

residents in a neighborhood was 37 percent for both FRMs and ARMs. A considerable number of loans in our sample are in tracts that were at least 60 percent black in 2000: almost half of ARMs and 41 percent of FRMs. This may be surprising considering that blacks made up just one quarter of New York City residents; but it likely reflects both the higher proportion of blacks in our non-prime sample (as noted earlier), and the fact that there are relatively high levels of residential racial segregation in New York City, so that black borrowers are more likely than non-black borrowers to live in neighborhoods that have a high concentration of black residents. In tracts that are predominantly non-black, FRMs are more prevalent than ARMs, while in predominantly black tracts, the majority of loans are ARMs.

Table 2 also considers the neighborhood share of non-prime loans at origination, calculated as the fraction of non-prime loans originated in LoanPerformance divided by total loans originated in HMDA, during the 2 years preceding the loan's origination month. The vast majority of loans in our sample are in tracts where fewer than 30 percent of mortgages were non-prime in the 2 years preceding origination.

The final two measures in table 2 are measured across all loan-months in our sample. While the majority of loan-month observations are in census tracts that experienced a foreclosure rate of less than one percent in the six months prior, about one in five loan-months were in neighborhoods where the foreclosure rate was 2 percent or more. The dynamic neighborhood REO rate is calculated as the number of properties listed as being in REO at any point in time during the six months preceding the month of analysis, divided by the total number of 1-4 family buildings in the census tract. Thus, the REO rate measures the *stock* of properties that were in REO at any point during the preceding six months while the foreclosure rate measures the *flow* of new properties into foreclosure during the preceding six months. About two thirds of loan-months were in census tracts where the REO rate was less than 1 percent, while 14 percent of loan-months experienced neighborhood REO rates above 3 percent.[15]

3.3. Borrower Race and Neighborhood Racial Composition

In our analyses below, we find that non-black home purchase borrowers in predominantly black neighborhoods have a risk of default that is at least as high as black borrowers in those same neighborhoods, even after controlling for an extensive set of borrower, loan, and neighborhood risk characteristics. Table 3 compares loan characteristics for black and non-black home purchase borrowers living in census tracts that are at least 60 percent black. For both ARMs and FRMs, the average FICO score is slightly lower, while average DTI and combined LTV are slightly higher for blacks. Based on findings from previous research, this suggests that black borrowers living in black neighborhoods are somewhat more likely to enter default than non-black borrowers living in black neighborhoods. However, black borrowers in black neighborhoods are considerably *more* likely to have a fully documented loan than non-black borrowers in those neighborhoods, and blacks are much less likely to take on additional mortgage debt. We might expect these factors to lower the default risk for blacks relative to non-blacks in predominantly black neighborhoods.

3.4. Default Rates

We examine the hazard of default for each month since origination, with default defined as 90 days of delinquency. This definition is used in much of the literature on mortgage default risk as it is entirely in the borrower's control and excludes the behavior of the lender or servicer. The hazard is simply the probability that a loan enters 90-day default, conditional on not having defaulted earlier. Figure 4 shows hazard rates up to December 2009, by the year the loan was originated. Two clear patterns emerge. First, ARMs tend to experience much higher default rates

---

[15] The higher stock of REOs compared with the flow of foreclosures reflects a large fraction of REO properties remaining in REO for long periods of time. For example, of properties that entered REO in 2007, fewer than half had exited within 12 months (Furman Center 2010).

on average than FRMs that were originated in the same year. For example, among 2004 originations, the default hazard in month 18 after origination was 0.4 percent for FRMs, compared to 1.0 percent for ARMs. Second, for both loan types, later originations are more likely to experience default. The underlying risk characteristics of loans may have been changing over time. In addition, adverse changes in the macroeconomic environment - in particular, rising unemployment and weak house price appreciation - may have contributed to the striking increase in default risk for loans originated in more recent years.

Figure 5 plots 90-day default hazards by borrower race for census tracts that are predominantly non-black (less than 40 percent black) and neighborhoods that are predominantly black (more than 60 percent black).[16] These displayed hazards do not control for any variables. In the non-black neighborhoods, black borrowers experience default rates that are higher than non-black borrowers, both for ARMs and FRMs. However, in black neighborhoods, black and non-black borrowers display similar default hazards.

4. Empirical Specification and Results

4.1. Empirical Specification

To examine the role of borrower, loan and neighborhood characteristics on default, we follow much of the literature on mortgage terminations and estimate semi-parametric Cox proportional hazard models of the form:

$$h_i(t) = h_0(t) \exp(\beta \text{ loan characteristics}_i + \gamma \text{ neighborhood characteristics}_i$$
$$+ \delta \text{ borrower characteristics}_i + \alpha \text{ calendar time and origination year fixed effects})$$

where $h_i(t)$ is the default hazard of mortgage i at time t, that is, the probability that mortgage i will experience a 90-day default at time t, conditional on not having previously defaulted. The

[16] "Black" includes only non-Hispanic blacks. "Non-black" includes all borrowers who reported a race, but did not report being black. Borrowers with missing race information are excluded.

proportional hazard model assumes that there is an underlying baseline hazard function $h_0(t)$ that is shared by all mortgages in the analysis sample. The model then allows time-varying explanatory variables to shift this baseline up or down proportionally, with α, β, γ and δ representing vectors of coefficient estimates. The Cox model provides no direct estimate of, and makes no assumptions about, the functional form of the baseline hazard, and is able to account for both right and left censoring of the longitudinal data. In our data, mortgage prepayments are treated as right-censored observations and there is also minor left-censoring as a few months typically elapse between the time of origination and the entry of the mortgage into the LoanPerformance database upon securitization.

The calendar time fixed effects in our models are in terms of quarter dummies. These will control for any city-, state- or nation- wide macroeconomic factors, including unemployment rates and city-wide house price movements. The origination year fixed effects are intended to pick up any city-wide systematic changes in mortgage characteristics over time, including average borrower risk and underwriting standards.

By estimating separate hazard models for hybrid ARMs and FRMs, we remove any endogeneity effects due to borrowers selecting into different product types. [17] Our results are displayed in tables 4, 5 and 6. We report hazard ratios (the exponential of the estimated coefficients) that can be interpreted as the proportional shift in the baseline hazard as a result of a unit change in the variable of interest. Hazard ratios greater than one indicate a positive effect, while those less than one indicate a negative effect. Our robust standard errors are clustered at the census tract level to account for any neighborhood level spatial correlation of residuals.

4.2. Adjustable Rate Mortgage Results

[17] In the hybrid ARM models discussed below, we also included an indicator variable for 3/27 ARMs. The coefficient on this variable implies that 3/27s have a default hazard that is approximately 18 percent lower than the 2/28s. When we limit the sample to only 2/28s (almost 3 in 4 of the hybrids), all the patterns that we describe below remain.

Loan and property characteristics. The first set of rows in table 4 show the hazard ratios associated with loan pricing terms. Virtually all of these terms are strongly statistically significant, as we would expect. ARMs with a starting interest rate that is 6 or more percentage points higher than the six-month LIBOR at the time of origination have a default hazard that is three times higher than ARMs with initial rate spreads of less than 2 percentage points above LIBOR (the reference category). The default hazard monotonically increases with the measures of relative rate. Similarly, default increases with the margin – the amount added to the six-month LIBOR to determine the rate that will apply at adjustment triggers. A margin of 7 percent or higher is associated with a roughly 50 percent higher default hazard compared with margins less than 5 percent. As well as the direct effect of higher mortgage payments on the likelihood of default, these loan pricing terms may reflect *ex ante* risk pricing by lenders, to the extent that our controls for borrower risk characteristics (discussed below) do not fully capture the lender's overall assessment of borrower risk.

The next set of explanatory variables in table 4 measure the size of the payment shock upon the initial adjustment of the interest rate. The reference category here is the months prior to the initial adjustment when there is no possible payment shock. In months 3 to 6 following the initial adjustment, a payment shock of greater than 30 percent is associated with a 60 percent higher hazard, while the effect of smaller payment shocks is monotonically increasing but insignificant. From 7 to 12 months after the initial adjustment, the default hazard is about 40 percent higher for those with payment shocks of less than 20 percent, about 70 percent higher for those with payment shocks of 20 to 30 percent, and over 90 percent higher for those with payment shocks over 30 percent. That even hybrid ARMs with less than a 20 percent increase in payments experience a significantly elevated default hazard following the initial adjustment is consistent with other research (Pennington-Cross & Ho, 2006). Indicator variables for 12 months or more after the initial adjustment are also included in the model, but the coefficients were small and insignificant and so are not displayed in the table.

Borrower risk characteristics generally have the expected effect on default behavior. Lower credit scores have a consistently positive effect on the default hazard. Borrowers with FICOs less than 530 have an approximately 90 percent higher default hazard than those with FICOs higher than 720 (the reference category). Borrowers with DTIs at origination that are over 45 percent have a 13 percent higher default hazard than those with lower DTIs.

Consistent with virtually all prior research, the current combined LTV has significant, large and monotonically increasing effects on default.[18] In model 1, which only controls for loan characteristics, the default hazard more than doubles once the current LTV is more than 90, compared with LTVs lower than 60 (the reference category). Once we start adding the neighborhood level controls in subsequent models, this increase in default hazard drops just slightly but is still strongly significant.

While we might expect that those who take out additional debt on the property to be more likely to default because of a higher debt burden, there is also a countering screening argument whereby only borrowers who are good risks will be able to secure additional financing, and so new debt may be negatively associated with default. Our results reveal a negative effect for new debt, suggesting that screening overwhelms the direct impact of having to shoulder a greater debt burden, however, the effect is statistically insignificant.[19]

---

[18] To estimate current LTV, we use the monthly dynamic loan balance information from LoanPerformance (the numerator) and adjust the property value (the denominator) using the appropriate community district level house price index. We do not have information on the dynamic loan balance of subsequent liens, so we assume that these are interest only mortgages where the principal does not change.

[19] We also tried specifications where the current LTV measure incorporated the level of any new additional debt. Since we do not know actual balances on home equity lines of credit, we assumed that the full amount of the credit line is used. The coefficients on these revised current LTV measures displayed a similar pattern to the current LTV measures shown in the table.

We now turn to other loan characteristics. Having a coborrower on the loan lowers the rate of default, probably because a coborrower diversifies the effect of income shocks. Original loan balances, entered in the model as a third order polynomial, are positively associated with significantly higher default rates (coefficients not displayed). Home purchase loans have higher default rates than do refinances, possibly reflecting the fact that refinancers have longer housing tenure, and also cannot be first time mortgage borrowers. Surprisingly, we find that owner-occupiers have elevated default rates compared with investors, though as we noted earlier, owner-occupancy is self-reported and may be unreliable. Further decomposition of this owner-occupier indicator by documentation level reveal that most of the positive effect that owner-occupation has on the default hazard is for no- or low- documentation loans. However, as shown in the table, full documentation loans as a group are not significantly less likely to default than loans without full documentation.

