Joint Center for Housing Studies

Mortgage Choice:
An Empirical Analysis Using Data From 2002

Michael LaCour-Little

Graduate School
of Design

John F. Kennedy
School of Government

HARVARD UNIVERSITY

Joint Center for Housing Studies

Harvard University

Mortgage Choice:
An Empirical Analysis Using Data From 2002

Michael LaCour-Little

BABC 04-6

February 2004

Abstract

What factors affect mortgage product usage? How important are borrower credit scores in determining mortgage outcomes? Do demographic factors also play a role? What product features best meet the needs of low-to-moderate income and first-time homebuyers? Will the amendments to the Home Mortgage Disclosure Act effective in 2004 provide additional useful information in understanding mortgage choice?

This paper examines home purchase mortgage choice employing data from 2002. Conventional, FHA, subprime, and special programs for low-to-moderate income borrowers are analyzed. Empirical results show that credit characteristics and other financial circumstances drive product choice, with some idiosyncratic factors such as the availability of reduced documentation programs and shorter time to closing helping to explain subprime product use. In contrast, demographic factors have little effect. Given difficulties defining subprime, we caution that data forthcoming as a result of HMDA changes will provide little help in understanding this relatively new market segment.

Introduction

Fostering home ownership in the United States has long been an important public policy objective. The Department of Housing and Urban Development (HUD) regularly reports homeownership rates and policymakers periodically set goals for increasing them. While the overall rate currently stands at 67.4 percent, up from 64 percent in 1989, there is considerable variation across population segments. Seventy-three point four percent of all white households were homeowners as compared to 47.2 percent of African-Americans and 45.4 percent of Hispanics. Geographically, the home ownership rate in suburban areas is 73.8 percent as compared to 50.7 percent in central cities, and the rate among households earning less than the median family income is 50.8 percent (source: www.hud.gov).

While historically based on Jeffersonian ideals of property and citizenship, the economic premise is that homeownership creates positive externalities. Homeowners make better citizens, since ownership promotes social capital formation. Homeowners have incentives to improve neighborhood quality since that quality is capitalized into house prices [Dipasquale and Glaeser (1999)]. Homeownership is thought to be associated with a variety of positive social and economic outcomes. Rohe, McCarthy, and Van Zandt [2000] and McCarthy, Van Zandt, Rohe [2001] review this research. Among social benefits identified, for example, Green and White [1997] report that children of homeowners have better outcomes when compared to children residing in renter households. In contrast, other researchers have focused on the effect of home ownership on labor mobility (see, for example, Green and Hendershott [2001]).

The effect of homeownership on wealth accumulation, particularly for lower-income households, has been another important strand of research. Retsinas and Belsky [2002] address an array of related questions. Focusing on investment returns, Duda and Belsky [2002] examine real gains and losses on homeownership in four cities over the period 1982-1999. In a related analysis, after examining both cross-sectional and temporal variations in housing appreciation rates, Case and Marynchenko [2002] comment "one cannot conclude that homeownership is in general a good or bad strategy for accumulating wealth".

Still another line of research has focused on constraints limiting homeownership. Early work on this topic includes Linneman and Wachter [1989], who reported a reduction in the effect of both income and wealth constraints with the mortgage design innovation that occurred in the early 1980s. In related work, Zorn [1993], examined the effect of constraints using survey data

from 1986 and found them binding on 46% of all households, though effects were much greater on renters than on current homeowners. In more recent work, Rosenthal [2002] reports that eliminating credit-underwriting constraints could potentially increase overall homeownership rates by 4 percentage points. Of course, completely relaxing underwriting standards would impose substantial costs on both lenders and those borrowers who experience foreclosure. Indeed, Deng, Quigley, and Van Order [1996] estimate that if expected losses in such programs were not priced into note rates, the losses from default alone could exceed ten percent of the funds made available for loans.

During the 1990s, the mortgage industry responded to increased demand and policymakers incentives with an expanded set of product offerings. These may be generally classified into two main types: (1) subprime mortgage products; and (2) LMI-targeted mortgage products. Another contract design innovation during the 1990s was the hybrid adjustable rate mortgage (ARM), under which the note rate is initially fixed (generally for 2, 3, 5, or 7 years) and then converts to a one-year adjustable instrument. This structure allowed borrowers to target expected holding periods relatively more precisely and obtain pricing off the short end of the yield curve. This mortgage type is generally held in portfolios by financial intermediaries rather than sold in the secondary market.

Loan purpose has traditionally been an important element in mortgage design. FHA-insured loans were originally designed to allow households with relatively small down payments to purchase moderately priced housing. In contrast, subprime mortgages were originally designed as refinancing loans for the purpose of debt consolidation. Survey data used by Surette, Courchane, and Zorn [2003] to compare subprime to prime borrowers, shows that 57 percent of subprime loans were taken out for the purpose of debt consolidation or cash out refinancing. In contrast, only 16 percent of prime loans were for that purpose. This mix will likely vary across lenders and over time, as well. But as a general characterization, subprime refinance loans re-capitalize borrowers who may have current, or past, financial problems. In contrast, subprime home purchase loans allow borrowers who may have encountered financial difficulties in the past (such as a bankruptcy or foreclosure) or who are currently in a relatively more precarious financial position, often due to high leverage, to purchase a home. Subprime mortgages are now widely available and do not involve income or geographic limitations. In contrast, specially

targeted programs are often means tested (income cannot exceed a certain level) and/or are available only in designated areas.

The development of the subprime segment has been controversial, due in part to the difficulty in defining it, and concern at both the state and federal level with abusive practices allegedly occurring in it. For contrasting regulatory views on the topic, see HUD [2000] or OCC [2003]. Cutts and Van Order [2003] develop an economic theory of market segmentation in the presence of a secondary market that can explain some of the major facts about the subprime segment. But they do not attempt any general definition of subprime, except to note that it is "typically the riskiest and most difficult part of the market to evaluate". A variety of more specific definitions have been proposed.

Bank regulatory agencies generally identify a subprime borrower as having one or more of the following characteristics: two or more 30 day delinquencies in the last 12 months, one or more 60 day delinquencies in the last 24 months, a judgment, foreclosure, repossession or charge-off in the last 24 months, a bankruptcy in the last 5 years, a credit score of 660 or lower, or a debt-to-income ratio of 50% or greater (OCC, [2001]). Another approach is to identify subprime loans based on their note rates. Surette, Courchane, and Zorn [2003] and Pennington-Cross [2003] classify loans as subprime if their note rates are above the ninetieth percentile among all rates in the month of origination. HUD takes still another approach, identifying loans as subprime if originated by lenders on its "Manufactured Home and Subprime Lender List". The list is "periodically updated and revised in response to feedback from lenders, policy analysts, housing advocacy groups, and other users of the lists" according to HUD's web site (www.hud.gov). According to the HUD list, there were 178 predominately subprime mortgage lenders, as of 2001, down from 252 in 1999. Obviously, to the extent subprime loans are originated by non-specialists, the list approach misses part of the market.

Due in part to policymakers concerns about subprime lending, the Home Mortgage Disclosure Act was amended effective January 1, 2004, to require additional reporting, including certain pricing information. For loans that close after January 1, 2004, lenders must report the spread between the annual percentage rate (APR) on the loan and the yield on comparable Treasury securities, if that spread exceeds a certain threshold (three percent for first-lien and five percent for subordinate-lien loans). Other amendments to HMDA require reporting of lien position, whether the loan is secured by a manufactured home, and whether the loan is a "high

cost" loan, as defined under the federal Home Owner's Equity Protection Act (HOEPA). HOEPA forms the model for most of the high-cost lending statutes enacted at the state and local level in recent years. Under HOEPA, as most recently amended, a loan is "high cost" if its APR spread over the comparable Treasury rate exceeds 8 percent or if points and fees exceed 5 percent of the original loan amount and certain additional rules apply. It is widely anticipated that a reportable APR spread will emerge as the new *de facto* standard for identifying non-prime lending in 2005 and beyond.

Unfortunately, none of the definitions described is entirely satisfactory. To begin with, definitions based on note rate or APR alone ignore the economics of the loan origination process, which involves both fixed and variable costs. Moreover, especially at smaller loan sizes and for borrowers with limited funds available for closing costs and down payments, it is common for loan originators to fund closing costs and recoup those costs in a higher note rate. As a result smaller loans tend to bear higher note rates.

An additional problem is that any definition based solely on credit profile misses the "low doc" segment of the subprime market in which lenders offer products requiring little or no documentation of employment, assets, or income. We will see in the empirical section later in this paper that a significant fraction of subprime loans are "low-doc". Such products are particularly popular among the self-employed and others for whom recurrent income may be relatively less stable or hard to establish. Underwriting requirements typically require better credit profiles and larger levels of equity for such products, reducing credit risk. Another borrower benefit in this segment is more rapid underwriting and shorter time to closing. We will see as we turn to the empirical section that subprime loans close more quickly than prime loans.

Still another category of subprime lending is non-standard transactions that are not accepted, or competitively priced, under secondary market guidelines, such as high LTV jumbos, high LTV purchases of non-standard properties, high LTV cash-out refinancing, especially on non-owner-occupied properties, and high debt-to-income cases to borrowers with otherwise good credit. Lenders typically trade off better credit profiles for higher LTVs in such cases or, conversely, lower LTVs for relatively worse credit. As a result, it is probably fair to say that there is a relatively greater amount of risk-based pricing in the subprime market compared to the prime market. In the prime market all borrowers meeting minimum underwriting standards are eligible, in principle at least, for par rate pricing and prices do not vary as much with credit risk.

Specialized mortgage programs targeted at the low-to-moderate income (LMI) segment comprise still another product array, some lender-specific and others investor-specific. Some of these are conventional programs, such as those sponsored by the government sponsored enterprises (GSEs); for example, Fannie Mae's Community Home Buyer's Program and Freddie Mac's Affordable Gold Alt 97. These are generally subject to higher credit risk than traditional conventional mortgage lending yet seldom described as subprime. In general, they offer reduced down payment requirements and greater underwriting flexibility, compared to traditional conventional loan products, but are rationed in some form, either by means-testing, by maximum loan amount limits, or by restricting availability to specific geographies. WFHM's branded program is the Emerging Markets National Program (EMNP) and was developed in partnership with Freddie Mac. We compare and contrast underwriting standards for this program with those of traditional FHA lending later in the paper.

We turn now to the academic research on mortgage choice, especially the choice of contract type. After an introductory review, we focus on prior studies of the FHA-conventional and prime-subprime choice.

Literature Review

The literature on mortgage choice is extensive, so the review here is necessarily limited. Follain [1990] defines the topic broadly to include an array of problems faced by households, including the choice of how much to borrow (the LTV decision), if and when to refinance or default (the termination decision), and the choice of mortgage instrument itself (the contract decision). While our empirical focus here is on contract choice, we will see as we turn to the econometric issues that many of these choices are inter-related.

The demand for mortgage debt is derived from housing demand. From the economic perspective, the primary determinant of housing demand is user cost, which may be defined as the cost of using one unit of housing capital for one period. User cost is the sum of the opportunity cost of capital invested in housing, mortgage financing costs, and depreciation in the housing unit, less tax benefits attributable to the tax deductibility of property taxes and mortgage interest and appreciation in capital value. If the user cost of owning is less than the user cost of renting, households will have an economic incentive to become homeowners.

