MODELING THE PERFORMANCE OF FHA-INSURED LOANS: BORROWER HETEROGENEITY AND THE EXERCISE OF MORTGAGE DEFAULT AND PREPAYMENT OPTIONS

Report

to

Office of Policy Development and Research
U.S. Department of Housing and Urban Development
451 Seventh St., SW Room 8212
Washington, D.C. 220410

by

Yongheng Deng and Stuart Gabriel[†]
ydeng@usc.edu and sgabriel@marshall.usc.edu

Unicon Research Corporation
1640 Fifth Street, Suite 100
Santa Monica, CA 90401

May, 2002

[†] The authors are Assistant Professor and Lusk Chair and Director, Lusk Center for Real Estate, School of Policy, Planning and Development and Marshall School of Business, University of Southern California. This project was funded by the U.S. Department of Housing and Urban Development under Contract: C-OPC 18484 Task Order 6. The authors are grateful to Dr. Robert Cotterman at Unicon Research Corporation for data preparation and to Lihong Yang for excellent research assistance. Any opinions expressed or implied are those of the authors and do not necessarily reflect the views or policies of the U.S. Department of Housing and Urban Development or the U.S. Government.

MODELING THE PERFORMANCE OF FHA-INSURED LOANS: BORROWER HETEROGENEITY AND THE EXERCISE OF MORTGAGE DEFAULT AND PREPAYMENT OPTIONS

by

Yongheng Deng and Stuart Gabriel

Abstract

Although mortgage loans to lower-income and higher credit risk borrowers are characterized by elevated default probabilities, those risks may be mitigated by their slower prepayment speeds. Loans to higher credit risk borrowers may prepay more slowly (when prepayment option is "in the money") owing to difficulties in borrower access to mortgage credit, problems of mortgage qualification, limited borrower knowledge of mortgage refinance options, or damped residential mobility. Because the prepayment risk premium is substantial, the differentially slower prepayment speeds of loans to higher risk borrowers may have important implications for loan profitability and for efforts to expand homeownership among those groups.

This paper estimates an option-based hazard model to simultaneously assess the competing risks of mortgage default and prepayment. In so doing, the analysis seeks to assess the differential default and prepayment probabilities among higher credit risk FHA mortgage borrowers. The empirical model derives from option theory and employs well-specified proxies for the mortgage put and call options in the default and prepayment equations. Further, given the availability of high quality micro data, the estimating equations control for borrower credit worthiness (credit scores) and other common underwriting variables among the approximately 40 contemporaneous and time-invariant indicators of borrower, loan, and locational risk.

Results of the analysis strongly support the predictions of option theory in explaining the exercise of default and prepayment options among mortgage borrowers. The estimates confirm that the call option (based on valuation of the mortgage at current and contract interest rates) is positive and highly significant in the exercise of the prepayment option; similarly, the value of the put option (probability of negative equity) also is positive and highly significant in the exercise of the default option. Results further suggest that a higher probability of negative equity reduces the risk of mortgage prepayment. Such an outcome is indeed plausible, in that households with poor equity positions may be less willing to exercise the refinance option owing to equity values that may be insufficient to refinance the remaining loan balance. Results further point to reduced consumer refinance propensity in more concentrated and less competitive loan markets. Among FHA borrowers, the initial loan-to-value ratio is negatively associated with prepayment propensity and positively associated with default propensity. As would be expected, higher credit score borrowers are less likely to exercise the default option, whereas lower credit score borrowers are less likely to prepay. Relative to white borrowers, estimates suggest that black and Hispanic borrowers are statistically less likely to exercise the prepayment option. Overall, results indicate the importance of slower prepayment speeds among higher default risk borrowers. For the investor in FHA-insured mortgage pools, the estimated cumulative probability of mortgage termination over five years among high default risk borrowers (as defined on the basis of borrower credit scores, levels of liquid assets, LTV ratio, and mortgage payment ratios), at 37%, remains well below that value, at 60%, for low-default risk borrowers. Investor recognition of this mortgage performance advantage should facilitate their greater willingness to acquire such loans and at more competitive pricing.

