U.S. HOUSEHOLD MORTGAGE TERMINATIONS BETWEEN 2007-2009: LOCK-IN EFFECT, REFINANCING AND MOBILITY

by

Melinda Tímea Antal

Submitted to

Central European University

Department of Economics

In partial fulfillment of the requirements for the degree of Master of Arts in Economics

Supervisors: Associate Professor Gábor Kézdi

Budapest, Hungary

2011

ABSTRACT

In this thesis I estimate the impact on refinancing and moving decisions of the so-called lock-in effect caused by the recent declines in home prices. During 2007 and 2009 falling house prices have caused major capital losses to many homeowners, preventing them from refinancing or moving, and "locking" them into their old mortgage contracts. The diminished home equity is thought to be the main reason behind low prepayment and mobility rates, because under current circumstances borrowers could only get new loans that would be insufficient to prepay the existing loans and provide new down payments. Using a two year panel data from the American Housing Survey, I find that the magnitude of this lock-in effect increased significantly compared to the results presented in the literature for previous time periods, diminishing the baseline, two year probability of refinancing by 47.28 percent and the probability of moving by 34.92 percent. I also present significant evidence, that households' current income problems affect negatively the propensity to refinance, but "force" borrowers to terminate their existing mortgages by either defaulting or moving, hence increasing the average two year mobility rate. The final model is estimated as a multinomial probit model allowing for inter-choice correlation of unobservables.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my academic supervisor Prof. Gábor Kézdi for his continuous support and helpful suggestions to completing my thesis. I would also like to express my gratitude to my colleagues, Attila Agod and Zsolt Szekeres, who introduced me to the interesting world of risk analysis and institutional research, giving me inspiration to conduct this analysis; and also to Dávid Nagy for his precious and challenging remarks. Finally, I would like to thank my best friend Péter, for supporting me as much during thesis writing as ever before.

CONTENTS

LIST OF TABLES AND FIGURES

INTRODUCTION

The largest share of the US debt market is consisted of securities backed by fixed-rate residential mortgage [1] debt guaranteed by government sponsored enterprises (GSEs) and comprises about $5 trillion outstanding[2]. This enormous volume of private debt led almost every participant on the investments market in the past couple of years to focus constantly on how to deal with the risks associated to these instruments. Nevertheless, the most important aspect when analyzing the risks of a mortgage backed security (MBS) or a collateralized mortgage obligation (CMO) is the prepayment or termination risk. This is the risk associated with the early, unscheduled return of the principal – the amount borrowed, and causes uncertainty at the time of the purchase over the yield-to-maturity of such securities because the cashflows cannot be determined. Hence most of the research in this topic has been made by fixed-income and real-estate experts, mainly interested in pricing mortgage-backed derivatives.

There has been surprisingly few published works on mortgage decisions on the perspective of the household and my study aligns to this minority, aiming to give answers to questions such as: How did the so-called lock-in effect of declining house prices affect the probability of refinancing and mobility? What is the estimated magnitude of this effect right after the housing market collapsed? And lastly, how it changed over time, comparing the analyzed period's estimates to the results of other studies conducted using the same data source but in different years. To answer these, I will estimate probability models using individual household level data from the American Housing Survey (AHS). The micro level data, collected biannually,

[1] Fixed rate loan means that the interest rate does not change during the entire term of the loan, and is the opposite of adjustable rate mortgages, or ARMs.

[2] The US debt market's (bond market) size in 2010 reached $31.2 trillion outstanding. Nearly one third of this amount outstanding is comprised by mortgage-backed bonds, out of which $5 trillion in value were guaranteed by Government Sponsored Enterprises. Source: http://www.census.gov/compendia/statab/2011/tables/11s1191.pdf

allows controlling for individual household characteristics, addressing eventual concerns of heterogeneity.

In the recent years institutional researchers have been puzzled by the very low mortgage prepayment rates during and after the credit crunch, because their previous models, built mostly on interest rate movements, could not fit the data anymore. As presented in Figure 2, starting with the second half of 2008 the FHLMC (Federal Home Loan Mortgage Corporation, also called Freddie Mac) mortgage rates on 30 year fixed rate mortgages reached all-time lows. However the monthly reports on MBS prepayment rates, expressed in SMM[3] (single monthly mortality) rates were not soaring as it has been expected based on previous experiences. Although the declining interest rates have created large incentives to refinance, a large fraction of households paid interest rates on old, fix rated mortgages that highly exceeded the currently available rates. To show this, Figure 1 summarizes the distribution of the interest rate differentials paid by borrowers in every other year, from 2001 until 2009[4], as reported by the AHS respondents. For each spread, more generally, for the difference over the current mortgage rate, the figure shows the fraction of households that pay more than the prevailing rate. We can see that the distributions of the 2007 spreads and the 2009 ones are at the two extremes: by 2009, due to the large decreases in the 30 year FHMLC rates the fraction of households paying 1 percent more over the market rates almost tripled.

Hence, the refinancing boom, which started in 2001 and lasted until the end of 2004, with interest rate levels not even as low as in 2009, didn't seem to repeat[5]. This time other effects seemed to take over and drive the prepayment functions. The most important reasons for the

[3] The single monthly mortality – SMM is the percentage of the principal amount of mortgages that are prepaid in the given month. It is a metric designed to track prepayments in the mortgage pool. (source: www.investopedia.com)

[4] The illustration has been inspired by (Campbell 2006), and could serve as a continuation of the original.

[5] The spike we see in Figure 2 at the beginning of the year 2010 is because of the massive buy-outs of the delinquent loans, made by the U.S. government in an attempt to prevent the total collapse of the GSEs.

significantly lower prepayment rates during and after the crisis were identified to be the lock-in effect, when homeowners with an increased Loan-to-Value (LTV) ratio (due to massive declines in their home prices) did not meet the requirements of the lender institutions and could not cheaply refinance their loans; the disappearance of the cash-out effect – due to decreasing homeowner equity cash-out refinancing[6] became almost impossible; the hold-off effect, when people, expecting that home prices will return to previous high levels, chose not to sell their properties, thus slowing housing turnover; and the securitization process effect, caused by changes in the practices of the participants of the primary and secondary mortgage market, tightening the underwriting standards, qualifications and loan size limits.

In this analysis emphasis will be put on the first, the lock-in effect, because it has been thought by many in the literature ((Clapp, et al. 2001), (Chan 2001), (Engelhardt 2003)) to dominate over the other effects listed above; and also because the impact of this effect has spread across all US regions after the housing bubble burst. Figure 3, illustrating the Case Schiller House Price Indices during the last two decades in 10 major metropolitan areas, shows massive declines in home prices starting with 2007. These declines induced numerous situations in which the outstanding debt of the borrower exceeded the value of the property serving as collateral, hence creating incentives to rather default than to continue paying the monthly mortgage payments. Also, having a high loan-to-value ratio (large outstanding debt to low house price) reduces significantly the possibility to qualify for a new loan, locking in the borrower into its current mortgage contract, and needlessly into the current residence.

Besides the lock-in effect of declining house prices I will also analyze various household level factors which affect the decisions of terminating a mortgage, such as the income of the family, the characteristics of the primary – main – loan and even some demographical attributes

[6] The cash-out effect, or home equity extraction, has attracted the attention of the Federal Reserve Board for its possible impact on consumer spending (Greenspan and Kennedy 2005).

or recent changes in the family structure. Three subgroups of households have been selected for the empirical analysis to be able to see the relative magnitude of the marginal effects of these factors on the probability of choosing to refinance over keeping the current mortgage, or on the probability to move over to stay in the same house. Termination probabilities are estimated first separately, with a probit model, then also jointly with a multinomial probit model believing that the decisions of refinancing, moving or staying in current-state may not be independent of each other. Most of the similar works have been using multinomial logistic (Cunningham and Capone 1990) or simple logistic regressions (Archer, Ling and McGill 1996) to model the termination probabilities, but as argued in the next sections the multinomial probit might be a better method to estimate these seemingly dependent outcomes. The binary probit models are estimated mainly for benchmarking reasons, and also to bring justification on why the multinomial model is a better choice. At the end I compare the estimation results with Campbell's (2006, p.1582-1583) findings, who modeled refinancing and moving probabilities with simple probit regressions in the 2001-2003 period, and with the results presented by Ferreira, Gyourko and Tracy (2008) who estimated mobility rates using a twenty years long AHS panel data (1985-2005).