Early theoretical work on mortgage demand, such as Jones [1993] and Brueckner [1994], focused on the opportunity cost of capital invested in housing. Absent frictions, if the return on other investments exceeds the return on housing, all households demand 100 percent loan-to-value ratios. Conversely, if the return on housing exceeds that of other investments, then the optimal strategy is 100 percent equity, i.e. no mortgage debt usage. Where investment returns are uncertain, however, the optimal mortgage size is ambiguous. In a world where mortgage debt is readily available and mortgage interest is tax deductible, an important comparison is between the after-tax cost of mortgage debt and the expected after-tax rate of return on housing. Hence, tax rates and tax policy have a significant impact on the demand for housing and, hence, the demand for mortgage debt. Empirical work on the demand for mortgage debt, as a function of household characteristics, has been much more limited. Ling and McGill (L&M) [1998] address the question using American Housing Survey data from the 1980s, modeling mortgage and housing demand as jointly determined. Consistent with the theoretical findings, L&M show that the tax savings associated with the interest deduction affect the level of mortgage debt usage, noting that lower income households who do not itemize lose the associated tax benefit of the mortgage interest deduction. Further discussion of taxes is outside of the scope of our work here but we refer the reader to Follain and Dunksy [1997] for a thorough treatment.

Related work on contract choice has focused on the choice of rate versus points (Brueckner [1995], Stanton and Wallace [1998]), and the choice of ARM versus FRM (Dhillon, Shilling, and Sirmans (1987), Brueckner and Follain [1988]). The consensus in these areas seem to be that points signal reduced borrower mobility, i.e. longer expected housing tenure, and that the ARM-FRM choice is both a function of expected mobility and the slope of the yield curve, which produces the savings between short and long rates. Consequently, in flat yield curve environments ARM demand falls off. It is also generally accepted that a preference for ARMs is related to higher cost housing markets since low initial rates stretch borrower budgets farther under payment ratio tests. ARMs tend to be held in portfolio by financial institutions whereas FRM tend to be securitized (see Ambrose and LaCour-Little [2001) on ARM performance patterns]. We turn now to a more detailed review of prior research studies more directly related to our topic here, namely, FHA versus conventional choice and prime versus subprime choice.

Gabriel and Rosenthal (G&R) [1991] provide an interesting starting point, since they used data from more than 20 years ago, a very different time period characterized by high rates

of inflation and interest rates. G&R used micro data from the 1983 Survey of Consumer Finances (SCF) to study borrower choice between FHA and conventional loans among borrowers who obtained a newly originated loan between 1978 and 1983. To make a government-insured loan a meaningful choice, loan size was restricted to FHA limits (then $67,500). G&R did not have access to any credit related variables but attempted to proxy for default risk. Among proxies they include expected housing appreciation rates, a proxy for borrower equity based on reported household wealth, and average unemployment rates in the borrower's occupation. They also include demographic variables, such as borrower age, whether the property is located in a central city, and whether the borrower is non-white. The probability model is estimated with a probit regression on a 340 household sample. Their results suggest that minorities are more likely to choose FHA loans after some weak controls for default risk. That conclusion, however, may have simply reflected differential underwriting across conventional and FHA products and limited funds for down payments among minority borrowers, or other omitted credit risk variables.

Hendershott, LaFayette, and Haurin (HL&H) [1997] examine both the FHA-conventional choice and the FRM-ARM in a nested logit framework. They use data from the 1984 American Housing Survey to study the choices of 819 young homeowners. At this time, ARMs were not a viable option for FHA borrowers. They find that mortgage choice is driven by down payment and monthly payment constraints and the desire to reduce mortgage insurance costs. HL&H do not include credit or demographic factors in their empirical analysis.

Turning to more recent work, Pennington-Cross and Nichols (PC&N) [2000] also address the FHA-conventional choice using a much larger and more complete data set, including credit information. PC&N combine information from four distinct sources to construct a data set of loans originated during 1996, including 21,246 FHA and 26,246 conventional loans. PC&N note that the two segments clearly have different credit score distributions: the mean FICO score for conventional borrowers is 717 while the mean for FHA borrowers is 665, with some convergence at higher LTV levels. PC&N then estimate four specifications of a logisitic regression model of mortgage choice controlling for a variety of financial, credit, and demographic variables, including MSA fixed effects. Their most complete specification includes FICO score (in continuous form), a limited set of credit bureau attributes, estimated LTV generated through an instrumental variables approach, an estimate of borrower permanent

income, and an estimate of the differential cost of mortgage insurance under the two programs for high LTV categories.   PC&N find that credit history plays an important role in product selection; for example, an increase of only 10 points in FICO score decreases the probability of FHA choice by 2.8%. In contrast to Gabriel and Rosenthal, who reported greater use of FHA by a combined category of non-white minority groups without controlling for credit history, PC&N find no effect for African-American borrowers but a significant preference for FHA among Hispanic borrowers.

Turning to the choice of prime versus subprime, the literature is much more limited, partly because the subprime market is relatively new. Surette, Courchane, and Zorn (SC&Z) [2003] use survey data to examine differences in both circumstances and experience for subprime compared to prime borrowers, using the rate-based definition of subprime previously described together with the HUD lender name approach, producing a sample used in regressions of 4,571 loans. Like PC&N, they find that credit score classifies borrowers reasonably well, e.g. 87 percent of prime borrowers have FICO scores greater than 620 whereas only 38 percent of borrowers in the subprime group are in this score range. SC&Z claim that adding demographics to risk variables in a logistic regression increases model explanatory power (as measured by an increase in the K-S statistic of 1.76 points). Beyond demographic factors, SC&Z show that adverse life events, such as marital dissolution, a major medical expense, unemployment, or change in income, all contribute to subprime product usage. In an interesting extension, SC&Z also analyze outcomes and transitions among product types, noting that approximately 40 percent of subprime borrower transition to prime via refinancing. Based on this finding it may be helpful to think of subprime product usage as episodic, more like a spell of unemployment, than a fixed borrower characteristic. This perspective is reinforced by the view, prevalent on Wall Street and elsewhere in the mortgage market, that the curing of temporary credit problems produces very rapid prepayment speeds on subprime mortgages.

Pennington-Cross, Yezer, and Nichols (PCY&N) [2000] address much the same question as that considered here: the choice between FHA, conventional, and subprime, using the same data as is used in PC&N [2000] to study the FHA-conventional choice. But they use the HUD list to identify subprime loans, probably leading to significant under-estimation, since only 612 out of 48,105 (1.3 percent) are classified as subprime using this criterion. In their data, the mean FICO score of prime borrowers is 717, whereas FHA and subprime borrowers have means of

665 and 669, respectively, while FHA borrowers have much higher LTVs than do subprime borrowers. PCY&N conclude that credit risk factors and the relative cost of mortgage insurance are the predominate factors in sorting borrowers into conventional, FHA, and subprime categories, although there are some unexplained demographic residual effects as well. PCY&N test both an ordered and multinomial logit and find that the multinomial has better explanatory power. As in PC&N, LTV is treated as endogenous and estimated via instrumental variables.

In summary, research on FHA-conventional choice to date has focused on the effects of income and wealth constraints, loan size limits, differences in underwriting criteria, and the relative cost difference of FHA versus non-government mortgage insurance. There are mixed results on the question of whether demographic characteristics are important and the finding in some research that minority borrowers are more likely to choose FHA is consistent with a relatively weaker financial picture and limited funds for down payments given FHA's relatively more liberal underwriting standards. Results from research on subprime lending has required category definition and emphasized borrower sorting based on credit characteristics again with mixed evidence on the effect of demographics. It is clear, however, that many subprime borrowers, unlike FHA borrowers, have sufficient funds to make relatively large down payments. But unlike prime borrowers, they have weaker credit histories and often carry higher current debt burdens. In all of these analyses, the "low doc" (sometimes called "Alt-A") phenomenon has remained unexplored.

Econometric Issues

A variety of difficult econometric questions arise in mortgage choice model specification. Fundamentally, we would like to estimate a general demand function, but typically observe only realized outcomes, conditional on the borrower's decision to apply to the particular lender and the lender's subsequent decision to approve. Moreover, the lender's decision to approve is conditioned on assessment of default risk and the borrower's assessment of default risk may affect choice of loan terms. And since borrowers choose many of the loan terms, in particular, the amount to borrow (equivalently, the loan-to-value ratio), this variable cannot be treated as exogenous to contract choice. A similar argument can be made about reduced documentation. Furthermore, available research on other dimensions of the mortgage choice question (and, of course, common sense) indicates that prospective borrowers choose based on price, so lower

rates are preferred to higher rates *ceterus paribus*. But given the prepayment option embedded in most mortgage contracts and the usual upward sloping yield curve, borrowers with relatively shorter expected housing tenure may prefer adjustable rate products over fixed rate products, so what constitutes a lower rate may vary by household. Moreover, for home purchase loans, the borrower is typically subject to purchase contract time constraints and must obtain a financing commitment expeditiously or risk losing the option to purchase the house. Consequently, borrowers would be expected to consider both the probability of approval and time required to close when choosing loan terms.

To address the LTV issue, we use an instrumental variable approach, following Pennington-Cross [2000]. In particular, we follow Ambrose, LaCour-Little, and Sanders [2002] and estimate loan-to-value ratio jointly with house price then use categorical transformations of the resulting fitted value as a right-hand side variable in the mortgage choice models. To address the reduced documentation issue, we follow an analogous process, estimating the probability of a reduced documentation level and employing that fitted value as a right-hand side variable in the mortgage choice model. The full set of equations estimated is provided in a technical appendix that follows the tables and precedes the graphs.

Data

Our data comes from the 2002 loan origination records of Wells Fargo Home Mortgage (WFHM), the nation's leading home mortgage lender. WFHM originated over $300 billion in mortgage loans during calendar year 2002 and is the industry leader in FHA lending, the leader in lending to ethnic minority groups, and to low-and-moderate income segments. WFHM has a relatively small, though growing, market share in the subprime segment, which anecdotally includes an above industry average percentage of home purchase loans.

Since our focus here is on home ownership, we restrict our data to loans originated for the purpose of home purchase, and exclude closed loans acquired from other lenders ("correspondent loans" in industry parlance) and those originated by mortgage brokers. We further restrict the population from which we sample to owner-occupied housing, both single-family and single housing units within condominium properties, and exclude manufactured housing, cooperative apartments, second homes, and 2-4 family dwelling units. We also exclude some specialized programs, such as VA loans and renovation loans, since they employ unique

underwriting standards. Some of these exclusions may be material for the LMI segment of the market; for example, center cities often contain a disproportionate number of 2-4 family dwelling units and in rural areas of the south and west, manufactured housing is a popular choice for many lower income households.   Since we wish to make FHA loans a meaningful choice in our analysis, we follow the earlier research and limit loan sizes to the FHA loan limits as they vary by MSA. This restriction produces some under-sampling of high-cost housing markets.

