CENTER FOR ECONOMIC RESEARCH DISCUSSION PAPER NO. 00-03.

Credit Risk and the Subprime Mortgage Market

Anthony Pennington-Cross
Research Institute for Housing America

Anthony Yezer
The George Washington University

Joseph Nichols
Standard & Poor's *DRI*

October 2000

DEPARTMENT OF ECONOMICS
Funger Hall, 2201 G Street, NW
Washington, District of Columbia 20052
Phone: (202) 994-6150
FAX: (202) 994-6147

Credit Risk and the Subprime Mortgage Market

October 12, 2000

Anthony Pennington-Cross[•]
Research Institute for Housing America
1919 Pennsylvania Avenue, NW
Washington, DC 20006
Voice: 202-557-2839
E-mail: anthony_pennington-cross@housingamerica.org

Anthony Yezer
Department of Economics
The George Washington University
Washington, DC 20052
Voice: 202-994-6755
E-mail: yezer@gwis2.circ.gwu.edu

Joseph Nichols
Standard & Poor's *DRI*
1200 G Street NW
10th floor
Washington, DC 20005
Voice: 202-383-2024
E-mail: jbnichols@dri.mcgraw-hill.com

[•] Address all correspondences to Anthony Pennington-Cross. The authors are the research director at the Research Institute for Housing America and a professor at The George Washington University. This work is supported in part by the Department of Housing and Urban Development Contract # DU100C000018441. This paper represents the authors approach and opinions only.

Credit Risk and the Subprime Mortgage Market

Anthony Pennington-Cross (Research Institute for Housing America)
Anthony Yezer (The George Washington University), &
Joseph Nichols (Standard & Poor's *DRI*)

Abstract

During the 1990s the role of subprime lending in the mortgage market has grown from a small and rarely considered segment into a highly visible, growing, and controversial part of the market. While public controversy exists regarding what role subprime lending should take there has been little evidence showing who and how homebuyers use subprime lending. This paper examines this issue by using a model of mortgage selection (subprime, FHA, or prime) for FHA eligible loans. The results show that borrowers who have had problems managing their financial responsibility and those who carry substantial non-real estate debt are more likely to use subprime lending. But, the subprime market does not primarily provide mortgages to traditionally 'underserved' households and neighborhoods. Instead it serves those with enough wealth to compensate for other deficiencies in their mortgage application.

Credit Risk and the Subprime Mortgage Market

I. Introduction

While it is very difficult to accurately measure the size of the subprime mortgage market, Home Mortgage Disclosure Act (HMDA) data used in conjunction with a list of lenders that specialize in subprime lending shows that in the early 1990s subprime mortgage lending was a small part of the mortgage market (0.74%)[1]. In addition, or perhaps because of the level of market presence market, subprime lending was an afterthought in the minds of most borrowers, large lenders, and their regulators. But by the end of the 1990s, subprime lending, and risk based pricing in general had become a more visible part of the mortgage market as measured. By this time, according to HMDA, the subprime market had grown to almost 9% of the total mortgage market, 10.9% of refinances, and 4.9% of home purchase originations.[2] This dramatic increase in explicit risk based pricing of mortgages, has help to fuel a growing debate about the behavior, consequences, and role of subprime lending in the mortgage market. Unfortunately there is very little information available to the public to inform this debate. But, as demonstrated by Pennington-Cross and Nichols (2000), the credit risk, as measured by downpayment, payment to income ratio, and credit history, of a borrower is on average worse for lower income borrower. It should follow

[1] Data tabulation from the Harvard Joint Center's State of the Nation's Housing State of the Nation's Housing report Table A-11. These statistics are calculated from HMDA and exclude manufactured housing. We define subprime lending as subprime specialist as identified by HUD based industry sources, denial rates, refinance share, and lender name.

[2] While this increase in subprime lending seems impressive it is likely to be biased upward due to changes in the structure of the subprime market (what types of institutions are doing the lending) and changes in HMDA coverage. For instance, mortgage bankers were not covered under HMDA until 1993. The extent of mortgage banker reporting has also increased over time as they adjusted

that subprime lending is more active in lower income neighborhoods. In contrast to these perceptions, in 1998 almost 48% of subprime loans were made to moderate and high income borrowers. Part of the answer to this puzzle is that subprime lending is currently not oriented towards borrowers who have low downpayments. It is the Federal Housing Authority (FHA) at the Department of Housing and Urban Development (HUD) that has been the major provided of low down payment loans for households with moderate and low income. Instead, subprime lenders typically require larger downpayments to compensate for other weaknesses in the mortgage application. Typical weaknesses include poor credit history, an unwillingness to document income, and high debt burdens. Subprime lenders have identified a segment of mortgage applicants who are not candidates for a prime mortgage, but have accumulated enough wealth to provide a downpayment - high risk borrowers with assets.

While regulators, newspapers, and lenders debate the benefits and costs of subprime lending, the more basic question of how and when households use subprime lenders remains unanswered. This paper will try to answer this question by estimating a model of home purchase mortgage selection, i.e. the choice among conventional prime, FHA, and subprime mortgages, for FHA eligible borrowers using a large and unique data set that includes household credit history, income stream, debt outstanding, and wealth or downpayment. The sample is restricted to those borrowers who are eligible to get an FHA insured loan to help

their data systems to the demands of HMDA. Lastly, when depositories acquired mortgage banking subsidiaries, along with their subprime business, all these loans become reported.

focus the analysis on lower and moderate income households and the relationship between private and government activity in the mortgage market.

For most borrowers, FHA insurance costs more than private mortgage insurance, but underwriting standards are more lenient, making FHA insurance attractive to a substantial proportion of buyers. In addition, the conventional market includes both prime and subprime lending. Subprime lending standards are generally less stringent than prime and, on some dimensions, less demanding than FHA lending standards. However, the costs of subprime mortgages are substantially higher than prime mortgages and usually higher than FHA mortgages. In this sense, the mortgage market can be viewed as an ordered ranking of lending standards and costs. Conventional prime borrowers must meet the highest lending standards but are rewarded with lower costs, while conventional subprime borrowers meet the most lenient lending standards but pay the highest costs. Between the two conventional alternatives lie FHA mortgage insurance.