This study uses data from a time period which hasn't been explored yet by many others, and finds evidence of the so-called lock-in effect of a greater magnitude than other articles in this field. Also presents significant marginal effects of variables, like constraint caused by liquidity problems, which previously, in the articles used as comparisons, yielded insignificant coefficient estimates. The remainder of the paper is organized as follows. Section I presents the conceptual framework in which the empirical models can be built, Section II describes the econometric tools and methods used to predict the probability of the different mortgage terminations, while Section III summarizes the data and sampling methods. The estimation results are discussed in Section IV, with concluding remarks presented in the last section.

I. CONCEPTUAL FRAMEWORK

1.1. OPTION-EMBEDDED MORTGAGE CONTRACTS

In the theoretical literature researchers like Hendershott and Van Order (1987) or Kau and Keenan (1995) have applied option-pricing methodologies to value mortgages, because a standard fixed-rate mortgage contract behaves actually as an annuity and can be viewed as a callable puttable amortizing bond, with two options embedded: the option to prepay and the option to default. The option to default is considered equivalent to a put option, allowing the termination of the debt by "selling" the right of ownership of the collateral to the lender.

The option of the borrower to prepay or to refinance is often modeled as a call option, because basically it allows the borrower (the buyer of the call option) to repay (to buy the underlying instrument of the option) to the lender (the "writer" of the option) the total amount of the principal outstanding of the mortgage at a certain time. Usually the borrower pays a fee for this right, either at the moment of the origination or when the option has to be exercised. The strike prices of these options are both equal to the unpaid mortgages balances. The call option is considered to be in-the-money (ITM) when the interest rate differential between the mortgage rate at origination and the prevailing volatile rate on the market is large enough to cover the transaction costs and the value of the lost call option of not being able to be exercised at a future time. Defining the efficient magnitude of this interest rate differential is non-trivial and we can find many attempts for this in the literature. Agarwal et al. (2006) estimate the spread that justifies refinancing at 1.1-1.4% for mortgages of around $140,000 to $200,000 in size, while Archer, Ling and McGill (1996) approximate the value of the call option of refinancing by the ratio of the

market value to the book value of the primary mortgages, considering the option in-the-money if this ratio exceeds one[7].

Using the ratio proposed by Archer et al. was not feasible in this study, because of the little variation across the sample – almost 90% of the households in each subsample was considered to be in the money. Even though the data is described only in the forthcoming section, we have to be aware of its limitations since the beginning, and this is one of the cases. For instance, there are many observations where the self-reported interest, the monthly payments, the term of the mortgage and the principal simply didn't "add up", meaning that either the reported mortgage rate was too low given the value of the principal, or the payments were too high. Campbell (2006) addresses this issue by correcting for implausibly low self-reported rates. However, using his approach and correcting for rates which are two percentage points below the average prevailing FHLMC rate during the mortgage origination period gives other concerns: the variation across the observations is reduced substantially, as the replaced rates would be the same for many households originating mortgages in the same years. Therefore, to capture the households' transaction costs this study includes the spread between the self-reported rates and the lowest FHLMC rate during 2007-2009 period (4.88%), as an alternative of an in-the-money call option's value. As presented later on, in the empirical findings section, the spread factor will have a high explanatory power in the case of refinancers, but will not influence the probability of moving.

[7] If this ratio exceeds one means that the discount rate (the prevailing current market rate) used to calculate the market value is less than the rate at origination which has been used as a discount rate to calculate the book value.

1.2. THE HIERARCHY OF CHOICES AND THE PROBABILITY OF A TERMINATION

The decision to terminate a mortgage involves actually three events: prepayment, default and moving. Prepayment generally is defined as an early payoff[8] since the loan is being paid off before its maturity. Refinancing (to get lower monthly payments, to shorten duration, to cash-out or for other reasons presented in Table 1) is just one type of early payoff. Some categorize the moving event as an early payoff too, because in the U.S. mortgages cannot be assumed by the new owners of the property, hence must be refinanced when a borrower moves. Using the AHS data we can clearly distinct the two events from each other, therefore refinancing and moving will be modeled as separate outcomes. The default event is defined as the event when a borrower loses the title of ownership over the property due to failure to make the mortgage payments. This event, however, cannot be observed directly from the AHS surveys as we don't get information about the reason which triggered a move, we can only observe that the household is present in the 2007 survey but in 2009 a different household responded to the interview questions.

This limitation imposed by the data needlessly drove this study into the direction of modeling default and moving as one single outcome, over two other alternatives: refinancing for economic incentives and staying in current situation (not moving and not refinancing). The dependency between the probability of moving and default or that of moving and refinancing is unquestionable then. For this reason the probability of these separate outcomes will be modeled simultaneously, with a multinomial probit specification, using borrower, mortgage and house related variables.

To sustain the validity of modeling default and moving as a single outcome, as well the usage of a multinomial model over a simple binary outcome model, I will present the probability

[8] There is another type of prepayment, the curtailment, which is not being considered here due to its relatively insignificant share in value compared to the early payoffs. Curtailment happens when the borrower pays over the scheduled amount but without totally paying back its debt.

model for mortgage terminations derived by Archer et al. (1996). They described the total probability of mortgage terminations for all reasons in time t as an equation of the hierarchy of choices:

$$\lambda_{T_t} = \lambda_{D_t} + (1 - \lambda_{D_t}) [\lambda_{M_t} + (1 - \lambda_{M_t})\lambda_{P.NM_t}] \quad (1)$$

where: λ_{D_t} is the probability of default at time t;

λ_{M_t} is the probability of moving – hence terminating at time t;

$\lambda_{P.NM_t}$ is the probability of terminating the mortgage contract by prepayment conditional on not moving[9].

Where the probability of moving is dependent on the vector of household, mortgage and property characteristics (x_t), the vector of potential income effects gained with the move (y_t) and the value of the option to prepay an in-the-money call or to preserve the option for a later time if interest rates have gone up since the origination (C_t):

$$\lambda_{M_t} = \lambda_{M_t}(x_t, y_t, C_t) \quad (2)$$

Because the data in use doesn't allow for measuring potential income effects in the case of moving, the probability of moving λ_{M_t} will just depend on the vector of household and property characteristics (x_t) and on the value of the call option, C_t. The value of the call option, the option to prepay, can be described in general as:

$$C_t = C(x_t, IRspread_t, P, \lambda_{t+1}) \quad (3)$$

where: x_t is a vector of household, mortgage and property characteristics;

[9] The borrower can prepay and neither default or move, or can move but not default while when defaulting it has to both move and prepay.

$IRspread_t$ is the spread between the current prevailing market rate for similar mortgage contracts and the origination rate;

P is a vector of parameters influencing the interest rate movements, and are considered as same across all households;

λ_{t+1} is a vector of the moving probabilities from time t+1 to the end of the mortgage term.

For simplicity and because of the limitations imposed by the data, in this study C_t depends only on the $IRspread_t$. This means that the probability to prepay conditional on not moving here can be stated as:

$$\lambda_{P.NM_t} = \lambda_{P.NM_t}(C_t, x_t) = \lambda_{P.NM_t}(IRspread_t, x_t) \quad (4)$$

And the probability to move is also

$$\lambda_{M_t} = \lambda_{M_t}(C_t, x_t) = \lambda_{M_t}(IRspread_t, x_t) \quad (5)$$

where x_t , C_t and $IRspread_t$ are the same as above. Equations (4) and (5) are estimated with a simple binary choice probit model. Archer et al. only focuses on the probability of prepayment, arguing that if moving behavior would also be considered the interdependency between the moving equation (2) and the call option equation (3) have to be confronted. C_t and λ_{M_t} are arguments which are present in the equations of one another so they can be logically determined only by using simultaneous equations. To address the difficulty posed by this interdependency, various articles treated the termination choices as a nested or a multinomial probability alternative, like Cunningham and Capone (1990), Philips, Rosenblatt and Vanderhoff (1996) or Clapp, et al. (2001). As presented in the next section, this study uses multinomial probit for modeling the interdependent mortgage terminations.

II. ECONOMETRIC MODELS FOR THE PROBABILITY OF MORTGAGE TERMINATIONS

This section presents the general setup of the empirical models which will be used to estimate the probability of the two observed termination decisions (moving or prepaying) versus the base outcome – not moving and not prepaying. For benchmarking reasons these will be estimated first separately, as binary choice models, then also interdependently, as a multinomial choice model. Estimating separate binary choice models treats the two decisions as independent, which is of course wrong. As argued previously, the need for a multinomial probability model arises from the fact that these decisions might not be fully independent of each other; hence the probabilities estimated solely as binary outcomes raise concerns of biasedness. To estimate the interdependent probabilities of the outcomes, an alternative-specific multinomial probit model should be used. However, given the data, there is no variable available that can be considered fully alternative specific and for situations like this the next best choice is to use the multinomial probit setup.