Subject to the usual problem of occasional missing values, we have near-complete micro-level data for each loan in the sample. In addition to product choice, location, loan amount, note rate, and transaction dates, we observe all major credit quality indicators, including loan-to-value ratio (*LTV*), debt-to-income ratio (*DTI*), borrower credit score at time of origination (*FICO*), and credit bureau attributes. Borrower demographic characteristics available include income, age, gender, and race, some of which are reported under provisions of the Home Mortgage Disclosure Act (HMDA). Based on the tract reported, we append additional 2000 census data, including relative income (*PCTMED*) and minority composition (*PCTMIN*). We follow the regulatory approach and define a loan as LMI if made to a household with income less than 80 percent of area median, regardless of location, or to a household buying in an area where median income is less than 80 percent of the area median, regardless of household income.   Table 1 reports descriptive statistics for the total starting population of 173,514 loans, then segmented into conventional versus FHA-insured (the two objective criteria publicly available through HMDA). The data consists of 137,040 conventional loans (79 percent) and 36,474 FHA-insured loans (21 percent).   Recall that we are restricting the population to loans within the FHA loan limits; hence, the share of loans that are FHA is somewhat higher than would be the case in the entire national mortgage market, since there are few FHA loans in high-cost housing markets.

Some differences between conventional and FHA-insured loans and borrowers are immediately evident from Table 1. To begin with, conventional loans are larger (mean size $183,000) compared to FHA (mean size $117,000). Differences in house values are even greater (mean size of $242,000 for conventional versus $120,000 for FHA). This difference is reflected in the average loan-to-value ratio (80.0 percent for conventional and 97.6 percent for FHA). The original LTV (while not used in the regressions) is labeled *ORIGLTV* and reported in the descriptive tables. Average note rates, however, are almost identical (6.63 percent conventional versus 6.61 percent for FHA).   Table 1 also shows that FHA borrowers are much younger

(average 34 versus 40, for conventional), have fewer years of employment history (5 years versus 7 years, for conventional), and weaker credit (average FICO 670 versus 724, for conventional). These patterns are very similar to those shown in prior studies.

As previously discussed, definition of subprime loans is inherently difficult. To bring the magnitude of this problem into sharper focus, we further segment the conventional loan category in Table 2, limiting the population to those loans for which we have APR available. The first panel uses the regulatory definition that focuses on credit factors, the second uses relative note rate (rates higher than the ninetieth percentile of all loans in the origination month), the third uses the new HMDA reportable APR spread definition, and the fourth uses WFHM internal management reporting systems (the category we adopt for use in further analysis). In the fifth panel, we apply the regulatory definition to the FHA-insured category and observe that about half of FHA loans would qualify as subprime. Combining the FHA and conventional categories and selecting either the most restrictive or least restrictive definition produces a range of estimates for the subprime category. Some simple calculations indicate that depending on definition used, the data contain somewhere between 8,101 and 45,711 "subprime" loans. This lack of clarity over the category creates major challenges to conducting meaningful research and crafting effective regulation. Perhaps it is time for a new lexicon discarding the term "subprime" entirely.

FHA-insured loans have been widely studied and are well understood, but the WFHM special targeted product (EMNP) deserves further description. The product is both similar to, and distinct, from FHA. As with other targeted programs, EMNP does limit eligibility, specifically by restricting borrower income to a maximum of 120 percent of HUD median family income for the area, but this requirement can be waived for properties in center cities. In terms of underwriting, EMNP offers higher loan limits (the GSE conforming loan limits) compared to the FHA loan limits, higher maximum loan-to-value ratios (100% for borrowers with a minimum FICO score of 620 and 97% for borrowers with FICO less than 620), and higher allowable payment ratios. In contrast, FHA does not have a stated FICO standard. The maximum allowable front-end back-end ratios under FHA are currently 29/41 while the EMNP program allows 45/45, if income is documented, and 40/40 if stated income is used. Up to 25 percent of income can be stated income under the EMNP. Neither EMNP nor FHA requires reserves (verifiable liquid assets over and above closing costs including funds for down payment).

EMNP offers both fixed-rate and hybrid ARM structures; in contrast, FHA currently offers only the one-year ARM design. EMNP offers both borrower-paid and lender-paid mortgage insurance options; under the latter, the note rate is increased to cover costs. We note in passing that subprime typically employs lender-paid mortgage insurance, which automatically results in a slightly higher note rate relative to conventional, even for borrowers with comparable risk profiles.

We begin with a full population of 173,514 loans, after exclusions previously noted. We then draw random samples from the population to create approximately equally sized choice-based sub-samples, generally about 500 loans per choice. This sampling procedure introduces bias in the intercept term only in logistic models and is readily corrected. Maddala [1983] provides the correction formula. Since our data contains a number of interesting sub-populations, we present separate analyses of these groups by repeatedly taking random draws. Among these are loans to first-time homebuyers and loans to low-and-moderate income households. When analyzing these sub-populations, we first select the sub-population then randomly re-draw equal size choice-based samples, so the different sub-groups may, but do not necessarily, overlap. Each of these regressions is preceded by descriptive statistics on the sample used. Although not used in the regressions, we include mean values of the various subprime definitions, to further illustrate the point made in Table 2.

Regression Results

In the interest of brevity, we do not report the two intermediate regressions: prediction of loan-to-value ratio (jointly estimated with house price) and the probability of a reduced documentation loan. We note in passing, however, that the later is highly related to self-employment status (with a bivariate correlation greater than 0.5). LTV estimation follows the existing literature closely.

We estimated two specifications of a multinomial logit (MNL) model of mortgage choice, defined as EMNP, FHA, and subprime, with all choice probabilities relative to conventional. These are estimated on samples from the two sub-groups of interest, LMI households and first-time homebuyers. For convenience in interpreting coefficients and to capture non-linearity in relationships, most variables are re-coded into categorical format, so that

14

reported odds-ratios are directly comparable[1].   Each MNL contains two specifications, one with and one without demographic factors and is preceded by a table of descriptive statistics. Hence, Table 3 shows descriptive statistics for the LMI sample, Tables 4-5 show results with and without demographics for the LMI sample, Table 6 shows descriptive statistics for the first-time homebuyer sample, and Tables 7-8 show MNL results with and without demographics loans for first-time homebuyers.

Tables 3-5 focuses on LMI households, defined by either borrower or tract-relative income. Table 3 presents descriptive statistics while Tables 4-5 present results of the MNL model with and without demographic factors. From Table 3, we observe that house prices are lower and LTVs are higher than the overall population (shown in Table 1). For LMI households, average annual incomes are uniformly below $52,000, regardless of product choice. Average credit scores differ significantly; however, for example, mean values are 729, 662, 670, and 621, for conventional, FHA, EMNP, and subprime, respectively. Moreover, low credit scores are concentrated in the subprime category, with 55% of subprime borrowers below 620, whereas only 5% of conventional borrowers had scores in that category. But subprime borrowers do not tend to have high LTVs; in contrast, both FHA and EMNP borrowers have average LTVs in excess of 95%.

Turning to the MNL results, Table 4 presents model results including demographic factors while Table 5 excludes them. Among LMI households, first-time homebuyers show a strong preference for FHA and EMNP over both conventional and subprime products; here, the coefficient on EMNP is actually larger than for FHA, suggesting that this new product innovation has been very successful in reaching its target customer. But it is credit risk factors that appear to effectively separate borrowers into various product choices and borrower age is the only demographic factor showing strong effects (positive for younger borrowers selection of FHA and EMNP products and uniformly negative for older borrowers selection of all products, relative to conventional). Contrary to popular claims, location in a low- or moderate-income area (LOW_INC and MOD_INC) is strongly negatively related to subprime product use and there is no statistically significant association with FHA or EMNP use, either. This is consistent with the viewpoint that it is borrower, not area, characteristics that affect mortgage product

---

[1] We did not do this in a few cases due to sparse cell problems that produced quasi-complete point separation problems in estimation.

choice. Having a credit score of less than 620 is a very strong predictor of subprime use and we observe monotonically declining coefficient magnitudes as credit scores increase. In all of these specifications, credit score of 700 or higher is the reference category.

A closing time within 30 days of application (*CLOSE_30*) also proved to be an important variable in predicting subprime product usage. We speculate that this may be due to borrowers having initially been turned down for prime products and then facing time constraints given home purchase contract provisions. In contrast, borrowers with shorter closing times avoid FHA and EMNP products, perhaps because of relatively greater documentation requirements, compared to conventional products. Likely related to this is the highly significant variable on documentation level (*PROB_FULLALT*[2]), which seems to effectively separate subprime from FHA and EMNP alternatives.

In Table 5, all demographic variables are excluded from the specification. Coefficients are very stable; for example, the coefficient on credit score less than 620 (FICO_620) changes from .89 to .88 for EMNP, from 1.72 to 1.75 for FHA, and from 2.66 to 3.17 for subprime. As another example, debt to income for FHA is 0.0272 in Table 4 and 0.0266 in Table 5, a minimal change in magnitude.

The second analysis, shown in Tables 6-8, focuses on first-time homebuyers. Table 6 provides descriptive statistics while Tables 7-8 present MNL model results with and without inclusion of demographic factors. As with LMI households, first time homebuyers are purchasing relatively less expensive houses ($157,000, $117,000, $114,000, and $125,000 average values for conventional, FHA, EMNP, and subprime, respectively). Credit scores show a predictable pattern (mean values of 726, 665, 684, and 632 for users of conventional, FHA, EMNP, and subprime products, respectively). In contrast, average LTV ratios are in the low eighties for conventional and subprime but in the high nineties for FHA and EMNP.

Table 7 shows MNL model results, including demographic factors. We again observe a generally monotonically declining effect of credit score on product choice as credit score increases. For example, the coefficient on subprime is largest (among the three alternative choices relative to conventional) for borrowers with credit scores below 640; between credit scores of 640 and 680, FHA is the largest. Subprime is larger again in the 680-700 category,

---

[2] The variable is coded so that a higher value indicates a higher probability of a full doc loan, so a negative sign is consistent with greater use of low doc and no doc loans in subprime.

consistent with reduced documentation program use for relatively higher credit score borrowers. The coefficient on EMNP approaches that of FHA in the score range of 640-680, indicating that this conventional product competes particularly well with FHA in this segment. As noted in other samples, the coefficient on short time to close is large and positive for subprime. We speculate that this could occur because borrowers turn to subprime after having failed to qualify for prime products and need to close quickly to satisfy home purchase contract terms.

As with the LMI sample, demographic factors are generally not statistically significant, except for borrower age (younger borrowers show preferences for FHA and EMNP products and avoid subprime). African-American borrowers do, however, show preferences for EMNP and FHA products, relative to conventional. Likewise, borrowers residing in above-average minority neighborhoods (*PCTMIN30_PLUS*) show a preference for the EMNP product, whereas neither FHA nor subprime is statistically significant. This result again underscores the success of the EMNP product in reaching traditionally under-served market segments. Area income levels are not statistically significant in this model specification. The short time to closing variable, here coded into even finer categorical variables, continues to be highly significant in predicting subprime use and negatively related to FHA and EMNP choice. The strong effect of documentation level (*PROB_FULLALT*) described previously persists here, too, indicating that borrowers who cannot, or do not wish, to completely document income and assets as required by conventional and government programs turn to subprime product options.