Two major data innovations are included in the construction of the data set used in this analysis. First, credit history is included. This allows us to examine the relative importance to borrowers of avoiding credit problems, increasing downpayments, and maintaining an adequate income stream when applying for a mortgage. Second, to our knowledge this is the first time that subprime loans have been included in this framework of first lien home purchase mortgage originations.

The literature has found that downpayment size and income are important determinants of both tenure choice (Haurin 1991, Linneman and Wachter 1989) and mortgage choice (Hendershott, LaFayette, and Haurin 1997 and 1995). However, these studies have not included the effects of the third constraint, the borrower's credit history. They have also ignored the role of subprime financing. The inclusion of borrower credit history could significantly alter the effects of income and downpayment constraints on mortgage choice. For example, Steinbach (1998) has shown that subprime lenders routinely allow a very high monthly payment to income (PTI) ratio if the borrower has excellent credit and equity.

Prior studies have also been limited by surveys with fairly small sample sizes (for example - American Housing Survey (AHS), Survey of Consumer Finance, National Longitudinal Survey, Survey of Consumer Credit). Hendershott, Lafayette, and Haurin (1997), using the AHS, have 581 observations. Surveys typically depend on the households to self-report financial information. Self-reporting may introduce a bias due to an unwillingness to disclose unfavorable credit history or a lack of knowledge. Self-reporting is particularly problematic for those with the most severe credit problems, precisely the group most active in the subprime market.

This paper addresses shortcomings in the current literature by including a borrower's credit history from a credit reporting bureau, explicitly including subprime financing along with prime and FHA financing, and using a very large sample of more than 48,000 observations from 39 MSAs.

II. Models of Mortgage Choice[3]

We follow the well-known models of Hendershott, LaFayette, and Haurin (1997) and Gabriel and Rosenthal (1991). The following model incorporates credit history measures, as part of the household's characteristics, into the LTV and PTI constraints. Household *j* maximizes utility over a set of choices including tenure status, mortgage type, and housing quantity, subject to budget and lending constraints:

$$\underset{h,x}{Max}\, U_j \left\langle h_j , x_j \middle| Z_j \right\rangle \tag{1.}$$

subject to,

$$Y_j = Ch_j + Px_j \tag{2.}$$

where (h,C) and (x,P) are respectively the levels and prices of housing and levels and prices of non-housing consumption, Y is household income, Z is a vector of household characteristics, and j indexes the households. The household solves this problem and finds the optimal amount of housing $\left\{ h * \left\langle Y_j , C, P \middle| Z_j \right\rangle \right\}$ and non-housing goods $\left\{ x * \left\langle Y_j , C, P \middle| Z_j \right\rangle \right\}$.

Within the housing tenure choice, borrowers also face a choice of prime, FHA, or subprime financing. Because prime mortgage financing, including any required mortgage insurance payments, is priced below FHA mortgage insurance and subprime mortgages are generally priced highest, we expect that borrowers will choose, in order, prime, FHA, and lastly subprime financing.

But, borrowers are constrained by lending requirements. In fact, it could be argued that in the short run many households have very little choice (a corner solution) among mortgage types because they are so heavily constrained by a credit history that cannot be altered or wealth insufficient to meet loan-to-value and monthly-payment-to-income ratios. Lenders use standards (PTI, LTV, and credit history) to limit the maximum credit risk presented by any borrower. Because FHA lending standards are more lenient than prime lending standards, wealth and income constrained borrowers are more likely to use FHA mortgage financing than prime financing. Subprime financing can be less strict than both FHA and prime financing regarding maximum front- and back-end PTI ratios (Steinbach, 1998 and Sub-Prime Funding Corp. Underwriting Manual, 1998). Credit history also plays a large role in the qualification process. For instance, in Freddie Mac's Affordable Gold program (Gold Measure Worksheet version 2.0) applicants with low FICO scores (less than 600 points) need very large downpayments, low PTIs, and a short term on the mortgage to qualify. Although subprime lenders allow 60% debt ratios and even current bankruptcies, they may also require a 30% downpayment to mitigate for the perceived risks of high PTI and poor credit history. Subprime lenders even have low documentation lending programs such as "No Income Verification" or "No Ratio" for borrowers with good credit history and a strong asset base. The mortgage market provides mortgage credit to a wide variety of borrowers because lenders can use a variety of approaches to compensate for weaknesses of an application. This flexibility is

---

[3] The basic model used here is an extension of Pennington-Cross and Nichols (2000), which examined the FHA-conventional choice.

greatest in subprime lending where credit scores can compensate for low downpayments and equity can compensate for having unverifiable income.[4]

The percentage downpayment constraints for prime (*c*), FHA (*f*), subprime (*s*) are

$$D_j^l(Z_j)V_j \leq W_j \quad l = c, f, s \tag{3.}$$

where V_j is the property value, W_j is household wealth and D_j is the minimum downpayment percentage. Payment to income limits are also conditioned on household characteristics (Z_j) and are

$$M_j \leq g_j^l(Z_j)Y_j \quad l = c, f, s \tag{4.}$$

where M_j is the monthly mortgage payment, and g_j is the maximum housing expense to income ratio allowed by the lender. In general, prime lenders are more conservative with respect to g_j and D_j (requiring greater downpayments and higher income ratios) than FHA and subprime lenders for a given set of characteristics.

Conditioned on tenure and mortgage type (fixed or adjustable), the household selects prime, FHA, or subprime financing. Household utility, U_l, can be written as $U_l = V_l + \varepsilon_l$, where V_l is the indirect utility function, a function of measurable characteristics and defined as $V_l = \beta' X_l$, where X_l is a matrix of variables that Gabriel and Rosenthal (1991) and Hendershott, LaFayette, and Haurin (1997) found to be important. This includes the relative price of mortgage insurance, permanent income, the magnitude of the value constraint (see

[4] The characteristics of the subprime lending environment are characterized fromWeicher (1997), Steinbach (1998), and the Sub-Prime Funding Corp. Underwriting Manual downloaded from www.allstatecapital.com/manual2.html on 3/17/98.

equations 3 and 4), and household characteristics (indicators of credit behavior, demographics, and location).