2.1. THE BINARY CHOICE MODELS

In terms of individually modeling household relocations or refinancing this study will not use different estimation methods for the two different decisions. It is important to emphasize, that these binary choice models will only serve as benchmarks and the main goal is to capture which factors affect the probabilities of the positive outcomes. The magnitude of the estimated coefficients will be treated with caution, rather focusing more on the signs of the coefficients.

Let's consider both refinancing and moving as a positive outcome of a decision. For each household we assume that this decision is based on comparing the indirect utility associated with deciding on a change over the current situation – not moving and not refinancing. A decision

takes place if the monetized value of the indirect gain exceeds the transaction costs involved with the decision. These transaction costs obviously are different, depending of which outcome we consider: when refinancing, usually households face a one-time prepayment penalty which has either been incorporated at the beginning in the price of the mortgage or has to be paid as a lump-sum at the moment of prepayment. The transaction costs of moving can be much greater than a prepayment penalty. First, because mortgage owners have to pay a similar prepayment penalty that refinancers do, then they also face large costs on logistics depending on the size of the family, the location of the new residence, etc.

There are also different motivations behind refinancing and moving. While the AHS data provides answers on why some of the borrowers opted to refinance their primary mortgages, the motivation behind a moving remains unknown, since we don't have any information about those who moved after the 2007 interview period. Table 1 contains information about the share of the different reasons for refinancing, as answered by those who prepaid their primary mortgages between 2007 and 2009. It is interesting to see that the majority indicated the reason for their decision as to get a lower interest, while some simply wanted to lower their monthly payments. Surprisingly, there are a few who refinanced to extract cash out of their equity, which has been thought to be almost impossible after the housing bubble burst.

On the other hand, the motivations for moving can be job-related such as reducing the commute time to the working place, or taking a new job on a different labor market; can be quality-of-life related, when a household moves because there is a desire for a different type or size of property, different set of neighborhood amenities, or different set of natural and cultural amenities. Defaulting and foreclosure is also a possible and non-negligible reason why a household cannot be found in the two consecutive samples.

All these motivations and factors involved in the decision making process can be summarized by a latent index U_i^*, which captures the net change in indirect utilities less the transaction costs of the decision. This latent index is a solution to the omitted costs of moving from equation (2). Household i is assumed to decide on refinancing/moving when this index is positive, and to remain in the current state otherwise.

The linear specification of U_i^* is simply:

$$U_i^* = X_i\beta + \varepsilon_i \quad (5)$$

where X_i is the vector of observed factors that affect household's decisions and ε_i is a random error term, which here is assumed to have a normal distribution. The U_i indicator can be observed for each household and it is set up to take a value of one if the household moves/prepays or zero otherwise. So the binary outcome dependent variable can be written like:

$$U_i = \begin{matrix} 1\ if\ U_i^* > 0, & positive\ outcome - household\ refinances/prepays \\ 0\ otherwise, & household\ does\ not\ move\ nor\ prepays \end{matrix} \quad (6)$$

The probability P_i that a household refinances, or alternatively that moves, can be characterized as follows:

$$P_i = \Pr(U_i^* > 0) = \Pr(X_i\beta + \varepsilon_i > 0) \quad (7)$$

This probability is the same as $\lambda_{P.NM_t}$(the lambda probability of prepaying conditional on not moving) in equation (4), as well as stands for λ_{M_t} (the lambda probability of moving) from equation (5). The vector of coefficients (β) will be estimated with a binomial probit model using the individual data we have on the observed factors (X_i) and outcome (U_i). The variables used in

vector X_i represent different characteristics of the borrower household, the mortgage contract, the property and will be presented in the forthcoming chapter.

2.2. THE MULTINOMIAL CHOICE MODEL

As mentioned earlier, mortgage terminations have different causes but they should not be considered as independent of each other. To model one without the other can result in a biased view of the whole set of choices and associated probabilities. To address this dependency among choices many articles, like Philips, Rosenblatt and Vanderhoff (1996), Cunningham and Capone (1990), Clapp, et al. (2001) and others, use the classical multinomial logit to estimate default and prepayment probabilities. The predominant use of the multinomial logit in analyzing polychotomous observations can be explained by the fact that it is an attractive and easy-to-interpret method. By assuming the stochastic parts of all options to be i.i.d. (independently and identically distributed random variables), the choice probabilities of the logit model have the form of a logistic function, hence the estimated results have a straight forward interpretation.

Quin and Wang (2007), one of the very few works breaking up with the classical approach, uses a multinomial probit model to estimate mortgage terminations and performance. They argue that assuming i.i.d. distribution for the stochastic parts is problematic, as it leads to the i.i.a. (independence of irrelevant alternatives) property[10]. In this study I will follow Quin and choose the multinomial probit over the logit, because of the strength of the i.i.a. property and other restrictions imposed by a multinomial logit model.

The multinomial probit model is used with discrete dependent variables that take more than two outcomes which don't have a natural ordering. The stochastic error terms are assumed

[10] As explained by (Qin and Wang 2007): "Take the famous red bus and blue bus problem for an example. If only travel time matters for travelers and without presence of blue bus, half of travelers choose train and the other half choose red bus. Fit a Logit model for this problem, and use it to predict how many people are going to choose train when there is blue bus available in addition to train and red bus. The prediction would be one third, which contradicts intuition." (p.6)

to have a multivariate normal distribution that is heteroskedastic and correlated. Cameron and Trivedi (2005) describe the multinomial probit model as an m-choice multinomial model, with the latent utility of the *j*th choice given by:

$$U_j = V_j + \varepsilon_j, \quad j = 1, 2, \ldots, m \quad (8)$$

where the errors are joint normally distributed, with

$$\varepsilon \sim \mathcal{N}\,[0, \Sigma], \quad (9)$$

where the $m \times 1$ vector $\varepsilon = [\varepsilon_1 \ldots \varepsilon_m]'$. And usually the $V_j = \boldsymbol{x}'_j\beta$ or $V_j = \boldsymbol{x}'\beta_j$. The greatest challenge in estimating a multinomial probit model lies in the fact the there is no closed-form expression for the choice probabilities as an *m*-choice model requires a numerical evaluation of an (*m*-1)-variate integral. For a three-choice MNP model, which is being estimated in the case of this study, the individual choice probabilities can be expressed like:

$$p1 = \Pr[y = 1] = \int_{-\infty}^{-\tilde{V}_{31}} \int_{-\infty}^{-\tilde{V}_{21}} f(\tilde{\varepsilon}_{21}, \tilde{\varepsilon}_{31}) d\tilde{\varepsilon}_{21} d\tilde{\varepsilon}_{31} \quad (10)$$

where $f(\tilde{\varepsilon}_{21}, \tilde{\varepsilon}_{31})$ is a bivariate normal with as many as two free covariance parameters and $\tilde{V}_{21}$ and $\tilde{V}_{31}$depend on regressors and parameters β. Here $p1$ is equivalent to equation (1), specifically λ_{T_t} conditional on $\lambda_{P.NM_t} = 1$, alternatively conditional on $\lambda_{M_t} = 1$.

For larger models than a four-choice MNP model the alternative is to use simulation methods, as the trivariate normal integral is the limit for numerical methods. However, since the computational tools and statistical packages are so elevated we can estimate with a very small time input even larger models, the challenge is more in the correct evaluation and explanation of the results.

III. DATA

3.1. THE AMERICAN HOUSING SURVEY

The data used for the analysis is from the 2007 and 2009 American Housing Survey (AHS) conducted by the U.S. Census Bureau for the Department of Housing and Urban Development (HUD). The AHS is the largest, regular housing sample survey in the United States, collecting national data every other year from a fixed sample of about 50,000 homes plus new constructions. It contains extensive micro-level data on apartments, single-family homes, mobile homes, vacant homes, family composition and demographic characteristics, income, housing and neighborhood quality, housing costs, equipment, fuels, size of housing unit, recent movers, etc. Considering all other possible sources, this dataset fits best for the purposes of the analysis because also includes detailed information about mortgages. Interviewers ask specific questions, like the number and amount of mortgages a household owns, mortgage interest rates and payments, origination dates, term structures or motives for previous refinancing decisions. Combining these with the demographic characteristics of the occupants, one can get the largest, publicly available sample size for a micro level analysis of mortgage related behavior, conducted biannually since 1985. Thus, the AHS is a panel data set with a long time dimension, out of which the last two available survey years will be used.