Table 8 reports results for first-time homebuyers excluding demographics. Results are again very similar both in sign and magnitude. The pattern of declining credit score coefficients as credit scores increase is repeated. Using debt to income as the comparison variable, its coefficient in Table 7 for FHA is 0.0428 and 0.0417 in Table 8. Comparable values for subprime are 0.0626 from Table 7 and 0.0617 in Table 8.

Two graphs follow the tables. Each of these shows the average conditional probabilities of product choice as a function of credit score category. Graph 1 is for LMI borrowers while Graph 2 is for first-time homebuyers. The pattern shown in the graphs is remarkably consistent across the two distinct samples, further reinforcing the conclusion that credit profile effectively sorts borrowers into product categories, and that LMI borrowers have similar product preferences and usage patterns and do not experience differential treatment. Results show that the probability of obtaining a conventional loan for a low credit score borrower is very low, less

than 10%; in contrast, for a high credit score borrower, the probability is 60-80%. One may also note that FHA products reach lower credit score borrowers relatively effectively, suggesting that recent policy proposals to expand FHA insurance in this segment might well be effective.

Conclusions

In this paper, we examined mortgage product choice using data on retail home purchase loans originated during 2002. In addition to demonstrating the difficulty of defining subprime lending, we found that credit scores are highly predictive in sorting borrowers into the four types of products examined. We also found evidence of several idiosyncratic factors that help explain subprime product use; in particular, high levels of debt, an apparent desire to close the transaction quickly, and the availability of reduced documentation programs that are associated with self-employment.

After controlling for credit characteristics and other factors, borrower and area demographics appear to play little role in product choice. For example, African-American or Hispanic borrower race was not statistically significant in subprime choice in either of the multinomial logit regression results reported. Moreover, the increase in model explanatory power (as measured by percentage reduction in the AIC statistic [not reported]) attributable to inclusion of all demographic factors is generally in the 2-3 percent range. Among demographic factors that do matter, borrower age appears to be the most important, with younger households showing a marked preference for FHA and EMNP products, no doubt due at least in part to their low down payment requirements. In contrast, subprime borrowers tend to be middle-income and middle-aged, much like borrowers choosing conventional loan products, but with much weaker credit histories and higher levels of current debt.

For first-time homebuyers and LMI households, FHA and EMNP appear to be the preferred choices, no doubt due largely to the low down payment requirements. This suggests carefully designed conforming products can compete effectively and help increase home ownership rates among traditionally underserved segments of the population. Subprime loans provide additional flexibility mainly for households who have sufficient liquidity for a relatively large down payment, but who often have high levels of current debt, a weaker credit history, or who wish a loan with reduced documentation requirements, often because they are self-employed. The quick time to close characteristic of subprime loans suggests that this category

may also play a valuable role when borrowers discover late in the home purchase process that they cannot qualify for a prime product. In general, all of these product offerings make credit more widely available than does strict reliance on the traditional conventional product, which is characterized by excellent credit, limited leverage, a large down payment, and ample time to document and close the loan.

A number of caveats apply to the evidence on mortgage product usage presented here. Results are, of course, conditional on both data and methodology. Our data employs home purchase loans only and the selection process has deliberately excluded a number of property types and loan programs that may be important in meeting the housing needs of lower income households. Moreover, we have under-sampled high cost housing markets, in order to make FHA choice a viable financing option in all cases. Lower income households face particularly difficult circumstances in those markets. We have also deliberately excluded loans originated by mortgage brokers, who today originate roughly half of all loans. In addition, the data is from a single lender, so there may be other market participants for which the patterns described do not apply. Only with continued research, preferably with multi-lender databases, can all the determinants of mortgage product usage be more completely understood.

Unfortunately, and perhaps despite expectations to the contrary, changes to HMDA effective in 2004 seem unlikely to provide much useful new information. This is because, as has been shown, the correlation between APR spread and subprime (however defined) is imperfect and none of the risk factors that so effectively sort borrowers into different categories will be reported. Given the lack of precision in defining subprime, it may be time for a new vocabulary that better reflects the complexity of alternative mortgage products rather than the simplistic categories that have been employed to date.

References

Ambrose, Brent and Michael LaCour-Little. 2001. Prepayment Performance of Adjustable Rate Mortgages Subject to Initial Year Discounts: Some New Evidence. Real Estate Economics 29(2): 305-328, 2001.

Brueckner, Jan K. 1995. Borrower Mobility, Adverse Selection, and Mortgage Points. Journal of Financial Intermediation 3: 416-441.

Brueckner, Jan K. 1994. The Demand for Mortgage Debt: Some Basic Results. Journal of Housing Economics 3: 251-262.

Brueckner, Jan K. and James R. Follain. 1988. The Rise and Fall of the ARM: An Econometric Analysis of Mortgage Choice. Review of Economics and Statistics 70, 92-102.

Cutts, Amy and Robert Van Order. 2003. On the Economics of Subprime Lending. Journal of Real Estate Finance and Economics, forthcoming.

Deng, Yongheng, John Quigley, and Robert Van Order. 1996. Mortgage Default and Low Downpayment Loans: The Costs of Public Subsidy Regional Science and Urban Economics, 26(3-4), 263-285, (1996).

Dipasquale, Denise and Edward L. Glaeser. 1999. Incentives and Social Capital: Are Homeowners Better Citizens? Journal of Urban Economics 45: 354-384.

Dhillon, Upinder, James D. Shilling and C.F. Sirmans 1987. Choosing Between Fixed and Adjustable Rate Mortgages. Journal of Money, Credit, and Banking 19: 260-267.

Duda, Mark, and Eric S. Belsky. 2002. Asset Appreciation, Timing of Purchases and Sales, and Returns to Low-Income Homeownership. In Retsinas and Belsky, editors, *Low-Income Homeownership: Examining the Unexamined Goal*. Washington, DC: Brookings Institution Press.

Follain, James R. 1990. Mortgage Choice. AREUEA Journal 18(2): 125-144.

Follian, James and Robert Dunsky. 1997. The Demand for Mortgage Debt and the Income Tax.
Journal of Housing Research 8(2) : 155-199.

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

Green, Richard and Michelle White. 1997. Measuring the Benefits of Homeownership: Effects
on Children. Journal of Urban Economics 41: 441-461.

Green, Richard and Patric Hendershott. 2001. Home Ownership and Unemployment in the U.S.
Urban Studies 38: 1509-1520.

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

HUD, 2000. Unequal Burden: Income and Racial Disparities in Subprime Lending in America

Jones, Lawrence D. 1993. The Demand for Home Mortgage Debt. Journal of Urban Economics
33(1): 10-28.

Ling, David C. and Gary A. McGill. 1998. Evidence on the Demand for Mortgage Debt by
Owner-Occupants. Journal of Urban Economics, 44 (3): tbd.

Linneman, Peter and Susan Wachter. 1989. The Impacts of Borrowing Constraints on Home
Ownership. AREUEA Journal 17(4): 389-402.

Maddala, G.S. 1983. Limited-Dependent and Qualitative Variables in Econometrics. New
York: Cambridge University Press.

McCarthy, George, Shannon Van Zandt, and William Rohe. 2001. The Economic Benefits and
Costs of Homeownership: A Critical Assessment of the Research. Research Insitute for
Housing America, Working Paper 01-02.

Office of the Comptroller of the Currency. 2001. Expanded Guidance for Subprime Lending Programs. OCC 2001-6.

Office of the Comptroller of the Currency. 2003. Economic Issues in Predatory Lending. Global Banking & Financial Analysis Working Paper, July 30, 2003.

Pennington-Cross, Anthony. 2003. Credit History and the Performance of Prime and Nonprime Mortgages. Journal of Real Estate Finance and Economics 27(3), forthcoming.

Pennington-Cross, Anthony and J. Nichols. 2000. Credit History and the FHA-Conventional Choice. Real Estate Economics 28(2): 307-336.

Pennington-Cross, Anthony J Nichols, and A.M. Yezer. 2000. Credit Risk and Mortgage Lending: Who Uses Subprime and Why? Research Institute for Housing America, Working Paper 00-03.

Retsinas, Nicolas and Eric S. Belsky. 2002. *Low-Income Homeownership: Examining the Unexamined Goal*. Washington, DC: Brookings Institution Press.

Rohe, William M., George McCarthy, and Shannon Van Zandt. 2000. The Social Benefits and Costs of Homeowenrship: A Critical Assessment of the Research. Research Institute for Housing America, Working Paper 00-01.

Rosenthanl, Stuart S. 2002. Eliminating Credit Barriers: How Far Can We Go? In Retsinas and Belsky, editors, *Low-Income Homeownership: Examining the Unexamined Goal*. Washington, DC: Brookings Institution Press.

Stanton, Richard and Nancy Wallace. 1998. Mortgage Choice: What's the Point? Real Estate Economics 26 (2): 173-206.

Surette, Brian, Marsha Courchane, and Peter Zorn. 2003. Subprime Borrowers: Mortgage Transitions and Outcomes. Journal of Real Estate Finance and Economics, forthcoming.

Zorn, Peter. 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.

List of Tables, Graphs, and Appendices

Number	Content
1	Descriptive Statistics: Full Population
2	Percentage Subprime by Type and Subprime Definition
3	Descriptive Statistics: LMI Sample
4	MNL LMI Sample
5	MNL LMI Sample, No Demographics
6	Descriptive Statistics: First-time Homebuyers Sample
7	MNL First-time Homebuyers Sample
8	MNL First-time Homebuyers Sample, No Demographics
9	Graphs: Mean Conditional Probability of Product Choice
10	Equation Appendix

Table 1: Descriptive Statistics – Full Population

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
Transaction Information	noteamt	Loan Note Amount	Conventional	136,937	182856.17	138236.27
			FHA	36,467	116913.84	43921.59
	saleamt	Sales Price Amount	Conventional	136,878	242491.64	212683.89
			FHA	36,469	120111.42	45691.93
	apr	Annual Percentage Rate	Conventional	126,176	6.80	1.17
			FHA	31,255	7.14	0.99
	noterate	Loan Note Rate	Conventional	136,947	6.63	0.95
			FHA	36,472	6.61	0.76
	term	Loan Term in Months	Conventional	137,040	345.92	47.90
			FHA	36,474	357.44	20.96
Credit Score	origfico	FICO score	Conventional	135,294	723.50	58.87
			FHA	34,710	669.50	65.29
Other Underwriting	cens_income	HMDA Reportable Borrower Income in 1000s	Conventional	137,040	85.91	112.24
			FHA	36,474	51.90	56.46
	debt_to_income	Debt to Income Ratio	Conventional	134,802	36.22	10.37
			FHA	36,369	38.48	8.37
	house_to_income	Housing Debt to Income Ratio	Conventional	134,746	24.95	9.22
			FHA	36,324	24.92	8.35
	yrsemply	Years Employed	Conventional	109,724	7.30	7.10
			FHA	28,595	5.13	5.38
	nodoc	Indicator for No Documentation Loans	Conventional	137,040	0.02	0.13
			FHA	36,474	0.00	0.07
	origltv	Loan to Value Ratio	Conventional	136,567	80.09	14.80
			FHA	36,462	97.61	4.56
Borrower Demographics	bor1_age	Borrower 1 Age	Conventional	137,021	40.13	11.99
			FHA	36,467	33.52	10.02
	black	Indicator for Race = Black	Conventional	137,040	0.04	0.18
			FHA	36,474	0.08	0.27
	hispanic	Indicator for Race = Hispanic	Conventional	137,040	0.07	0.25
			FHA	36,474	0.12	0.32
	native	Indicator for Race = Native American	Conventional	137,040	0.00	0.06
			FHA	36,474	0.01	0.08

