Mortgage Default and Prepayment Risks among Moderate– and Low-Income Households

Working Paper: September 2010

By
Roberto G. Quercia
Anthony Pennington-Cross
Chao Yue Tian

Center for Community Capital
The University of North Carolina at Chapel Hill

Prepared with financial support from The Ford Foundation

Center *for* Community Capital
Research and analysis on the transformative power of capital

UNC
COLLEGE OF
ARTS & SCIENCES

Mortgage Default and Prepayment Risks among Moderate and Low Income Households

Roberto G. Quercia

University of North Carolina at Chapel Hill

quercia@email.unc.edu

Anthony Pennington-Cross

Marquette University

anthony.pennington-cross@marquette.edu

Chao Yue Tian

University of North Carolina at Chapel Hill

tianc@email.unc.edu

September 2010

Abstract

To assess the sustainability of affordable housing credit, a unique sample of community reinvestment loans is analyzed. Conditional probability (hazard) of default tends to be higher and prepayment, lower for lower income groups. However, after controlling for observed mortgage and borrower characteristics, the hazards converge and even reverse in order of magnitude. Furthermore, very low-, low-, and moderate-income groups react with distinct patterns to changes in the loan-to-value ratio and the local unemployment rate. Finally, more financially stretched borrowers (those with high debt-to income ratios) seem to initiate the default option more aggressively as home equity declines.

I. Introduction

For the last 20 years, through multiple administrations and regardless of political leaning, there has been a consistent public policy effort to increase access to homeownership. Homeownership rates peaked at approximate 69 percent in 2004 through 2005 and have since decline to approximately 67 percent as of the second quarter of 2010 (Current Population Survey/Housing Vacancy Survey, Series H-111 Reports, and Bureau of the Census)[1] The positive impact of moving a family from a rental unit to an owned home can come from many different sources, from changes in peer groups to more civic pride and a stronger physical and emotional connection to the area and local schools (See for example, Haurin, Parcel, and Haurin 2002, Aaronson 2000, and Manturuk, Lindblad, and Quercia 2010).

However, as a result of the subprime bust, the near collapse of primary and secondary mortgage markets, property markets, financial markets, and finally the labor markets, the desirability of extending home buying credit to low and moderate income families is again being questioned. Cumulative default rates on subprime loans originated in 2006 surpassed 35 percent and on prime loans surpassed 10 percent (Amromin and Paulson 2010). These default rates represent households that failed as homeowners. Due to the significant failure rate, some have begun to question whether homeownership is the best way for some households to invest and consume shelter and housing services (Florida 2010).

[1] Available as of September 2010 on: http://www.census.gov/hhes/www/housing/hvs/qtr210/files/q210press.pdf

Public policies, such as the Community Reinvestment Act (CRA), the Affordable Housing Goals for Fannie Mae and Freddie Mac, and the Federal Housing Authority, encouraged the financial community to increase access to credit beyond traditional boundaries. Due to these policies and many other factors, access to credit and hence homeownership was expanded to include more households with financial constraints, but safety and soundness concerns still imposed some limitations. Eventually, subprime lending practices extended credit to almost any borrower, even if borrowers could not document income and down payment sources.

As a response to the crisis, it is understandable that many began to question whether the extension of credit was imprudent and could be at least partly to blame for the collapse of the mortgage market. Researchers fall on both sides of the argument. Some contend that the problem was the fact that low income homeowners are riskier borrowers (McArdle 2009, Bhutta 2009); others contend that these borrowers are as risky as other borrowers (and sometimes a better risk (Van Order and Zorn 2002, Mills and Lubuele 1994, Calem and Wacther 1999). Finally, others contend the problem was not the borrower but the risky subprime mortgage products themselves compared with the less risky community reinvestment products (Ding, Quercia, Li, and Ratcliffe 2010). The recent literature has focused on the performance of high cost subprime mortgages (Danis and Pennington-Cross 2008 and Pennington-Cross and Ho 2010). However, the study of the performance of community lending mortgages has been limited due to data availability problems.

Calem and Wachter (1999) examine the loan performance of a program administered by one bank in Philadelphia. The data set consists of 2,390 loans originated between 1988 and 1994. They find lower credit score and high debt-to-income ratio are important factors predicting delinquency. They also find that homes that are very expensive relative to other homes in the same census tract present higher delinquency risk. Quercia, Stegman, Davis and Stein (2002) examine the performance of a sample of approximately 1,000 community reinvestment loans and they find credit score to be important in predicting default and income, but not significantly related to default.

A number of studies compare loan performances for different income groups and neighborhoods. In an early study, Mills and Lubuele (1994) examine neighborhood differences in loan performance using loans from affordable lending programs provided through members of the National Association of Affordable Housing Lenders. They study over 2,000 loans and find that loans originated in low and moderate income neighborhoods have better loan performance than loans originated in other neighborhoods. Van Order and Zorn (2000) compare the performance of loans originated to different income groups and in different neighborhoods using a sample of conventional mortgages consisting of over 400,000 loans originated between 1975 and 1983. However, the dataset does not contain credit score information. They find that although defaults are higher in low-income neighborhoods, different neighborhoods respond to loan-to-value ratio very similarly. However, they find a perhaps unanticipated pattern of defaults: in particular, loans with borrower income to neighborhood income ratios of 81 percent to 120 percent are less likely to default than loans to borrowers whose ratio is less than 60 percent. Borrowers with ratios greater than 200 percent have similar default probability as the lowest income group (whose ratio is less than 60 percent). Van Order and Zorn (2002) compare the performance of loans borrowed by low income households and minority households with others. They model prepay and default using two different large data sets. They conclude that the performance of loans originated to low income and minority borrower is at least no worse than other loans given that low prepayment risks cancel out the high default risks.

In this study, we build on this prior work and examine the default and prepayment propensities of low and moderate income borrowers using a unique sample of community reinvestment loans. The analysis dataset includes detailed borrower, loan and performance information on a large sample of 16,000 loans. The dataset allows us to examine risk propensities for different low income groups. For instance, we are able to examine whether loans originated to households with extremely low income levels, such as to borrowers with incomes less than 40 percent of area median income, exhibit higher levels of default and whether these borrowers are more or less sensitive to the traditional economic and financial drivers of default and prepayment such as interest rates, house prices, and credit scores.

In general, consistent with prior work, we find that lower income groups tend to have higher default and lower prepay hazards than their higher income counterparts. However, after controlling for controlling for determinants of observed risk in a number of simulations, we find that the effect is reversed and the difference in loan performance among income groups becomes much smaller. Moreover, while loan-to-value ratio and local unemployment rate have a larger effect on the default decisions of low income groups, $refi$, which measures the degree to which the refinance option is in the money, has a greater impact on higher income groups. Finally, because of its absence in much of the literature on mortgage performance in both the subprime and prime market segments (for example, Pennington-Cross 2003) we also examine the importance of ability to pay, as measured by the debt payment to income ratio (the debt-to-income ratio). An earlier empirical examination of default by Berkovec, Canner, Gabriel, and Hannan (1998), which ignored the competing risk of prepayment, used 1987, 1988, and 1989 Federal Housing Authority (FHA) loans and included measures of the debt-to-income ratio. The results were inconsistent and indicated default probabilities for loans with debt-to-income ratios of 41 to 53 percent were the lowest and those over 65 percent were the highest.

The remainder of the study is divided into 4 sections. Next, we discuss the empirical strategy followed by a description of the database used in the analysis. We discuss the study findings in section IV and derive implications in section V.

II. Empirical Strategy

Our loan performance analysis is based on a competing risk, proportional hazard framework (for example, see Deng, Quigley, and Van Order 2000 and more recently Pennington-Cross and Ho 2010). The two termination events of loan default and prepay are jointly modeled while addressing the data censoring issue. The model starts with the definition of hazard rate λ. The hazard rate of default (prepay) at time t is the probability of a loan to default (prepay) at time t given that it has survived up to time t. Let λ_i^r be the hazard rate of default ($r = D$) or prepay ($r = P$) for loan i. Conditional on covariates (risk determinants) $X_i(t)$, the hazard rates are defined as

$$\lambda_i^r(t \mid X_i(t)) = \lim_{\Delta t \to 0} \frac{\Pr(t < T_i^r < t + \Delta t \mid T_i^r \geq t, X_i(t))}{\Delta t} \tag{1}$$

and we further assume that there is some unobserved heterogeneity, which is assumed to be independent of observed characteristics, with joint density $g(\theta_D, \theta_P)$ so empirically, the hazards are specified as

$$\lambda_i^r(t \mid X_i(t), \theta_D, \theta_P) = \exp(\lambda_0^r(t) + X_i(t) * \beta_r + \theta_r) \tag{2}$$

where λ_0^r is the baseline hazard for risk r, and β_r is the impact of observed factors on risk r. For estimation simplicity, we assume the same set of risk determinants for both default and prepayment. We define default as the first 90-day delinquency observed on a mortgage and prepayment as the loan is paid off.