Along many benefits, the AHS also presents some weaknesses. The main problem is on the data quality side, as participation in the interviews is voluntary, and respondents may refuse the interview or give just partial answers. The mortgage data is especially sensitive in this matter, as a detailed review of Lam and Kaul (2003) revealed: they show that as mortgages age the respondents reports about the original principal value, monthly payments or interest rates differ from one survey to the next. This issue is overcome here by omitting those observations which

presented inconsistent answers in the two consecutive years, but it can still cause inefficient coefficient estimates.

Another challenging thing about the AHS survey is that it follows housing units rather than occupants. Hence the observations from 2007 and 2009 needed to be matched following a whole set of strict criteria to ensure that the data set of non-terminators and refinancers are the same households in both years. Although at the beginning there was a relatively large sample size (65,419 observations), after applying the filters for a more robust analysis, the usable sample gets diminished to almost its tenth (6,691 obs.). The sample selection process is briefly described in the following part.

3.2. SAMPLE

As mentioned above, the individual households can be identified only by matching the unique control numbers assigned to each housing unit. Based on these control ids there are 65,419 housing units which were present in both years' datasets, but only 27,864 households are surely the same. The rest were excluded because either did not indicate that at least some of the same household members lived in the unit at the 2007 survey; or reported a move prior to the 2009 survey. For stability purposes several other checks have been conducted to compare personal level characteristics, such as age of householder should have changed by two, while race and gender of householder shouldn't have changed. Out of these households only 14,100 had a mortgage in 2007 and 12,069 were still traceable in the 2009 mortgage sample.

Since the aim of this analysis is to explain individual mortgage refinancing behavior within this two year time period, three groups had to be identified in the data: those who refinanced, those who moved (and needlessly refinanced) and those who served as a base group of comparison for these – the non-movers and non-refinancers (simply referred as non-terminators in the following). The household has been considered a refinancer if the origination year of the

primary mortgage reported in 2007 is less or equal to 2007 while the one reported in 2009 is after 2007; also, if the purchase year of the residence is the same in both surveys. Those households were considered as movers which were present in the 2007 dataset, but in the 2009 survey the interviewed occupants reported to be different households. The base group of non-terminators was carefully selected, having the same mortgage and demographical characteristics in both years, hence diminishing somewhat the above mentioned shortcoming of the AHS data quality.

For computational and robustness reasons several observations were omitted, which didn't confirm to the following: the mortgage contract has to be a 15, 20 or 30 year fixed rate, single family mortgage, with an original principal between $10,000 and $934,140[11], originated after 1989, with a property value serving as collateral between $35,000 and $1 million (as reported in 2009) and having a monthly mortgage payment not less than $100. This way those who have very specific contracts or who could be considered outliers in any of these terms were eliminated. Also, those who had very small remaining balances were omitted, because they would have terminated due natural and obvious reasons, coming to the end of their loans.

Finally, the sample used in this analysis is consisted of 6,691 observations. There are 5,014 non-mover and non-refinancer households; 874 who refinanced between 2007 and 2009; and 803 who moved after the 2007 survey. The descriptive statistics for the whole sample as well as for each group of outcomes can be found in Table 2.

3.3. VARIABLES

In the following section, after a general overview, the emphasis will be put on those variables which were derived from the raw survey information. These may need special explanations to be fully interpretable and understandable when analyzing the empirical results presented in the next chapter. Generally, most of the variables were measured at the 2007 survey

[11] This value has been set as a topcoded value by the AHS.

date except the dependent binomial variables – whether or not the household refinanced, moved or continued the monthly mortgage payments during the selected period. This way controlling for time zero is already achieved, and there is no need to include special variables for time fixed effects. As mentioned previously, Table 2 contains the descriptive statistics on all the variables considered; as well Appendix presents the whole set of formulas used to calculate the different indicators and financially relevant variables.

The two most important variables in this work are LIQCONSTR and COLLCONSTR. They serve as indicators (dummies) for income and collateral constraints and were first presented by Archer et al. (1996)[12]. According to their approach, a household is considered collateral constrained if the Loan-to-Value ratio of its debt exceeds 90 percent, while liquidity constrained are those whose yearly mortgage payments over their total income would exceed a threshold of .28. This study uses the same thresholds in constructing these indicators.

The indicator COLLCONSTR will serve to measure the lock-in effect of the declining house prices in 2009. It depends on the loan-to-value ratio (LTV), one of the key risk factors considered by lenders when analyzing a borrower's qualifications for a mortgage. Typically, assessments with high LTV ratios are considered riskier; therefore if the mortgage gets accepted the loan will usually cost more to the borrower. Officially it is calculated as the ratio of the mortgage amount over the appraised value of the property. Given the data, here we can calculate it as the ratio of the total outstanding debt of the household over the current self-reported value of the property.

The total outstanding debt is the sum of the book value of the primary mortgage and the book value of all other existing mortgages and home equity loans on the collateral. If we account

[12] Despite the long time since their article has been published, many researchers use the same indicators as a best practice to capture such effects. Campbell (2006) uses dummies for loan and income problems derived by the same methodology, while Ferreira et al. (2008) uses a similar indicator for capturing the effects of negative home equity.

for more than one mortgage loan, then we have to realize that these post-origination "junior liens" raise the contemporaneous LTV ratio and reduce borrower equity. This is similar to a company's new stocks issuance, and it is labeled as "equity dilution".

The value of the property (HOUSEVAL07 and HOUSEVAL09), as subjectively reported by the interviewed household, can be considered to be an accurate measure on average of the current fair market value of the residence. Archer, Ling and McGill (1996) argue that market value estimates by households may be more important for a behavioral analysis of termination decisions than „true" market values, since the average U.S. house owner generally overestimates its house value. However, this error is usually uncorrelated with property characteristics; rather it has to do with plans of moving in the near future or cashing-out from the equity.

The book value of a mortgage (BOOKVAL) is the original face value[13] of the mortgage less the amount of principal repayment, or the mortgage amount outstanding at a particular point in time. Here, BOOKVAL has been calculated as the future value of the principal (in 2007) less the future value of the payments made between origination and 2007.

The LIQCONSTR indicator is one if the total mortgage payments in 2007 exceeded the 28 percent of the total family income reported for previous year. Originally this indicator was calculated based on yearly mortgage payments that would be required by a refinanced 30 year fixed rate mortgage contract (at the prevailing market rate). Here instead, the actual, self-reported mortgage payments are considered, allowing variation both across mortgage payment burdens and total income.The value of the call option (C_t) as presented in the Conceptual Framework here depends only on the interest rate spread (SPREAD), calculated as the difference between the self-reported original interest rate and the lowest FHMLC rate between 2007 and 2009, which was 4.88 percent.

[13] At origination, the book value and face value are the same.

IV. EMPIRICAL RESULTS

4.1. DESCRIPTIVE STATISTICS

Before starting to analyze the estimated coefficients and marginal effects obtained with the two regression methods, I will present some of the most relevant differences in the descriptive statistics of the three subsamples, as summarized in Table 2. We can see that more than 20% of the whole sample households were considered collateral constrained using their self-reported home values in 2009 as an approximation of the value of their collateral. However, there is a striking difference in this share if we compare those who didn't choose to refinance, nor to move, and those who refinanced: 22.5% of the non-terminating households were facing the burdens of the depressed values of their homes serving as collaterals, compared to "just" the 13.6% of the refinancers. This clearly indicates that there were significant differences among those who filed for a new mortgage during the 2007-2009 period and those who stayed away from the opportunity which arise on the market. The lock-in effect caused by declining home prices is obvious then: the average Loan-to-Value ratio is the largest for the non-terminators, who did not meet the requirements, imposed by the lenders, and could not cheaply refinance their loans.