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
	asian	Indicator for Race = Asian	Conventional	137,040	0.04	0.20
			FHA	36,474	0.02	0.13
	raceoth	Indicator for Race = Other	Conventional	137,040	0.01	0.12
			FHA	36,474	0.01	0.11
	white	Indicator for Race = White	Conventional	137,040	0.73	0.44
			FHA	36,474	0.70	0.46
	race_na	Indicator for Race = NA	Conventional	137,040	0.10	0.30
			FHA	36,474	0.06	0.24
Area Demographics	cens_inclvl	Income relative to HUD MSA Median Family Income	Conventional	137,040	143.24	181.41
			FHA	36,474	92.59	100.94
	cens_pctmed	Census Level Median Income to Decennial MSA Median	Conventional	136,969	120.22	39.08
			FHA	36,470	103.10	26.22
	cens_pctmin	Percent Minority Population in Census Tract	Conventional	137,006	21.76	21.33
			FHA	36,471	23.64	23.04
	low_inc	Indicator if Cens_Inclvl =< 50	Conventional	137,040	0.09	0.29
			FHA	36,474	0.14	0.35
	mod_inc	Indicator if Cens_Inclvl > 50 and <= 80	Conventional	137,040	0.19	0.39
			FHA	36,474	0.35	0.48

Table 2: Percent Subprime by Type and Subprime Definition

Loan Category	Definition	N	% Subprime
Conventional	Regulatory	126179	23.79
	Relative Note Rate	126179	9.72
	HMDA reportable APR spread	126179	6.42
	WFHM	126179	7.19
FHA	Regulatory	31255	50.21

Table 3: Descriptive Statistics: LMI Sample

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
Transaction Information	saleamt	Sales Price Amount	Conventional	457	164448.72	105271.21
			EMNP	405	109260.37	51927.45
			FHA	389	113824.78	45898.86
			Subprime	366	110185.86	76751.83
	noteamt	Loan Note Amount	Conventional	457	125812.09	75664.53
			EMNP	405	107544.90	50971.82
			FHA	389	111423.04	44705.94
			Subprime	366	90568.48	61900.28
	noterate	Loan Note Rate	Conventional	457	6.44	0.74
			EMNP	405	6.84	0.76
			FHA	389	6.61	0.80
			Subprime	366	9.09	1.48
	apr	Annual Percentage Rate	Conventional	345	6.62	0.91
			EMNP	254	7.21	1.16
			FHA	245	7.08	1.06
			Subprime	366	7.69	3.94
Credit Score	loanfico	FICO score	Conventional	457	728.90	55.24
			EMNP	405	670.18	52.08
			FHA	389	662.88	67.69
			Subprime	366	621.30	60.82
	fico_620	FICO Bucket < 620	Conventional	457	0.05	0.21
			EMNP	405	0.14	0.35
			FHA	389	0.27	0.45
			Subprime	366	0.55	0.50
	fico620_640	FICO Bucket 620-640	Conventional	457	0.02	0.15
			EMNP	405	0.17	0.38
			FHA	389	0.10	0.29
			Subprime	366	0.09	0.29
	fico640_660	FICO Bucket 640-660	Conventional	457	0.05	0.22
			EMNP	405	0.16	0.37
			FHA	389	0.14	0.35
			Subprime	366	0.09	0.28
	fico660_680	FICO Bucket 660-680	Conventional	457	0.06	0.24

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			EMNP	405	0.14	0.35
			FHA	389	0.08	0.27
			Subprime	366	0.09	0.29
	fico680_700	FICO Bucket 680-700	Conventional	457	0.08	0.27
			EMNP	405	0.12	0.33
			FHA	389	0.11	0.31
			Subprime	366	0.06	0.23
Other Underwriting	cens_income	HMDA Reportable Borrower Income in 1000s	Conventional	457	51.29	43.16
			EMNP	405	41.05	17.81
			FHA	389	41.17	16.06
			Subprime	366	43.06	23.47
	ltv	Loan to Value Ratio	Conventional	457	78.99	18.25
			EMNP	405	98.30	6.32
			FHA	389	97.61	4.92
			Subprime	366	83.36	9.20
	house_to_income	Housing Debt to Income Ratio	Conventional	457	27.06	9.73
			EMNP	405	28.71	8.66
			FHA	389	27.84	8.23
			Subprime	366	28.13	11.13
	debt_to_income	Debt to Income Ratio	Conventional	457	35.44	11.67
			EMNP	405	34.23	8.34
			FHA	389	37.26	9.62
			Subprime	366	41.80	11.44
	high_DTI	High Debt to Income Indicator	Conventional	457	0.07	0.25
			EMNP	405	0.01	0.11
			FHA	389	0.06	0.25
			Subprime	366	0.26	0.44
	ltv_8090	2SLS - LTV Bucket	Conventional	457	0.67	0.47
			EMNP	405	0.46	0.50
			FHA	389	0.46	0.50
			Subprime	366	0.25	0.43
	ltv_9095	2SLS - LTV Bucket	Conventional	457	0.09	0.29
			EMNP	405	0.41	0.49
			FHA	389	0.32	0.47
			Subprime	366	0.28	0.45

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
	ltv_95plus	2SLS - LTV Bucket	Conventional	457	0.03	0.17
			EMNP	405	0.11	0.31
			FHA	389	0.19	0.39
			Subprime	366	0.46	0.50
Other_Misc	close_30	Closed 1 to 30 days	Conventional	457	0.44	0.50
			EMNP	405	0.44	0.50
			FHA	389	0.36	0.48
			Subprime	366	0.53	0.50
	homebuy1st	First Time Homebuyer	Conventional	457	0.56	0.50
			EMNP	405	0.85	0.35
			FHA	389	0.83	0.37
			Subprime	366	0.64	0.48
	selfemploy	Self Employment Indicator	Conventional	457	0.05	0.21
			EMNP	405	0.01	0.10
			FHA	389	0.01	0.10
			Subprime	366	0.09	0.29
	prob_fullalt	Estimated Prob. of Full/Alt Doc.	Conventional	457	0.89	0.14
			EMNP	405	0.95	0.07
			FHA	389	0.94	0.07
			Subprime	366	0.89	0.12
Borrower Demographics	bor1_age	Borrower 1 Age	Conventional	457	41.46	14.21
			EMNP	405	33.92	10.64
			FHA	389	33.44	9.63
			Subprime	366	39.03	11.08
	native	Indicator for Race = Native American	Conventional	457	0.00	0.05
			EMNP	405	0.00	0.07
			FHA	389	0.01	0.07
			Subprime	366	0.00	0.00
	asian	Indicator for Race = Asian	Conventional	457	0.03	0.18
			EMNP	405	0.02	0.13
			FHA	389	0.04	0.19
			Subprime	366	0.03	0.16
	black	Indicator for Race = Black	Conventional	457	0.05	0.22
			EMNP	405	0.12	0.33
			FHA	389	0.10	0.30

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			Subprime	366	0.13	0.33
	hispanic	Indicator for Race = Hispanic	Conventional	457	0.10	0.30
			EMNP	405	0.26	0.44
			FHA	389	0.18	0.38
			Subprime	366	0.16	0.37
	other	Indicator for Race = Other	Conventional	457	0.01	0.11
			EMNP	405	0.01	0.09
			FHA	389	0.03	0.17
			Subprime	366	0.01	0.12
	race_na	Indicator for Race = NA	Conventional	457	0.08	0.28
			EMNP	405	0.02	0.15
			FHA	389	0.06	0.24
			Subprime	366	0.21	0.41
Area Demographics	low_inc	Indicator if Cens_Inclvl =< 50	Conventional	457	0.21	0.41
			EMNP	405	0.23	0.42
			FHA	389	0.21	0.41
			Subprime	366	0.27	0.44
	mod_inc	Indicator if Cens_Inclvl > 50 and <= 80	Conventional	457	0.51	0.50
			EMNP	405	0.54	0.50
			FHA	389	0.58	0.49
			Subprime	366	0.45	0.50
	pctmin30_plus	Census Tract > 30% Minority	Conventional	457	0.40	0.49
			EMNP	405	0.55	0.50
			FHA	389	0.48	0.50
			Subprime	366	0.57	0.50
Subprime Definitions	subprime_occ	OCC Subprime Definition	Conventional	457	0.23	0.42
			EMNP	405	0.53	0.50
			FHA	389	0.58	0.49
			Subprime	366	0.82	0.38
	subprime_top10	Subprime top 10% Loan Volume	Conventional	457	0.06	0.23
			EMNP	405	0.19	0.39
			FHA	389	0.04	0.20
			Subprime	366	0.87	0.33
	apr_over_hmda	HMDA APR spread Subprime Definition	Conventional	428	0.01	0.10
			EMNP	352	0.06	0.23

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			FHA	349	0.01	0.12
			Subprime	290	0.73	0.45

Table 4: Multinomial Logit Model, LMI Sample

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Intercept	Intercept	EMNP	-5.9570	1.0537	20.8722	<.0001
		FHA	-4.7088	1.0065	19.4521	<.0001
		Subprime	.7909	0.8600	9.2027	0.0024
Credit Score	fico_620	EMNP	0.8887	0.4168	4.5470	0.0330
		FHA	1.7150	0.4013	18.2605	<.0001
		Subprime	2.6629	0.4552	34.2257	<.0001
	fico620_640	EMNP	1.8943	0.4123	21.1129	<.0001
		FHA	1.5884	0.4258	13.9160	0.0002
		Subprime	2.3352	0.4900	22.7113	<.0001
	fico640_660	EMNP	1.4803	0.3168	21.8316	<.0001
		FHA	1.4941	0.3184	22.0240	<.0001
		Subprime	1.8247	0.3823	22.7829	<.0001
	fico660_680	EMNP	1.1767	0.2992	15.4633	<.0001
		FHA	0.8060	0.3138	6.5959	0.0102
		Subprime	1.7154	0.3612	22.5529	<.0001
	fico680_700	EMNP	0.9789	0.2708	13.0693	0.0003
		FHA	0.8626	0.2708	10.1493	0.0014
		Subprime	1.0002	0.3468	8.3193	0.0039
Other Underwriting	house_to_income	EMNP	0.0189	0.0105	3.2269	0.0724
		FHA	-0.0138	0.0104	1.7543	0.1853
		Subprime	0.00240	0.0115	0.0435	0.8348
	debt_to_income	EMNP	-0.0122	0.00937	1.6814	0.1947
		FHA	0.0272	0.00902	9.0953	0.0026
		Subprime	0.0933	0.0105	79.2529	<.0001
	ltv_8090	EMNP	1.1470	0.3953	8.4215	0.0037
		FHA	0.4959	0.3271	2.2988	0.1295
		Subprime	4.5070	0.7578	35.3695	<.0001
	ltv_9095	EMNP	2.0257	0.4603	19.3642	<.0001
		FHA	1.1468	0.4091	7.8570	0.0051
		Subprime	5.8516	0.8237	50.4721	<.0001
	ltv_95plus	EMNP	1.8644	0.5779	10.4093	0.0013
		FHA	1.2615	0.5296	5.6746	0.0172
		Subprime	7.1921	0.8948	64.6084	<.0001