Previous studies have employed non-parametric (for example, Deng, Quigley, and Van Order 2000) and parametric (for example, Pennington-Cross and Ho 2010) estimations specifying the baseline hazards. We employ a flexible non-parametric baseline. We model the baseline hazards with local regression, motivated by Cleveland (1979) and others, to smooth the product limit estimators of prepay and default (Kaplan-Meier hazards) described in Greene (2003). We choose the smoothing parameters so that they maximize the AIC information criteria (Cohen 1999). In addition, local regressions are assumed to be convex at the local level.

Prepayment and default events are assumed to be independent and the survival function S_i can be defined as:

$$
\begin{aligned}
& S_i\left(t \mid X_i(t), \theta_D, \theta_P\right) \\
& = \Pr\left(T_i \geq t \mid X_i(t), \theta_D, \theta_P\right) \\
& = \Pr\left(T_i^D \geq t \mid X_i(t), \theta_D, \theta_P\right) * \Pr\left(T_i^P \geq t \mid X_i(t), \theta_D, \theta_P\right) \\
& = \exp\left(-\int_0^t \left[\lambda_i^D\left(s \mid X_i(t), \theta_D, \theta_P\right) + \lambda_i^P\left(s \mid X_i(t), \theta_D, \theta_P\right)\right] ds\right)
\end{aligned} \tag{3}
$$

The log likelihood, LL, is expressed in discreet time assuming risk determinants are constant within each time interval.

$$
LL = \sum_{uncensored} \log \lambda_i^r\left(t \mid X_i(t), \theta_D, \theta_P\right) + \sum_{all} \log S_i\left(t \mid X_i(t), \theta_D, \theta_P\right) \tag{4}
$$

We allow for M groups of loans and they have distinct prepayment and default probabilities. Unobserved heterogeneities are assumed to follow a discrete probability distribution and points of support or mass points p_m sum to one. We follow Pennington-Cross and Ho (2010) and use a logistic transformation so that p_m lie in [0, 1]:

$$p_m = \frac{e^{q_m}}{\sum e^{q_m}} \tag{5}$$

where $q_m \in (-\infty, +\infty)$ and q_1 is normalized to 0 .

We break our sample into four income groups according to the ratio of the borrower household income over county area median income. Using the estimates from the individual analysis, we simulate default and prepayment hazards for individual income groups and illustrate the resulting changes in probabilities for each risk type for each income group.

III. Data

The loan data for the analysis come from a sample of community reinvestment home purchase loans that are part of the Community Advantage Program (CAP)[2] a program created in 1998 by Self-Help Venture Fund, in partnership with Fannie Mae. CAP provides a conforming secondary market outlet for community reinvestment loans. With funding from the Ford Foundation, Self-Help purchases loans and sells them to Fannie Mae while retaining liability for ten years.

[2] See Riley, Ru, and Quercia (2009) for a brief overview of CAP data.

To qualify for CAP, a borrower must fit one of these criteria:

1) Income at or below 80 percent of area median income (AMI),
2) Income between 80 and 115 percent of AMI, and
    a) The home is in a low-income census tract or
    b) The home is in a minority census tract (tract minority composition is greater than 30 percent), or
    c) The borrower is a minority.

The Self-Help portfolio excludes loans whose purchase price exceeds the appraisal price by more than 10 percent, and whose original loan-to-value ratio expressed as a percentage exceeds 125 percent.

The underwriting guidelines of individual lenders are rather flexible toward the borrowers that CAP intends to serve. For example, 2008 guidelines show maximum loan to value ratio allowed of at least 97 percent for all lenders. The data reveals that down payments required range from several hundred dollars to 3 percent for most lenders.

Since its inception, CAP has purchased over 50,000 mortgages national wide (although states like North Carolina have significantly higher representation). The median borrower has income of $ 30,792, or 60% of their AMI, and the median loan amount is $ 79,000 . As is typical with community reinvestment lending, the majority of CAP loans have prime loan features: thirty-year fixed-rate amortizing loans[3] with prime-level interest

[3] There are around 9 percent adjustable rate mortgages in CAP loans. Among fixed rate mortgages, close to 99 percent of them are 30 year fixed rate loans. Table 1 provides a more detailed breakdown by loan types.

rates (Figure 1 compares interest rates on CAP loans and 30-year conventional mortgage rate by the Freddie Mac/Federal Reserve), no prepayment penalties, no balloons, escrows for taxes and insurance, documented income, and standard prime-level fees. Most CAP loans in our sample are originated between 1997 and 2006. Table 2 provides a summary of loans by origination year. As of June 2010, only 4.2% of CAP mortgages had experienced foreclosure since the program started.

We drop the following loans:

1) Adjustable rate mortgages. We drop them since Self-Help has not reported their interest rate history on a quarterly basis.
2) Manufactured housing (the the majority of CAP loans are made to purchase single family homes).
3) Loans originated earlier than 1990s. We drop them to keep the hazard estimation more consistent within the sample.
4) Loans to borrowers with less than 5,000 dollars of annual household income.
5) Loans to borrowers with a ratio of household income to county median income of less than 10 percent or more than 125 percent.

We are left with 16,283 loans whose loan and borrower information is complete. In our survival analysis, we choose to analyze the performance on a quarterly basis and model loan history (loan age) of up to 48 quarters.

Insert Figure 1 here:

Insert Table 1 here:

Insert Table 2 here:

We include a number of loan characteristics, debt-to-income ratio (*dti*), borrower credit score at origination (*credit score*), current loan-to-value ratio (*cltv*), in our hazard analysis. The version of the debt-to-income ratio used in this paper is commonly called the front end ratio and is the ratio of mortgage payments to household income. Households with larger debt burdens should be more susceptible to any income or other debt payment shocks, thus triggering delinquency and potentially default. The impact of high debt burdens on prepayments is more ambiguous. Again, prepayments may be more likely if they are used to refinance in the presence of a negative shock but high debt burdens may also make it harder to qualify to refinance. To examine the issue further, an indicator for high debt-to-income ratio (*hdti*) is interacted with various variables to test whether households that are stretched thin in terms of their monthly ability to make mortgage payments are more or less sensitive to the incentives to default and prepay a mortgage.

Because credit history is negatively associated with mortgage default and positively associated with mortgage prepayment (Pennington-Cross 2003), we include credit score at origination (*credit score*). High loan-to-value ratio is associated with higher probabilities of default (Kau, Keenan, and Kim 1994) and we proxy current loan-to-value ratio (*cltv*) using the outstanding balance on the loan and an estimate of current house value provided by Fannie Mae.

We also construct a measure of the net present value gain from refinancing a fixed rate mortgage ($refi$), commonly described as the value of the refinance option. At time t, the gain from refinancing is the percentage reduction in discounted value of all future mortgage payments if holding current mortgage, PV_c or refinance, PV_r:

$$refi_t = \left[\frac{PV_{ct} - PV_{rt}}{PV_{ct}}\right] \tag{6}$$

where

$$PV_{jt} = \sum_{m=0}^{RMT} \frac{P_{jt}}{(1+d_t)^m} \tag{7}$$

where $j = c, r$, RMT is remaining mortgage term in months that varies with t, discount rate d_t is the ten-year T-Bill rate, and

$$P_{jt} = i_{jt} Q \left[\frac{(1+i_{jt})^{RMT}}{(1+i_{jt})^{RMT} - 1}\right] \tag{8}$$

where i_{ct} is the mortgage interest rate on current mortgage and i_{rt} is the 30-year fixed conventional mortgage rate from federal reserve (reported by Freddie Mac). We expect prepayment hazards to rise with $refi$.

We also include macroeconomic indicators, such as local unemployment rate, mortgage rate volatility, and house price volatility. The county level unemployment rate is used to proxy for labor market conditions. Higher unemployment rates should indicate a higher probability that the borrower has lost their job or has a lower income stream making it more difficult to make mortgage payments. Unemployment may also increase the use of "distressed" prepayments but it also makes it harder to meet underwriting requirements to refinance. We capture interest rate volatility with the use of eight-quarter forward looking moving variance of 30-year fixed rate conventional mortgage rate. We expect that more volatility of interest rates will reduce refinance probabilities since borrowers may wait for interest rates to lower even further. Similarly, house price volatility increases the value of delaying the default (Kau and Kim 1994 and Kau and Keenan 1995). If borrowers believe that prices will increase in the future and extinguish the financial gain from defaulting they will default sooner; or, if they believe prices may drop even further increasing the size of the gain, they may use a strategy of delaying default. We use eight-quarter forward looking moving variance of Federal Housing Finance Agency's House Price Index to capture house price volatility.