The liquidity problem, indicating whether a household had to spend more than 28% of its total yearly income on mortgage payments, doesn't seem to be so pronounced across the whole sample, but it affects a significantly larger share of those who opted to move versus those who refinanced. This means that the movers were facing more pronouncedly diminishing yearly incomes, which is in line with the larger unemployment rates we could see for the respective regions. When searching for explanations on why they have decided to move, we can think of diminishing yearly incomes due to unemployment or worsening business environment, and that they could not bare any longer the payment burdens of their mortgages. Other notable

differences across the three subsamples are found at those variables which indicate the wealth of a household: those who refinanced have reported considerably higher property values in both survey years, have had larger outstanding debt balances and earned significantly more money than the rest of the families. The average of years spent in residence is also much higher for the refinancers, eight and half years versus 5.8 years of the whole sample. This could indicate that especially those were interested in refinancing who got their mortgages around 2000, when rates were at record heights. Figure 2 shows that the 30 year Freddie Mac (FHLMC) commitment rate for fixed mortgages was more than 8% in mid-2000, while by the end of the analyzed period this went to a low of 5%.

4.2. ESTIMATION RESULTS

Table 3 presents the final estimation results of the binary choice probability models (Models (1) and (2)), as well as the estimated results for the multinomial choice probability model (Model (3)). Using an alternative specific multinomial probit estimation method without defining actual alternative specific variables, theoretically, one should get similar estimates as using the multinomial probit. Table 5 summarizes the estimated marginal effects using the asmprobit command with the same list of variables and sample data. The results of the variables of interest don't change significantly in their magnitudes and signs, leading to the conclusion, that the usage of the multinomial probit is a valid. Also, as presented in Table 3, the explanatory variables resulted estimated coefficients which are slightly higher in magnitudes in the case of the multinomial probit, indicating that the simple binary choice models might have been biased towards zero. Otherwise the signs and significance levels remain unchanged compared to the benchmark models, while the marginal effects decrease in their magnitudes, as well as the base probabilities of the reference households to terminate their mortgages either way. The columns headed "Probability Estimates" report the probability for the reference household and the change in this probability caused by a unit change in a binary variable and a one-standard-

deviation change in a continuous variable. All regressions include region and time fixed effects, control for gender, marital status and house location. In the reference household the household head is a married, non-white male living in an urban area.

The most important findings of this paper are related to the constraint indicators on collateral and liquidity and will be presented based on the results in Model (3). The reference household has been estimated to have a base probability for refinancing of 13.62 percent, while for moving 11.34 percent. Looking at the estimated marginal effects, the collateral constrained indicator has the highest impact for both outcomes simultaneously. Holding BOOKVAL and HOUSEVAL07 constant, households with higher loan-to-value ratios than 90% have 6.44 percentage points lower probability to refinance, respectively 3.96 percentage points lower probability to move. This means that taking two households having the same amount of debt outstanding and the same house values in 2007, the one which faced a larger decrease in its property value until 2009 will refinance with a lower probability. Also, this borrower will less likely decide or have the possibility to move. These results are the most important proofs to sustain the validity of the lock-in effect caused by declining property prices. The probability estimates represent a decline of 47.28 percent in the average refinancing rate and a 34.92% decline in the average mobility rate. These relatively large net declines imply that during the credit crisis (2007-09) the lock-in effect dominated the impact of other factors, such as liquidity problems or incentives to refinance given the large interest rate spreads.

Holding constant the book value of the mortgage (BOOKVAL), the estimated marginal effects of LIQCONSTR show that if a household passes the .28 threshold in its yearly payment per income ratio, the probability to refinance decreases by 3.3 percentage points, while the probability to move (possibly to default) increases with 3.5 percentage points. The alternative signs for this explanatory variable can be explained solely by variation in incomes: substituting LIQCONSTR with the log of total income, holding BOOKVAL constant, households with

higher total yearly incomes would refinance with a higher probability and would decide to move with a lower probability. This finding indicates, that some households were "forced" to move and terminate their mortgages because a decrease in their yearly incomes.

Further results show that one standard deviation change in the interest rate spread is estimated to increase the probability of refinancing by 1.7 percentage points, holding all household specific variables constant, but doesn't affect significantly the decision of a movement. Also, according to the estimation results, the demographic variables WHITE and HHDIVORCED have a large positive effect on the refinancing base rates, not affecting significantly though mobility. A positive change in the household's size increases the probability of moving nearly as much as the lock-in effect decreases it, with more than 3.6 percentage points. However the residence being situated in the rural area seems to be the largest impediment of a moving decision, diminishing the base probability estimate with more than 4 percentage points. This accounts for a net decrease in moving probability of as much as 36.07% compared to the base reference households living in an urban area, and having an average probability of moving of 11.34%.

4.3. COMPARING ESTIMATED MARGINAL EFFECTS WITH SIMILAR WORKS

Table 4 summarizes the results of two similar articles ((Campbell 2006) and (Ferreira, Gyourko and Tracy 2008)) to help comparing the estimated signs and magnitudes of the coefficients of most important explanatory variables in this study – the collateral constraint and the liquidity problem indicator. These specific articles were selected, because they use two different time periods which might be interesting in comparison of the time interval used in this study.

Campbell (2006, p.1582-1583), using AHS data from 2001 to 2003, shows a higher base probability for refinancing and a much lower mobility rate for the reference households. The higher average refinancing probability is simple to explain: between 2001 and 2003 the house prices were rapidly climbing and mortgage commitment rates were continuously declining, hence creating numerous opportunities for borrowers to refinance their old mortgages. On the other hand, the much lower reference mobility rate may be explained with the absence of the defaulters from the sample of movers, since default rates in that period were extremely low. The variables of interest enter in the equation with the predicted negative signs, but only the loan (collateral constraint) problem indicator is economically or statistically significant. Campbell argues that the weakness of these effects may be due to relaxed standards for mortgage lending in the years prior the refinancing boom; or that the rise of house prices has reduced the fraction of households with insufficient home equity (high LTV ratios).

Ferreira, Gyourko and Tracy (2008) concentrated on the probability of moving using a twenty year long AHS panel data, from 1985 to 2005. They found that over this period the average mobility rate was around 12 percent, and that households who have negative equity positions in their house have two-year mobility rates that are 5.6 percentage point lower than similar households with positive equity in their house. Regarding the income effect, they also found that a negative change in the family income increases the probability of moving by almost 1 percentage points. They conclude that over the past two decades the lock-in effect for owner occupied borrowers dominated the impact of foreclosures induced by negative equity.

Compared to these studies, the results obtained from the 2007-2009 period show that constraints on collateral and income simultaneously played a major role in the diminishing termination rates. The lock-in effect causes higher net declines in the average two-year probability of refinancing than ever before, while the households' liquidity problems seem to finally have explanatory power on the probability of mortgage terminations.

CONCLUSION

In this study I estimate mortgage probability terminations after the recent housing market downturn in the United States, using a two period panel data of the American Housing Survey from 2007 respectively 2009. The household level observations allow for identifying three possible outcomes regarding mortgage decisions: refinancing, defined as early payoff of the outstanding loans; moving, defined as either terminating by default or by prepaying the old debt and moving to a different home; and as a base outcome – staying in the current mortgage setup and residence. After showing, that these three outcomes are not fully independent from each other, I estimate the mortgage termination probabilities by using a multivariate probit model, allowing for inter-choice correlation of the unobservables.

The most important findings are related to the effect of the declining house prices within the analyzed period, proving that households which suffered major home equity losses have refinanced or moved with a significantly lower probability. This phenomenon of being constrained by diminishing collateral value is called lock-in effect, as it ties borrowers to their current mortgages, not being able to qualify for a new, more favorable loan. A more interesting and somewhat unexpected result is that of the liquidity constraint – or income problem, which proved to significantly diminish the probability of refinancing but increased the probability of moving by almost 4%. This confirms what has been argued in the first section that in this specific data sample defaults might have triggered a part of the mobility. The other results support the option based theory, which states that households will refinance with a higher probability if there is a higher spread between the prevailing market rate and the rate at origination.

The timeliness of this topic in unquestionable because substantially lower household mobility is likely to have various social costs, including poorer labor market matches, diminished support of local public goods and lesser maintenance or reinvestment in the homes. Unexplored

refinancing opportunities will force households to spend a large fraction of their income on mortgage payments for a longer time, which could generally affect consumption, schooling decisions for children or future investment decisions.

BIBLIOGRAPHY

Agarwal, Sumit, John C. Driscoll, and David I. Laibson. "When should borrowers refinance their mortgages?" *Working paper, Brown University and Harvard University*, 2006.

Archer, Wayne R., David C. Ling, and Gary A. McGill. "The effect of income and collateral contstraints on residential mortgage terminations." *Regional Science and Urban Economics* 26 (1996): 235-61.