Using the well-developed characteristics of the conditional discrete choice model {McFadden (1981)} we can estimate the probability that a household will choose prime, FHA, or subprime lending. Two different estimators are used. First we estimate a multinomial logit model. This estimator does not order the choices made and requires that the same arguments explain the probability of making each choice. Second, we use the apparent ordering of mortgage choices (prime, FHA, and subprime) and estimate an ordered logit model.

III. Credit History, Mortgage, and Demographic Data

Table 1 provides descriptions, mean, minimum, maximum, and the standard deviation of each variable. For instance, the mean FICO score is 693 with a minimum of 406 and a maximum of 826 providing a good breadth of credit history experiences. The data in this study came from four sources. First was the F-42 database of the Department of Housing and Urban Development (HUD) Federal Housing Authority (FHA), which contains detailed loan information and household characteristics for FHA loans, but no credit history. Second was a real estate transaction database from Experian, which has detailed loan information and household identifiers (address of the property, the amount of the loan, the value of the property, the LTV, and the type of the loan) but no information on household characteristics. It contains a census of conventional loans in each county covered by Experian. This database was built from property transfer records at the local level. Third was individual borrower credit history from Experian. This credit history was matched to FHA and conventional loans by name, social security number and property address, with all identifying information subsequently deleted. Fourth was the Home Mortgage Disclosure Act (HMDA) data. HMDA data was matched by loan amount, census tract, and lender identification to conventional- Experian loans, to provide income and racial characteristics of households securing conventional loans.[5]

To separate the subprime and prime conventional loans a list of subprime lenders that report to HMDA created by the Office of Policy, Development, and Research (PD&R) in the United States Department of Housing and Urban Development (Randal Scheessele (1998)) was used. This list was created from trade publications and may therefore not include all subprime lenders who report to HMDA. In addition, not all subprime lenders report to HMDA. The probability of reporting for HMDA purposes increases with lender size. Lastly, the list is unable to separate prime from subprime lending by HMDA reporters that traditionally originate both types of loans. Measurement error may include some conventional loans categorized as prime that may actually be subprime and some loans categorized as subprime that could actually be prime loans.

The sample includes fixed rate loans originated between February 1996 and July 1996, excluding loans for multifamily properties, refinancing, non-owner occupancy, and loans made to investors. The loans were matched by Experian to credit history files archived on March 31, 1996 by address, name, and social

[5] The credit history from Experian matched to the loan information at an 84 percent rate. Conventional loans matched successfully to HMDA 52 percent of the time. Appendix 1 provides

security number. This date was chosen to insure that the credit data did not include information on the new mortgage, but was as current as possible. Observations with missing or obvious data coding errors were excluded. [6]

A stratified sampling scheme varied sampling rates inversely with the FHA market share in each MSA. In subsequent statistical analysis, the effects of the sample stratification were offset by weighting each observation inversely to its sampling probability. Specifically, conventional loans were sampled at 1/3 of the FHA sampling rate. The final sampling rate of FHA loans was, on average, 20.5% with a minimum of 6.1% in Chicago, IL and a maximum of 76.6% in Toledo, OH. The average sample of loans in an MSA was 1,233 with a maximum of 2,391 in Philadelphia, PA and a minimum of 309 loans in Raleigh, NC. The total sample of 48,105 observations, of which 26,247 are prime loans, 21,246 are FHA insured loans, and 612 are subprime loans, contain data on loan terms, borrower demographics, and credit history.

Downpayment, Income, and Credit History

Because FHA lending standards require very low downpayments and even insure mortgages with negative equity after insurance premiums have been financed, we would expect mean FHA LTVs to be very high. In contrast, prime lenders generally require larger downpayments and even subprime lenders

---

additional information on the matching process.

[6] Incomplete data was defined as having missing values for one of more of the key variables used in the analysis (mortgage amount, property value, date of closing for the mortgage, interest rate, term of the mortgage, indicator for a first time home buyer, purpose of the loan, and the name, social security number, income, and assets of the borrower). Some variables were not missing data but instead contained data entry errors (e.g., LTVs greater than 300% or income of $20). The following set of conditions was used to identify any observations containing obvious data errors: FICO scores greater than 850 or less than 360, LTV greater than 110% or less than 20%,

typically do not finance mortgages with less than 5% down. In fact, subprime lenders require borrowers with poor credit history to provide large downpayments to compensate for the perceived higher risk of default and delinquency. Therefore, it is not surprising that Table 2 shows that the average downpayment for subprime loans was 16.2% well above FHA average of 5.7%. In addition, prime borrowers have better PTIs and credit (FICO) scores. Note that subprime borrowers lie between FHA and prime borrowers on average in terms of LTV, PTI, and credit scores.

While it is clear that FHA serves borrowers who are wealth constrained and cannot afford substantial downpayments, the borrowers using subprime lenders apparently are diverse and not easily characterized. Perhaps the answer lies in the ability of the subprime lender to use discretion and unique lending programs that may not require that the borrower's income be verified or that none of the standard ratios (LTV or PTI) be used in the screening process. While a borrower who does not provide documentation supporting a steady income stream could not qualify for prime or FHA financing, it does not imply that the borrower has little wealth or a poor credit history.

IV. Model Specification and Results

The choice model is estimated for a sample of 48,105 households that purchased homes in 39 MSAs from February through July 1996. Because it can be argued that LTV and mortgage choices are jointly determined, LTV is

---

annual income of borrower greater than $1,000,000 or less that $1,000, age of borrower less than 18, and a loan amount less than $5,000.

estimated using instrumental variables. The predicted LTVs are then used to generate any variables that are affected by LTV.[7]

Specification

The following specification, taken from Hendershott, LaFayette, and Haurin (1997) and Gabriel and Rosenthal (1991), is used to estimate the conditional prime, FHA, subprime choice model:

$$C_j = b_0 + b_1 F_j + b_2 \Theta_j + b_3 D_j + b_4 L_j + e_j \qquad (5.)$$

where F_j is a matrix of financial-monetary variables, Θ_j is a matrix of credit history variables, D_j is a matrix of demographic variables, L_j is a matrix of location specific variables, and e_j is a normally distributed error term. These matrixes are discussed in turn below and Table 2 provides summary statistics for each explanatory variable as well as a brief description and the sources of data.