I. Introduction

Recent years have witnessed ongoing research and policy debate as regards the micro-foundations of mortgage loan performance. In part, studies have sought to test options theory as regards the competing risks of mortgage terminations (see, for example, Deng, Quigley, and Van Order [2000]) or Deng and Quigley [2001]). Analyses have further endeavored to exploit micro-data to distinguish mortgage performance across borrower and loan characteristics (see, for example, Berkovec, Canner, Gabriel and Hannon [1998], Archer, Ling, and McGill [2001], Van Order and Zorn [2001], and Cotterman [2001][1]). Those studies have been viewed with particular importance, given concerted government efforts to enhance mortgage availability and homeownership opportunity among traditionally underserved and higher risk groups.

Despite methodological and data advances, recent analyses of the micro-foundations of mortgage loan performance suffer from numerous limitations. Prior studies (including Berkovec et al [1998], Deng et al [2000, 2001], and Archer et al [2001]) often did not benefit from important information on borrower creditworthiness, as typically is summarized in mortgage applicant credit scores. Similarly, earlier papers were circumscribed by limited availability of other borrower, loan, and locational information critical to mortgage loan performance, including indicators of borrower wealth, mortgage payment-to-income and debt-to-income ratios, neighborhood, market concentration and other controls (see, for example, Deng et al [2000, 2001], Archer et al [2001], and Van Order and Zorn [2001]). For the most part, prior analyses were limited to the specification of mortgage performance controls at the time of loan origination and did not include contemporaneous updates to such controls (notably including mortgage put and call options) over the life of the mortgage. Finally, prior analyses often focused on models of default probability without consideration of prepayment risk or vice-versa. Clearly, the literature suggests that default and prepayment behavior are competing risks, and hence should be jointly modeled (see, for example, Deng,

[1] Cotterman (2001) uses micro-data to study the effects of borrower and neighborhood characteristics (including race, ethnicity, and income) on the default of FHA mortgages. He found that those African-American neighborhoods

Quigley, and Van Order [2000].)

This paper applies a competing risk framework to model the micro-foundations of FHA-insured mortgage performance. The FHA data are well suited to analyses of loan default, given the inclusion in the program of large numbers of relatively higher credit risk borrowers. Less well appreciated is the possibility that those same higher risk borrowers may prepay their mortgages more slowly (when prepayment option is "in the money"), given problems of access to mortgage finance, difficulties in mortgage qualification, limited borrower knowledge of mortgage refinance options, or reduced residential mobility. To the extent the prepayment risk of mortgages originated among "high-risk" borrowers is relatively damped, it should be reflected in the pricing of those loans. Indeed, from a mortgage pricing perspective, the reduced prepayment risk associated with lower-income and higher credit-risk FHA borrowers may serve to mitigate the higher default probabilities associated with those borrowers.

This paper estimates an option-based hazard model to simultaneously assess the competing risks of FHA mortgage default and prepayment. In so doing, it seeks to assess differential default and prepayment probabilities among higher credit risk FHA mortgage borrowers. The empirical model derives from option theory and employs well-specified proxies for the mortgage put and call options in the default and prepayment equations. Further, given the availability of high quality micro data, the estimating equations control for borrower credit worthiness (credit scores) and other common underwriting variables among the approximately 40 contemporaneous and time-invariant indicators of borrower, loan, and locational risk.

Results of the analysis strongly support the predictions of option theory in explaining the exercise of default and prepayment options among FHA mortgage borrowers. The estimates confirm that the intrinsic value of call option is positive and highly significant in the exercise of the prepayment option; similarly, the value of the put option (probability of negative equity) also is positive and highly significant in the exercise of the default option. Results further suggest that a higher probability of negative equity reduces

---

with high concentrations of lower income households are associated with higher rates of default, whereas individual borrower race or income alone is unrelated to default.

the risk of mortgage prepayment. Such an outcome is indeed plausible, in that households with poor equity positions may be less willing to exercise the refinance option owing to equity values that may be insufficient to refinance the remaining loan balance.