In terms of building characteristics, 2-4 family properties have lower default hazards compared to other building types. Building age is significantly and monotonically related to default: properties that are 50 years or older have default hazards that are up to 60 percent higher than those built in the last two years (coefficients not displayed).

Non-racial neighborhood characteristics. In model 2 of table 4 (and all subsequent models), we add a variety of neighborhood characteristics. It is noteworthy that apart from current LTV, the loan and property coefficients are virtually unchanged when these neighborhood characteristics are added. This is an important finding that confirms the validity of most existing research on mortgage defaults that do not have the wealth of very local neighborhood level information that we are able to use here. The LTV coefficients are smaller in magnitude when we add the neighborhood characteristics as local house price appreciation is likely correlated with these neighborhood characteristics.

The estimates on the neighborhood characteristics themselves highlight important spatial patterns in default. First, we include several census tract level demographic measures from the 2000 Census. As expected, we find that lower median income neighborhoods have higher rates of default, and loans against properties in neighborhoods with more high school graduates have lower rates of default. Using poverty rates instead of median income in these models gave similar results. Neighborhoods with more non-native born residents are also associated with lower rates of default, but differentiating these immigrant neighborhoods by their dominant race or ethnic group (black, Hispanic or Asian) did not reveal any clear differences among them.

The next set of rows in table 4 show the effects of neighborhood foreclosure and REO activity. We find a positive and generally monotonically increasing effect of both foreclosure notices and the fraction of properties held by lenders (REOs) within the prior 6 months on the default hazard in a given month. Because foreclosure and REO rates are positively correlated (an REO must have originally generated a foreclosure notice), it is somewhat inappropriate to think of these two sets of coefficients in isolation, but rather, they should be considered together. A foreclosure rate of over 3 percent in the surrounding census tract increases the default hazard by over one quarter, compared with mortgages in neighborhoods where the foreclosure rate is less than 1 percent. In addition, REO rates of greater than 3 percent increase the default hazard by at least 12 percent, compared to neighborhoods with REO rates of less than 1 percent.

We also included the non-prime share of mortgages originated in the census tract during the two years prior to the loan's origination. In model 2, this is positively associated with default. ARMs in tracts where this non-prime share is more than 30 percent have a default hazard that is over 25 percent higher than loans where the non-prime share was less than 10 percent. However, we should note that in models 3 and 4 when we add explanatory variables for neighborhood racial composition, the non-prime share estimates are reduced in magnitude and no longer significant.

To capture potential differences in the underwriting standards that prevailed at the neighborhood level at the time of the loan's origination, we also ran models that included mortgage application denial rates in the census tract in the six months prior to the loan's origination.[20] Higher mortgage denial rates at origination may indicate more stringent underwriting standards at the time of a loan's origination and are thus associated with lower default rates. We found that while these coefficients were negative, they were small and insignificant once all the other controls were included. To capture the ability of borrowers to avoid default by selling their property, we also tried including mortgage application denial rates in the census tract within the six months prior to the month of observation (as opposed to at origination). Higher current denial rates may indicate more stringent underwriting standards for potential buyers, which would hinder the ability of the borrower to sell the house, and thus result in higher default rates. These coefficients were also insignificant. The inclusion of these additional denial rate variables did not change the magnitudes or significance of our other results reported in table 4 and these alternate models are not shown in the table.

Other census tract level variables that we tried in alternate models included crime rates from the city's police records (property crime, violent crime and all crimes)[21] and the fraction of residents who had moved into the tract within the last two years from the 2000 Census. None of these variables were statistically significant.

Neighborhood racial composition. Model 3 of table 4 shows that residing in a census tract with higher proportions of black residents is associated with higher default rates. Almost one third of hybrid ARMs in our sample are made to borrowers living in tracts that are over 80 percent black. For borrowers in these tracts, the default hazard is over 25 percent higher than that of

[20] The HMDA mortgage application denial rate is: the number of denied applications, divided by the sum of loans originated, denied applications, and approved applications that were not accepted by the applicant.
[21] Crime rates were calculated as the number of crimes reported in the census tract in the six months preceding the month of analysis, divided by the tract's population in the 2000 Census.

borrowers in tracts with fewer than 20 percent black residents (the reference category). We do not see significant or consistent patterns of Hispanic or Asian population shares on default outcomes.

Comparing models 2 and 3, we see that the loan and property coefficients do not change as we add more neighborhood characteristics. However, the addition of the racial composition variables reduces the magnitude of the foreclosure rate coefficients, reflecting the positive correlation between foreclosure rates and predominantly black neighborhoods.[22] And as noted above, the inclusion of the neighborhood race variables reduces the magnitude and significance of the share of non-prime loans that we found in model 2, although the estimates remain monotonically increasing.

Borrower race. In model 4, we add the race of the primary borrower.[23] Specifically, we interact being a black borrower with the neighborhood share of black residents, with the reference category being non-black borrowers in neighborhoods with fewer than 20 percent black residents. The estimates display a somewhat unexpected pattern. In neighborhoods that are more than 40 percent black, the estimates on the share of black residents are virtually identical across borrower race; i.e., borrower's own race does not matter. Only in neighborhoods with fewer than 40 percent black residents do black borrowers have significantly higher default hazards. In model 5, we repeat model 4, but remove the variables on neighborhood racial composition, leaving in only the primary borrower's race. These results reinforce the interpretation that being black, by itself, has a limited effect on default: the hazard is only 10 percent higher for black versus white borrowers (the reference category).[24] Overall, the pattern of coefficients in models 4 and 5 suggest that the neighborhood share of black residents is as important in explaining default as the borrower's own

[22] We also interacted these racial composition measures with the foreclosure measures but found no effect.
[23] Race and/or ethnicity were missing for 15 percent of ARMs. The models in table 4 included indicator variables to capture this category, but our results do not change if we simply omit these observations from the sample. Our results also do not change if we define "black" to also include Hispanic black, or if we allow "black" to include instances where the coborrower is black (almost 90 percent of coborrowers are of the same race and ethnicity as the primary borrower).
[24] Note that this is a different reference category than that used in model 4.

race. Results for other borrower and neighborhood races and ethnicity did not yield clear patterns, although we find that Hispanic white borrowers tend to have lower default hazards. We also included the gender of the primary borrower in models 4 and 5 (coefficients not displayed) but the coefficients were small and insignificant.

We also estimated model 4 with the addition of community district fixed effects (community district boundaries are shown in figure 2). The effect of foreclosure rates when we include these fixed effects is slightly smaller than the ones displayed in table 4, but they are still significant at 1 percent, and substantial in magnitude. For example, a foreclosure rate of over 3 percent increases the default hazard by 19 percent relative to mortgages in neighborhoods where the foreclosure rate is less than 1 percent. All the other coefficients, including those on the neighborhood racial compositions, were virtually unchanged. It is worth emphasizing that a community district is still relatively small in size: the average land area is just over four square miles. These results suggest that most of the estimated neighborhood effects in table 4 are due to variation at the very local neighborhood level.

Finally, we have estimated all our models separately for home purchases and refinance loans. In table 5, we present the neighborhood foreclosure, REO, racial composition and borrower race coefficients associated with model 4 from table 4. The estimated impact of foreclosure rates are similar across the two samples, however, high REO rates have a larger and more significant effect on the default hazard for home purchases compared to refinances. This is consistent with refinance borrowers being less sensitive to visible deterioration in their surroundings as they have greater neighborhood attachment stemming from their longer housing tenure.

We find that for home purchases, the neighborhood racial composition results are even stronger in magnitude. Both black and non-black borrowers in neighborhoods that are over 80 percent black have default hazards that are about two thirds higher than the reference group (non-black borrowers in neighborhoods with fewer than 20 percent black residents). By contrast, the models for refinance loans do not display this pattern, and in fact, the neighborhood race effects are all insignificant.

4.3. Fixed Rate Mortgage Results

Table 6 displays our results for fixed rate mortgages. In general, the patterns are not so dissimilar to those for the hybrid ARMs. There are, however, a few notable differences. The magnitude of the coefficients on FICO score and current LTV are larger, which is perhaps unsurprising because there is no equivalent to the payment shocks that are very important in explaining the defaults for ARMs. FRM borrowers who take out additional new debt have a 12 percent higher default hazard than those who do not, whereas this variable had an insignificantly negative effect on ARM defaults. This suggests that being able to secure additional financing does not provide much more information on FRM borrower risk, possibly because they are better selected or because underwriting standards are more stringent for FRMs. Instead, the pure effect of a larger debt burden dominates, resulting in a positive coefficient on taking on new debt.

The effect of foreclosures and REOs remain similar in magnitude and significance as for the ARMs. In terms of neighborhood racial composition, the coefficients on predominantly black neighborhoods are larger than for ARMs. In model 3, tracts with more than 40 percent black residents have a default hazard that is at least 30 percent higher than tracts with fewer than 20 percent black residents (the reference category). Model 4 shows that while borrowers of all races have higher default hazards in more black neighborhoods, non-black borrowers tend to have even higher default hazards than do black borrowers in those neighborhoods. Relative to the reference category (non-black borrowers in 0-20 percent black neighborhoods), non-black borrowers in 80-100 percent black neighborhoods have default hazards that are 52 percent higher, while black borrowers in these mostly black neighborhoods have default hazards that are only 28 percent higher.

In model 5, where we remove neighborhood racial composition, the coefficient on being a black borrower is itself small and insignificant.

4.4. Interpretation of the Neighborhood Results

Our results indicate that neighborhood characteristics play a significant role in default outcomes, beyond the effects of loan and individual borrower characteristics, including the borrower's own race. We find that as the rate of foreclosure notices filed and the number of REOs in the neighborhood increases, the hazard of default increases, even after controlling for a host of other variables. We also find that home purchase borrowers living in predominantly black neighborhoods are more likely to enter default, with non-black borrowers in these neighborhoods at least as likely to default as black borrowers. In this section, we consider some possible interpretations of these results.

Neighborhood foreclosure rates and REO activity. The most direct interpretation of our findings on foreclosures and REOs is that they are in and of themselves important and that there are contagion effects at play. Neighbors may share information about the efficacy of default or the foreclosure process, leading other neighbors struggling with mortgage payments and negative equity to enter into default. High neighborhood foreclosure rates may also reduce the stigma associated with defaulting on a mortgage, making it more likely that others default as well. Survey results indicate that homeowners who know people who have defaulted are more willing to default if they are underwater than homeowners who have not been exposed, although a large majority of survey respondents still believe that strategic defaults are "immoral" (Guiso et al., 2009).