In this paper, we address the effect of income difference on loan performance. In addition to analyzing the full sample, we also look at each income group. We partition the sample according to the income over county median income at origination. After considering the income distribution and key qualifications of CAP borrowers the sample is divided into four groups: those with annual income of 10 to 40 percent, 40 to 60 percent, 60 to 80 percent and 80 percent and above AMI. Figure 2 gives the distribution of the ratio.

Insert Figure 2 here:

Table 3 includes the definitions of variables used in the analysis. Table 4 summarizes the variables in the full sample and Table 5 summarizes the variables for each individual income group.

Insert Table 3 here:

Insert Table 4 here:

Insert Table 5 here:

The CAP loan and borrower characteristics can be compared with the Loan Performance data reported in Danis and Pennington-Cross (2008) and Pennington-Cross and Ho (2010). As a first look at the performance and characteristics of CAP loans, we plot out the Kaplan-Meier hazards and local regressions. Figure 3 presents results based on estimations of the default hazards for mortgages. Figure 4 presents results based on estimations of the prepayment hazards.

Insert Figure 3 here:

Insert Figure 4 here:

An examination of these Figures reveals important differences. For CAP loans, default peaks at the 7^{th} and 8^{th} quarters and prepayment peaks at the 11^{th} quarter. These are rather early peaks compared to the subprime fixed rate mortgages reported in Pennington-Cross and Ho (2010)[4]. For that loan sample, both default and prepayment do not peak until loan

---

[4] Default defined as Real Estate Owned and prepay defined as balance becomes zero and in the prior

age is well beyond 4 years. There is a sharper decline in the conditional incidence rates, especially for prepayment, for CAP loans. The peak loan ages for Self-Help are closer to the standard default assumption and standard prepayment model published by the Public Securities Association that are set to peak at 30 months, however, the sharp decline in prepayment rate right after the peak is unique for CAP loans. Given the unique shapes of baseline hazards, we choose to model the baseline under a non-parametric framework rather than imposing a baseline shape a priori.

CAP loan borrowers often have impaired credit, similar to borrowers targeted by subprime lenders. The average credit score at origination is 677 for CAP loan borrowers while the average credit score from Loan Performance fixed rate subprime mortgages is 664 in Pennington-Cross and Ho (2010), and 649 in Danis and Pennington-Cross (2008)[5].

On average, the loans are observed for 18 quarters (more than 4 years) while the Loan Performance sample in Pennington-Cross and Ho (2010) is close to 19 months. Borrowers with higher income-to-AMI ratios have, on average, higher income, lower debt-to-income ratios, larger size loans and larger down payments. Borrowers with the highest income-to-AMI ratio (80 to 125 percent) have on average lower credit score at origination than 40 to 60 and 60 to 80 percent income-to-AMI groups. This may indicate that lenders may be more flexible with regard to credit history when issuing loans to

month the loan was either current or delinquent.

[5] Mortgages analyzed by Pennington-Cross and Ho (2010) were originated from 1998 to 2005 and mortgages were originated from 1996 to 2003 in Danis and Pennington-Cross (2008).

higher income groups.

Compared with the Loan Performance samples, CAP loans have much higher current loan-to-value ratios especially given the greater average loan age. Except for the highest income-to-AMI group, CAP loans tend to have similar or even higher current loan-to-value ratio as compared to the hybrid sample of Loan Performance (Pennington-Cross and Ho 2010).

IV. Estimation Results

Loan performance across income groups

The hazard estimation results for the full sample and for each income group are presented in Table 6 and 7. In the full sample, higher debt-to-income ratio increases both default and prepayment probabilities. The relative impact of debt-to-income ratio on prepayment probability is more pronounced for higher income groups. Keep in mind that the highest income group in our sample is the modest income group. However, the impact on default is not significant for individual income groups. Since the study of debt-to-income ratio on default is limited in the literature, we will explore the impact of debt-to-income ratio in more detail later.

Insert Table 6 here

Insert Table 7 here

Consistent with prior studies, credit score and loan-to-value ratio seem to be very strong indicators for default and prepay. Credit scores (the prior ability to pay financial obligations in a timely fashion) are negatively associated with default and positively associated with prepay, while loan-to-value ratio exhibits the opposite effects. For individual groups, the relative magnitudes of the impact of credit score on default and prepay decrease with income, except for the highest income group. The relative impact of loan-to-value ratio on default decreases with income while its impact increases with income expect for the lowest income group.

The variable $refi$ is a strong and positive indicator for both default and prepay. This is consistent with Pennington-Cross and Ho (2010). The relative magnitudes of the impact of $refi$ on both default and prepay increase for the higher income groups, except for prepayment in the highest income group.

Local unemployment rate is positively related to default and negatively related to prepay. The relative magnitude of impact on default is greatest for the lowest income group and shows virtually no significant impact for the highest income group. The relative impact on prepay increases with income, except for the highest income group.

For the most part, greater variation in future mortgage interest rates increases default and deters prepayment. Its relative impact on default increases with income, except for the highest income group and it also increases with income for prepay except for the income group between 60 to 80 percent of AMI.

Greater variation in local house prices has very little impact on default in the full sample. In the individual group estimates, it is positively and significantly related with default for the highest income group. It is positively related to prepay in the full sample but insignificant in the individual sample estimations.

Estimated Default and Prepayment Hazards

To complement the above analyses, we next examine differences in loan performance between different income groups in more detail through simulations of default and prepayment hazards under different scenarios. First, we compile the default and prepay hazards for each income group by loan age (up to 48 quarters) using their own mean characteristics and their own model estimates. The default and prepayment patterns are presented in Figure 5 and 6 respectively. Default patterns of individual groups over time decrease with income except for the highest income group. The default rate difference of different income groups varies as the loans age over time. At the 8th quarter (around the peak of default of different income groups), the highest default rate is over 30 percent higher than the next highest default rate and is over 70 percent higher than the lowest default rate.

The prepayment hazards are highest for the lowest income group and highest for the highest income group. The differences are at their largest near the peak of the hazard in the 12th quarter. Specifically, in the 12th quarter, the highest prepayment rate is 70 percent higher than the lowest. These results provide some of the defining characteristics of very low income lending – the loans are more likely to default but also less likely to prepay.

Insert Figure 5 here:

Insert Figure 6 here:

Figures 5 and 6 do not control for observed borrower or mortgage characteristics. To help determine whether the differences are driven by the modeled relationship (coefficient estimates) or the different characteristics of the groups, Figures 7 and 8 simulate the baselines for each group using their individual model estimates but the characteristics of the moderate income group (1.25>Income/AMI > 0.8). The simulated default and prepay patterns are presented in Figures 7 and 8 respectively. In general the baseline hazards are much more similar, and the ordering of the results is different. For example, after controlling for borrower and loan characteristics, the lowest income group now has the lowest default baseline and almost the lowest prepayment baseline hazard. This indicates that once observed characteristics are controlled, lower income groups may even outperform higher income group (default) rates within our sample. In addition, the simulated prepayment results using the lower income group estimates produce higher prepayment rates.

In summary, after controlling for observed borrower and loan characteristics, the lowest income group (0.1<Income/AMI<0.4) defaults at the lowest probability and prepays at almost the lowest probability. These results indicate that the lowest income borrowers have loans that are the least likely to terminate in a neutral economic and financial market.

Insert Figure 7 here

Insert Figure 8 here

Next, we compare the impacts on default patterns of individual risk factors. We set the age of the loans to the12th quarter for all models and normalize the probability to 1 for the initial value of each risk factor. Figures 9, 10, and 11 illustrate the default patterns by credit score, loan-to-value ratios, $refi$, and unemployment respectively.

Insert Figure 9 here

Insert Figure 10 here

Insert Figure 12 here

The comparisons show that higher credit scores have similar impact on the default hazards of all income groups. In contrast, loan-to-value ratios have the greatest impact on default probabilities for the lowest income group and the lowest impact for the highest income group. The unemployment rate has the greatest impact on the lowest income group and barely any impact on default patterns of the highest income group. This result likely reflects unobservable factors such as the amount of wealth a household has to help it continue making mortgage payments after losing a job or earning less money.