Results further point to the importance of other borrower, loan, and market characteristics in the estimation of mortgage termination risks. For example, findings indicate reduced consumer refinance propensity in more concentrated and less competitive loan markets. Among FHA borrowers, the initial loan-to-value ratio is negatively associated with prepayment propensity and positively associated with default propensity. As would be expected, higher credit score borrowers are less likely to exercise the default option, whereas lower credit score borrowers are less likely to prepay. In that regard, the 5-year cumulative probability of prepayment is 22 percent higher among borrowers with scores in excess of 740 than among those with scores below 620. Borrowers with a larger number of dependents as well as older and lower income borrowers are less likely to exercise the prepayment option. Relative to white borrowers, estimates suggest that black and Hispanic borrowers are statistically less likely to exercise the prepayment option. Among white borrowers, for example, the 5-year cumulative probability of prepayment of 55% is more than 2-1/2 times in excess of the 20% rate estimated for similarly credit worthy blacks. Indeed, computation of cumulative prepayment rates by race and credit worthiness illustrates the strikingly lower prepayment propensities of black borrowers, relative to their white and Asian counterparts.

Overall, results indicate the appropriateness of the competing risk specification and illustrate the importance of slower prepayment speeds among higher credit risk borrowers. As is evidenced below, the substantially elevated default probabilities of higher credit risk groups are more than offset by damped prepayment propensities among those borrowers, resulting in significantly lower loan termination propensities overall. Indeed, among high default risk borrowers, at 5 years post loan origination, loan termination probabilities via prepayment are about 4 times those emanating from loan default. For the investor in FHA-insured mortgage pools, the estimated cumulative probability of mortgage termination over five years among high default risk borrowers (at 37%) remains well below that value (at 60%) for low-default risk borrowers. Investor recognition of this mortgage performance advantage should facilitate their greater willingness to acquire such high credit risk FHA-insured loans and at more competitive price.

The plan of the paper is as follows. Section II presents the basic model and estimation strategy. Section III describes the FHA database whereas section IV discusses estimation and simulation results. Conclusions and implications for mortgage pricing are discussed in section V.

II. Methodology.

Recent research on mortgage borrowers indicates that the exercise of mortgage prepayment and default options are behaviorally distinct, but not independent. For example, one cannot calculate accurately the economic value of the default option without considering simultaneously the financial incentive for prepayment (Deng, Quigley, and Van Order [2000]). Furthermore, risk preferences and other idiosyncratic differences across borrowers may vary widely. However, appropriate modeling of prepayment and default risks is crucial to the pricing of mortgages and to an understanding of the economic behavior of homeowners.

The analysis applies a proportional hazard framework to assess the competing risks of mortgage termination by prepayment and default. The model derives from option theory and predicts that well-informed mortgage borrowers in a perfectly competitive market will exercise the default or prepayment option in order to increase their wealth. In this context, FHA loan borrowers will exercise the default option when the market value of the mortgage equals or exceeds the market value of the collateral. Similarly, borrowers can increase their wealth by refinancing their loans when the market value of the mortgage exceeds the par value of the mortgage. However, these two options compete against each other. For example, when an individual decides to exercise the default option, she is making the decision to forego future exercise of the prepayment option. Kau et al (1995) have outlined the theoretical relationships among the options, and Schwartz and Torous (1993) have demonstrated their practical importance. Furthermore, empirical evidence shows that certain borrower characteristics that have strong association with one option may have the opposite association with the other option. For example, a lower-income borrower with a poor credit history may have higher default risks but lower refinancing risks due to credit problems and/or liquidity constraints that typically affect the ability to qualify for a new loan.