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Other Misc	close_30	EMNP	0.0391	0.1546	0.0640	0.8002
		FHA	-0.2803	0.1552	3.2627	0.0709
		Subprime	0.4274	0.1776	5.7895	0.0161
	homebuy1st	EMNP	1.2062	0.1842	42.8799	<.0001
		FHA	1.0634	0.1803	34.7717	<.0001
		Subprime	-0.1368	0.1928	0.5030	0.4782
	selfemploy	EMNP	-1.0212	0.6038	2.8605	0.0908
		FHA	-1.6014	0.6592	5.9024	0.0151
		Subprime	-4.2178	0.5730	54.1779	<.0001
	yrsemply	EMNP	-0.0141	0.0171	0.6784	0.4102
		FHA	-0.0154	0.0166	0.8652	0.3523
		Subprime	0.0283	0.0163	3.0271	0.0819
	prob_fullalt	EMNP	1.4201	1.1180	1.6133	0.2040
		FHA	1.4148	1.0719	1.7422	0.1869
		Subprime	-14.1171	1.4036	101.1565	<.0001
Borrower Demographics	age18_34	EMNP	0.6034	0.1657	13.2542	0.0003
		FHA	0.7026	0.1647	18.1955	<.0001
		Subprime	0.1559	0.1952	0.6376	0.4246
	age55_64	EMNP	-0.6563	0.3671	3.1966	0.0738
		FHA	-0.3961	0.3552	1.2432	0.2649
		Subprime	-0.2245	0.3556	0.3984	0.5279
	age65_	EMNP	-1.0761	0.4742	5.1499	0.0232
		FHA	-2.1540	0.6586	10.6979	0.0011
		Subprime	-1.4278	0.4745	9.0553	0.0026
	asian	EMNP	-1.0411	0.4838	4.6319	0.0314
		FHA	-0.2108	0.4070	0.2683	0.6045
		Subprime	-0.5370	0.5197	1.0677	0.3015
	black	EMNP	0.4423	0.3148	1.9744	0.1600
		FHA	0.3231	0.3201	1.0187	0.3128
		Subprime	0.4894	0.3489	1.9669	0.1608
	hispanic	EMNP	0.5477	0.2285	5.7447	0.0165
		FHA	0.1965	0.2369	0.6882	0.4068
		Subprime	-0.2450	0.2750	0.7937	0.3730
	other	EMNP	-0.3520	0.7301	0.2325	0.6297
		FHA	1.1529	0.5949	3.7555	0.0526

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
		Subprime	0.8440	0.7032	1.4405	0.2301
	race_na	EMNP	-0.7778	0.3530	4.8564	0.0275
		FHA	-0.2581	0.3014	0.7337	0.3917
		Subprime	1.1577	0.2905	15.8842	<.0001
Area Demographics	low_inc	EMNP	0.3130	0.2573	1.4800	0.2238
		FHA	0.1466	0.2595	0.3192	0.5721
		Subprime	-0.7519	0.3086	5.9340	0.0149
	mod_inc	EMNP	0.0101	0.2030	0.0025	0.9601
		FHA	0.1590	0.2029	0.6140	0.4333
		Subprime	-0.6429	0.2379	7.3028	0.0069
	pctmin30_plus	EMNP	0.3168	0.1855	2.9174	0.0876
		FHA	0.1731	0.1820	0.9045	0.3416
		Subprime	0.3107	0.2085	2.2208	0.1362
Misc Control Variables	fico_miss	EMNP	0.3301	0.3990	0.6844	0.4081
		FHA	-0.1930	0.3927	0.2416	0.6231
		Subprime	-0.7491	0.4277	3.0675	0.0799
	dti_miss	EMNP	1.2724	0.5761	4.8780	0.0272
		FHA	1.6569	0.5794	8.1784	0.0042
		Subprime	3.2766	0.6338	26.7274	<.0001

Table 5: Multinomial Logit Model, LMI Sample, No Demographics

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Intercept	Intercept	EMNP	-6.1337	1.0686	21.8117	<.0001
		FHA	-4.7232	0.9963	19.9833	<.0001
		Subprime	0.2244	0.8255	6.1203	0.0134
Credit Score	fico_620	EMNP	0.8791	0.3973	4.8963	0.0269
		FHA	1.7474	0.3837	20.7443	<.0001
		Subprime	3.1745	0.4282	54.9551	<.0001
	fico620_640	EMNP	1.9479	0.4040	23.2475	<.0001
		FHA	1.6258	0.4172	15.1837	<.0001
		Subprime	2.7286	0.4749	33.0059	<.0001
	fico640_660	EMNP	1.5727	0.3059	26.4290	<.0001
		FHA	1.5967	0.3079	26.8848	<.0001
		Subprime	2.1416	0.3668	34.0966	<.0001
	fico660_680	EMNP	1.2180	0.2877	17.9217	<.0001
		FHA	0.8499	0.3052	7.7554	0.0054
		Subprime	1.9516	0.3461	31.7943	<.0001
	fico680_700	EMNP	1.0011	0.2618	14.6240	0.0001
		FHA	0.9193	0.2624	12.2714	0.0005
		Subprime	1.2347	0.3379	13.3553	0.0003
Other Underwriting	house_to_income	EMNP	0.0277	0.00964	8.2527	0.0041
		FHA	-0.0101	0.00947	1.1335	0.2870
		Subprime	-0.00623	0.0104	0.3582	0.5495
	debt_to_income	EMNP	-0.0142	0.00892	2.5396	0.1110
		FHA	0.0266	0.00859	9.6126	0.0019
		Subprime	0.0880	0.00997	77.9008	<.0001
	ltv_8090	EMNP	1.1453	0.3746	9.3473	0.0022
		FHA	0.5505	0.3118	3.1170	0.0775
		Subprime	4.5205	0.7546	35.8879	<.0001
	ltv_9095	EMNP	2.0672	0.4314	22.9581	<.0001
		FHA	1.1717	0.3837	9.3234	0.0023
		Subprime	5.5201	0.8197	45.3532	<.0001
	ltv_95plus	EMNP	2.0916	0.5401	14.9952	0.0001
		FHA	1.3772	0.4952	7.7353	0.0054
		Subprime	6.7488	0.8885	57.6984	<.0001

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Other Misc	close_30	EMNP	0.0254	0.1503	0.0285	0.8659
		FHA	-0.2649	0.1513	3.0660	0.0799
		Subprime	0.4261	0.1722	6.1202	0.0134
	homebuy1st	EMNP	1.3511	0.1764	58.6339	<.0001
		FHA	1.2548	0.1729	52.6424	<.0001
		Subprime	-0.0552	0.1819	0.0920	0.7616
	selfemploy	EMNP	-0.6450	0.5903	1.1939	0.2745
		FHA	-1.3524	0.6536	4.2819	0.0385
		Subprime	-3.9518	0.5721	47.7157	<.0001
	yrsemply	EMNP	-0.0330	0.0163	4.0834	0.0433
		FHA	-0.0386	0.0160	5.8407	0.0157
		Subprime	0.0299	0.0156	3.6558	0.0559
	prob_fullalt	EMNP	1.8760	1.1219	2.7958	0.0945
		FHA	1.7489	1.0582	2.7315	0.0984
		Subprime	-13.3823	1.4037	90.8873	<.0001
Misc Control Variables	fico_miss	EMNP	0.4606	0.3834	1.4434	0.2296
		FHA	-0.2997	0.3776	0.6301	0.4273
		Subprime	-0.8618	0.4043	4.5444	0.0330
	dti_miss	EMNP	1.2873	0.5395	5.6922	0.0170
		FHA	1.5026	0.5475	7.5311	0.0061
		Subprime	2.8527	0.5972	22.8174	<.0001

Table 6: Descriptive Statistics: First-time Homebuyer Sample

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
Transaction Information	saleamt	Sales Price Amount	Conventional	538	156830.72	73993.70
			EMNP	493	114260.61	50851.41
			FHA	442	116914.88	44441.86
			Subprime	406	125184.20	71792.91
	noteamt	Loan Note Amount	Conventional	538	124728.27	54586.72
			EMNP	493	113443.73	50442.78
			FHA	442	114266.49	43555.85
			Subprime	406	101993.94	56586.54
	noterate	Loan Note Rate	Conventional	538	6.26	0.74
			EMNP	493	6.89	0.74
			FHA	442	6.43	0.77
			Subprime	406	8.88	1.54
	apr	Annual Percentage Rate	Conventional	467	6.46	0.95
			EMNP	434	7.20	1.36
			FHA	334	6.92	1.02
			Subprime	406	9.32	1.65
Credit Score	loanfico	FICO score	Conventional	538	726.03	54.38
			EMNP	493	684.33	50.66
			FHA	442	665.30	59.86
			Subprime	406	631.75	63.78
	fico_620	FICO Bucket < 620	Conventional	538	0.03	0.18
			EMNP	493	0.08	0.27
			FHA	442	0.24	0.43
			Subprime	406	0.49	0.50
	fico620_640	FICO Bucket 620-640	Conventional	538	0.05	0.22
			EMNP	493	0.12	0.33
			FHA	442	0.14	0.34
			Subprime	406	0.12	0.32
	fico640_660	FICO Bucket 640-660	Conventional	538	0.04	0.19
			EMNP	493	0.16	0.37
			FHA	442	0.10	0.31
			Subprime	406	0.07	0.25
	fico660_680	FICO Bucket 660-680	Conventional	538	0.07	0.26

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			EMNP	493	0.15	0.36
			FHA	442	0.11	0.31
			Subprime	406	0.05	0.22
	fico680_700	FICO Bucket 680-700	Conventional	538	0.09	0.28
			EMNP	493	0.12	0.33
			FHA	442	0.11	0.31
			Subprime	406	0.11	0.31
Other Underwriting	cens_income	HMDA Reportable Borrower Income in 1000s	Conventional	538	55.62	28.79
			EMNP	493	45.31	17.38
			FHA	442	47.18	20.89
			Subprime	406	50.65	26.85
	ltv	Loan to Value Ratio	Conventional	538	82.00	14.30
			EMNP	493	99.35	2.83
			FHA	442	97.79	3.83
			Subprime	406	83.06	9.65
	house_to_income	Housing Debt to Income Ratio	Conventional	538	25.35	9.25
			EMNP	493	25.92	8.02
			FHA	442	24.98	7.75
			Subprime	406	26.93	9.63
	debt_to_income	Debt to Income Ratio	Conventional	538	35.60	10.26
			EMNP	493	36.91	6.42
			FHA	442	38.48	8.13
			Subprime	406	39.63	10.09
	high_DTI	High Debt to Income Indicator	Conventional	538	0.05	0.21
			EMNP	493	0.01	0.09
			FHA	442	0.04	0.19
			Subprime	406	0.19	0.39
	ltv_8090	2SLS - LTV Bucket	Conventional	538	0.71	0.46
			EMNP	493	0.56	0.50
			FHA	442	0.52	0.50
			Subprime	406	0.33	0.47
	ltv_9095	2SLS - LTV Bucket	Conventional	538	0.10	0.30
			EMNP	493	0.34	0.48
			FHA	442	0.31	0.46
			Subprime	406	0.26	0.44