Cameron, Colin A., and Pravin K. Trivedi. *Microeconomics - Methods and Applications*. Cambridge: Cambridge University Press, 2005.

Campbell, John Y. "Household Finance (A Presidential Address)." *The Journal of Finance*, August 2006: 1553-1604.

Caplin, Andrew, Charles Freeman, and Joseph Tracy. "Collateral damage: Refinancing constraints and regional recessions." *Journal of Money, Credit and Banking 29*, 1997: 496-516.

Chan, Sewin. "Spatial lock-in: Do falling house prices constrain residential mobility?" *Journal of Urban Economics 49(3)*, 2001: 567-586.

Clapp, John M., Gerson M. Goldberg, John P. Harding, and Michael LaCour-Little. "Movers and Shuckers: Interdependent Prepayment Decisions." *Real Estate Economics* 29, no. 3 (2001): 411-450.

Cunningham, Donald F., and Charles A. Capone. "The relative termination experience of adjustable to fixed-rate mortgages." *The Journal of Finance* XLV, no. 5 (1990): 1687-1703.

Engelhardt, Gary V. "Nominal loss aversion, housing equity constraints, and household mobility: Evidence from the United States." *Journal of Urban Economics 53(1)*, 2003: 171-195.

Ferreira, Fernando, Joseph Gyourko, and Joseph Tracy. "Housing Busts and Household Mobility." *NBER Working Paper Series, w14310*, 2008.

Greenspan, Alan, and James Kennedy. "Estimates of home mortgage originations, repayments, and debt on one-to-four family residences." *Finance and Economics Discussion Series - 41, Federal Reserve Board*, 2005.

Hendershott, Patric H., and Robert Van Order. "Pricing mortgages: An interpretation of the models and results." *Journal of Financial Services Research*, 1987: 77-111.

Kau, James B., and Donald C. Keenan. "An overview of the option-theoretic pricing of mortgages." *Journal of Housing Research* 6, no. 2 (1995): 217-244.

LaCour-Little, Michael. "Mortgage Termination Risk: A review of the recent literature." *Journal of Real Estate Literature* 16, no. 3 (2008): 295-326.

Lam, Ken, and Bulbul Kaul. *Analysis of Housing Finance Issues Using the Americal Housing Survey (AHS)*. Washington, DC: U.S. Department of Housing and Urban Development, Office of Policy Development and Research, 2003.

Levin, Alexander, and Andrew Davidson. "Prepayment Risk- and Option-Adjusted Valuation of MBS." *The Journal of Portfolio Management* 31, no. 4 (2005): 73-85.

Philips, Richard A., Eric Rosenblatt, and James H. Vanderhoff. "The probability of fixed- and adjustable-rate mortgage termination." *Journal of Real Estate Finance and Economics* 13, no. 2 (1996): 95-104.

Qin, Jun, and Wei Wang. "A Model for Mortgage Termination and Performance." *Countrywide Security Corporation*, 2007.