Foreclosure and REO concentrations could also be serving as proxies for omitted variables. One possibility is local economic conditions. Although the calendar time fixed effects control for city-wide macroeconomic conditions, there could still be substantial variation by neighborhood. In particular, some neighborhoods may have more defaults and foreclosures because of correlated income and employment shocks. Especially if residents have similar socioeconomic backgrounds or human capital, job losses in one industry or occupation may affect many residents in the same neighborhood. This employment effect could be related to house prices: Mayer et al. (2009) argue that upon job loss, homeowners in neighborhoods where housing prices are rising are more likely to sell their home than to default. Another possibility is that foreclosure and REO activity are proxying for unobserved borrower characteristics that are correlated with default risk. For example, our investor indicator could be inaccurate as it is usually self-reported, and investors may be more likely to both enter default and to choose investment properties in higher risk housing markets that *ex post* turn out to have higher foreclosure and REO rates. However, because tracts with high foreclosure and REO rates are geographically concentrated in particular community districts[25], the fact that controlling for community district fixed effects in our hazard models only slightly reduces the estimated impact of foreclosures and REOs suggests that these omitted variable interpretations are valid, but are probably limited in importance.

Measurement error in house price appreciation may cause some of the apparent effect of foreclosures and REO. Because the foreclosure and REO variables are measured at the census tract level, while our house price indices are at the larger community district level, less house price appreciation in higher foreclosure tracts compared with the community district average would lead to systematically overestimated housing values and underestimated current LTVs for loans in high foreclosure neighborhoods. Indeed, since borrowers with positive equity can avoid default by selling their property, the fact that current LTVs ranging from 60 to 90 are positive and significant in our default models suggests that our LTV measure is underestimated. Perhaps if we had enough sales observations to construct good repeat sales indices at the census tract level, our current LTV

[25] For example, during the first six months of 2008, 55 percent of tracts experiencing a foreclosure rate of at least 3 percent were located in just five community districts.

measure would be more accurate and the foreclosure effects would be reduced.[26] Foreclosures and REOs could be causing diminished house price appreciation via several channels. First, they increase the housing supply and drive down market prices. Second, they may generate negative externalities such as the visible deterioration of properties that lead to lower property values in high foreclosure tracts relative to others in the same community district.[27] Further, foreclosure rates may reflect expectations about future house price depreciation that are not already captured in price indices based on recent sales transactions.

In sum, we interpret our finding that default hazards increase with higher recent foreclosure and REO activity in the neighborhood to be reflecting some combination of a contagion effect and a proxy effect for very local house price fluctuations. To a lesser extent, the foreclosure and REO rates may also serve as a proxy for neighborhood economic conditions or unobserved borrower characteristics.

Neighborhood racial composition. We find that while there are elevated default rates among black borrowers in non-black neighborhoods, borrowers in predominantly black neighborhoods also have higher default rates, regardless of race, even after controlling for a host of other borrower, loan, and neighborhood characteristics. The most obvious interpretation of this finding is that the share of black residents is proxying for unobserved borrower and loan characteristics that are correlated with higher default rates. There are several plausible candidates for these unobserved characteristics.

First, there may be a systematic difference between individuals living in predominantly black neighborhoods and those living elsewhere that we cannot account for, even though we already control for a variety of individual risk characteristics, as well as tract level variables such as education

[26] In a study of Los Angeles County, Aragon et al. (2010) argue that even repeat sales indices at the zip code level are poor predictors of individual property values.
[27] For example, in a study of Massachusetts, Campbell et al. (2009) conclude that the primary effect of foreclosures on house prices comes from crime and vandalism.

and income. In particular, default risk may be related to the financial reserves the borrower could draw upon to weather shocks such as job loss, uninsured expenses related to health crises, or family instability. If residents of primarily black neighborhoods are systematically less likely to have significant financial reserves or if they are more likely to experience these adverse shocks, then they will have a higher risk of default. However, we find neighborhood effects only for home purchase and not refinance borrowers, which makes it less likely that the neighborhood effects are explained by location-related unobserved characteristics, which are likely to apply to all borrowers in the neighborhood.

Second, there may be unobserved risk characteristics that apply specifically to our sample of borrowers. In particular, unobserved variation in lending and underwriting practices across census tracts could lead to systematically different borrower characteristics in certain tracts. For example, if underwriters used less stringent or improper standards for mortgage applicants in black neighborhoods, then that will lead in those neighborhoods to a selection of borrowers with relatively higher default risk, even after controlling for FICO scores and other observable risk factors. We might expect refinance borrowers to be less likely to have these borrower risk unobservables as they have undergone earlier screening by another party (the underwriter of their previous mortgage), and they have demonstrated that they can carry a mortgage, at least up to the time of their refinance. Thus, this explanation could be consistent with no neighborhood effects for refinance loans.

Third, there may be unobserved aspects of the loan that we cannot control for, despite our rich set of loan specific variables. In particular, we are not able to account for any upfront fees that were paid at origination. If borrowers in black neighborhoods pay higher upfront fees, they will be less able to overcome income shocks and are thus more likely to default, all else equal. We also cannot discern whether the loans in our data were originated through a broker. There is some evidence that borrowers in black neighborhoods are more likely to use a mortgage broker, perhaps due to their more limited financial sophistication, a lack of trust for lending institutions staffed by "outsiders", or just limited access to local bank branches that makes it more difficult and costly to shop around for a mortgage. The incentives of a broker are in many cases poorly aligned with those of the borrower (and possibly the underwriter and lender as well), and others have found that loans involving brokers are more likely to enter default (Coulton et al., 2008; Laderman & Reid, 2008). This might be because the information given to the underwriter by the broker (and hence recorded in our data) is more likely to be inaccurate (for example, an inflated appraisal), or because aggressive marketing tactics used by brokers make it more difficult for borrowers to evaluate the broker's offerings and more likely to accept an inappropriate loan. That our findings do not apply to refinance borrowers might be due to their being less susceptible to aggressive marketing tactics as they are not first time homebuyers and may be more financially literate as a result of their prior homeowning and borrowing experience.

Another possible interpretation of our neighborhood race findings is that the share of black residents is proxying for local economic conditions. However, predominantly black tracts are geographically concentrated in particular community districts[28] and adding community district fixed effects does not change the coefficients on neighborhood racial composition. This reduces the likelihood that the share of black residents is just a proxy for local unemployment or other local concerns.

Finally, we consider two possible ways measurement error may influence our neighborhood race findings. First, similar to the discussion above for foreclosures and REOs, it is possible that predominantly black census tracts experience less house price appreciation than the rest of the surrounding community district, leading to a systematic underestimate of the current LTV for loans in predominantly black neighborhoods. Of course, this raises the question of why these particular

[28] For example, 78 percent of tracts that have over 80 percent black residents are located in just five community districts.

neighborhoods have poorer housing price performance. To investigate this measurement error possibility further, we reestimated model 4 of table 4 for two separate samples: ARMs in non-black neighborhoods (less than 40 percent black residents) and ARMs in black neighborhoods (more than 60 percent black residents). We find that the coefficients on current LTV for the black neighborhood sample are smaller than for the non-black, while statistical significance remains at one percent for both samples. Underestimated current LTV in black neighborhoods would have implied the opposite result, suggesting that this source of measurement error cannot be solely responsible for our findings.

A second source of possible measurement error could come from misreported borrower race, such that the neighborhood racial composition becomes a proxy for the individual's race and the coefficients on individual race become small and insignificant. In the HMDA data, race is usually self-reported, but if the individual does not specify a race, the lender can record a subjective assessment, which could lead to misreporting error. To the extent that black borrowers believe that there is racial discrimination in lending, they may have an incentive to misreport their race, especially if the loan application is not made in person. However, we cannot find any empirical evidence that suggests that HMDA race classifications are systematically in error.

In sum, we think that the most likely explanation of our neighborhood racial composition findings is that borrowers and loans in black neighborhoods have, on average, substantially worse unmeasured characteristics than those in non-black neighborhoods. This interpretation is consistent with the non-prime mortgage industry (brokers, lenders, underwriters, etc.) treating black neighborhoods differently in terms of marketing and underwriting practices or loan terms in ways that are unobserved in our data.[29] However, our results do not allow us to discern the reasons for

[29] Our results on non-prime loan concentrations provide some additional evidence consistent with this idea of differential treatment. Model 2 of table 4 shows that neighborhoods with higher non-prime shares have higher default hazards, even after controlling for a large set of loan and borrower characteristics. One way to think about this is that if

this differential treatment. It is possible that some non-prime lenders and brokers "targeted" minority neighborhoods with inappropriately priced loans and improper underwriting.[30] Alternatively, our results are also consistent with institutional pressures to extend more credit to these neighborhoods, to boost homeownership rates, or simply to tap new markets, by employing strong marketing tactics or more lenient underwriting standards.

5. Conclusion

Our rich data set allows us to improve upon the existing literature by assessing the impact that borrower characteristics, the type of loan and its terms, and the characteristics of the neighborhood (measured at the census tract level) have on the probability that a non-prime mortgage will default.

Similar to existing research on mortgage defaults using loan level data, we find that the current LTV, borrower credit scores and debt-to-income ratios at origination, interest rates, loan size, ARM margins and payment shocks upon rate adjustment are all significant in explaining default behavior. We further show that, besides LTV, the magnitude of the estimates on these loan and borrower risk characteristics do not change when we add a variety of census tract level controls to capture very local neighborhood characteristics such as foreclosure and REO activity, racial composition, and the share of existing non-prime mortgages. This is an important finding that

---

we could include an indicator variable for "differential treatment," it would have a positive coefficient in the hazard models. Instead, the non-prime share of mortgages may be proxying for this variable. While the coefficients on the share of non-prime loans remain positive once we include the neighborhood racial composition variables in model 3 of table 4, they become insignificant. This may suggest that the share of black residents is an even better proxy for this missing indicator.

[30] The targeting hypothesis has been stated most starkly in litigation some cities have filed against various lenders. Baltimore, for example, has accused Wells Fargo of "target[ing] these kinds of predatory practices ['charging excessive fees; charging excessively high interest rates that are not justified by borrowers' creditworthiness; requiring large prepayment penalties while deliberately misleading borrowers about the penalties; using deceptive sales practices to wrap insurance products into mortgages; convincing borrowers to refinance mortgages into new loans that only benefit Wells Fargo; deceiving borrowers into believing that they are getting fixed rate loans when they are really getting adjustable rate loans'] at African-American neighborhoods and residents." Mayor and City Council of Baltimore v. Wells Fargo N.A., Case No. 1:08-cv-00062-JFM (D.Md. April 7, 2010).

confirms the validity of most existing research on mortgage defaults that only have data on the loan and borrower risk variables, and not the rich neighborhood level information that we are able to use here.

Regarding the neighborhood characteristics themselves, our main finding is that they importantly affect default rates, even after controlling for the usual set of loan and borrower risk characteristics. First, default rates increase as the rate of foreclosure notices and the number of REOs in the neighborhood increases. We argue that this may reflect a contagion effect, or that foreclosure and REO rates may be serving as proxies for weak neighborhood housing market conditions that are not already captured in the community district level prices indices and, to a lesser extent, as proxies for very local economic conditions or unobserved borrower characteristics. At first blush, it seems hardly surprising that defaults are higher in high foreclosure and REO areas. However, what we have shown is that this effect is happening at an extremely local level. Our estimates remain large and significant when we include fixed effects for the 56 community districts in our New York City sample.