This paper follows Deng, Quigley, and Van Order (2000) in application of an option-based hazard model to simultaneously estimate the competing risks of FHA loan default and prepayment. In this model, T_p and T_d are discrete random variables representing the duration of a mortgage prior to termination by the mortgage holder in the form of prepayment or default, respectively. Following the Cox model, the joint survivor function conditional on ξ_p, ξ_d, r, H, Y, and X can be expressed in the following form:[2]

$$\begin{aligned} &S\left(t_p, t_d \mid r, H, Y, X, \xi_p, \xi_d, \theta\right) \\ &= \exp\left\{-\xi_p \sum_{k=1}^{t_p} \exp\left(\gamma_{pk} + \beta_{p_1}^{'} g_{pk}\left(r, H, Y\right) + \beta_{p_2}^{'} X\right)\right. \\ &\qquad \left. - \xi_d \sum_{k=1}^{t_d} \exp\left(\gamma_{dk} + \beta_{d_1}^{'} g_{dk}\left(r, H, Y\right) + \beta_{d_2}^{'} X\right)\right\}. \end{aligned} \tag{1}$$

In this formulation $g_{jk}\left(r, H, Y\right)$ are time-varying variables measuring the financial values of the prepayment and default options ($j = p, d$). The equation includes empirical measures of the intrinsic values of prepayment option (the call option) as well as the default option (the put option). The relevant interest rates and property values are r and H, respectively, whereas Y is a vector of other variables that also are relevant to an empirical description of the market values of the default and prepayment options.

Following Deng, Quigley and Van Order (2000), the "*Call Option*" for each individual FHA loan borrower is defined as:

$$Call\_Option_{i,k} = \frac{V_{i,m} - V_{i,r}^{*}}{V_{i,m}}, \tag{2}$$

where

[2] The proportional hazard model introduced by Cox and Oakes (1984) provides a framework for considering the contingent claims model empirically and for measuring the effect of financial options on the behavior of mortgage holders.

$$V^*_{i,r} = \sum_{S=1}^{TM_i - k_i} \frac{P_i}{\left(1+r_i\right)^s},$$
$$V_{i,m} = \sum_{S=1}^{TM_i - k_i} \frac{P_i}{\left(1+m_{\tau_i+k_i}\right)^s}, \qquad (3)$$

r_i is mortgage note rate, TM_i is the mortgage term, k_i is the mortgage duration after origination at time τ_i, $m_{\tau_i+k_i}$ is the market interest rate, and P_i is the monthly mortgage payment.

Typically, we cannot measure directly from the micro data the extent to which the default option is "in the money" without knowing the entire path of individual house values. We can, however, estimate the probability of exercise of the "put option" based on the initial loan-to-value ratio and the diffusion process of house prices. Specifically, the "*put option*" variable is defined as:

$$Put\_Option_{i,k} = \Phi\left(\frac{\left(\log V_{i,m} - \log M_{i,k}\right)}{\sqrt{\omega^2}}\right), \qquad (4)$$

where $\Phi(\cdot)$ is cumulative standard normal distribution function, ω^2 is an estimated variance, $V_{i,m}$ is defined previously, and the market value M_i of property *i*, purchased at a price of C_i at time τ_i and evaluated k_i quarters thereafter is

$$M_{i,k} = C_i\left(\frac{I_{j,\tau_i+k_i}}{I_{j,\tau_i}}\right), \qquad (5)$$

where the term in parentheses follows a log normal distribution.

The vector *X* is comprised of other non-option-related variables, including both time-varying and time constant determinants of mortgage performance. Time-varying variables include the unemployment rate of the MSA, the log value of a Herfindahl-Hirschmann Index of metropolitan mortgage lending market concentration, interactions of put and call options with borrower credit scores, and interactions of black and Hispanic households with the log value of the Herfindahl-Hirschmann Index.[3] Time-invariant variables include categorical measures of borrower credit score, borrower race/ethnicity, borrower