Financial-Monetary Variables

One consideration for the homebuyer is the relative cost of the mortgage. We focus on the costs to the homebuyer that are derived from differences in mortgage insurance rates and interest rates. For each buyer we construct the present discounted value of interest and mortgage insurance payments for each mortgage option. For mortgage insurance fees we assume payments stop when equity reaches 20% and that mortgage payments are made on time with no house price appreciation. The borrower's credit is graded using the system reported by the Sub-Prime Funding Corp's Underwriting Manual. We rely on credit history variables such as late payment rates on revolving, installment, and mortgage

[7] See Pennington-Cross et al (2000) for details of the estimation technique.

credit as well as indicators of judgments, liens, or bankruptcy. In this fashion we estimate what the best available interest rate would be from a subprime lender. Using estimates of interest rate spreads generated by Wall Street firms (Weicher 1997) and the Mortgage Guaranty Insurance Corporation survey of credit terms and interest rates (Steinbach 1998), rates are increased over prime rates by 200 basis points for B rated borrowers, 300 basis points for C rated borrowers and 500 for D rated borrowers. Because we estimate that over 95% of FHA borrowers financed the upfront mortgage insurance premiums in 1996, we assume this is true for everyone when calculating the cost of an FHA insured mortgage. To measure the relative cost of prime mortgage insurance versus FHA insurance (p_c/p_f) we create the ratios of the present discounted value of the insurance fees. To measure the relative costs of FHA mortgage financing and subprime mortgage financing, we create a ratio of the discounted interest costs for FHA mortgage financing to the discounted interest costs of subprime mortgage financing (p_f/p_s). The specification uses these ratios to test the importance of relative prices in the mortgage choice framework.

A measure of the permanent income (y_j) of the individual is estimated from the cross-section of homebuyers and follows the basic method used by Zorn (1993). A simple model of current income provides parameter estimates for age variables that are used to estimate a stream of income through the age 65. This stream is discounted at the rate of 7% and transformed into an annuity (a coupon bond) that matures when the individual is 65 years old. The annuity provides the estimated value of the individual's permanent income.[8]

The amount of debt (d_j) is created from the credit history data and is defined as the sum of current revolving debt and non-real estate installment loans. It is expected that increases in the non-real estate debt burden will make it more difficult for borrowers to qualify for the lower cost mortgage.

The value constraint (v_j) indicates if the household can purchase the desired amount of housing $\left\{h^*\left\langle Y_j, C, P \middle| Z_j\right\rangle\right\}$ or if the household is constrained by income and/or downpayment constraints. If $h^{\max}$ is the maximum amount of housing that the household can purchase, given the prime lending standards in equations 3 & 4, then when $h^*\left\langle Y_j, C, P \middle| Z_j\right\rangle \succ h^{\max}\left\langle Y_j, C, P \middle| Z_j\right\rangle$ the household is value constrained. We follow, in spirit, Haurin (1991) and Hendershott, LaFayette, and Haurin's (1997) approach. The literature typically refers to these types of constraints as credit constraints, but we rename these value constraints in order to differentiate them from the effects of credit history. It is expected that value constrained households are more likely to choose FHA and subprime financing.

The utility maximizing amount of housing that a household would like to own, in the absence of any mortgage finance constraints, is determined by

[8] Since we do not have data on assets, income is estimated up to retirement age or 65 years old and it is assumed that there is no par or face value payment at term (i.e., no retirement savings). A log-log form is used. Explanatory variables in the model are age(+), marital status(+), a dummy variable for location in an underserved census tract as defined by HUD(-), dummy variables for Black(-), Indian(-), Asian(-), and Hispanic(-) racial backgrounds, and MSA dummy variables(+/-). All the economic variables are significant at the 95% level except the Asian racial dummy. The adjusted R^2 is 0.21.

maximizing the utility function (equation 1) subject to the budget constraint (equation 2) or $\left\{h*\left\langle Y_j, C, P \middle| Z_j \right\rangle\right\}$. This ignores the income and wealth constraints (equations 3 & 4) imposed by lending standards. To determine the unconstrained demand, we estimate a reduced form house price equation over unconstrained homeowners, defined as households who purchase a home with downpayments greater than or equal to 30% of the value of the home, PTIs less than 20%, and FICO scores above 700.[9] Using the estimated non-constrained coefficients, the desired house price is calculated for all remaining homeowners. If the estimated house price is greater than the actual house price, the homeowner is defined as value constrained (v_j=1).

Credit History Variables

A variety of credit measures are tested. The Fair Isaac FICO score (f_j), one of the more common aggregate credit measures available, is used as a summary variable in the analysis.

Using Freddie Mac's Gold Measure Worksheet we create the following more detailed credit history variables: any_j is 1 if the borrower has any delinquencies or derogatory information ever or less than five credit lines have ever been open, otherwise any_j is 0; rev_j is 1 if the borrower does not have a revolving credit line or if total revolving balance is greater than \$500, otherwise rev_j is 0; few_j is 1 if the borrower has less than 3 credit lines open ever, otherwise few_j is 0; del_j is 0,1,2,3, or 4 if the borrower has respectively 0-10, 11-15, 16-40, 41-60, or >60 percent of credit lines ever 30 days delinquent or worse; pub_j is 1 if

[9] See Pennington-Cross (2000) for estimation details.

there are any public record items on the credit report, otherwise pub_j is 0 ; and inq_j is the number of inquiries in the past 6 months divided by two. All of these variables have been designed so that positive values indicate worse credit history and are expected to increase the probability of selecting FHA or subprime financing.