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
	ltv_95plus	2SLS - LTV Bucket	Conventional	538	0.02	0.14
			EMNP	493	0.07	0.25
			FHA	442	0.14	0.35
			Subprime	406	0.38	0.49
Other_Misc	close1_10	Closed 1 to 10 days	Conventional	538	0.01	0.12
			EMNP	493	0.02	0.13
			FHA	442	0.01	0.09
			Subprime	406	0.08	0.27
	close10_20	Closed 10 to 20 days	Conventional	538	0.14	0.35
			EMNP	493	0.12	0.32
			FHA	442	0.07	0.25
			Subprime	406	0.25	0.43
	close20_30	Closed 20 to 30 days	Conventional	538	0.29	0.45
			EMNP	493	0.26	0.44
			FHA	442	0.23	0.42
			Subprime	406	0.26	0.44
	homebuy1st	First Time Homebuyer	Conventional	538	1.00	0.00
			EMNP	493	1.00	0.00
			FHA	442	1.00	0.00
			Subprime	406	1.00	0.00
	selfemploy	Self Employment Indicator	Conventional	538	0.06	0.23
			EMNP	493	0.03	0.17
			FHA	442	0.01	0.12
			Subprime	406	0.09	0.29
	prob_fullalt	Estimated Prob. of Full/Alt Doc.	Conventional	538	0.87	0.13
			EMNP	493	0.92	0.08
			FHA	442	0.93	0.08
			Subprime	406	0.88	0.14
Borrower Demographics	bor1_age	Borrower 1 Age	Conventional	538	37.55	13.00
			EMNP	493	32.59	10.48
			FHA	442	32.35	9.18
			Subprime	406	36.81	10.89
	age18_34	Age Bucket 18-34	Conventional	538	0.49	0.50
			EMNP	493	0.65	0.48
			FHA	442	0.66	0.48

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			Subprime	406	0.46	0.50
	age55_64	Age Bucket 55-64	Conventional	538	0.06	0.23
			EMNP	493	0.02	0.13
			FHA	442	0.03	0.18
			Subprime	406	0.05	0.23
	age65_	Age Bucket 65+	Conventional	538	0.05	0.22
			EMNP	493	0.02	0.13
			FHA	442	0.00	0.05
			Subprime	406	0.01	0.12
	native	Indicator for Race = Native American	Conventional	538	0.00	0.04
			EMNP	493	0.01	0.08
			FHA	442	0.02	0.13
			Subprime	406	0.00	0.07
	asian	Indicator for Race = Asian	Conventional	538	0.07	0.25
			EMNP	493	0.00	0.06
			FHA	442	0.02	0.14
			Subprime	406	0.02	0.13
	black	Indicator for Race = Black	Conventional	538	0.03	0.16
			EMNP	493	0.09	0.28
			FHA	442	0.10	0.30
			Subprime	406	0.06	0.25
	hispanic	Indicator for Race = Hispanic	Conventional	538	0.08	0.27
			EMNP	493	0.17	0.38
			FHA	442	0.11	0.31
			Subprime	406	0.15	0.35
	other	Indicator for Race = Other	Conventional	538	0.01	0.11
			EMNP	493	0.01	0.09
			FHA	442	0.02	0.14
			Subprime	406	0.00	0.07
	race_na	Indicator for Race = NA	Conventional	538	0.09	0.29
			EMNP	493	0.06	0.23
			FHA	442	0.03	0.18
			Subprime	406	0.23	0.42
Area Demographics	low_inc	Indicator if Cens_Inclvl =< 50	Conventional	538	0.15	0.35
			EMNP	493	0.16	0.36

Category	Variable	Description	Loan Type	N	Mean	Std. Dev
			FHA	442	0.13	0.34
			Subprime	406	0.15	0.36
	mod_inc	Indicator if Cens_Inclvl > 50 and <= 80	Conventional	538	0.30	0.46
			EMNP	493	0.42	0.49
			FHA	442	0.39	0.49
			Subprime	406	0.32	0.47
	pctmin30_plus	Census Tract > 30% Minority	Conventional	538	0.25	0.44
			EMNP	493	0.34	0.47
			FHA	442	0.29	0.45
			Subprime	406	0.37	0.48
Subprime Definitions	subprime_occ	OCC Subprime Definition	Conventional	538	0.21	0.41
			EMNP	493	0.42	0.49
			FHA	442	0.54	0.50
			Subprime	406	0.74	0.44
	subprime_top10	Subprime top 10% Loan Volume	Conventional	538	0.02	0.15
			EMNP	493	0.22	0.42
			FHA	442	0.03	0.16
			Subprime	406	0.80	0.40
	apr_over_hmda	HMDA APR spread Subprime Definition	Conventional	538	0.01	0.07
			EMNP	493	0.08	0.27
			FHA	442	0.00	0.07
			Subprime	406	0.70	0.46

Table 7: Multinomial Logit Model, First-time Homebuyers

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Intercept	Intercept	EMNP	-3.8902	0.9679	8.7079	0.0032
		FHA	-3.9440	1.1129	11.7692	0.0006
		Subprime	1.2899	0.7145	19.2915	<.0001
Credit Score	fico_620	EMNP	0.3762	0.4378	0.7382	0.3902
		FHA	2.5652	0.4179	37.6800	<.0001
		Subprime	2.8775	0.4394	42.8777	<.0001
	fico620_640	EMNP	0.6842	0.3158	4.6949	0.0303
		FHA	1.5711	0.3220	23.8004	<.0001
		Subprime	1.7351	0.3517	24.3318	<.0001
	fico640_660	EMNP	1.5333	0.3095	24.5353	<.0001
		FHA	1.7461	0.3263	28.6288	<.0001
		Subprime	1.5685	0.3716	17.8135	<.0001
	fico660_680	EMNP	1.0136	0.2480	16.6973	<.0001
		FHA	1.0634	0.2642	16.2022	<.0001
		Subprime	0.5091	0.3377	2.2728	0.1317
	fico680_700	EMNP	0.7184	0.2278	9.9491	0.0016
		FHA	0.9145	0.2425	14.2192	0.0002
		Subprime	1.2015	0.2750	19.0948	<.0001
Other Underwriting	house_to_income	EMNP	0.000012	0.0106	0.0000	0.9991
		FHA	-0.0341	0.0110	9.5278	0.0020
		Subprime	0.00411	0.0120	0.1178	0.7314
	debt_to_income	EMNP	0.00526	0.00942	0.3120	0.5764
		FHA	0.0428	0.00970	19.4332	<.0001
		Subprime	0.0626	0.0107	34.0420	<.0001
	ltv_8090	EMNP	1.2125	0.3535	11.7629	0.0006
		FHA	0.9606	0.3759	6.5303	0.0106
		Subprime	2.1872	0.4998	19.1479	<.0001
	ltv_9095	EMNP	1.9166	0.4376	19.1845	<.0001
		FHA	0.9557	0.4635	4.2513	0.0392
		Subprime	2.9907	0.5868	25.9797	<.0001
	ltv_95plus	EMNP	2.3359	0.6083	14.7441	0.0001
		FHA	1.1293	0.6163	3.3573	0.0669
		Subprime	4.5576	0.7127	40.8881	<.0001

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Other Misc	close1_10	EMNP	-0.2529	0.5306	0.2273	0.6336
		FHA	-0.4384	0.5570	0.6194	0.4313
		Subprime	1.9226	0.4696	16.7639	<.0001
	close10_20	EMNP	-0.2431	0.2078	1.3678	0.2422
		FHA	-0.8425	0.2426	12.0594	0.0005
		Subprime	0.8675	0.2217	15.3098	<.0001
	close20_30	EMNP	-0.3233	0.1592	4.1245	0.0423
		FHA	-0.5872	0.1685	12.1422	0.0005
		Subprime	-0.1955	0.1934	1.0214	0.3122
	selfemploy	EMNP	-0.0361	0.5961	0.0037	0.9517
		FHA	-0.3195	0.6480	0.2431	0.6220
		Subprime	-3.3904	0.5578	36.9432	<.0001
	yrsemply	EMNP	-0.0431	0.0149	8.4049	0.0037
		FHA	-0.0160	0.0145	1.2162	0.2701
		Subprime	-0.00032	0.0150	0.0005	0.9829
	prob_fullalt	EMNP	0.6840	1.1978	0.3260	0.5680
		FHA	1.2019	1.3444	0.7993	0.3713
		Subprime	-10.8048	1.1727	84.8844	<.0001
Borrower Demographics	age18_34	EMNP	0.5627	0.1563	12.9569	0.0003
		FHA	0.6379	0.1642	15.0846	0.0001
		Subprime	0.00864	0.1807	0.0023	0.9618
	age55_64	EMNP	-0.3680	0.3998	0.8475	0.3573
		FHA	-0.00036	0.3886	0.0000	0.9993
		Subprime	0.0598	0.3769	0.0252	0.8738
	age65_	EMNP	-0.6181	0.4469	1.9129	0.1666
		FHA	-2.4210	1.0448	5.3693	0.0205
		Subprime	-0.8622	0.5720	2.2724	0.1317
	native	EMNP	1.7780	1.2445	2.0413	0.1531
		FHA	2.6610	1.1793	5.0913	0.0240
		Subprime	1.1066	1.4813	0.5580	0.4551
	asian	EMNP	-2.8798	0.7533	14.6165	0.0001
		FHA	-1.4255	0.4362	10.6780	0.0011
		Subprime	-1.2294	0.5111	5.7868	0.0161
	black	EMNP	0.9345	0.3482	7.2026	0.0073
		FHA	0.9405	0.3553	7.0073	0.0081

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
		Subprime	0.6043	0.4024	2.2549	0.1332
	hispanic	EMNP	0.7006	0.2325	9.0828	0.0026
		FHA	0.1452	0.2596	0.3126	0.5761
		Subprime	0.1406	0.2742	0.2628	0.6082
	other	EMNP	-0.3051	0.6730	0.2056	0.6502
		FHA	0.8786	0.5680	2.3925	0.1219
		Subprime	-0.2580	0.7827	0.1086	0.7417
	race_na	EMNP	-0.3596	0.2643	1.8514	0.1736
		FHA	-0.9606	0.3118	9.4902	0.0021
		Subprime	1.1889	0.2409	24.3519	<.0001
Area Demographics	low_inc	EMNP	0.1431	0.2332	0.3766	0.5394
		FHA	0.2686	0.2493	1.1613	0.2812
		Subprime	-0.3931	0.2693	2.1304	0.1444
	mod_inc	EMNP	0.2786	0.1637	2.8955	0.0888
		FHA	0.2424	0.1725	1.9732	0.1601
		Subprime	-0.3088	0.1991	2.4049	0.1210
	pctmin30_plus	EMNP	0.3044	0.1688	3.2518	0.0713
		FHA	0.1640	0.1799	0.8308	0.3621
		Subprime	0.2224	0.1935	1.3213	0.2504
Misc Control Variables	fico_miss	EMNP	1.0390	0.5584	3.4623	0.0628
		FHA	-0.2025	0.5630	0.1294	0.7191
		Subprime	-1.4968	0.6227	5.7786	0.0162
	yrsemp_miss	EMNP	0.00412	0.5430	0.0001	0.9940
		FHA	-0.9392	0.8427	1.2421	0.2651
		Subprime	0.3799	0.5560	0.4670	0.4944