Our findings that neighborhood foreclosure and REO rates are associated with higher default rates raise concerns about how local governments, community development organizations, and lenders should target foreclosure avoidance programs such as loan modifications. For example, there may be a "tipping point" for the concentration of foreclosures. If so, efforts to prevent foreclosure are most viable in neighborhoods that have a sufficient number of foreclosures to merit intervention, but have not yet reached that tipping point.

Second, we find that living in tracts with more black residents has a large positive effect on default for home purchase loans, even after controlling for a rich set of loan and borrower characteristics, including the borrower's own race. We argue that this result likely reflects unobserved loan and borrower characteristics that are correlated with higher default risk being more important in black neighborhoods. This is consistent with a differential treatment of black neighborhoods by the mortgage industry, but our data cannot conclusively prove that black neighborhoods were "targeted" for unsustainable non-prime loans that were inappropriate in pricing or other terms, or were subject to low quality underwriting practices. Nevertheless, our findings raise concerns as to whether such loan products should be marketed at all, particularly in vulnerable neighborhoods, or if they should be accompanied by disclosures about their propensity to end in default. Regardless of whether targeting took place, our findings suggest that policy-makers should take into account the disproportionate harm that the current foreclosure crisis may have in predominantly black neighborhoods.

In closing, our finding that neighborhood characteristics have significant effects on default suggests that policy-makers should take neighborhood context into account in designing their responses to the foreclosure crisis, and in shaping the regulation of mortgage products. Rather than think of borrowers and loans solely on an individual basis, it may be more appropriate to think of more systemic neighborhood level efforts to address the current crisis.

References

Apgar, William, and Mark Duda. 2005. "Collateral Damage: The Municipal Impact of Today's Mortgage Foreclosure Boom." *Homeownership Preservation Foundation*.

Aragon, Diego, Andrew Caplin, Sumit Chopra, John V. Leahy, Yann LeCun, Marco Scoffier, and Joseph Tracy. 2010. "Reassessing FHA Risk." NBER Working Paper 15802.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan. 1994. "Race, Redlining, and Residential Mortgage Loan Performance." *Journal of Real Estate Finance and Economics*, 9(3): 263-94.

Campbell, John Y., Stefano Giglio and Parag Pathak. 2009. "Forced Sales and House Prices." NBER Working Paper 14866.

Coulton, Claudia, Tsui Chan, Michael Schramm, and Kristen Mikelbank. 2008. "Pathways to Foreclosure: A Longitudinal Study of Mortgage Loans, Cleveland and Cuyahoga County, 2005-2008." *Center on Urban Poverty and Community Development, Case Western Reserve University.*

Courchane, Marsha J., Brian J. Surette, and Peter M. Zorn. 2004. "Subprime Borrowers: Mortgage Transitions and Outcomes." *Journal of Real Estate Finance and Economics*, 29(4): 365-92.

Danis, Michelle A., and Anthony N. Pennington-Cross. 2005. "The Delinquency of Subprime Mortgages." Federal Reserve Bank of St. Louis Working Paper 2005-022A.

Demyanyk, Yuliya. 2009. "Quick Exits of Subprime Mortgages." *Federal Reserve Bank of St. Louis Review*, 2: 79-93.

Ding, Lei, Roberto G. Quercia, Janneke Ratcliffe, and Wei Lei. 2008. "Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models." Center for Community Capital University of North Carolina Working Paper.

Firestone, Simon, Robert Van Order, and Peter M. Zorn. 2007. "The Performance of Low-Income and Minority Mortgages." *Real Estate Economics,* 35(4): 479-504.

Foote, Christopher L., Kristopher S. Gerardi, Lorenz F. Goette, and Paul Willen. 2009. "Reducing Foreclosures." Federal Reserve Bank of Boston Public Policy Discussion Paper 2009-02.

Furman Center for Real Estate and Urban Policy. 2010. "Foreclosed Properties in NYC: A Look at the Last 15 Years." http://furmancenter.org/files/publications/ Furman_Center_Fact_Sheet_on_REO_Properties.pdf.

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

Grover, Michael, Laura Smith, and Richard M. Todd. 2008. "Targeting Foreclosure Interventions: An Analysis of Neighborhood Characteristics Associated with High Foreclosure Rates in Two Minnesota Counties." *Journal of Economics and Business*, 60(1-2): 91-109.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2009. "Moral and Social Constraints to Strategic Default on Mortgages." NBER Working Paper 15145.

Harding, John, Thomas J. Miceli, and C.F. Sirmans. 2000. "Do Owners Take Better Care of Their Housing Than Renters?" *Real Estate Economics*, 28(4): 663-81.

Harding, John, Eric Rosenblatt, and Vincent Yao. 2009. "The Contagion Effect of Foreclosed Properties." *Journal of Urban Economics*, 66(3): 164-78.

Hartley, Daniel A. 2010. "The Effect of Foreclosures on Owner-Occupied Housing Prices: Supply or Dis-Amenity?" Paper presented at the Federal Reserve Bank of Cleveland Policy Summit, Cleveland, OH.

Haughwout, Andrew, Christopher Mayer, and Joseph Tracy. 2009. "Subprime Mortgage Pricing: The Impact of Race, Ethnicity and Gender on the Cost of Borrowing." *Brookings-Wharton Papers on Urban Affairs*: 33-63.

Haughwout, Andrew, Richard Peach, and Joseph Tracy. 2008. "Juvenile Delinquent Mortgages: Bad Credit or Bad Economy?" *Journal of Urban Economics*, 64(2): 246-57.

Immergluck, Dan, and Geoff Smith. 2006. "The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values." *Housing Policy Debate*, 17(1): 57-79.

Ioannides, Yannis M. 2003. "Interactive Property Valuations." *Journal of Urban Economics*, 53(1): 145-70.

Jiang, Wei, Ashlyn Aiko Nelson, and Edward Vytlacil. 2009. "Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency." Columbia University Business School Working Paper.

Laderman, Elizabeth, and Carolina Reid. 2008. "Lending in Low- and Moderate-Income Neighborhoods in California: The Performance of CRA Lending During the Subprime Meltdown." Federal Reserve Bank of San Francisco Working Paper 2008-05.

Lee, Kai-Yan. 2008. "Foreclosure's Price-Depressing Spillover Effects on Local Properties: A Literature Review." Federal Reserve Bank of Boston Community Affairs Discussion Paper 2008-01.

Leonard, Tammy, and James Murdoch. 2009. "The Neighborhood Effects of Foreclosure." *Journal of Geographical Systems*, 11(4): 317-32.

Lin, Zhenguo, Eric Rosenblatt, and Vincent W. Yao. 2009. "Spillover Effects of Foreclosures on Neighborhood Property Values." *Journal of Real Estate Finance and Economics,* 38(4): 387-407.

Mayer, Christopher, Karen Pence, and Shane M. Sherlund. 2009. "The Rise in Mortgage Defaults." *Journal of Economic Perspectives*, 23(1): 27-50.

Okah, Ebiere, and James Orr. 2010. "Subprime Mortgage Lending in New York City: Prevalence and Performance." Federal Reserve Bank of New York Staff Report 432.

Pedersen, Camille, and Lucy Delgadillo. 2007. "Residential Mortgage Default in Low- and High-Minority Census Tracts." *Family and Consumer Sciences Research Journal*, 35(4): 374-91.

Rogers, William, and William Winter. 2009. "The Impact of Foreclosures on Neighboring Housing Sales." *Journal of Real Estate Research,* 31(4): 455-479.

Schloemer, Ellen, Wei Li, Keith Ernst, and Kathleen Keest. 2006. "Losing Ground: Foreclosures in the Subprime Market and Their Cost to Homeowners." *Center for Responsible Lending.*

Schuetz, Jenny, Vicki Been, and Ingrid Gould Ellen. 2008. "Neighborhood Effects of Concentrated Mortgage Foreclosures." *Journal of Housing Economics,* 17(4): 306-319.

Van Order, Robert, and Peter Zorn. 2000. "Income, Location and Default: Some Implications for Community Lending." *Real Estate Economics*, 28(3): 385-404.

Table 1: Loan and Borrower Characteristics

	Adjustable Rate Mortgages	Fixed Rate Mortgages
Average loan amount	$352,461	$360,614
Owner-occupier	0.909	0.880
Average interest rate at origination	7.4	6.9
Average relative interest rate at origination[1]	4.3	0.9
Average ARM margin	5.9	
Payment shock at first ARM adjustment:		
<20%	0.975	
20-30%	0.014	
>30%	0.011	
Average FICO at origination	621	665
Debt-to-income at origination > 45%[2]	0.424	0.361
Average combined LTV at origination	74.4	69.4
Took on additional mortgage debt[3]	0.051	0.108
Full documentation	0.382	0.419
Has coborrower	0.206	0.284
Primary borrower race/ethnicity[4]:		
White	0.224	0.311
Black	0.461	0.368
Hispanic white	0.152	0.141
Asian	0.093	0.111
Other	0.070	0.068
Number of loans	30,307	29,414

Source: LoanPerformance

Table indicates the fraction of records in each category, unless otherwise noted.

[1] For ARMs: interest rate minus the six-month LIBOR rate at origination. For FRMs: interest rate minus the Freddie Mac average interest rate for prime 30-year FRMs at origination. Expressed in percentage points.

[2] DTI is missing for 19% of ARMs and 38% of FRMs.

[3] Additional mortgage debt was secured that totaled at least 5% of the original loan amount.

[4] Race/ethnicity is missing for 15% of ARMs and 20% of FRMs.

Table 2: Neighborhood Characteristics

		Adjustable Rate Mortgages	Fixed Rate Mortgages
Census tract demographics from 2000 Census			
Median income:	<$30,000	0.230	0.211
	$30,000-$40,000	0.261	0.253
	>$40,000	*0.509*	*0.535*
% High school graduates:	*<70%*	*0.465*	*0.421*
	70-80%	0.342	0.326
	>80%	0.193	0.252
% Non-native born	*<20%*	*0.119*	*0.129*
	20-40%	0.466	0.445
	40-60%	0.351	0.353
	>60%	0.064	0.074
% Black	*0-20%*	*0.296*	*0.424*
	20-40%	0.105	0.097
	40-60%	0.104	0.082
	60-80%	0.162	0.136
	80-100%	0.333	0.261
% Hispanic	*0-20%*	*0.611*	*0.618*
	20-40%	0.190	0.192
	40-100%	0.199	0.190
% Asian	*0-20%*	*0.913*	*0.883*
	20-40%	0.076	0.097
	40-100%	0.011	0.020
Census tract loan composition at origination			
Share of non-prime loans[1]	*<10%*	*0.079*	*0.141*
	10-20%	0.332	0.359
	20-30%	0.529	0.445
	>30%	0.060	0.054
Number of loans		30,307	29,414
Census tract foreclosure and REO activity (dynamic)			
Recent foreclosure rate[2]	*<1%*	*0.526*	*0.555*
	1-2%	0.292	0.246
	2-3%	0.111	0.113
	>3%	0.071	0.085
Recent REO rate[3]	*<1%*	*0.648*	*0.675*
	1-2%	0.132	0.113
	2-3%	0.077	0.068
	>3%	0.142	0.144
Number of loan-months		503,579	946,922

Table indicates fraction of records in each category. Italics indicate category omitted from hazard models in tables 4-6.