Demographic Characteristic Variables

Demographic characteristics are represented by dummy variables indicating borrower race (Black b_j, Indian i_j, Asian a_j, Hispanic h_j), and marital status (m_j). A spatial segregation version of the gini coefficient (g_j) is also included to measure the extent of racial segregation in each MSA. A 0 indicates complete racial integration of the group, while a 100 indicates complete segregation of the racial group. Racial segregation data is collected from the Census Bureau.

Location Variables

A variety of location variables are used to describe the type of market the loan was made in. Variables used to describe the housing market include: a dummy variable indicating that the purchase is made in an "underserved" census tract (uns_j) as defined by HUD, the one-year percent change in Freddie Mac's reported repeat sales home price index (Dp_j), and the standard deviation of Dp_j for the last 10 years (σDp_j). Variables from Bureau of Labor and Statistics reflect the condition of the local labor market and are the average unemployment rate (u_j) for the last 5 years for the MSA and the change in the unemployment rate in the last year (Du_j). Other variables measuring area housing cost and the FHA loan limit include are a dummy indicating whether HUD defined the MSA as a high cost area (hc_j) and the ratio of FHA's loan limit divided by DRI's estimate of the median house price for the MSA (ll/hp_j).

In general, indicators of increased risk associated with a location may increase the probability that a borrower will use FHA or subprime financing. This should especially be true for indicators that could affect the ability of the borrower to pay the mortgage in the future such as changes in the probability of being unemployed. In addition, the last two variables (hc_j and ll/hp_j) are used to indicate FHA's role in the market. For instance, it is expected that in high cost areas moderate income borrowers may be under more financial stress and therefore may use FHA more frequently. Although this study only includes loans under the FHA loan limit, the fraction of the market FHA defines as eligible for FHA insurance could affect how often borrowers select FHA over subprime or prime mortgage financing. For instance, if FHA covers only a small part of the market lenders may not be able to generate enough FHA business to cover the fixed costs of using the FHA program. In addition, if FHA covers only the bottom part of the housing market, the structures being sold may have a difficult time meeting FHA habitability requirements.

Estimation

Two sets of results are reported. Table 3 provides the estimated coefficients from the multinomial logit estimation and Table 4 provides the ordered logit results. The general specification is as follows:

$$C_j = \boldsymbol{b}_0 + \boldsymbol{b}_1 F_j + \boldsymbol{b}_2 \Theta_j + \boldsymbol{b}_3 D_j + \boldsymbol{b}_4 L_j + \boldsymbol{e}_j \tag{6.}$$

where F_j is a matrix of financial-monetary variables, Θ_j is a matrix of credit history variables, D_j is a matrix of demographic variables, L_j is a matrix of location specific variables, and $\boldsymbol{e}_j$ is a normally distributed error term as discussed above. For each of the estimation techniques (multinomial and ordered) two specifications are reported - one with the FICO score and the other with more detailed credit history. Table 4 shows that ordering is statistically valid as shown by the mu of index, but the multinomial approach has better explanatory power. The log of likelihood is provided as a relative goodness of fit measure and t statistics indicate the significance of each parameter estimate with critical values of approximately 1.95 and 1.65 for the 5% and 10% levels. Tables 5 and 6 provide estimated marginal effects of the explanatory variables calculated at their means. All results discussed refer to the multinomial specification with FICO scores unless otherwise noted.

Financial costs play an important and varied role in the choice of prime, FHA, and subprime mortgage financing. For instance, homebuyers who are value constrained are more likely to use FHA than prime and subprime financing. Borrowers with higher permanent income are more likely to use prime financing while borrowers carrying a lot of non-real estate debt are more likely to use FHA and subprime financing. But for all measures the magnitude of the responses is always substantially higher for FHA and conventional choices. For instance, Figure 1 shows that as the amount of non-real estate debt increases from the mean of $10,842 to $48,000 the probability of selecting prime financing drops from

77% to 53% while FHA's increases from 21% to 45% and subprime decreases from 1.77% to 1.50%.

As the cost of conventional mortgage insurance increases relative to FHA mortgage insurance, borrowers are less likely to use prime financing and more likely to choose FHA financing. This results is consistent for both the multinomial and ordered logit models. But the results are not so consistent for the relative cost of FHA and subprime lending. The order logit estimation finds the expected result that as the interest cost of FHA financing increases relative to subprime, borrowers are more likely to use subprime financing and less likely to use FHA financing. But the multinomial estimates find the opposite results. In addition, when the full array of credit history indicators are included the relative cost of FHA and subprime is no longer statistically significant. This may indicate that some households that use subprime lenders can't respond to prices because they are being constrained by their credit history or other non-price rationing mechanisms.

While Figure 1 shows that the amount of non-real estate debt can more than double the probability of using FHA, the changes in credit scores dwarfs this effect. Figure 2 shows that a decrease in a borrower's FICO score from a mean of 693 to 406, the lowest recorded score, increases the probability of choosing FHA from 21% to 71%. Over the same range the probably of using prime financing decreases from 77% to 26% and increases for subprime from 1.77% to 3.10%. Again, while mortgage choice is sensitive to a borrower's summary credit score, subprime loans never become an important alternative for most borrowers. Figure

2 shows that FHA, while widely recognized as a low downpayment option, is the primary mortgage selection for households with low credit scores.

The detailed credit history variables show that FHA is a more likely choice no matter how the borrower's credit history is tarnished. In contrast, borrowers are more likely to use subprime when a high fraction of outstanding credit lines are delinquent or when there are negative public record items on their credit reports. Borrowers that are more than 30 days late on 60% or more of their loans are more than twice as likely to use FHA and subprime financing as compared to those that are least 30 days delinquent on less than 10% of their loans.

The borrower demographic results indicate that, even after controlling for borrower income, debt, and credit history, racial groups behave differently. For instance, Black, Indians, and Hispanics are more likely to use FHA or subprime financing than Whites. In contrast, Asians are less likely to use FHA but more likely to use subprime financing than Whites.