In summary, lower income households are more sensitive to labor market and housing market conditions. Therefore, we should expect that default rates should rise more rapidly for low and very low income households when house prices decline, when more very low down payment loans are originated, and when the labor market deteriorates.

A Closer Look at Debt-to-Income Ratio

As discussed earlier, we find significant impacts of the debt-to-income ratio on the default and prepayment propensities in the full sample but these impacts were statistically insignificant when the sample is divided into separate income groups. Since the availability of debt-to-income ratio information in the analysis dataset offers a unique opportunity to study the impact of income on loan performance, we examine more narrowly whether borrowers with different debt-to-income burdens respond to other risks differently. This additional analysis is especially relevant for low income lending if low and moderate income borrowers have fewer resources and thus may be more likely to be overburdened by debt than their higher income borrowers. Not surprisingly, the data used in this analysis (Table 5) indicate that the debt-to-income ratio is positively correlated with income. The highest income group (1.25>Income/AMI>0.8) has an average debt-to-income of 22 percent and the lowest income group (0.1<Income/AMI<0.4) has an average debt-to-income of 33 percent.[6]

To simplify the analysis, we consider debt-to-income ratios to be high if they exceed 31 percent or 38 percent or higher (*hdti*)[7]. Figure 12 shows the distribution of debt-to-income ratio in the sample. There is no indication in the distribution of a mass of scores around an enforced underwriting standard. The estimation results are presented in Tables 8 and 9. The coefficients can be interpreted as additive.

Insert Figure 12 here

Insert Table 8 here

Insert Table 9 here

The coefficient estimates for *hdti* itself and many interaction terms are largely insignificant, regardless of the definition of debt-to-income ratio considered. However, some patterns are apparent. For instance, we see a greater impact of current loan-to-value ratio on default for borrowers with high debt to income ratios, although the estimate using the 31 percent cutoff is barely significant at about 90 percent. Taken together, these

[6] These debt-to-income ratios represent the income used by underwriters to qualify a loan. If underwriters only include documented income until it exceeds or meets a required ratio then the reported income for the higher income households may be understated. If the actual income was available instead of the documented income it may be that the debt-to-income ratio would be even lower for higher income households.

[7] Traditionally, Fannie Mae guidelines set a debt-to-income limit of 38% for community mortgages (see for example Fannie Mae seller service guide for 2006, Part X, Chapter 3, Section 304: Underwriting Community Lending Mortgages (08/31/02): "The borrower's total debt-to-income ratio should not exceed 38 percent - unless a higher ratio is adequately offset by another Contributory Risk factor that decreases the likelihood of mortgage default." The current government sponsored loan modification program, the Home Affordable Modification Program (HAMP), requires participating lenders to bring debt-to-income ratios to 31 (see U.S. Department of Treasury 2009).

results show that households saddled with greater housing burdens may respond more aggressively to equity position and other default triggers. In addition, a high debt-to-income ratio is associated with a lower propensity to prepay the loan.

Conclusions

Even prior to the Great Recession, there was an ongoing debate about the desirability of extending credit to low and moderate income families because of the belief that such extensions of credit would be too costly. Part of this debate was based on the contention that the higher default risks exhibited by these borrowers could be compensated by lower prepayment risks. This study contributes to the understanding of default and prepayment risks among low and moderate income households by carrying out a comprehensive analysis based on a competing risk proportional hazard model using a unique sample of over 16,000 community reinvestment loans that includes detailed loan and borrower characteristics including debt-to-income ratio.

Overall, consistent with prior studies, we find that low and very low income groups exhibit higher default but lower prepayment probabilities. The estimated default and prepay hazards (conditional quarterly probability) tend to peak much earlier than those reported for subprime loans elsewhere in the literature.

However, after controlling for observed loan and borrower characteristics, the distance between probabilities become smaller and the order of magnitude is reversed. The predicted probability of default is lowest for the lowest income group (0.1<Income/Area Median Income < 0.4) and almost the lowest for the probability of prepayment. While the default patterns of different income groups by credit score are very similar, the default patterns vary greatly for changes in the equity position and the unemployment rate. Lower income groups tend to be more sensitive to equity position and local unemployment rate changes. Overall, these findings suggest that while the expected propensity of a low income loan to terminate is relatively low in adverse economic conditions in the labor market or the housing market loans made to low income households should be expected to deteriorate more quickly.

Our results also offer evidence that debt-to-income ratio is important in determining default patterns for this sample of low income borrowers. Borrowers with higher debt-to-income ratio are found to be more sensitive to equity position. A higher loan-to-value ratio is associated with a higher default risk for households burdened by large mortgage payments. These findings provide some empirical support for current efforts to modify loans to reduce housing payments to 31 percent of household income. They are consistent with prior work by Ding and Quercia (2009) who encourage decision makers to tailor loan modifications to the unique characteristics of the borrower, loan, and market, including the use of meaningful principal reduction.

Acknowledgements

We would like to thank Anthony Yezer, David Ribar, and Janneke Ratcliffe for helpful comments, and Sarah Riley for help with data.

References

Amromin, G., and A. L. Paulson. 2010. Default Rates on Prime and Subprime Mortgages: Differences and Similarities. *Profitwise News and Views* September: 1-14.

Aaronson, D. 2000. A Note on the Benefit of Homeownership. *Journal of Urban Economics* 47 (3): 356–369.

Berkovec, J., G. Canner, S. Gabriel, T. Hannan. 1998. Discrimination, Competition, and Loan Performance in FHA Mortgage Lending. *Review of Economics and Statistics* 80(2) 241–250.

Bhutta, N. 2010. GSE Activity and Mortgage Supply in Lower-Income and Minority Neighborhoods: The Effect of the Affordable Housing Goals. *The Journal of Real Estate Finance and Economics* 2010 May: 1-24.

Calem P. S. and S. M. Wachter. 1999. Community Reinvestment and Credit Risk: Evidence from an Affordable-Home-Loan Program. *Real Estate Economics* 27 (1): 105-134.

Cohen, R. A. 1999. An Introduction to PROC LOESS for Local Regression. Proceedings of the 24th SAS Users Group International Conference, Paper 273.

Danis, M. A., and A. Pennington-Cross. 2008. The delinquency of subprime mortgages. *Journal of Economics and Business* 60(1-2): 67-90.

Deng, Y., J. M. Quigley, and R. Van Order. 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica* 68(2): 275-307.

Ding, L. and R. G. Quercia. 2009. Tailoring Loan Modifications: When is Principal Reduction Desirable. Working Paper. Center for Community Capital, University of North Carolina, Chapel Hill.

Ding, L., R. G. Quercia, W. Li, and J. Ratcliffe. 2010. Risky Borrowers or Risky Mortgages: Disaggregating Effects Using Propensity Score Models. *Journal of Real Estate Research*. Forthcoming

Florida, R. 2010 Homeownership Is Overrated: Today's economy requires a more mobile workforce. *Wall Street Journal Online*, June 7th http://online.wsj.com/article/SB10001424052748703559004575256703021984396.html.

Greene, W. 2003. *Econometric Analysis* 5th edition. Upper Saddle River: NJ

Haurin, D., T. Parcel, and R. Haurin. 2002. Does Homeownership Affect Child Outcomes? *Real Estate Economic* 30(4): 635-666.

Kau, J., D. Keenan, and T. Kim. 1994. Default Probabilities for Mortgages. *Journal of Urban Economics* 35 (3): 278-296.

Kau, J., and Kim, T. 1994. Waiting to Default: The Value of Delay. *Journal of American Real Estate and Urban Economics Association* 22 (3): 539-51.

Kau, J., and D. Keenan. 1995. An Overview of the Option-Theoretic Pricing of Mortgages. *Journal of Housing Research* 6 (2): 217-244.

Manturuk, K., M. Lindblad, and R. G. Quercia. 2010. Homeownership and Civic Engagement in Low-Income Urban Neighborhoods: A Longitudinal Analysis. Working Paper. Center for Community Capital, University of North Carolina, Chapel Hill.

McArdle, M. 2009. Rethinking the CRA. *The Atlantic*, Jun 26th, downloaded from http://www.theatlantic.com/business/archive/2009/06/rethinking-the-cra/20207/

Mills, E. M. and L. S. Lubuele. 1994. Performance of Residential Mortgages in Low- and Moderate-Income Neighborhoods. *The Journal of Real Estate Finance and Economics* 9: 245-260.

Pennington-Cross, A. 2003. Credit History and the Performance of Prime and Nonprime Mortgages. *The Journal of Real Estate Finance and Economics* 27(3): 279-302.