[1] The number of non-prime loans originated in the Loan Performance database / Total loans originated in HMDA during the 6 months preceding the loan's origination.

[2] The number of *lis pendens* filed in the 6 months preceding the analysis month / The stock of buildings.

[3] The number of properties listed as REO in the 6 months preceding the analysis month / The stock of buildings.

Table 3: Comparison of Black and Non-Black Home Purchase Borrowers in Predominantly Black Neighborhoods

	Adjustable Rate Mortgages		Fixed Rate Mortgages	
	Black	**Non-Black**	**Black**	**Non-Black**
Average loan amount	$365,896	$363,570	$372,620	$381,339
Owner-occupier	0.912	0.870	0.840	0.858
Average interest rate at origination	7.4	7.4	7.3	7.2
Average ARM margin	5.9	5.9		
Payment shock at first ARM adjustment:				
<20%	0.977	0.978		
20-30%	0.011	0.015		
>30%	0.012	0.007		
Average FICO at origination	658	667	690	701
Average debt-to-income at origination	35.8	33.6	23.4	20.9
Full documentation	0.311	0.211	0.296	0.234
Average combined LTV at origination	85.9	84.9	84.0	82.2
Took on additional mortgage debt	0.050	0.063	0.081	0.124
Number of loans	3,341	1,602	1,178	636

Source: LoanPerformance

Table indicates the fraction of records in each category, unless otherwise noted.

Predominantly black neighborhoods are census tracts that have at least 60 percent non-Hispanic black residents in the 2000 Census.

Table 4: Hazard Models of 90-Day Default for Adjustable Rate Mortgages

		Model 1			Model 2			Model 3			Model 4			Model 5		
Loan characteristics																
Relative interest rate at origination:																
	2-4	1.256	(0.063)	**	1.243	(0.063)	**	1.236	(0.063)	**	1.241	(0.063)	**	1.240	(0.063)	**
	4-6	1.777	(0.106)	**	1.723	(0.103)	**	1.699	(0.102)	**	1.707	(0.103)	**	1.713	(0.102)	**
	>6	3.484	(0.265)	**	3.329	(0.258)	**	3.279	(0.256)	**	3.294	(0.256)	**	3.299	(0.253)	**
ARM Margin:	5-6%	1.145	(0.047)	**	1.126	(0.047)	**	1.122	(0.046)	**	1.121	(0.047)	**	1.127	(0.047)	**
	6-7%	1.310	(0.050)	**	1.291	(0.050)	**	1.283	(0.050)	**	1.283	(0.050)	**	1.294	(0.050)	**
	>7%	1.503	(0.086)	**	1.473	(0.085)	**	1.468	(0.085)	**	1.465	(0.085)	**	1.477	(0.085)	**
3-6 months post-adjustment and:																
	payment shock <20%	1.169	(0.100)		1.173	(0.100)		1.177	(0.101)		1.177	(0.101)		1.177	(0.100)	
	payment shock 20-30%	1.216	(0.137)		1.199	(0.135)		1.204	(0.136)		1.206	(0.136)		1.207	(0.136)	
	payment shock >30%	1.631	(0.203)	**	1.625	(0.203)	**	1.627	(0.203)	**	1.626	(0.203)	**	1.623	(0.203)	**
7-12 months post-adjustment and:																
	payment shock <20%	1.397	(0.117)	**	1.400	(0.116)	**	1.406	(0.117)	**	1.405	(0.117)	**	1.405	(0.117)	**
	payment shock 20-30%	1.735	(0.171)	**	1.697	(0.167)	**	1.703	(0.167)	**	1.713	(0.168)	**	1.717	(0.168)	**
	payment shock >30%	1.939	(0.228)	**	1.919	(0.226)	**	1.920	(0.225)	**	1.928	(0.225)	**	1.924	(0.225)	**
FICO at origination:	680-720	1.205	(0.058)	**	1.208	(0.057)	**	1.202	(0.057)	**	1.207	(0.058)	**	1.209	(0.058)	**
	650-680	1.312	(0.061)	**	1.311	(0.060)	**	1.307	(0.060)	**	1.304	(0.061)	**	1.304	(0.061)	**
	620-650	1.499	(0.069)	**	1.499	(0.070)	**	1.489	(0.070)	**	1.489	(0.070)	**	1.489	(0.070)	**
	590-620	1.490	(0.079)	**	1.505	(0.080)	**	1.494	(0.079)	**	1.492	(0.080)	**	1.491	(0.080)	**
	560-590	1.783	(0.099)	**	1.816	(0.101)	**	1.810	(0.101)	**	1.806	(0.101)	**	1.797	(0.100)	**
	530-560	1.889	(0.110)	**	1.937	(0.115)	**	1.928	(0.114)	**	1.920	(0.113)	**	1.913	(0.113)	**
	<530	1.864	(0.123)	**	1.919	(0.129)	**	1.902	(0.128)	**	1.899	(0.128)	**	1.895	(0.127)	**
Debt-to-income at origination >45%		1.127	(0.029)	**	1.130	(0.029)	**	1.132	(0.029)	**	1.129	(0.029)	**	1.128	(0.029)	**
Current combined LTV:	60-70%	1.178	(0.054)	**	1.127	(0.051)	**	1.111	(0.051)	*	1.113	(0.051)	*	1.126	(0.051)	**
	70-80%	1.423	(0.061)	**	1.327	(0.056)	**	1.302	(0.055)	**	1.307	(0.056)	**	1.328	(0.056)	**
	80-90%	1.618	(0.076)	**	1.478	(0.070)	**	1.443	(0.069)	**	1.447	(0.069)	**	1.475	(0.070)	**
	90-95%	2.201	(0.124)	**	1.980	(0.112)	**	1.935	(0.110)	**	1.938	(0.110)	**	1.975	(0.112)	**
	95-100%	2.340	(0.147)	**	2.094	(0.132)	**	2.048	(0.129)	**	2.053	(0.129)	**	2.089	(0.131)	**
	>100%	2.284	(0.142)	**	1.964	(0.125)	**	1.917	(0.123)	**	1.922	(0.123)	**	1.965	(0.125)	**

continued

	Model 1			Model 2			Model 3			Model 4			Model 5		
Took on additional mortgage debt	0.929	(0.039)		0.944	(0.039)		0.942	(0.039)		0.945	(0.039)		0.946	(0.039)	
Has coborrower	0.754	(0.024)	**	0.775	(0.024)	**	0.778	(0.024)	**	0.780	(0.025)	**	0.779	(0.025)	**
Home purchase	1.377	(0.097)	**	1.374	(0.095)	**	1.377	(0.096)	**	1.389	(0.097)	**	1.389	(0.096)	**
Owner-occupier	1.311	(0.054)	**	1.350	(0.057)	**	1.357	(0.057)	**	1.371	(0.058)	**	1.370	(0.057)	**
Full documentation	1.003	(0.156)		0.987	(0.151)		0.976	(0.148)		0.958	(0.143)		0.960	(0.144)	
Non-racial neighborhood characteristics																
Median income: <$30,000				1.195	(0.044)	**	1.172	(0.046)	**	1.146	(0.045)	**	1.166	(0.043)	**
$30,000-$40,000				1.077	(0.032)	*	1.067	(0.032)	*	1.062	(0.032)	*	1.074	(0.032)	*
% High school graduates: 70-80%				0.996	(0.030)		0.939	(0.030)	*	0.938	(0.030)	*	0.964	(0.029)	
>80%				0.977	(0.036)		0.915	(0.037)	*	0.903	(0.037)	*	0.925	(0.035)	*
% Non-native born: 20-40%				0.914	(0.036)	*	0.924	(0.037)	*	0.917	(0.036)	*	0.918	(0.035)	*
40-60%				0.870	(0.036)	**	0.893	(0.038)	**	0.886	(0.037)	**	0.871	(0.036)	**
>60%				0.659	(0.044)	**	0.717	(0.049)	**	0.721	(0.048)	**	0.684	(0.044)	**
Recent foreclosure rate: 1-2%				1.220	(0.035)	**	1.151	(0.034)	**	1.149	(0.034)	**	1.191	(0.034)	**
2-3%				1.303	(0.050)	**	1.207	(0.047)	**	1.206	(0.047)	**	1.263	(0.048)	**
>3%				1.392	(0.062)	**	1.270	(0.058)	**	1.268	(0.058)	**	1.341	(0.060)	**
Recent REO rate: 1-2%				1.076	(0.039)	*	1.056	(0.038)		1.059	(0.039)		1.070	(0.039)	
2-3%				1.070	(0.045)		1.047	(0.044)		1.047	(0.044)		1.061	(0.045)	
>3%				1.140	(0.041)	**	1.116	(0.041)	**	1.119	(0.041)	**	1.136	(0.041)	**
Share of non-prime loans: 10-20%				1.135	(0.061)	*	1.091	(0.060)		1.093	(0.060)		1.127	(0.061)	*
20-30%				1.182	(0.065)	**	1.069	(0.062)		1.066	(0.062)		1.142	(0.064)	*
>30%				1.257	(0.083)	**	1.139	(0.079)		1.137	(0.079)		1.217	(0.081)	**

continued

		Model 1	Model 2	Model 3	Model 4	Model 5
Neighborhood racial composition						
% Hispanic:	20-40%			1.023 (0.041)	1.045 (0.041)	
	40-100%			0.945 (0.045)	0.993 (0.047)	
% Asian:	20-40%			0.931 (0.048)	0.948 (0.049)	
	40-100%			0.930 (0.137)	0.951 (0.139)	
% Black:	20-40%			1.111 (0.049) *		
	40-60%			1.189 (0.054) **		
	60-80%			1.209 (0.062) **		
	80-100%			1.268 (0.060) **		
Borrower race/ethnicity						
Black borrower and neighborhood is:						
	0-20% black				1.258 (0.113) **	
	20-40% black				1.291 (0.097) **	
	40-60% black				1.183 (0.073) **	
	60-80% black				1.255 (0.074) **	
	80-100% black				1.278 (0.067) **	
Non-black borrower and neighborhood is:						
	20-40% black				1.068 (0.055)	
	40-60% black				1.175 (0.076) *	
	60-80% black				1.228 (0.086) **	
	80-100% black				1.310 (0.082) **	
Black borrower						1.098 (0.037) **
Hispanic white borrower					0.829 (0.036) **	0.804 (0.033)
Asian borrower					0.937 (0.043)	0.919 (0.041)
Number of loan-months				503,579		
Number of loans				30,307		
Log pseudolikelihood		-76861.238	-76653.936	-76621.404	-76588.244	-76609.448

Cox proportional hazard models of 90-day default, hazard ratios reported. See tables 1 and 2 for notes on explanatory variables and the left-out categories.