Location still plays a role in the selection mortgage choice also. In general prime financing is more likely when house prices are increasing and when the unemployment rate is going down in the MSA. In contrast, while the choice of prime and FHA financing is unresponsive to the volatility of house prices ($\sigma\Delta p_j$), the probability of choosing subprime financing increases from 1.77% to 2.9% when the volatility is increased from the mean of 2.3% to the maximum of 5.8%.

In locations considered high cost, the probability of choosing FHA is 6% higher. In addition, in areas where FHA sets the loan limit so that a large portion of the market is eligible for FHA mortgages, the probability of using FHA also increases. This is true in spite of the fact that this study includes only loans that are FHA eligible (loans under the FHA loan limit). These results support the hypothesis that when the FHA market is defined as only the bottom part of the market, it may have difficulty generating enough business for lenders to overcome the fixed costs of learning and staying up with FHA programs and/or that it may be difficult to find homes that meet FHA's habitability requirements in the lowest price portion of the market.

V. Implications for the Subprime Mortgage Market

One goal of this paper, as indicated in the introduction, is to identify how and when households use subprime lending. On this front the results are quite clear – households who exhibit characteristics of high credit risk are more likely to use subprime lending. These indicators include blemishes in credit history and high amounts of non-real estate debt. But there is little evidence to support the idea that subprime lending serves lower income households or households with little wealth to use as a downpayment. This is consistent with published subprime lending requirements which explicitly use downpayments to help compensate for poor credit history or high debt burdens. Beyond household risk indicators, homebuyers in locations with more unstable and deteriorating housing markets are also more likely to use subprime lending. After controlling for credit risk factors, we still find that Black and Asian homebuyers are more likely to use subprime lending. In fact Asians seem to be less likely than Whites to use FHA, thus avoiding government insurance, while more likely to use subprime. One potential explanation for this result is that more Asians are small business owners who report very little taxable and verifiable income. Or as indicated by Johnston et al (1997), some Asian households secure downpayments from relatives or other non-verifiable sources. But, as with the location results, the results of the racial variables can be interpreted in many different and equally compelling ways. This is true because some variables are likely to be missing from the specification and the results represent demand and supply factors simultaneously. But the simple observation that Black homebuyers are more likely to be using more expensive mortgages than Whites with similar credit risk factors including location and credit history is troubling and merits further investigation.

For the subprime mortgage market as a whole, the results confirm that subprime lenders have identified a segment of the high risk mortgage market that can be distinguished from that served by FHA. There is no evidence that subprime lenders serve homebuyers in 'undereserved' (as defined by HUD) areas or low income borrowers with few savings or assets. Instead subprime lenders serve those homebuyers who have not always fulfilled their financial obligations and have taken on lots of non-real estate debt, but are able to provide substantial downpayments to compensate for these shortcomings.

VI. Conclusions

Overall, the results of the mortgage choice model are consistent with initial expectations. Credit history plays an important role in the selection of prime, FHA, or subprime mortgage financing. Other measures of credit risk, such as income, non-real estate debt, and value constraints, are also very important determinants of FHA use, but play smaller role in determining use of subprime financing. In contrast, all choices are sensitive to the relative cost of different types mortgages. Demographic characteristics and location of the borrower play a role in mortgage choice, with minorities in general being more likely to use subprime financing. Also, when FHA sets its loan limits high relative to the median home price, homebuyers are more likely to use FHA financing, but the magnitude of costs, debts and credit history dwarf this effect.

Sensitivity tests show that no one indicator can make subprime a likely choice for any household. For subprime to be a likely choice it requires almost all households risk indicators be negative. It also may be very difficult to identify the characteristics that make subprime lending a viable option to borrowers because of underwriting flexibility that is not captured in this model. For instance, subprime lenders can make loans to people who do not want to document their income or indicate the source of the downpayments. But, our results do indicate that a homebuyer is more likely to use subprime lending when risk indicators such as credit history and location are worse.

References

Gabriel, S. and S. Rosenthal. 1991. Credit Rationing, Race, and the Mortgage Market. *Journal of Urban Economics* 29(3): 371-379.

Haurin, D. R. 1991. Income Variability, Homeownership, and Housing Demand. *Journal of Housing Economics* 1(1): 60-74.

Haurin, D.R., P. Hendershott and S. Wachter. 1996. Wealth Accumulation and Housing Choices of Young Households: An Exploratory Investigation. *Journal of Housing Research* 7(1): 33-57.

Hendershott, P., W. LaFayette and D. Haurin. 1997. Debt Usage and Mortgage Choice: The FHA-Conventional Decision. *Journal of Urban Economics* 41(2): 202-217.

Johnston, S., M. Katimin, and W. Milczarski. 1997. Homeownership Aspirations and Experiences: Immigrant Koreans and Dominicans in Northern Queens, New York City. Cityscape 3(1): 63-90.

Joint Center for Housing Studies of Harvard. 2000. The State of the Nation's Housing. Graduate School of Design, John F. Kennedy School of Government.

Linneman, P. and S. Wachter. 1989. The Impacts of Borrowing Constraints on Homeownership. *AREUEA Journal* 17(4): 389-402.

McFadden, D. 1981. Econometric Models of Probabilistic Choice, in Charles Manski and Daniel McFadden", (eds.), *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press: Cambridge, MA.

Pennington-Cross, A. and J. Nichols. 2000. Credit History and the FHA-Conventional Choice. *Real Estate Economics* 28(2): forthcoming.

Steinbach, G. 1998. Making Risk-Based Pricing Work. *Mortgage Banking* September: 11-20.

Weicher, J.C. 1997. *The Home Equity Lending Industry: Refinancing Mortgages for Borrowers with Impaired Credit*. The Hudson Institute: Indianapolis, IN.

Zorn, P. M. 1993. The Impact of Mortgage Qualification Criteria on Households' Housing Decisions: An Empirical Analysis Using Microeconomic Data. *Journal of Housing Economics* 3(1): 51-75.

Appendix 1: Matching Experian Real Estate Transactions to HMDA Data.