APPENDIX A: FIGURES AND TABLES

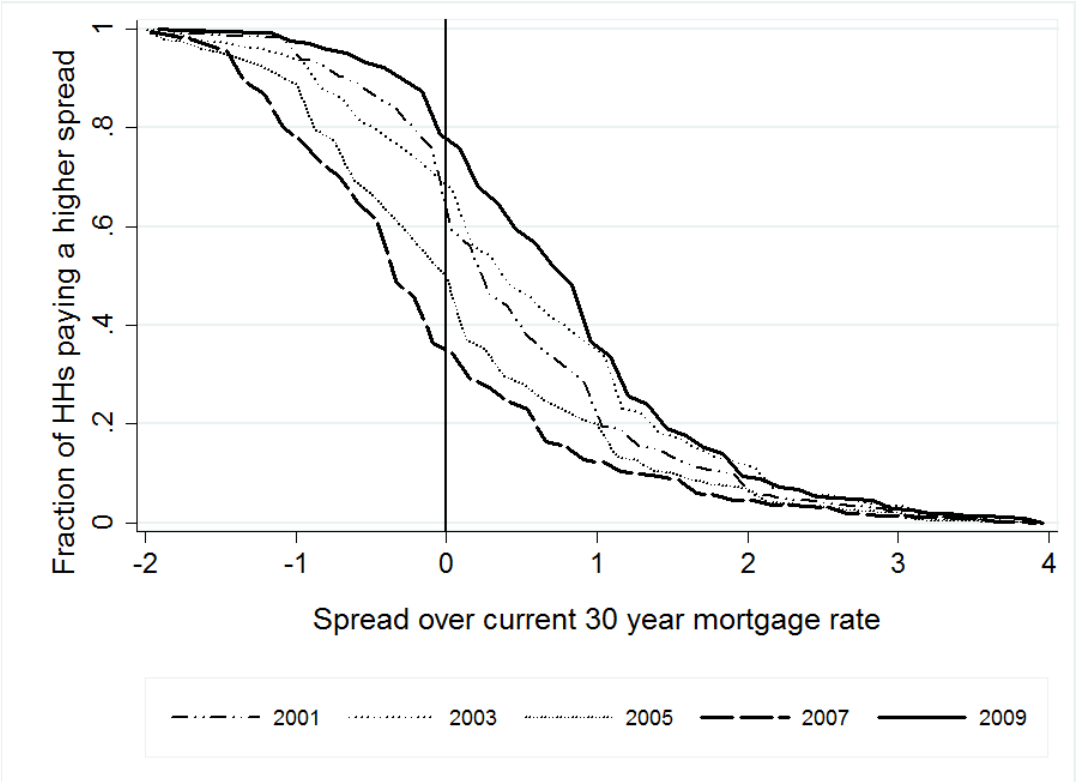

Figure 1. Distribution of self-reported mortgage spreads

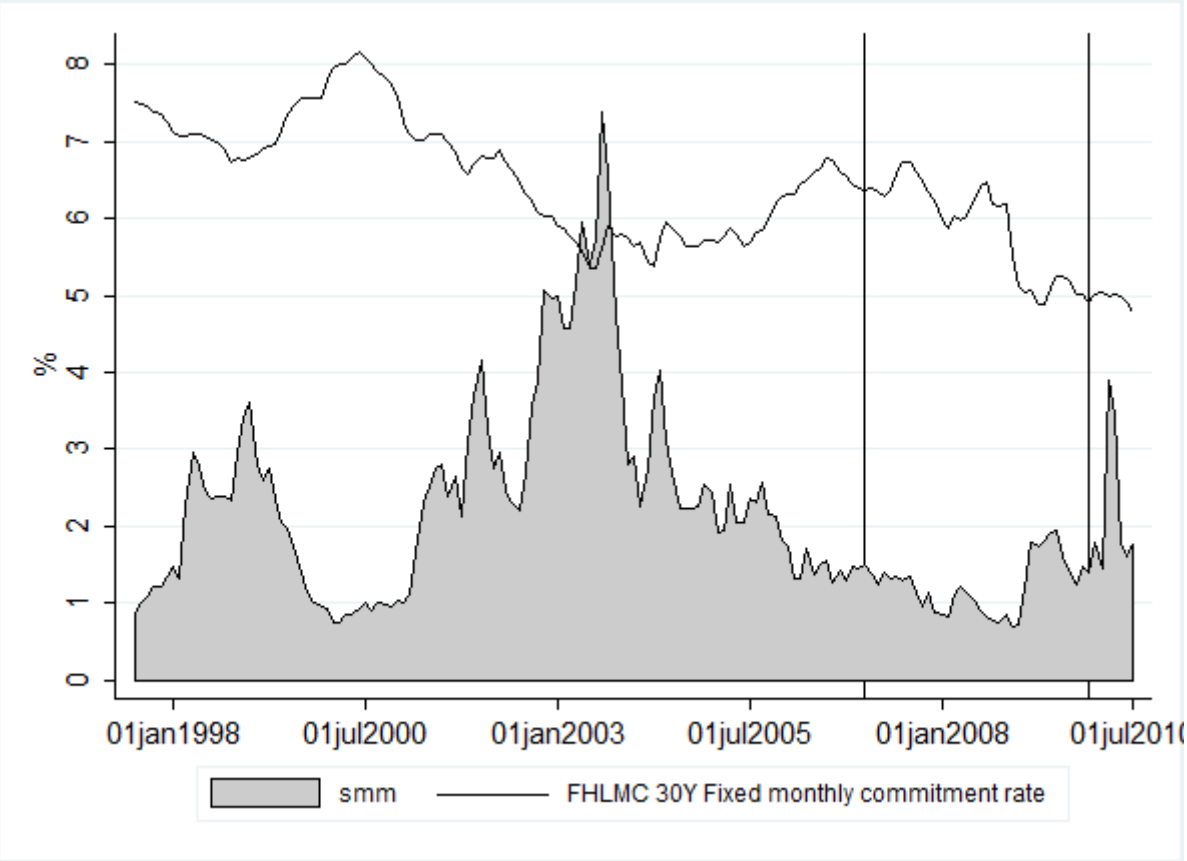

Figure 2. Interest rate environment and average mortgage prepayment rates

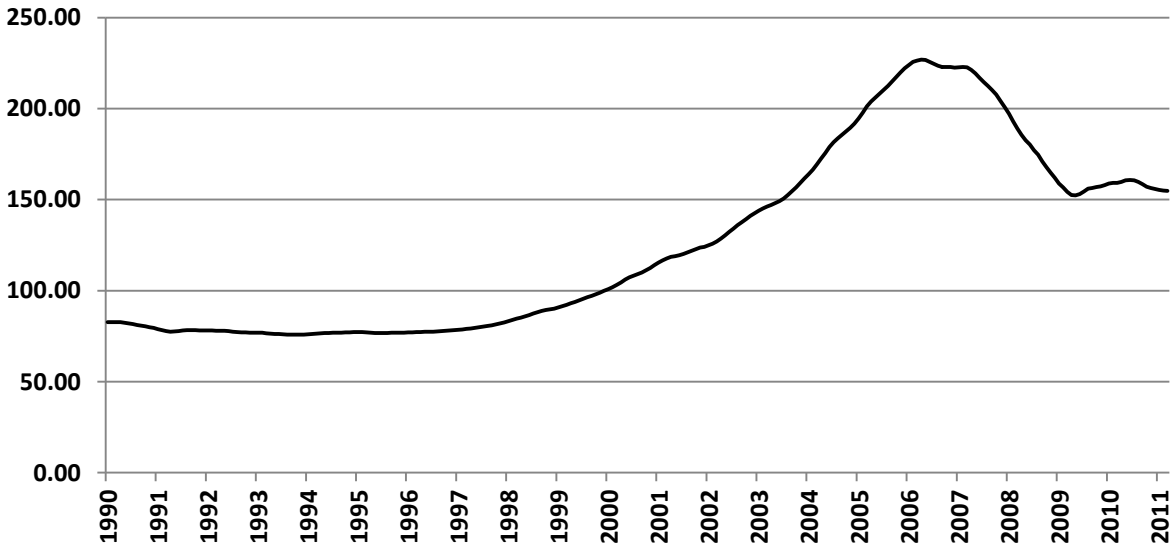

Figure 3. Seasonally adjusted Home Price Index levels(S&P/CaseSchiller indices)

Table 1. Refinancing reasons as answered in 2009 survey

	Refinancers	
Reason for refinancing the primary mortgage	Frequency	% of total answers
to renew or extend the primary mortgage	9	1.03%
to increase payments	16	1.83%
to receive cash	73	8.35%
to reduce monthly payments	134	15.33%
to get lower interest	642	73.46%
Observations	874	

Table 2. Descriptive statistics by decision groups

		Non-terminators		Refinancers		Movers		Whole sample	
		Mean	St.dev	Mean	St.dev	Mean	St.dev	Mean	St.dev
COLLCONSTR	Collateral constrained	0.225		0.136		0.183		0.208	
LIQCONSTR	Liquidity constrained	0.148		0.119		0.181		0.148	
ITM	Call option "In-the-Money" for primary mortgage	0.900		0.875		0.915		0.899	
AMMORT	Amount of first mortgage when acquired ($)	154,494	(94,244)	166,897	(89,762)	152,290	(95,178)	155,850	(93,871)
TERM1	Term of primary mortgage (years)	27.81	(5.079)	27.16	(5.690)	27.85	(5.089)	27.73	(5.168)
RATE1	Self-reported mortgage rate (%)	6.191	(1.271)	6.279	(1.203)	6.269	(1.294)	6.212	(1.265)
SPREAD	Spread btw. self-reported rate and current lowest	1.311	(1.270)	1.399	(1.203)	1.389	(1.294)	1.332	(1.265)
PMT	Monthly mortgage payments ($)	1,175	(720.7)	1,267	(652.1)	1,171	(704.0)	1,187	(710.7)
HOUSEVAL	Self-reported value of the residence in 2007 ($)	261,312	(210,435)	320,219	(248,846)	253,568	(191,209)	268,077	(214,618)
HOUSEVAL09	Self-reported value of the residence in 2009 ($)	224,572	(153,065)	281,959	(171,901)	230,811	(154,498)	232,817	(156,973)
BOOKVAL	Book value of primary mortgage ($)	132,763	(95,958)	147,187	(93,075)	134,336	(102,276)	134,836	(96,475)
LTV	Loan-to-value ratio (including second mortgages)	0.704	(0.547)	0.616	(0.439)	0.683	(0.537)	0.690	(0.534)
MARKETVAL	Market value of outstanding loan ($)	194,071	(126,804)	208,803	(119,788)	195,974	(128,781)	196,224	(126,227)
TOTALINC	Total income ($)	103,739	(98,447)	116,928	(99,577)	104,224	(100,570)	105,520	(98,936)
UNEMP	Difference in unemployment rate	4.659	(0.899)	4.719	(0.884)	4.752	(0.982)	4.678	(0.908)
RURAL	Rural area	0.315		0.327		0.237		0.307	
YEARSINRES	Years spent in residence	5.319	(5.196)	8.568	(7.989)	5.702	(6.009)	5.790	(5.837)
FIRSTTIME	First time buyer	0.398		0.309		0.448		0.392	
SEX	Male=1	0.608		1.364		1.402		1.392	
WHITE	White	0.864		0.907		0.874		0.871	
HHCOL	Householder has at least 2 years of college	0.544		0.590		0.540		0.550	
HHPLUS	Size of household increased	0.180		0.165		0.231		0.184	
HHMIN	Size of household decreased	0.154		0.177		0.157		0.158	
HHDIVORCE	Recently divorced	0.029		0.709		0.602		0.676	
HHMARR	Recently married	0.058		0.0593		0.0818		0.0610	
HHAGE	Age of householder's head	44.00	(12.33)	46.70	(11.62)	41.23	(12.28)	44.02	(12.31)
PER	Number of persons living in household	2.998	(1.472)	2.945	(1.416)	2.761	(1.434)	2.963	(1.462)
	Number of sample observations	5014		874		803		6691	

Table 3. Estimating the probability of mortgage terminations

	(1)		(2)		(3)			
	Whether a household refinanced between 2007-09		Whether a household moved between 2007-2009		Whether a household refinanced between 2007-2009		Whether a household moved between 2007-2009	
Dependent Variables	Coefficients	Probability Estimates	Coefficients	Probability Estimates	Coefficients	Probability Estimates	Coefficients	Probability Estimates
Reference household		0.1519		0.1306		0.1362		0.1134
Collateral constrained	-0.383***	-0.0794***	-0.275***	-0.0530***	-0.522***	-0.0644***	-0.405***	-0.0396***
	(0.065)	(0.0118)	(0.066)	(0.0115)	(0.087)	(0.0107)	(0.088)	(0.0103)
Liquidity constrained	-0.144**	-0.0320**	0.154**	0.0347**	-0.174*	-0.0330***	0.186**	0.0350**
	(0.068)	(0.0142)	(0.066)	(0.0157)	(0.090)	(0.0124)	(0.088)	(0.0140)
Interest rate spread	0.085***	0.0199***	0.016	0.00337	0.110***	0.0170***	0.030	0.000659
	(0.018)	(0.00416)	(0.018)	(0.00390)	(0.024)	(0.00371)	(0.025)	(0.00345)
Ln(Unemployment rate difference)	0.158	0.0372	0.158	0.0336	0.212	0.0269	0.229	0.0263
	(0.114)	(0.0268)	(0.113)	(0.0241)	(0.151)	(0.0237)	(0.151)	(0.0211)
Ln(Book value of primary mortgage)	0.126***	0.0297***	0.077**	0.0164**	0.171***	0.0242***	0.116***	0.0111*
	(0.033)	(0.00772)	(0.033)	(0.00693)	(0.044)	(0.00697)	(0.044)	(0.00613)
White	0.162**	0.0358**	0.076	0.0156	0.236**	0.0326**	0.114	0.00895
	(0.071)	(0.0147)	(0.070)	(0.0140)	(0.095)	(0.0131)	(0.094)	(0.0126)
Rural area	0.063	0.0151	-0.221***	-0.0447***	0.059	0.0190*	-0.281***	-0.0409***
	(0.048)	(0.0116)	(0.053)	(0.0102)	(0.065)	(0.0105)	(0.071)	(0.00890)
Ln(Market value of residence in '07)	0.186***	0.0437***	-0.080**	-0.0169**	0.238***	0.0418***	-0.086*	-0.0205***
	(0.038)	(0.00884)	(0.034)	(0.00722)	(0.051)	(0.00798)	(0.045)	(0.00640)
Positive change in HH size	-0.067	-0.0154	0.168***	0.0379***	-0.075	-0.0195*	0.216***	0.0362***
	(0.059)	(0.0133)	(0.058)	(0.0137)	(0.079)	(0.0116)	(0.077)	(0.0122)
Household head recently divorced	0.337***	0.0922**	0.242*	0.0582	0.451***	0.0707**	0.335*	0.0347
	(0.124)	(0.0384)	(0.135)	(0.0361)	(0.165)	(0.0332)	(0.176)	(0.0291)
Number of observations	4,791	4,791	4,647	4,647	5,415	5,415	5,415	5,415