Robust standard errors clustered by census tracts are reported in ( ). Statistical significance is indicated by: *5% and **1%.

All models include: a cubic in original loan balance; indicators for 3/27s, prepayment penalty in effect, cash-out refinance, second home, low documentation, property type (single & 2-4 family), building age (0-10, 11-50, >50 years & missing), >12 months post-adjustment and payment shock is <20, 20-30 & >30%; indicators for missing values of FICO, DTI, prepayment penalty; and fixed effects for origination year and calendar quarter. Models 4 and 5 also include indicators for Hispanic black, Hispanic other race, non-Hispanic other race, missing race, female and missing gender.

Table 5: Hazard Models of 90-Day Default for Adjustable Rate Mortgages -Home Purchases and Refinances

		All ARMs			Home Purchases			Refinances		
Recent foreclosure rate:	1-2%	1.149	(0.034)	**	1.136	(0.045)	**	1.167	(0.050)	**
	2-3%	1.206	(0.047)	**	1.202	(0.064)	**	1.197	(0.066)	**
	>3%	1.268	(0.058)	**	1.244	(0.076)	**	1.296	(0.084)	**
Recent REO rate:	1-2%	1.059	(0.039)		1.133	(0.055)	**	0.986	(0.049)	
	2-3%	1.047	(0.044)		1.149	(0.070)	*	0.954	(0.057)	
	>3%	1.119	(0.041)	**	1.198	(0.058)	**	1.015	(0.055)	
% Hispanic:	20-40%	1.045	(0.041)		1.122	(0.064)	*	0.953	(0.053)	
	40-100%	0.993	(0.047)		1.072	(0.075)		0.913	(0.064)	
% Asian:	20-40%	0.948	(0.049)		0.984	(0.069)		0.925	(0.071)	
	40-100%	0.951	(0.139)		0.948	(0.197)		0.975	(0.185)	
Black borrower and neighborhood is:										
	0-20% black	1.258	(0.113)	**	1.463	(0.169)	**	1.143	(0.155)	
	20-40% black	1.291	(0.097)	**	1.570	(0.174)	**	1.072	(0.104)	
	40-60% black	1.183	(0.073)	**	1.566	(0.139)	**	0.910	(0.081)	
	60-80% black	1.255	(0.074)	**	1.652	(0.133)	**	0.943	(0.077)	
	80-100% black	1.278	(0.067)	**	1.691	(0.136)	**	0.980	(0.073)	
Non-black borrower and neighborhood is:										
	20-40% black	1.068	(0.055)		1.058	(0.073)		1.125	(0.087)	
	40-60% black	1.175	(0.076)	*	1.318	(0.113)	**	1.018	(0.091)	
	60-80% black	1.228	(0.086)	**	1.345	(0.138)	**	1.145	(0.113)	
	80-100% black	1.310	(0.082)	**	1.643	(0.143)	**	1.020	(0.096)	
Hispanic white borrower		0.829	(0.036)	**	0.896	(0.051)		0.789	(0.051)	**
Asian borrower		0.937	(0.043)		0.938	(0.057)		1.008	(0.075)	
Number of loan-months		503,579			198,680			304,875		
Number of loans		30,307			11,931			18,376		
Log pseudolikelihood		-76588.244			-36244.342			-34292.662		

Cox proportional hazard models of 90-day default, hazard ratios reported.

See tables 1 and 2 for notes on explanatory variables and the left-out categories.

Robust standard errors clustered by census tracts are reported in ( ). Statistical significance is indicated by: *5% and **1%. All models include: a cubic in original loan balance; indicators for 3/27s, prepayment penalty in effect, second home, low documentation, property type (single & 2-4 family), building age (0-10, 11-50, >50 years & missing), >12 months post-adjustment and payment shock is <20, 20-30 & >30%, Hispanic black, Hispanic other race, non-Hispanic other race, female; indicators for missing values of FICO, DTI, prepayment penalty, race, gender; and fixed effects for origination year and calendar quarter. The refinance model also includes an indicator for.cash-out refinance.

Table 6: Hazard Models of 90-Day Default for Fixed Rate Mortgages

		Model 1		Model 2		Model 3		Model 4		Model 5	
Loan characteristics											
Relative interest rate at origination:											
	0-1	1.459	(0.087) **	1.456	(0.087) **	1.451	(0.087) **	1.456	(0.087) **	1.463	(0.087) **
	1-2	2.327	(0.148) **	2.286	(0.145) **	2.262	(0.144) **	2.273	(0.144) **	2.300	(0.146) **
	2-3	3.685	(0.281) **	3.546	(0.267) **	3.462	(0.260) **	3.483	(0.264) **	3.557	(0.269) **
	>3	4.690	(0.439) **	4.499	(0.423) **	4.422	(0.417) **	4.434	(0.418) **	4.487	(0.423) **
FICO at origination:	680-720	1.732	(0.084) **	1.722	(0.084) **	1.725	(0.084) **	1.727	(0.084) **	1.722	(0.084) **
	650-680	2.197	(0.113) **	2.166	(0.111) **	2.163	(0.111) **	2.169	(0.112) **	2.162	(0.111) **
	620-650	2.943	(0.152) **	2.918	(0.150) **	2.902	(0.149) **	2.897	(0.149) **	2.900	(0.150) **
	590-620	3.526	(0.216) **	3.487	(0.213) **	3.471	(0.212) **	3.469	(0.213) **	3.475	(0.213) **
	560-590	4.179	(0.298) **	4.168	(0.298) **	4.152	(0.297) **	4.105	(0.294) **	4.098	(0.294) **
	530-560	4.760	(0.397) **	4.702	(0.392) **	4.722	(0.392) **	4.645	(0.389) **	4.621	(0.388) **
	<530	5.183	(0.522) **	5.226	(0.532) **	5.292	(0.540) **	5.191	(0.533) **	5.135	(0.526) **
Debt-to-income at origination >45%		1.102	(0.040) **	1.094	(0.040) *	1.091	(0.040) *	1.088	(0.040) *	1.091	(0.040) *
Current combined LTV:	60-70%	1.247	(0.065) **	1.167	(0.061) **	1.150	(0.060) **	1.149	(0.060) **	1.166	(0.061) **
	70-80%	1.521	(0.082) **	1.383	(0.075) **	1.358	(0.073) **	1.354	(0.073) **	1.380	(0.074) **
	80-90%	1.988	(0.107) **	1.758	(0.096) **	1.722	(0.093) **	1.714	(0.093) **	1.750	(0.095) **
	90-95%	2.631	(0.176) **	2.261	(0.155) **	2.212	(0.151) **	2.208	(0.151) **	2.256	(0.155) **
	95-100%	2.816	(0.201) **	2.396	(0.175) **	2.338	(0.172) **	2.333	(0.171) **	2.389	(0.175) **
	>100%	3.532	(0.215) **	2.819	(0.181) **	2.732	(0.176) **	2.720	(0.175) **	2.807	(0.181) **
Took on additional mortgage debt		1.118	(0.044) **	1.127	(0.044) **	1.131	(0.044) **	1.132	(0.045) **	1.132	(0.044) **
Has coborrower		0.809	(0.027) **	0.826	(0.027) **	0.833	(0.028) **	0.828	(0.028) **	0.823	(0.028) **
Home purchase		1.094	(0.066)	1.167	(0.071) *	1.190	(0.073) **	1.205	(0.074) **	1.187	(0.072) **
Owner-occupier		1.040	(0.048)	1.062	(0.049)	1.064	(0.049)	1.069	(0.050)	1.070	(0.050)
Full documentation		0.770	(0.063) **	0.753	(0.061) **	0.735	(0.059) **	0.732	(0.060) **	0.747	(0.061) **

continued

		Model 1	Model 2			Model 3			Model 4			Model 5		
Non-racial neighborhood characteristics														
Median income:	<$30,000		1.022	(0.046)		0.963	(0.045)		0.962	(0.045)		1.016	(0.045)	
	$30,000-$40,000		0.977	(0.036)	**	0.960	(0.035)	**	0.958	(0.035)	**	0.975	(0.035)	**
% High school graduates:	70-80%		0.956	(0.035)		0.921	(0.036)	*	0.925	(0.036)	*	0.943	(0.035)	
	>80%		0.960	(0.045)		0.925	(0.045)		0.929	(0.045)		0.943	(0.045)	
% Non-native born:	20-40%		0.971	(0.042)		0.954	(0.040)		0.953	(0.040)		0.968	(0.041)	
	40-60%		0.978	(0.044)		0.956	(0.043)		0.957	(0.043)		0.974	(0.043)	
	>60%		0.865	(0.057)	*	0.876	(0.063)		0.884	(0.064)		0.881	(0.059)	
Recent foreclosure rate:	1-2%		1.216	(0.047)	**	1.144	(0.045)	**	1.142	(0.045)	**	1.203	(0.047)	**
	2-3%		1.273	(0.062)	**	1.177	(0.058)	**	1.174	(0.058)	**	1.253	(0.061)	**
	>3%		1.344	(0.076)	**	1.228	(0.071)	**	1.223	(0.071)	**	1.320	(0.075)	**
Recent REO rate:	1-2%		1.096	(0.051)	*	1.073	(0.050)		1.079	(0.050)		1.096	(0.051)	*
	2-3%		1.123	(0.060)	*	1.098	(0.060)		1.105	(0.060)		1.125	(0.060)	*
	>3%		1.153	(0.048)	**	1.122	(0.047)	**	1.129	(0.047)	**	1.159	(0.048)	**
Share of non-prime loans:	10-20%		1.142	(0.069)	*	1.111	(0.067)		1.118	(0.068)		1.147	(0.069)	*
	20-30%		1.240	(0.081)	**	1.143	(0.077)	*	1.148	(0.077)	*	1.229	(0.081)	**
	>30%		1.152	(0.093)		1.057	(0.088)		1.056	(0.088)		1.137	(0.093)	