Two key variables – race and income of borrower – were added to the Experian non-FHA home purchase information by finding the corresponding mortgages in the Home Mortgage Disclosure Act (HMDA) database.

The Experian database includes all non-FHA home-purchase mortgages made during the months of February-July 1996. The database covers approximately 400 counties and 1.3 million records. Each record includes the loan amount, census tract of the property, and lender name and identification number.

HMDA and Experian use different sets of lender codes, so the first step in the matching process was to create a crosswalk of HMDA and Experian lender codes. Lender codes (HMDA & Experian) were considered to be equivalent for a pair of lenders when, at least five times in a single county, a single loan in the Experian file for a given lender code and a single loan in the HMDA file for a given lender code had the same loan amount within the same census tract. After this process, Experian loans that had multiple matches with HMDA were visually inspected (sorted by zip code of lender and name of lender) to identify loans with the equivalent lender names.

This crosswalk between HMDA and Experian lender codes was then used to match HMDA and Experian loan records. A Loan was considered matched if it was the only loan that had the same loan amount and the same lender within a census tract. Table A1 summarizes the results of this matching process. It shows the percentage of Experian records matched and the percentages not matched for various reasons.

Note that the match rate reported here is different than the rate reported in the data section because this is the rate for all loans, and not just loans already matched to credit history. The columns are not mutually exclusive, so a loan could be counted in more than one column. The column definitions are as follows:

2. Experian loan record was missing Census Tract.
3. Experian loan record was missing lender code.
4. Nullified Match because there were two or more Experian loans that matched to a single HMDA loan.
5. The Experian loan had a non-zero Census tract that is not found in the corresponding county in the HMDA database.
6. No loans were found in the HMDA file that matched the Experian loan amount in the corresponding Census tract.
7. A unique match of loan amount within Census tract was found, but the Experian and HMDA lenders did not match.
8. Multiple HMDA loans were matched to the loan amount in a Census tract, but none of them matched the Experian lender.
9. The lender that made the Experian loan had two or more loans at this amount in this Census tract, in the HMDA database.

Table A1: Reason for No Match

MSA	match	2	3	4	5	6	7	8	9
Charleston, SC	52.2%	11.1%	6.0%	1.9%	3.3%	12.5%	4.5%	6.3%	6.2%
Charlotte, NC	52.4%	12.7%	3.5%	3.1%	1.5%	8.9%	4.3%	8.2%	5.2%
Chicago, IL	47.7%	8.6%	1.3%	2.5%	2.1%	15.8%	10.0%	10.0%	2.6%
Cincinnati, OH	58.2%	12.1%	3.2%	1.2%	2.0%	10.7%	5.5%	5.4%	2.2%
Cleveland, OH	68.0%	8.0%	2.3%	1.4%	0.1%	7.9%	4.2%	2.7%	5.8%
Columbus, OH	55.8%	14.5%	2.7%	1.8%	0.4%	8.4%	5.3%	6.6%	4.9%
Dayton, OH	61.3%	10.9%	4.4%	1.5%	0.2%	8.1%	5.0%	4.0%	5.3%
Denver, CO	48.2%	10.9%	3.2%	1.9%	8.0%	11.3%	5.8%	7.4%	3.9%
Ft. Lauderdale, FL	34.0%	15.5%	3.9%	1.2%	12.2%	4.4%	5.7%	15.8%	7.9%
Fresno, CA	31.4%	12.8%	8.9%	2.3%	2.8%	31.5%	4.7%	2.4%	4.7%
Jacksonville, FL	48.0%	10.5%	3.1%	2.2%	8.6%	7.6%	3.6%	6.1%	10.9%
Los Angeles, CA	44.6%	6.0%	4.6%	1.5%	1.4%	28.3%	6.2%	2.8%	5.0%
Miami, FL	35.9%	15.0%	5.3%	1.1%	6.6%	8.8%	7.7%	16.7%	3.8%
Middlesex-Somerset, NY	62.8%	15.3%	1.7%	1.3%	0.9%	12.7%	5.1%	4.6%	2.9%
Milwaukee, WI	53.6%	7.6%	2.8%	2.5%	3.5%	15.6%	6.5%	6.1%	2.5%
Nassau-Suffolk, NY	35.1%	6.4%	1.9%	2.2%	17.1%	26.3%	8.4%	4.3%	0.4%
New York, NY	38.5%	6.2%	3.1%	2.6%	9.4%	34.6%	4.9%	2.1%	0.8%
Newark, NJ	61.2%	14.2%	2.2%	1.6%	2.1%	14.2%	5.8%	3.5%	2.1%
Oakland, CA	45.9%	10.2%	3.7%	2.3%	2.0%	19.5%	4.6%	3.4%	8.8%
Orange County, CA	42.0%	13.0%	3.7%	2.2%	1.3%	21.2%	5.9%	4.2%	7.0%
Orlando, FL	36.8%	14.8%	6.9%	1.3%	11.1%	8.9%	6.7%	10.5%	4.6%
Philadelphia, PA	55.9%	11.4%	1.6%	1.9%	1.7%	19.1%	6.7%	4.6%	2.3%
Pittsburgh, PA	58.3%	7.0%	1.3%	2.5%	1.2%	19.0%	7.0%	3.1%	2.8%
Portland, OR	45.4%	11.4%	3.2%	3.4%	3.9%	16.4%	6.6%	5.3%	5.2%
Raliegh-Durham, NC	39.0%	9.2%	4.2%	2.1%	28.6%	5.8%	3.3%	4.2%	4.4%
Riverside, CA	31.0%	15.8%	8.8%	2.9%	7.4%	17.6%	5.6%	4.9%	7.7%
Sacramento, CA	37.2%	12.5%	5.6%	3.0%	5.6%	20.5%	5.0%	3.2%	8.9%
St. Louis, MO	48.4%	7.2%	2.6%	2.2%	3.3%	13.9%	9.9%	10.3%	3.2%
San Diego, CA	40.4%	11.3%	4.9%	2.2%	1.7%	21.4%	6.5%	4.7%	7.4%
San Francisco, CA	51.6%	5.3%	3.9%	2.4%	2.4%	20.6%	4.6%	2.3%	7.3%
San Jose, CA	49.3%	7.0%	3.3%	2.5%	0.9%	19.6%	5.9%	4.0%	7.9%
Sarasota, FL	44.7%	9.7%	5.8%	1.7%	15.1%	6.9%	4.5%	6.7%	5.6%
Seattle, WA	50.7%	9.1%	1.7%	3.0%	1.9%	16.9%	6.0%	4.9%	6.3%
Tacoma, WA	43.4%	10.7%	3.3%	1.3%	5.7%	18.6%	8.4%	5.1%	3.9%
Tampa, FL	39.4%	8.9%	6.0%	2.9%	20.1%	6.0%	3.6%	7.0%	6.6%
Toledo, OH	60.0%	9.4%	2.4%	2.6%	0.2%	9.3%	5.8%	5.3%	5.3%
Ventura, CA	47.2%	12.4%	3.7%	2.6%	1.8%	17.3%	5.0%	3.4%	7.5%
Washington, DC	60.8%	12.0%	2.1%	1.6%	0.9%	11.9%	5.0%	4.0%	2.9%
West Palm Beach, FL	39.7%	15.5%	3.1%	1.0%	11.6%	6.9%	7.3%	11.5%	3.9%
Wilmington, DE	44.6%	12.0%	1.6%	0.8%	0.0%	24.3%	6.7%	7.0%	2.7%