Pennington-Cross, A. and G. Ho. 2010. The Termination of Subprime Hybrid and Fixed Rate Mortgages. *Real Estate Economics* 38(3): 399-426.

Quercia R. G., M. A. Stegman, W. R. Davis, and E. Stein. 2002. Performance of Community Reinvestment Loans: Implications for Secondary Market Purchases. N. P. Retsinas and E. S. Belsky editors. *Low-Income Homeownership*. Joint Center for Housing Studies: Cambridge, Massachusetts and Brookings Institution Press: Washington, D.C.

Quercia, R. G., S. F. Riley, and H. Ru. 2009. Community Advantage Program Database: Overview and Generalization. *Cityscape*, 11(3).

U.S. Department of Treasury. 2009. Homeowner Affordability and Stability Plan: Executive Summary. Press Releases. U.S. Department of the Treasury. Downloaded from on http://financialstability.gov/latest/tg33.html

Van Order, R. and P. Zorn. 2002. Performance of Low-Income and Minority Mortgaes. N. P. Retsinas and E. S. Belsky editors. *Low-Income Homeownership*. Joint Center for Housing Studies: Cambridge, Massachusetts and Brookings Institution Press: Washington, D.C.

Table 1: Number of Loans by Loan Type

Loan Type	Number of Loans
AdjusTable-rate mortgage, 10 year term.	106
AdjusTable-rate mortgage, 15 year term.	263
AdjusTable-rate mortgage, 20 year term.	193
AdjusTable-rate mortgage, 25 year term.	120
AdjusTable-rate mortgage, 30 year term.	3,470
Fixed-rate mortgage, 10 year term	26
Fixed-rate mortgage, 15 year term	364
Fixed-rate mortgage, 20 year term	122
Fixed-rate mortgage, 25 year term	93
Fixed-rate mortgage, 30 year term	41,697
Fixed-rate mortgage, 40 year term	12
Total	46,466

Table 2: Number of Loans by Year Originated

Year	Number of Loans
1966	1
1968	1
1983	1
1984	3
1989	13
1990	43
1991	126
1992	439
1993	703
1994	935
1995	1,265
1996	1,579
1997	3,310
1998	4,390
1999	3,344
2000	5,457
2001	6,583
2002	4,789
2003	3,190
2004	3,355
2005	2,645
2006	2,010
2007	1,486
2008	800
2009	8
Total	46,476

Table 3: Variable Definitions

Variable	Definition
dti	Debt-to-income ratio. The fraction of combined income that goes toward mortgage payments
credit score	Borrower's credit score at origination
cltv	Current loan-to-value ratio
refi	Percentage reduction in present value of future payments if refinance into the market rate
unemp_rate	County level unemployment rate
varmrate	Variance of future national mortgage rate
varhpi	Variance of MSA level house price index reported by Federal Housing Finance Agency
hdti	Indicator for loans with high debt-to-income ratio
*hdti*fico*	Interaction between *credit score* and *hdti*
*hdti*cltv*	Interaction between *cltv* and *hdti*
*hdti*refi*	Interaction between *refi* and *hdti*
*hdti*unemp*	Interaction between *unemp_rate* and *hdti*
hdti varmrate*	Interaction between *varmrate* and *hdti*
hdti varhpi*	Interaction between *varhpi* and *hdti*

Table 4: Descriptive Statistics of Self-Help Mortgage Data -- Full Sample

		Full sample	
		Mean	Std Dev
Constant within Each Loan	*original balance*	$83,473	$33,436
	annual income	$30,930	$10,102
	credit score	677	64
	dti or front-end ratio	0.27	0.07
	income / ami	0.58	0.16
	Loans	16,283	
Varies within Each Loan	*refi*	0.090	0.076
	loan age	18.1	10.6
	cltv	0.786	0.209
	unemp_rate	5.536	1.750
	varmrate	0.116	0.105
	varhpi	0.0002	0.0004
	Observations	212,149	

Table 5: Descriptive Statistics of Self-Help Mortgage Data -- by Income Group

		Income/AMI (10% -40%)		Income/AMI (40%-60%)		Income/AMI (60%-80%)		Income/AMI (80%-125%)	
		Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
Constant within Each Loan	*original balance*	$58,793	$21,973	$78,357	$28,867	$94,564	$33,958	$104,111	$39,740
	annual income	$18,416	$4,148	$27,289	$5,938	$36,683	$7,326	$47,614	$11,092
	credit score	665	65	678	65	683	63	671	61
	dti	0.33	0.09	0.28	0.07	0.25	0.06	0.22	0.06
	income / ami	0.34	0.05	0.51	0.06	0.69	0.06	0.91	0.09
Varies within Each Loan	Loans	2,109		7,120		5,981		1,073	
	refi	0.104	0.077	0.089	0.076	0.083	0.074	0.097	0.075
	loan age	18.6	10.6	18.6	10.9	17.3	10.1	17.5	11.3
	cltv	0.741	0.236	0.779	0.203	0.806	0.194	0.831	0.244
	unemp_rate	5.616	1.718	5.551	1.770	5.438	1.710	5.798	1.868
	varmrate	0.116	0.105	0.116	0.105	0.117	0.106	0.119	0.107
	varhpi	0.0002	0.0003	0.0002	0.0004	0.0002	0.0004	0.0002	0.0003
	Observations	30,053		93,724		75,163		13,209	

Table 6: Completing Risk Results for Full Sample and Each Income Group -- Default

	Full sample		Income/AMI (10% -40%)		Income/AMI (40%-60%)		Income/AMI (60%-80%)		Income/AMI (80%-125%)	
	Coef	Std Err	Coef	Std Err	Coef	Std Err	Coef	Std Err	Coef	Std Err
dti	0.098*	0.022	0.004	0.057	0.064	0.035	0.058	0.042	-0.090	0.107
credit score	-0.628*	0.027	-0.712*	0.064	-0.641*	0.039	-0.615*	0.044	-0.695*	0.107
cltv	0.208*	0.013	0.326*	0.050	0.225*	0.017	0.221*	0.023	0.151*	0.055
refi	0.285*	0.029	0.129*	0.059	0.271*	0.040	0.363*	0.052	0.466*	0.125
unemp_rate	0.174*	0.021	0.214*	0.051	0.162*	0.032	0.190*	0.038	-0.013	0.103
varmrate	0.196*	0.024	0.156*	0.053	0.189*	0.035	0.240*	0.044	0.225*	0.103
varhpi	0.025	0.022	-0.025	0.059	0.038	0.031	0.025	0.040	0.235*	0.087
loc1	-0.589*	0.176	0.333*	0.106	-0.476*	0.184	-0.391	0.251	-0.194	0.192
loc2	-0.051	0.116	-3.212*	1.090	-0.001	0.194	-0.683*	0.182	-1.277	0.831
q1	.		.		.		.		.	
q2	-0.673*	0.205	-0.780*	0.305	-1.024*	0.340	1.048*	0.300	-0.997*	0.462
Loans	16,283		2,109		7,120		5,981		1,073	
Obs	212,149		30,053		93,724		75,163		13,209	
Loglike	-39,509		-5,030		-17,091		-14,554		-2,554	

Notes: * indicates significance at 95 percent. *loc1* and *loc2* are shift parameters of the two heterogeneity groups. *q1* and *q2* are logistic transformation parameters for the heterogeneity mass points. *q1* is normalized to zero.

Table 7: Completing Risk Results for Full Sample and Each Income Group -- Prepay

	Full sample		Income/AMI (10% -40%)		Income/AMI (40%-60%)		Income/AMI (60%-80%)		Income/AMI (80%-125%)	
	Coef	Std Err	Coef	Std Err	Coef	Std Err	Coef	Std Err	Coef	Std Err
dti	0.074*	0.015	0.137*	0.046	0.184*	0.022	0.199*	0.023	0.291*	0.058
credit score	0.224*	0.016	0.282*	0.055	0.216*	0.024	0.144*	0.024	0.177*	0.058
cltv	-0.340*	0.019	-0.403*	0.064	-0.336*	0.027	-0.365*	0.028	-0.556*	0.092
refi	0.437*	0.016	0.372*	0.054	0.436*	0.024	0.473*	0.024	0.313*	0.057
unemp_rate	-0.174*	0.016	-0.214*	0.052	-0.212*	0.025	-0.103*	0.023	-0.139*	0.061
varmrate	-0.174*	0.016	-0.149*	0.052	-0.197*	0.025	-0.139*	0.025	-0.220*	0.063
varhpi	0.040*	0.014	0.057	0.044	-0.013	0.024	0.018	0.021	0.001	0.056
loc1	0.290*	0.058	-1.509*	0.262	0.135	0.082	-1.712*	0.366	0.350*	0.132
loc2	-1.972*	0.314	0.434*	0.161	-2.634*	0.966	0.275*	0.070	-1.761*	0.591
q1	.		.		.		.		.	
q2	-0.673*	0.205	-0.780*	0.305	-1.024*	0.340	1.048*	0.300	-0.997*	0.462
Loans	16,283		2,109		7,120		5,981		1,073	
Obs	212,149		30053		93724		75163		13,209	
Loglike	-39,509		-5,030		-17,091		-14,554		-2,554	

Notes: * indicates significance at 95 percent. *loc1* and *loc2* are shift parameters of the two heterogeneity groups. *q1* and *q2* are logistic transformation parameters for the heterogeneity mass points. *q1* is normalized to zero.