Standard errors are reported in parentheses; the significance levels are denoted by: *** p<0.01, ** p<0.05, * p<0.1. The first two columns report probit regressions on refinancing and confirmed changes of address btw. 2005-2007. Column (3) reports on multinomial probit regression estimates, when all three outcomes are included simultaneously. The column headed "Probability Estimates" reports the probability for the reference household and the change in this probability caused by a unit change in a binary variable and a one-standard-deviation change in a continuous variable.

Table 4. Comparisons with selected literature

	Thesis (Model (3))				Campbell (2006)					Ferreira, Gyourko and Tracy (2008)		
	Whether a household refinanced between 2007-2009		Whether a household moved between 2007-2009			Whether a household refinanced between 2001-2003		Whether a household moved between 2001-2003			Whether a household moved between 1985-2005	
Dependent Variables	Coeff.	Prob. Estimates	Coeff.	Prob. Estimates		Coeff.	Prob. Estimates	Coeff.	Prob. Estimates		Coeff.	Prob. Estimates
Reference household		15.19		13.06			27.9		4.5			11.91
Collateral constrained	-0.522***	-0.0644***	-0.405***	-0.0396***	Loan problem	-0.179***	-4.6	-0.050	-0.5	Negative home equity (indicator)	N/A	-0.056**
	(0.087)	(0.0107)	(0.088)	(0.0103)		(0.071)		(0.088)				(0.021)
Liquidity constrained	-0.174*	-0.0330***	0.186**	0.0350**	Income problem	-0.076	-0.2	0.132	1.4	- change in log real household income	N/A	0.009**
	(0.090)	(0.0124)	(0.088)	(0.0140)		(0.099)		(0.130)				(0.003)
Sample size	5,415		5,415			5,190		5,735			58,363	

Sources: (Campbell 2006), (Ferreira, Gyourko and Tracy 2008)

Table 5. Marginal effects when using alternative specific multivariate probit estimation

		(1)	(2)	(3)
		Non-movers and non-refinancers	Refinancers	Movers
Dependent Variables		Probability Estimates	Probability Estimates	Probability Estimates
Probability estimate of reference household		0.7537	0.1342	0.1121
COLLCONSTR	Collateral constrained	0.1040**	-0.0657***	-0.0380*
		(0.0433)	(0.0196)	(0.0223)
LIQCONSTR	Liquidity constrained	-0.0023	-0.0328***	0.0354**
		(0.0327)	(0.0124)	(0.0175)
SPREAD	Interest rate spread	-0.0182***	0.0175***	0.0007
		(0.0046)	(0.0037)	(0.0034)
UNEMP	Ln(Unemployment rate difference)	-0.0542**	0.0285	0.0258
		(0.0295)	(0.0239)	(0.0210)
BOOKVAL	Ln(Book value of primary mortgage)	-0.0357***	0.0248***	0.0108*
		(0.0086)	(0.0069)	(0.0060)
WHITE	White	-0.0397**	0.0302**	0.0094
		(0.0174)	(0.0149)	(0.0125)
RURAL	Rural area	0.0205	0.0199*	-0.0404*
		(0.0527)	(0.0117)	(0.0226)
HOUSEVAL	Ln(Market value of residence in '07)	-0.0220**	0.0416***	-0.0197***
		(0.0095)	(0.0079)	(0.0063)
HHPLUS	Positive change in HH size	-0.0170	-0.0188	0.0357**
		(0.0301)	(0.0121)	(0.0158)
HHDIVORCE	Household head recently divorced	-0.1098***	0.0735	0.0373
		(0.0415)	(0.0492)	(0.0296)
Number of cases (total)		6,613	6,613	6,613
Number of obs.		19,839	19,839	19,839

Standard errors are reported in parentheses and the significance levels of coefficients and marginal effects are denoted by: *** p<0.01, ** p<0.05, * p<0.1. The column headed "Probability Estimates" reports the probability for the reference household and the change in this probability caused by a unit change in a binary variable and a one-standard-deviation change in a continuous variable. The results were obtained using the stata command "asmprobit" substituting real alternative-specific variables with the discrete numbers assigned to the choices.

APPENDIX B: DEFINITION OF VARIABLES

COLLCONSTR (Collateral constrained): dummy variable, taking a value of 1 if the book value of the outstanding total debt of the household exceeds 90 percent of the value of the property as reported in 2009.

LIQCONSTR (Liquidity constrained): dummy variable, taking the value of 1 if the household reported a total annual mortgage payment in 2007 that exceeds the 28 percent of its total yearly income.

AMMORT: The amount of the primary mortgage when acquired, or the face value of the loan, expressed in USD.

TERM1: Duration of the primary mortgage, expressed in years.

RATE1: Self-reported mortgage rate (%).

SPREAD: The difference between the self-reported mortgage rate and current lowest commitment rate on the market (4.88 if considering Freddie Mac rates).

PMT: Monthly mortgage payments, expressed in USD.

HOUSEVAL07: Self-reported value of the residence in 2007, expressed in USD.

HOUSEVAL09: Self-reported value of the residence in 2009, expressed in USD. In the case of movers, if the residence hasn't been occupied by 2009 and no reported house value is available for the second period, this observation is derived by appraising the reported 2007 value with the Case-Schiller House Price Index of the given SMSA region.

BOOKVAL: Book value of a mortgage, expressed in USD. The book value of the primary mortgages was calculated as follows:

$$FV\ of\ principal = AMMORT * \left(1 + \frac{\frac{RATE1}{100}}{12}\right)^{(2007 - origination\ year)*12}$$

$$FV\ of\ payments = -PMT * \frac{\left(1 + \frac{\frac{RATE1}{100}}{12}\right)^{(2007 - origination\ year)*12-1}}{\frac{\frac{RATE1}{100}}{12}}$$

$$BOOKVAL = FV\ of\ principal + FV\ of\ payments$$

LTV: The (book) loan-to-value ratio, including all existing debt on current equity, calculated as:

$$Book\ LTV = \sum_{i=1}^{n} \frac{BOOKVAL\ mortgage_i}{houseprice_{2009}}$$

TOTALINC: The sum of all household incomes, expressed in USD.

UNEMP: Difference in the unemployment rate on SMSA regional levels, between 2009 and 2007.

RURAL: Dummy variable, taking the value of 1 if the residence is situated in a rural area.

YEARSINRES: Number of years spent in current residence.

FIRSTTIME: The current residence is the first self-owned house of the respondent household.

SEX: Dummy variable, taking the value of 1 if the head of the household is a female, and zero otherwise.

WHITE: Dummy variable, taking the value of one if the head of the household is white.

HHCOL: Dummy variable, taking the value of one if the head of the household has at least 2 years of college education.

HHPLUS: Size of household increased between 2005 and 2007.

HHMIN: Size of household decreased between 2005 and 2007.

HHDIVORCE: The head of the household went from being married in 2005 to divorced, or separated, in 2007.

HHMARR: The head of the household went from being divorced, single or separated in 2005 to married in 2007.

HHAGE: Age of the householder.

PER: Number of persons living in the same household.