Appendix 2: Calculation of User Cost Measure

The user cost of ownership is defined as follows:

$$UC_j = (1-t_{y,j})(r_j + t_{p,s}) - \pi^e_m + \delta \quad (7.)$$

where t_y is the marginal income tax rate, r is the nominal mortgage rate (FHA rate is available on sample records and national average for the month of origination is used for conventional loans), t_p is the marginal property tax rate, π^e is the expected inflation in housing prices which is assumed to be myopic, δ is the economic depreciation rate which is defined as g*d, g being the structure land ratio which is assumed to be 0.83 and d being the depreciation rate which is assumed to be 0.017 following Linnenman & Wachter (1989), and *s*, *m*, and *j* indicate the variable is geographically defined at the state, MSA, and individual level. The basic simplifying assumptions are zero transactions costs, myopic expectations regarding future inflation rates, interest rates, and tax rates, and treatment of debt and equity as earning equal after-tax risk-adjusted rates of return.

For FHA borrowers, the marginal income tax rate (t_y) is estimated based on the characteristics of each individual. Each borrower is assigned to one of three filing status categories -- married, single, or head of household. All married persons are assumed to file jointly; non-married persons with dependents are assumed to file as head of household; and non-married persons with no dependents are assumed to file as single. The number of dependents is provided in the HUD F42 database. With this information the filing status is determined. Income levels are reduced by the deductions allowed by filing status, number of dependents, mortgage interest payments, and the estimated amount of state taxes paid. State taxes are based on the same information as federal taxes and the tax schedule of the state of residence. Total itemized deductions are defined as the sum of the interest rate deduction and state taxes. The federal taxable income is calculated using the minimum of itemized or standard deductions. In addition, a deduction of $10,000 is applied to all retirees (age greater than or equal to 65) to account for the non-taxable portion of social security benefits. Once the total federal taxable income is defined, the marginal tax rate is calculated using the appropriate schedule for the filing status of the borrower.

To estimate the marginal income tax rate of individuals buying non-FHA homes we use the CPS reported federal tax rate average by income class groups for homeowners as shown in the table below. This approach will introduce some error to the estimate but will be correct on average within each income group.

Property tax rates (t_p) are created at the state level for the last year available (1994) using state and local property tax revenues and estimates of the total valuation of property.

$$t_{p,s} = T_s/(KH_s * PH_s), \qquad (8.)$$

where T_s is the property tax revenues for the state and local governments, KH_s is the number of existing houses, PH_s is the median price of existing homes, and s is the state. Data on tax revenues are collected by DRI and are available from U.S. Department of Commerce, Bureau of the Census, Government Finances. The number of existing homes is collected from DRI and is available from U.S. Department of Commerce, Bureau of the Census. Median house prices were estimated by DRI and are derived from the Federal Housing Finance Board Mortgage Interest Rate Survey and median prices released by the National Association of Realtors.

Table A5: Estimated Marginal Tax Rate by Income Group

Income Group	marginal tax rate
$0 > t_v < 5{,}000$	0.15
$5{,}000 =< t_v < 10{,}000$	2.38
$10{,}000 =< t_v < 15{,}000$	5.27
$15{,}000 =< t_v < 20{,}000$	7.93
$20{,}000 =< t_v < 25{,}000$	10.86
$25{,}000 =< t_v < 30{,}000$	12.15
$30{,}000 =< t_v < 35{,}000$	14.48
$35{,}000 =< t_v < 40{,}000$	15.92
$40{,}000 =< t_v < 45{,}000$	17.19
$45{,}000 =< t_v < 50{,}000$	17.17
$50{,}000 =< t_v < 55{,}000$	18.16
$55{,}000 =< t_v < 60{,}000$	20.09
$60{,}000 =< t_v < 65{,}000$	23.75
$65{,}000 =< t_v < 70{,}000$	26.64
$70{,}000 =< t_v < 75{,}000$	27.57
$75{,}000 =< t_v < 80{,}000$	27.92
$80{,}000 =< t_v < 85{,}000$	28.07
$85{,}000 =< t_v < 90{,}000$	28.15
$90{,}000 =< t_v < 95{,}000$	28.10
$95{,}000 =< t_v < 100{,}000$	28.12
$100{,}000 =< t_v < 125{,}000$	28.65
$125{,}000 =< t_v < 150{,}000$	30.71
$150{,}000 =< t_v < 175{,}000$	31.42
$175{,}000 =< t_v < 200{,}000$	33.71
$200{,}000 =< t_v < 250{,}000$	37.34
$250{,}000 =< t_v$	38.24