Table 8: Completing Risk Results by Debt-to-Income Ratio -- Default

	Debt-to-Income Ratio Cutoff at 38%		Debt-to-Income Ratio Cutoff at 31%	
	Coef	Std Err	Coef	Std Err
dti	0.099*	0.027	0.058	0.033
credit score	-0.647*	0.029	-0.637*	0.031
cltv	0.206*	0.014	0.196*	0.015
refi	0.275*	0.030	0.310*	0.033
unemp_rate	0.174*	0.023	0.146*	0.026
varmrate	0.195*	0.025	0.190*	0.029
varhpi	0.027	0.024	0.055	0.030
hdti	0.009	0.123	0.142	0.083
*hdti*credit score*	0.223*	0.087	0.042	0.046
*hdti*cltv*	0.169*	0.054	0.036	0.022
*hdti*refi*	0.093	0.098	-0.070	0.055
*hdti*unemp_rate*	-0.030	0.077	0.078	0.045
*hdti*varmrate*	-0.017	0.084	0.013	0.049
*hdti*varhpi*	0.001	0.059	-0.067	0.045
loc1	-0.661*	0.210	-0.608*	0.165
loc2	-0.010	0.113	-0.096	0.121
q1	.		.	
q2	-0.649*	0.218	-0.727*	0.196
Loans	16,283		16,283	
Obs	212,149		212,149	
Loglike	-39,498		-39,494	

Notes: * indicates significance at 95 percent.
loc1 and *loc2* are shift parameters of the two heterogeneity groups.
q1 and *q2* are logistic transformation parameters for the heterogeneity mass points. *q1* is normalized to zero.

Table 9: Completing Risk Results by Debt-to-Income Ratio -- Prepay

	Debt-to-Income Ratio Cutoff at 38%		Debt-to-Income Ratio Cutoff at 31%	
	Coef	Std Err	Coef	Std Err
dti	0.092*	0.018	0.099*	0.022
credit score	0.226*	0.017	0.205*	0.019
cltv	-0.337*	0.020	-0.329*	0.022
refi	0.432*	0.016	0.406*	0.019
unemp_rate	-0.171*	0.017	-0.159*	0.019
varmrate	-0.172*	0.017	-0.161*	0.019
varhpi	0.034*	0.015	0.028	0.019
hdti	-0.175*	0.077	-0.131*	0.051
*hdti*credit score*	-0.007	0.059	0.058	0.034
*hdti*cltv*	-0.041	0.059	-0.030	0.037
*hdti*refi*	0.063	0.057	0.095*	0.034
*hdti*unemp_rate*	-0.038	0.061	-0.051	0.035
*hdti*varmrate*	-0.019	0.061	-0.047	0.035
*hdti*varhpi*	0.048	0.042	0.029	0.029
loc1	0.307*	0.061	0.309*	0.056
loc2	-1.941*	0.324	-1.998*	0.317
q1	.		.	
q2	-0.649*	0.218	-0.727*	0.196
Loans	16,283		16,283	
Obs	212,149		212,149	
Loglike	-39,498		-39,494	

Notes: * indicates significance at 95 percent.
loc1 and *loc2* are shift parameters of the two heterogeneity groups.
q1 and *q2* are logistic transformation parameters for the heterogeneity mass points. *q1* is normalized to zero.

Figure 1: Freddie Mac 30-Year Fixed Mortgage Rate vs. CAP Sample Mortgage Rates at Origination

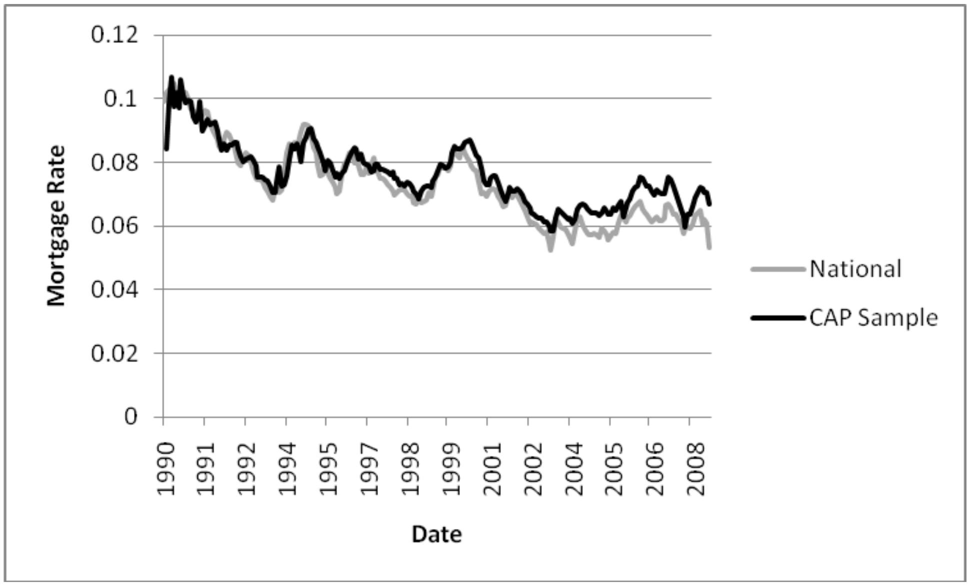

Figure 2: Distribution of Household Income over County Median Household Income at Origination