continued

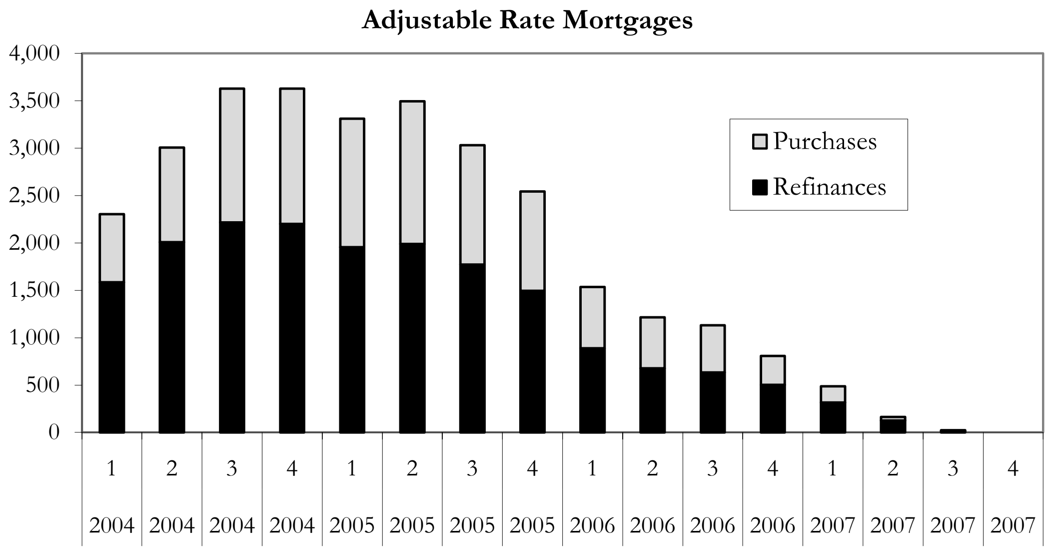
		Model 1	Model 2	Model 3	Model 4	Model 5
Neighborhood racial composition						
% Hispanic:	20-40%			1.011 (0.045)	1.018 (0.046)	
	40-100%			1.080 (0.059)	1.107 (0.063)	
% Asian:	20-40%			1.060 (0.058)	1.063 (0.058)	
	40-100%			1.160 (0.119)	1.182 (0.123)	
% Black:	20-40%			1.157 (0.062) **		
	40-60%			1.308 (0.073) **		
	60-80%			1.311 (0.077) **		
	80-100%			1.347 (0.076) **		
Borrower race/ethnicity						
Black borrower and neighborhood is:						
	0-20% black				1.267 (0.133) *	
	20-40% black				1.195 (0.128)	
	40-60% black				1.270 (0.118) **	
	60-80% black				1.290 (0.095) **	
	80-100% black				1.280 (0.083) **	
Non-black borrower and neighborhood is:						
	20-40% black				1.164 (0.075) *	
	40-60% black				1.333 (0.111) **	
	60-80% black				1.312 (0.110) **	
	80-100% black				1.522 (0.117) **	
Black borrower						1.043 (0.043)
Hispanic white borrower					0.902 (0.048)	0.902 (0.047) *
Asian borrower					0.963 (0.053)	0.967 (0.053)
Number of loan-months				946,922		
Number of loans				29,414		
Log pseudolikelihood		-52412.870	-52327.696	-52306.504	-52282.475	-52305.809

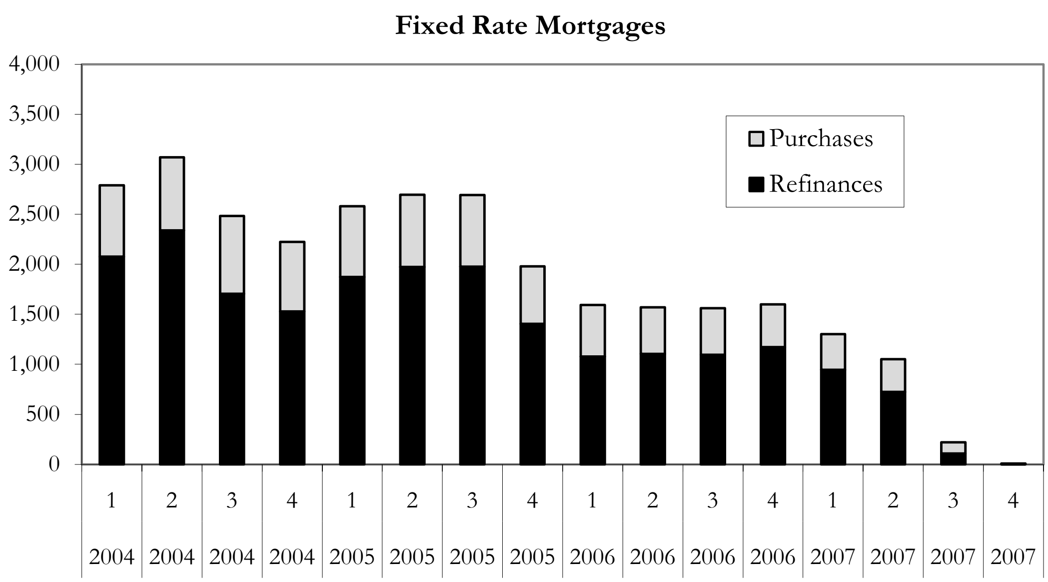
Cox proportional hazard models of 90-day default, hazard ratios reported. See tables 1 and 2 for notes on explanatory variables and the left-out categories.

Robust standard errors clustered by census tracts are reported in ( ). Statistical significance is indicated by: *5% and **1%.

All models include: a cubic in original loan balance; indicators for prepayment penalty in effect, cash-out refinance, second home, low documentation, property type (single & 2-4 family), building age (0-10, 11-50, >50 years & missing); indicators for missing values of FICO, DTI, prepayment penalty; and fixed effects for origination year and calendar quarter. Models 4 and 5 also include indicators for Hispanic black, Hispanic other race, non-Hispanic other race, missing race, female and missing gender.

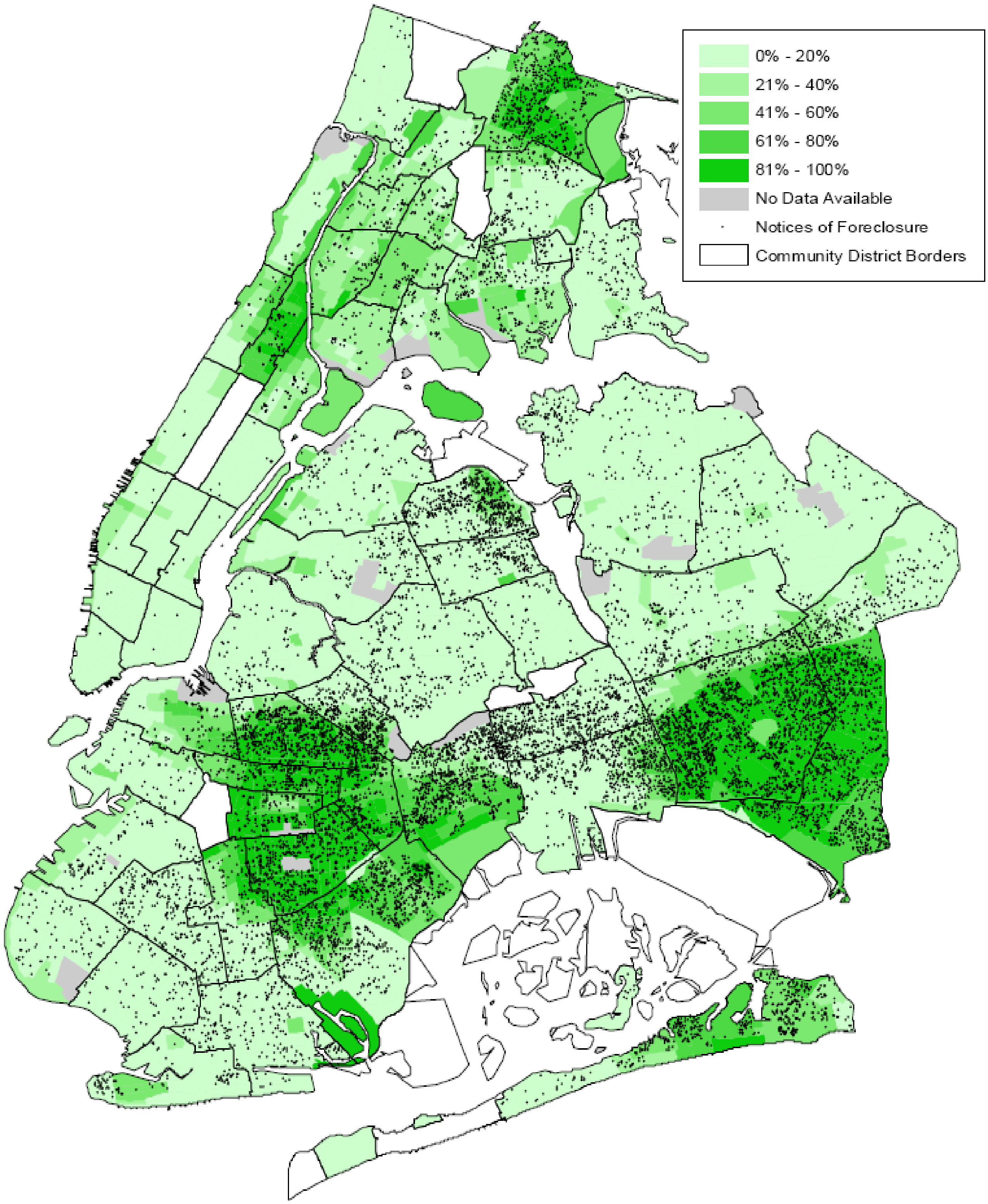
Figure 1: Originations in the Analysis Sample by Quarter





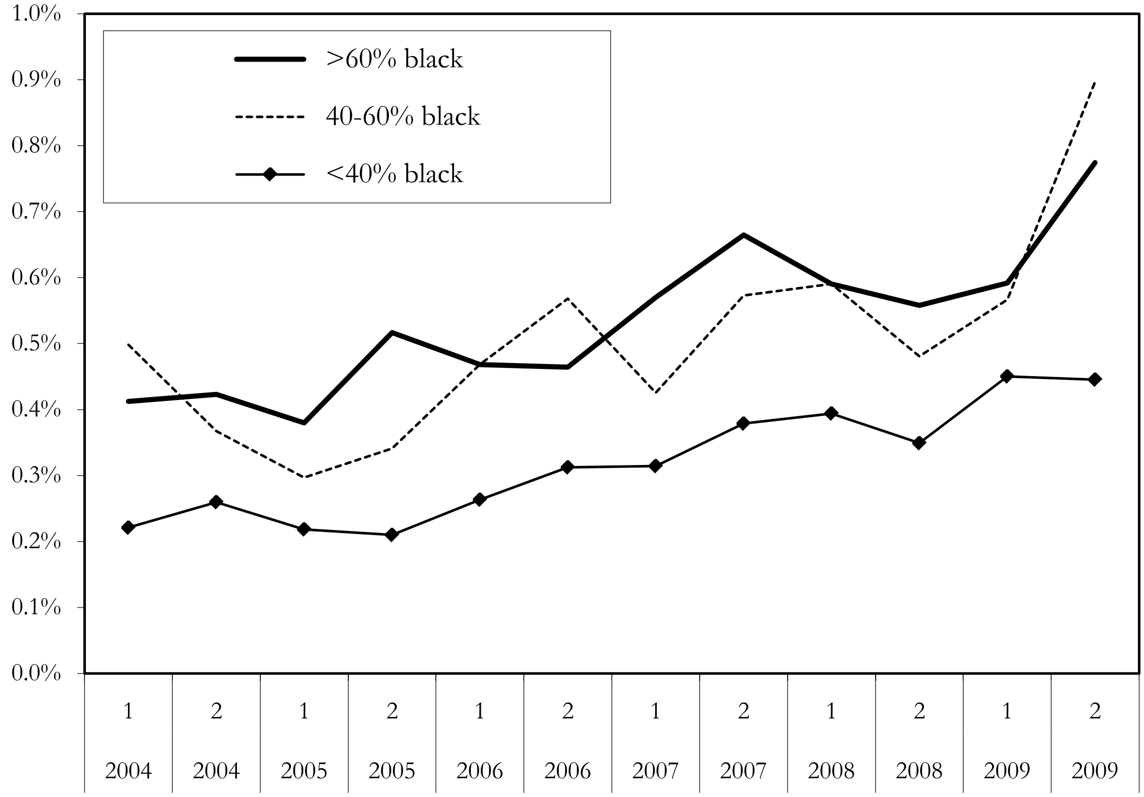
Source: LoanPerformance

Figure 2: Notices of Foreclosure and Percent Black Residents in New York City Census Tracts



Each dot represents a notice of foreclosure (*lis pendens*) occurring in 2009. Source: Public Data Corporation. Census tracts are shaded based on the percent of black residents in the 2000 US Census.