Table 8: Multinomial Logit Model, First-time Homebuyers, No Demographics

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Intercept	Intercept	EMNP	-3.0961	0.9355	4.8585	0.0275
		FHA	-3.6579	1.1470	9.4815	0.0021
		Subprime	1.4214	0.6188	27.9200	<.0001
Credit Score	fico_620	EMNP	0.1985	0.4167	0.2268	0.6339
		FHA	2.2770	0.3978	32.7702	<.0001
		Subprime	3.1098	0.4236	53.8868	<.0001
	fico620_640	EMNP	0.5766	0.3038	3.6028	0.0577
		FHA	1.3981	0.3083	20.5662	<.0001
		Subprime	1.9209	0.3455	30.9205	<.0001
	fico640_660	EMNP	1.3652	0.2924	21.7980	<.0001
		FHA	1.5089	0.3106	23.6002	<.0001
		Subprime	1.6530	0.3565	21.5009	<.0001
	fico660_680	EMNP	0.9861	0.2399	16.8884	<.0001
		FHA	1.0041	0.2571	15.2483	<.0001
		Subprime	0.6716	0.3287	4.1760	0.0410
	fico680_700	EMNP	0.7460	0.2193	11.5772	0.0007
		FHA	0.8836	0.2353	14.0964	0.0002
		Subprime	1.2994	0.2644	24.1617	<.0001
Other Underwriting	house_to_income	EMNP	0.00738	0.00917	0.6467	0.4213
		FHA	-0.0321	0.00957	11.2837	0.0008
		Subprime	-0.00146	0.0106	0.0189	0.8907
	debt_to_income	EMNP	0.00298	0.00914	0.1063	0.7444
		FHA	0.0417	0.00939	19.7334	<.0001
		Subprime	0.0617	0.0104	34.8758	<.0001
	ltv_8090	EMNP	1.3256	0.3374	15.4389	<.0001
		FHA	0.9332	0.3487	7.1621	0.0074
		Subprime	2.3720	0.5297	20.0530	<.0001
	ltv_9095	EMNP	2.3021	0.4096	31.5849	<.0001
		FHA	1.2011	0.4270	7.9113	0.0049
		Subprime	3.0017	0.6034	24.7464	<.0001
	ltv_95plus	EMNP	2.6844	0.5698	22.1960	<.0001
		FHA	1.4235	0.5713	6.2078	0.0127
		Subprime	4.4039	0.7158	37.8476	<.0001

Category	Variable	Choice	Estimate	Standard Error	Wald Chi Square	Probability
Other Misc	close1_10	EMNP	-0.0904	0.5171	0.0306	0.8612
		FHA	-0.2807	0.5493	0.2612	0.6093
		Subprime	1.8812	0.4620	16.5800	<.0001
	close10_20	EMNP	-0.2671	0.2023	1.7432	0.1867
		FHA	-0.8499	0.2386	12.6907	0.0004
		Subprime	0.8232	0.2152	14.6280	0.0001
	close20_30	EMNP	-0.3314	0.1548	4.5842	0.0323
		FHA	-0.5859	0.1643	12.7104	0.0004
		Subprime	-0.2063	0.1881	1.2029	0.2727
	selfemploy	EMNP	-0.2766	0.5652	0.2395	0.6246
		FHA	-0.3290	0.6404	0.2640	0.6074
		Subprime	-3.4260	0.5436	39.7255	<.0001
	yrsemply	EMNP	-0.0595	0.0137	18.9936	<.0001
		FHA	-0.0395	0.0133	8.8387	0.0029
		Subprime	0.000688	0.0136	0.0025	0.9598
	prob_fullalt	EMNP	0.2426	1.1424	0.0451	0.8318
		FHA	1.5575	1.3496	1.3319	0.2485
		Subprime	-10.8590	1.1464	89.7295	<.0001
Misc Control Variables	fico_miss	EMNP	1.4450	0.5443	7.0478	0.0079
		FHA	-0.0241	0.5480	0.0019	0.9650
		Subprime	-1.2296	0.5996	4.2062	0.0403
	yrsemp_miss	EMNP	-0.4257	0.5042	0.7129	0.3985
		FHA	-1.5011	0.7970	3.5474	0.0596
		Subprime	0.1148	0.5131	0.0500	0.8230

Equation Appendix

The model presented in this paper can be represented as:

$$\log(H_i) = \alpha_0 + \alpha_1 f_i + \alpha_2 c_i + \alpha_3 d_i + \alpha_4 L_i + \varepsilon_i \qquad (1)$$

$$L_i = \delta_0 + \delta_1 f_i + \delta_2 c_i + \delta_3 d_i + \delta_4 \log(H_i) + \xi_i \qquad (2)$$

$$\hat{L}_i = \hat{\delta}_0 + \hat{\delta}_1 f_i + \hat{\delta}_2 c_i + \hat{\delta}_3 d_i + \hat{\delta}_4 \log(H_i) \qquad (3)$$

$$P(r_i = 1 \mid f_i, c_i, d_i) = \frac{e^{(\gamma_0 + \gamma_1 f_i + \gamma_2 c_i + y_3 d_i)}}{1 + e^{(\gamma_0 + \gamma_1 f_i + \gamma_2 c_i + \gamma_3 d_i)}} \qquad (4)$$

$$P(p_{ij} = 1 \mid f_i, c_i, d_i, \hat{L}_i, r_i) = \frac{e^{(\beta_{j0} + \beta_{j1} f_i + \beta_{j2} c_i + \beta_{j3} d_i + \beta_{j4} \hat{L}_i + \beta_{j5} r_i)}}{1 + \sum_{k=1}^{J-1} e^{(\beta_{k0} + \beta_{k1} f_i + \beta_{k2} c_i + \beta_{k3} d_i + \beta_{k4} \hat{L}_i + \beta_{k5} r_i)}} \qquad (5)$$

Where H_i is the home price, f_i is a vector of financial-monetary and loan characteristic variables, c_i is a vector of credit history variables, d_i is a vector of demographic variables, L_i is the loan to value ratio, r_i is the probability of being full/alternative documentation type, and p_{ij} is the probability that a borrower chooses loan type j, with j $\in$ {FHA, Subprime, EMNP} and J = Conventional. Equations (1) and (2) are estimated simultaneously using instrumental variables, which generates the predicted values represented by equation (3). Equation (4) is estimated using logistic regression and equation (5) is estimated using multinomial logit regression.

Graph 1: Mortgage Choice and FICO Score - Retail, LMI, All Doc Types
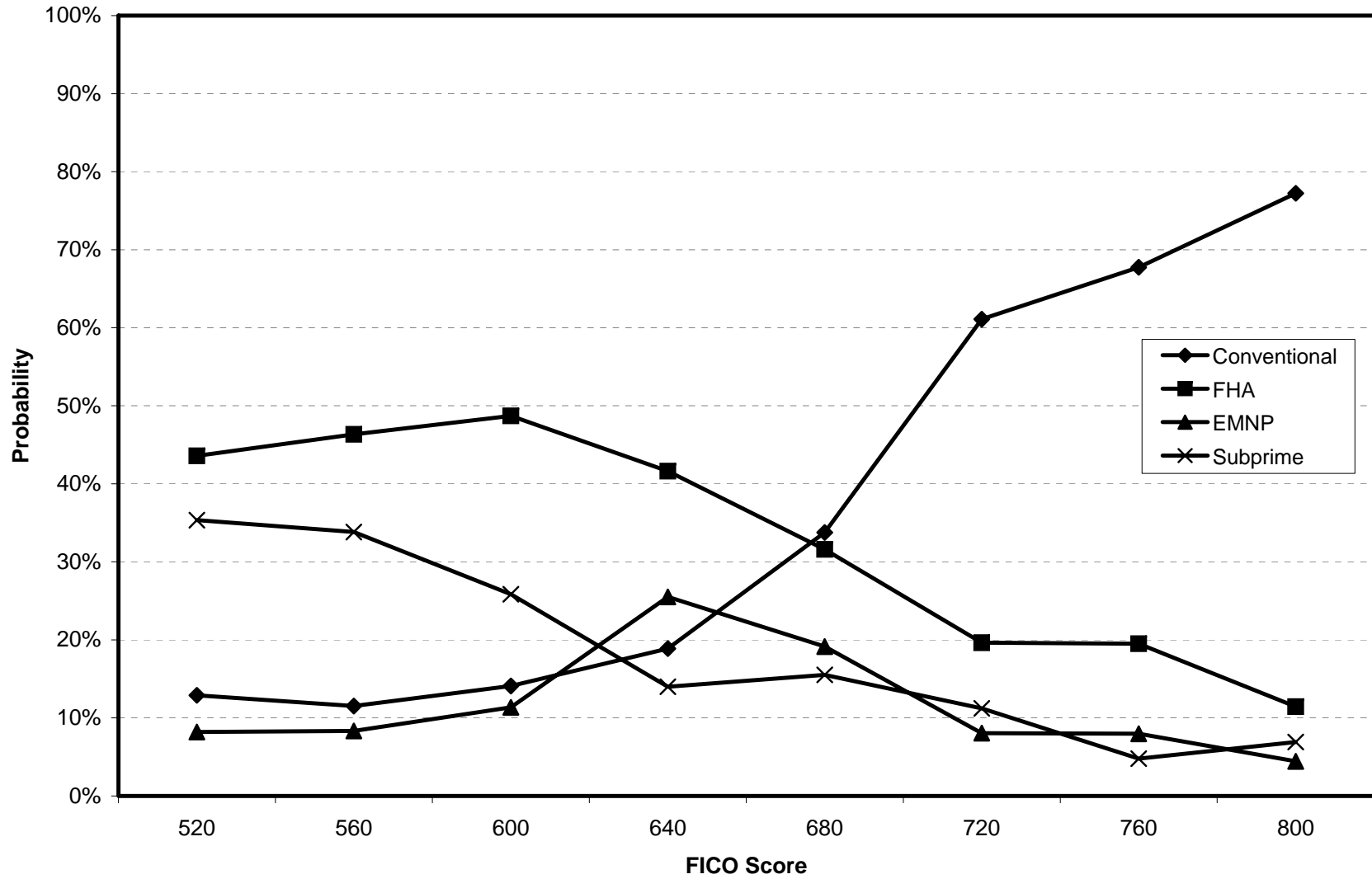
Graph 2: Mortgage Choice and FICO Score - Retail, First Time Homebuyer, All Doc Types