[3] The measure of market concentration employed in the analysis is the Herfindahl-Hirschmann Index, defined as the sum of the squared market shares of all providers of home purchase loans in the market, where a market is defined

housing expenditure-to-income ratio, borrower debt-to-income ratio, borrower gender and marital status, borrower age group, first-time homebuyer status, seller offer to buy down the mortgage rate, whether the mortgage is amortized in 30 years or less, whether the property is located in the central city, whether the property is located in a rural area, and whether the property is a new home. Other continuous controls include mortgage loan-to-value ratio at origination, log value of property appraisal value, number of dependents in borrower's household, log value of borrower liquid assets, and log value of household income. Also included among controls for mortgage performance are census tract level variables reflecting neighborhood racial/ethnic mix, proportion rental occupied stock, and ratio of census tract to MSA median income.

Unobserved error terms associated with the hazard functions for prepayment and default are denoted ξ_p and ξ_d, respectively. θ is a vector of parameters (*e.g.*, γ and β) of the hazard function. γ_{jk} are parameters of the baseline hazard function. The baseline may be estimated with a flexible form suggested by Han and Hausman (1990), such that:

$$\gamma_{jk} = \log\left[\int_{k-1}^{k} h_{0j}(s)\,ds\right], \quad j = p, d. \tag{6}$$

Alternatively, the form of the baseline may be imposed by employing some standard benchmark experience such as the mortgage "PSA and SDA experiences." [4]

The estimated competing risks of prepayment and default are then used to simulate the potential risks to FHA mortgage lending as derived from various borrower and loan characteristics, notably including

---

as a metropolitan statistical area. This measure of market concentration was used successfully by Berkovec, Canner, Gabriel, and Hannan (1998) to explain differences in competitive conditions across MSAs.

[4] The Public Securities Association (PSA) has defined a prepayment measurement standard that has been widely adopted by fixed-income securities analysts. This is a series of 360 monthly prepayment rates expressed as constant annual rates. The series begins at 0.2 percent in the first month and increases by 0.2 percent in each successive month until month 30, when the series levels out at 6 percent per year until maturity. (See Fabozzi [1995] for details.) The Bond Market Association (formerly the Public Securities Association) has also developed the Standard Default Assumption (SDA) that is widely used as a benchmark to measure loan default experience. The SDA series begins at 0.02 percent annual constant rate in the first month and increases by 0.02 percent in each successive month until month 30, when the series levels out at 0.6 percent per year for the next 30 months. Then the series declines by 0.0095 percent each month from month 61 to month 120. Afterwards, the default rate levels there until maturity. Prepayments and defaults are often reported as simple linear multiples of the PSA and SDA schedules, respectively. Therefore, by adopting PSA and SDA schedules as the baselines for prepayment and default functions, respectively,

loan-to-value and payment-to-income ratios as well as borrower liquid assets and credit scores. Further simulations are undertaken for hypothetically comprised high- and low- default and prepayment risk borrowers. As indicated below, total loan terminations from default and prepayment among higher credit risk borrowers are estimated to be substantially below those of low credit risk borrowers, suggesting enhanced investor profitability of those loans when such prepayment option is "in the money".

III. Data

The principal data utilized in this study consist of a large random sample of FHA-insured home purchase loans originated during the 1992-1996 period.[5] All loans are fully amortizing, most with thirty-year terms. The individual loan records contain information on a large number of loan, borrower, and property-related characteristics and also indicate termination date of each loan and reason for termination.[6] Attached to the loan record files are borrower credit scores at time of loan application as are measures of local housing market performance including house price appreciation and volatility.[7] Further, using a census tract indicator for each property location, each loan record file is matched to neighborhood socioeconomic and housing market indicators from the 1990 Census of Population and Housing. Other neighborhood or metropolitan area level variables, including unemployment rates, also are appended to the record file. FHA data on the race of the borrower and census measures of neighborhood racial composition enable assessment of race-related effects associated with the

---

the factors of proportionality estimated from the hazard model can be expressed simply as a percentage of the PSA and SDA experiences.