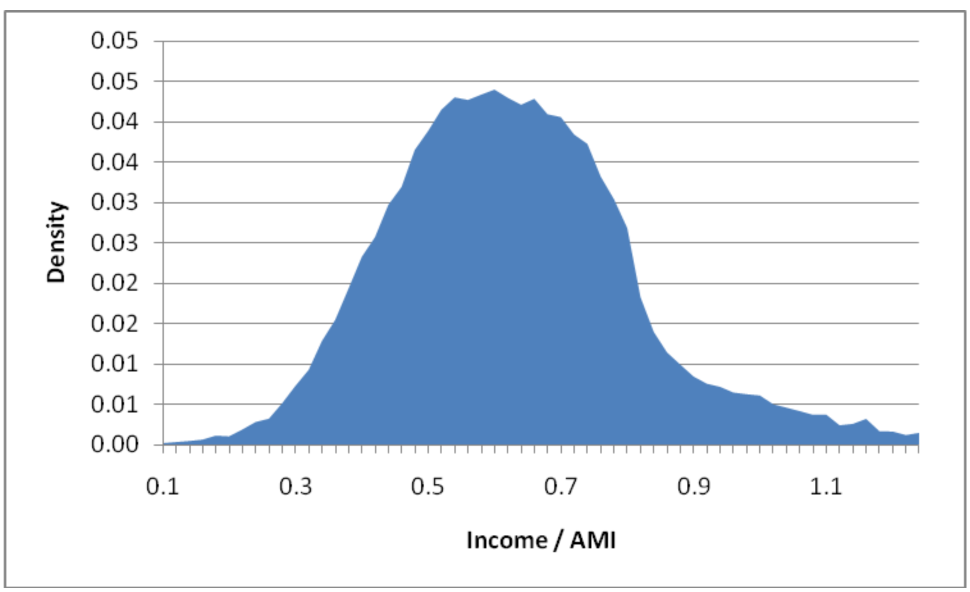

Figure 3 Non-Parametric Default Hazards

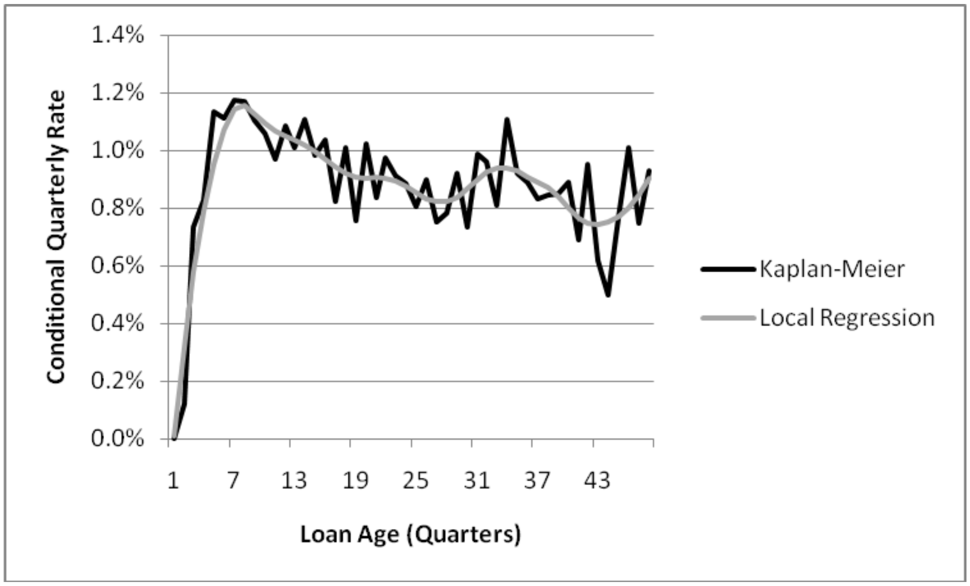

Figure 4 Non-Parametric Prepayment Hazards

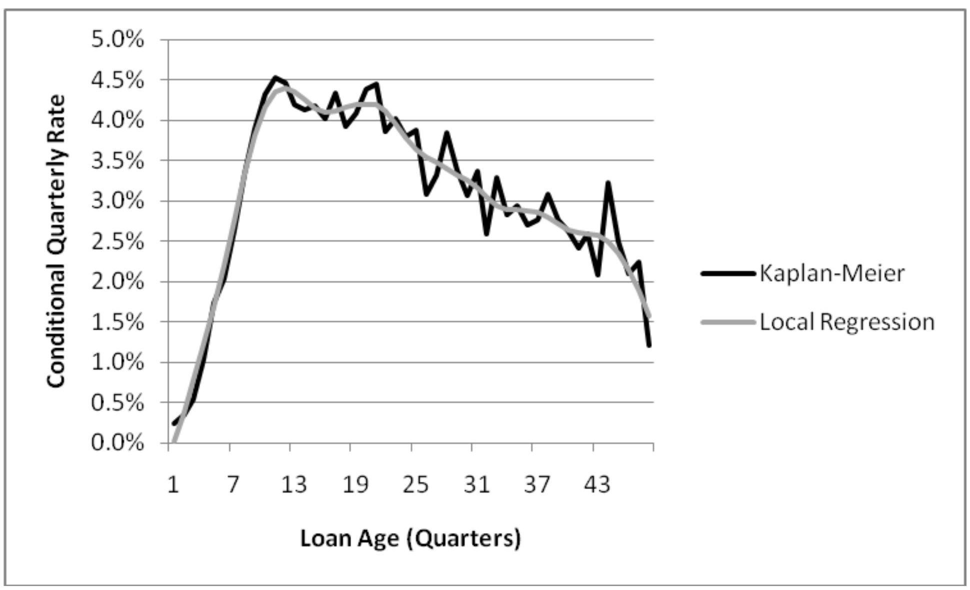

Figure 5: Default Pattern for Each Income Group (Own Characteristics, Own Model)

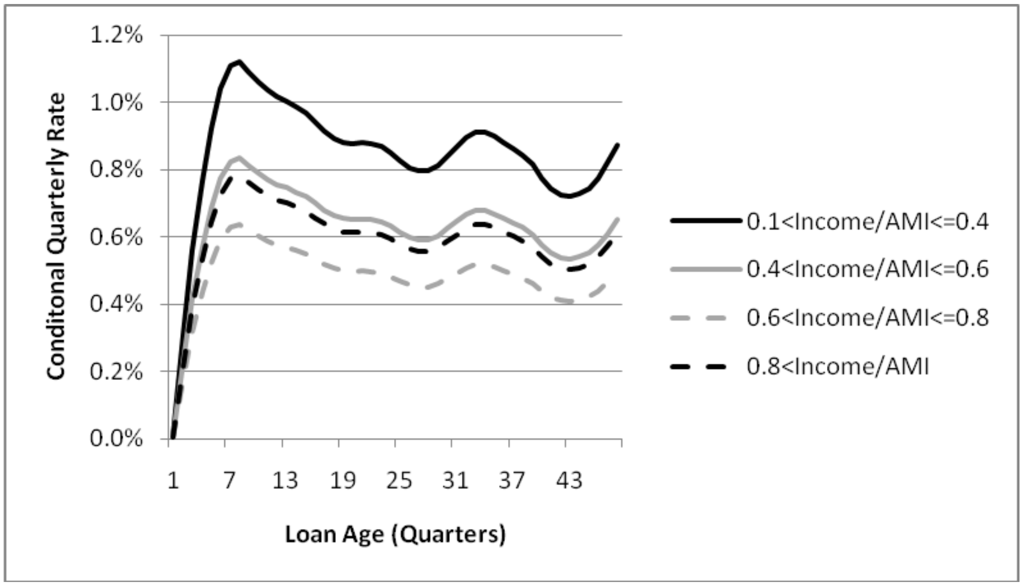

Figure 6: Prepay Pattern for Each Income Group (Own Characteristics, Own Model)

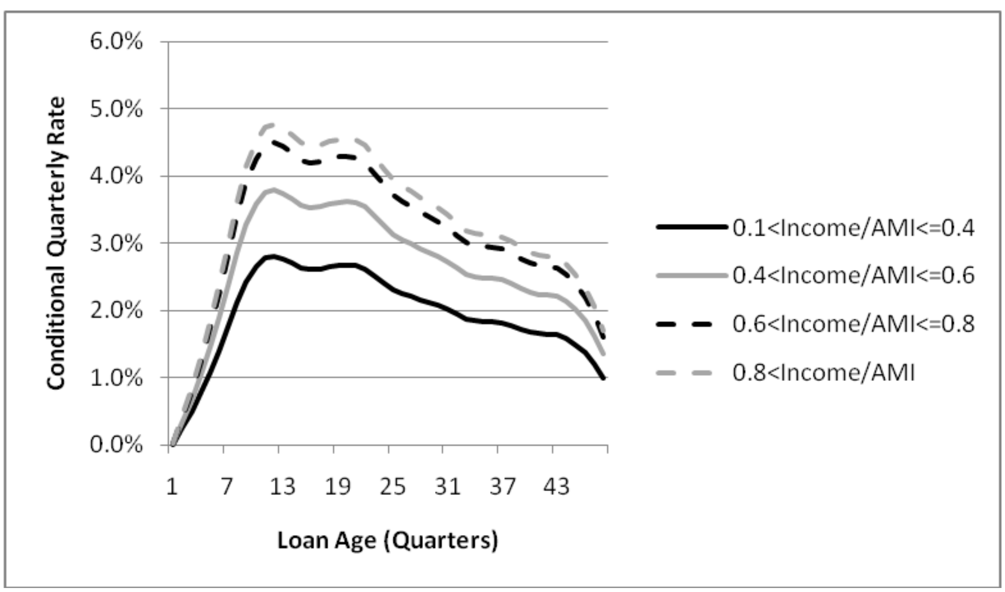

Figure 7: Default Pattern Predicted by Different Income Group Estimates for Income Group 4 (Income/AMI >0.8) Characteristics