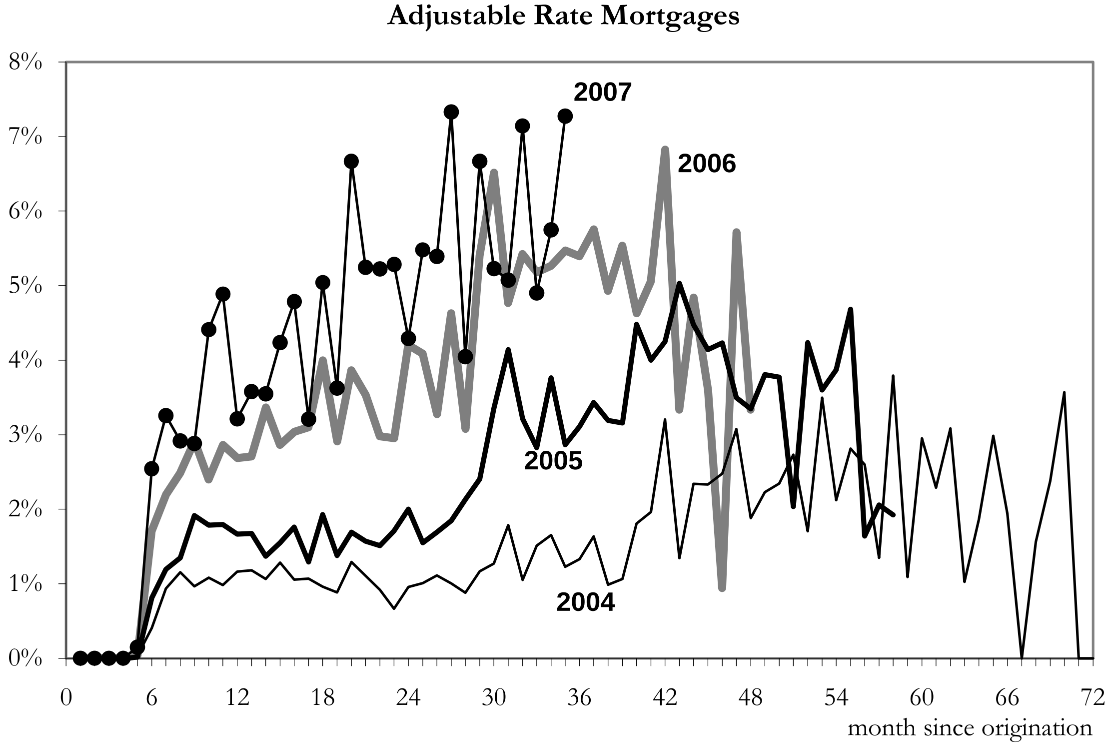
Figure 3: Census Tract Foreclosure Rates in New York City by Percent Black Residents

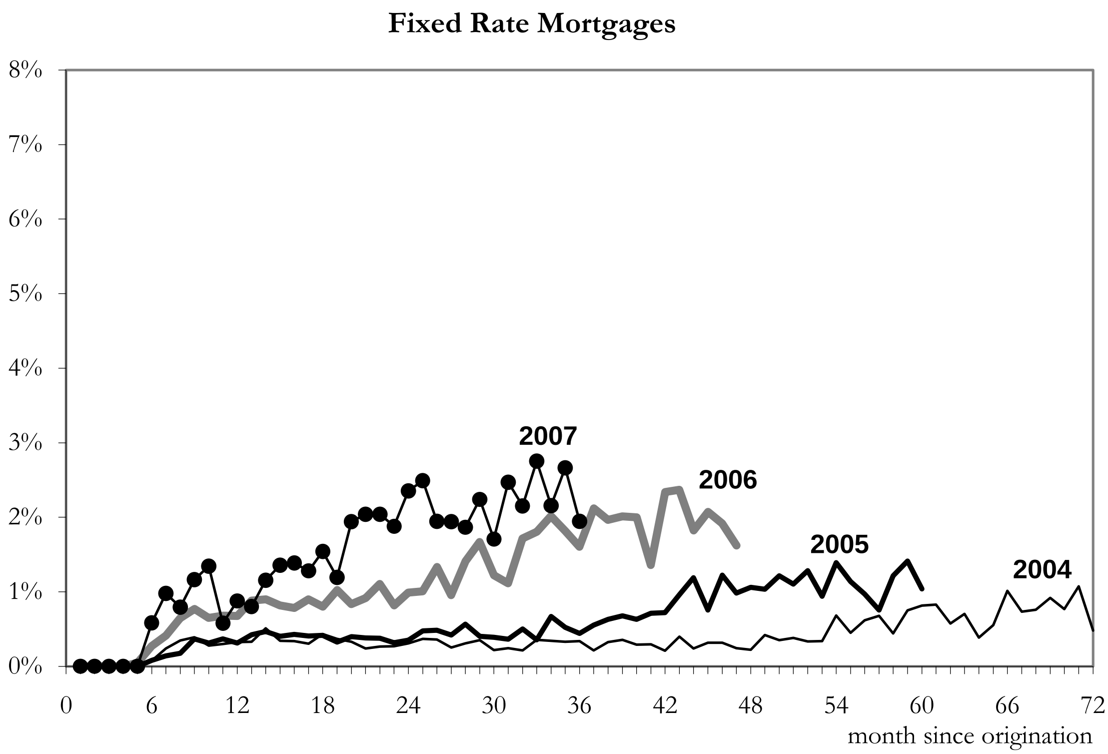


Percent of black residents in census tracts is from the 2000 US Census.

The foreclosure rate is the number of foreclosure notices issued in a census tract within each 6-month period, divided by the stock of buildings in that tract.

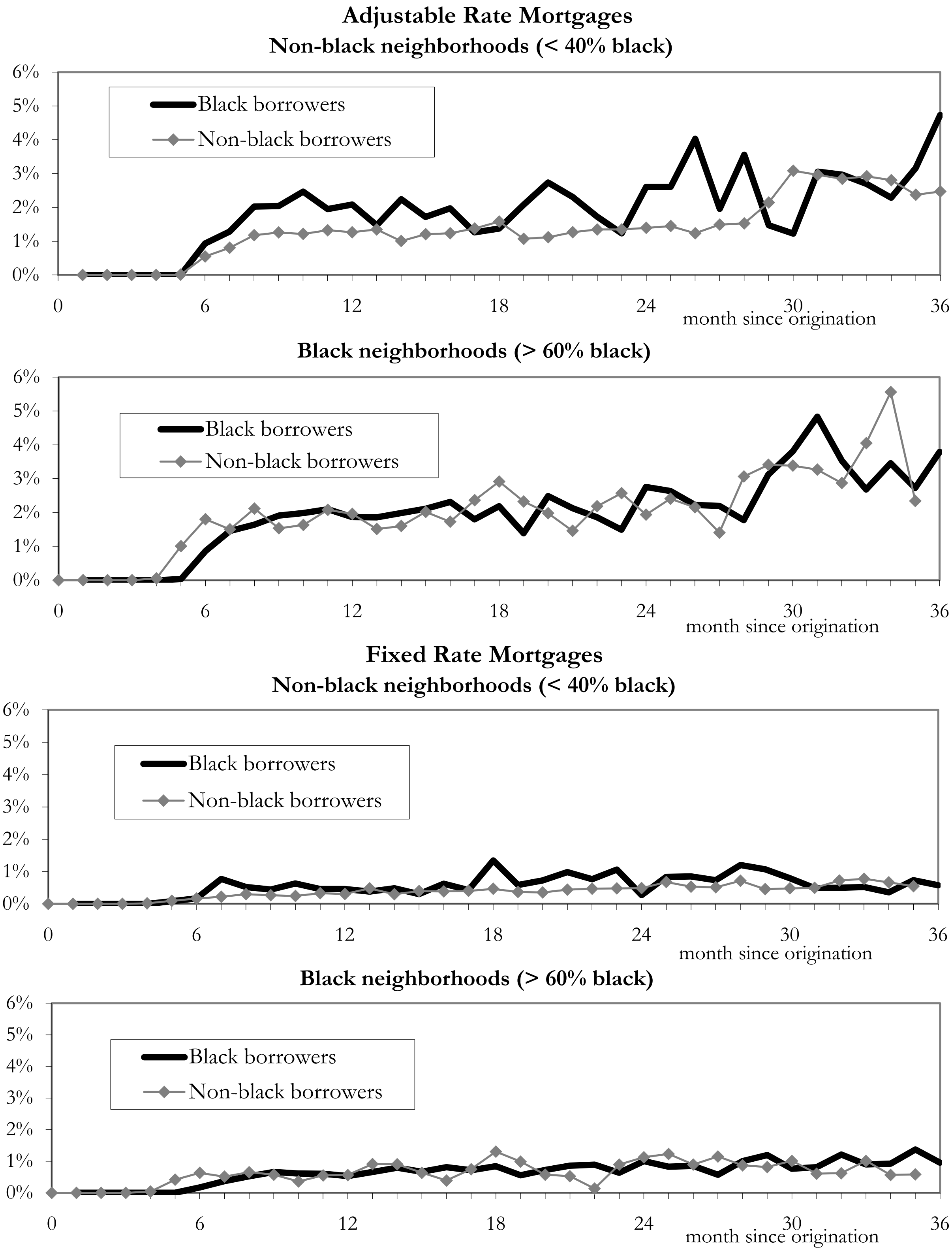
Figure 4: 90-Day Default Hazards by Month Since Origination and Origination Year





Source: LoanPerformance

Figure 5: 90-Day Default Hazards by Month Since Origination, Borrower Race and Census Tract Percent Black



Percent of black residents in census tracts is from the 2000 US Census.