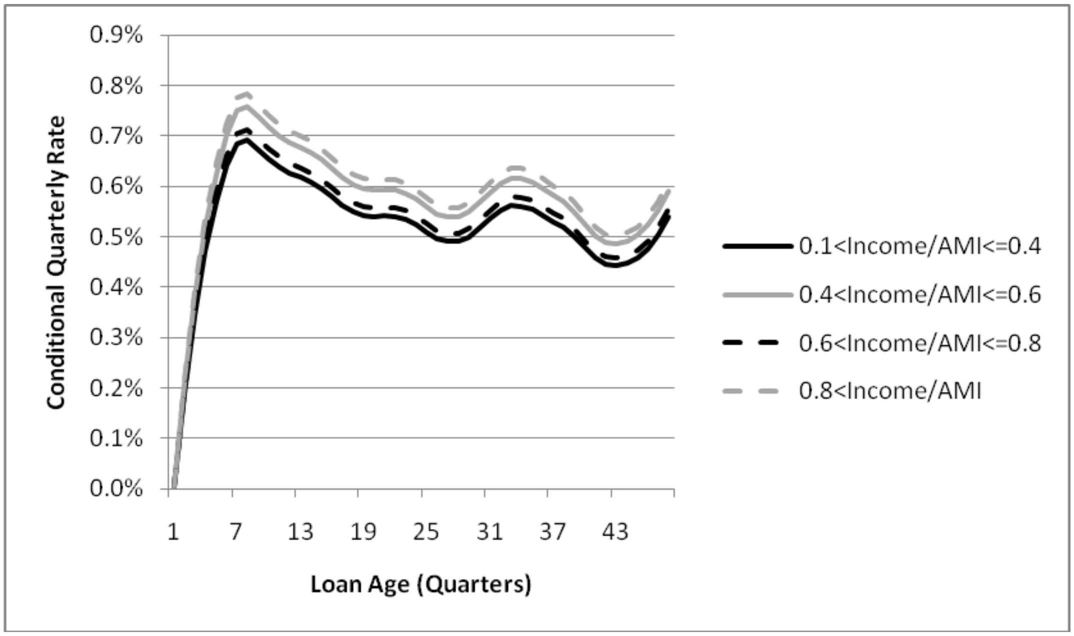

Figure 8: Prepay Pattern Predicted by Different Income Group Estimates for Income Group 4 (Income/AMI >0.8) Characteristics

Figure 9: Default Probability Changes (Normalized to 1) and Credit Score for Different Income Groups at Quarter 12 (Own Characteristics, Own Model)

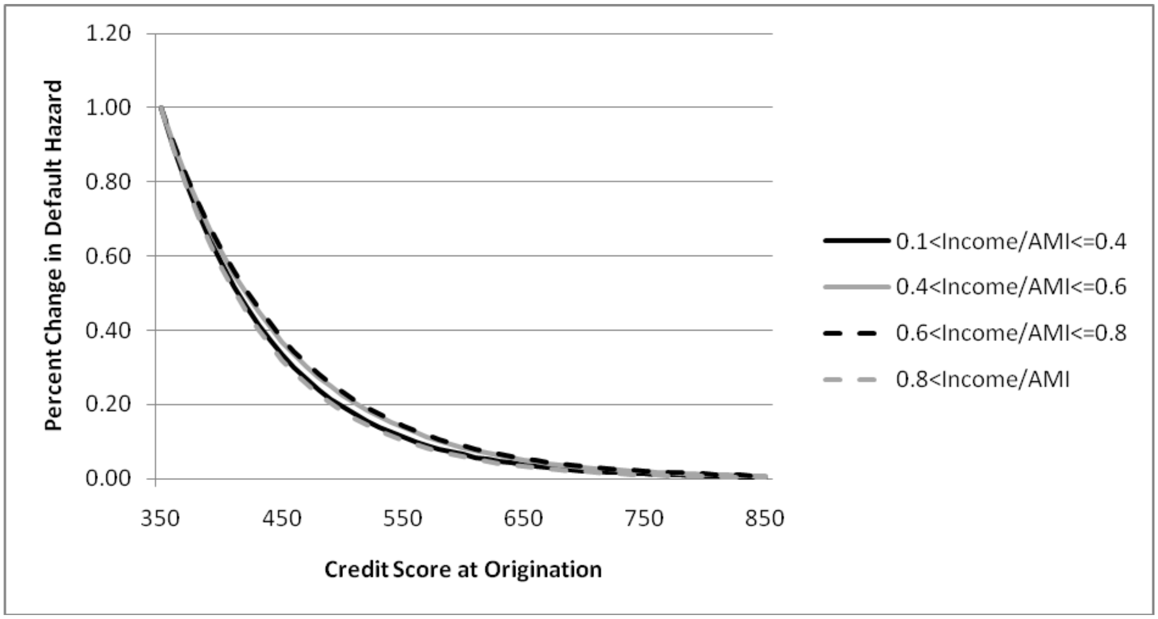

Figure 10: Default Porbability Changes (Normalized 1) and Loan-to-Value Ratio for Different Income Group at Quarter 12 (Own Characteristics, Own Model)

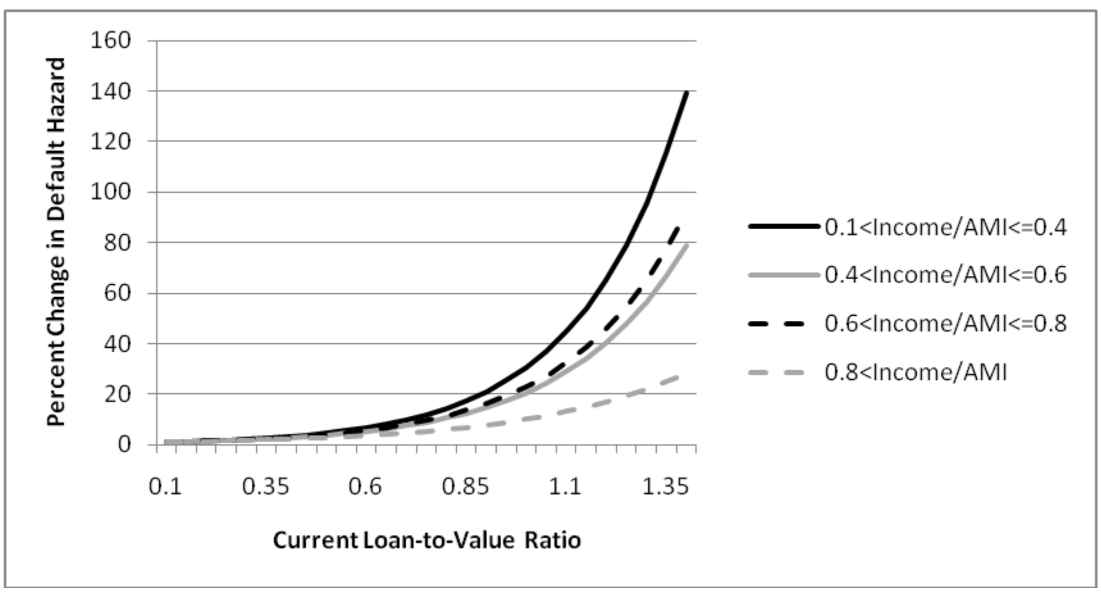

Figure 11: Default Probability Changes (Normalized to 1) and Unemployment for Different Income Group at Quarter 12 (Own Characteristics, Own Model)

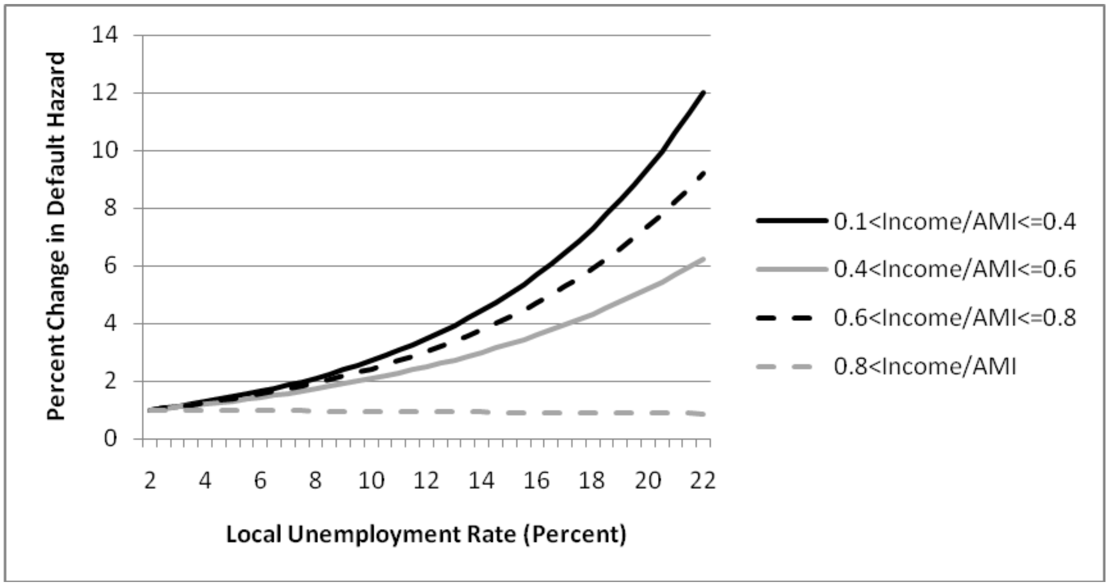

Figure 12: Distribution of Debt-to-income ratio at Loan Origination