[5] The final sample consists of 12,012 loans randomly selected from the 120,342 endorsed loans applications from 1992, 1994, and 1996. Loan origination dates are concentrated in those three calendar years but also spread out into other years. The 120,342 loan data provided by Unicon Research is a choice-based sample with weights that accounted for choice-based sampling and differential loan losses by race and loan status. For each of the application years, the weighted cumulative default rates for the loans comprising the sub-sample of 12,012 loans was found to be quite similar to that observed in the parent population. The individual loan files are observed on a monthly basis from month of origination through that of termination, maturation, or through the end of 2000 for active loans.

[6] As defined for this analysis, default outcomes include both lender foreclosure and situations where the borrower conveys title of the property in lieu of foreclosure (correct this definition as needed). Loan prepayment prior to completion of the amortization period is defined among loan termination options.

[7] Borrower credit score information is provided by Equifax and Trans Union. If the data provides both Equifax and Trans Union scores for an individual borrower, we take the average of the two scores. Numerous recent papers (see, for example, Avery et al (1996)) point to the importance of controls for borrower credit score in micro-analyses of mortgage default likelihoods.

performance of FHA-insured loans. The FHA data set encompasses nearly 300 different metropolitan areas, allowing for substantial variability in the structure of local lending markets.

The FHA-insured data are well-suited to analyses of loan performance, given the inclusion in the program of large numbers of relatively high credit-risk borrowers. Although both FHA and conventional mortgage applications are evaluated according to formal underwriting criteria, the FHA guidelines are less strict than those of conventional lenders, particularly as regards the level of borrower equity at the time of loan origination and the acceptable ratios of housing expense to income and total debt expense to income. Approximately 61 percent of the loans in the sample have loan-to-value ratios exceeding 95%. Similarly, the debt obligation ratios of the FHA borrowers in the sample are relatively high, averaging about 35% for the ratio of total debt payments-to-income and about 23% for the ratio of housing debt payments to income. First-time homebuyers and moderate-income borrowers comprise a large portion of the sample, and minorities are well represented as well.

Table I displays the means and variances of the time-invariant covariates, whereas Table II provides the same for time-varying covariates at origination and termination. As is evidenced in Table I, some two-thirds of FHA borrowers were first-time buyers; the average mortgage loan-to-value ratio among sampled loans was 94 percent. As would be expected, the majority of sampled loans were to married borrowers, aged 25-35, with housing expense-to-income ratios of 20-38%, debt-to-income ratios of 20-41%, and credit scores in the range of 620-740. As would be expected (Table II), among prepaid loans, the computed mean of the call option value at termination substantially exceeded that at time or loan origination. Owing to equity build-up over the loan period, the value of the put option at the time of loan origination (probability of negative equity) substantially exceeded that at time of loan termination.

IV. Empirical Results

Our competing risks analysis is based on a stratified sample of FHA loan data provided by HUD. A weighting procedure variable is used in the maximum likelihood estimation (MLE) procedure to correct the possible sample selection bias. The weight variable accounts for the choice-based sample selection across the race and loan status cells. More specifically, the weight is defined as the inverse of probability that the loan observation is being selected from a cell where it was sampled.[8] The competing risks of default and prepayment are estimated jointly.

Table III presents three variants of the competing risks model of FHA loan termination. Each model contains separate default and prepayment baseline functions that follow the SDA and PSA experiences.[9] Model 1 does not control directly for the values of the call and put options in the estimating equations. Further, that model excludes controls for mortgage borrower credit scores. Accordingly, the specification of Model 1 approximates that of most prior micro-data analyses of FHA mortgage default and provides a benchmark for the competing risks models discussed below. However, in contrast to most prior analyses, which were restricted to controls for local market conditions at the time of loan origination, Model 1 includes several time-varying proxies.[10] The time-varying covariates include SMSA level unemployment rates, log value of the Herfindahl-Hirschmann Market Concentration Index, and an interaction of black and Hispanic households and the Herfindahl Index terms. Model 2 extends Model 1 by including the values of both the call and put options in both risk equations. Model 3 extends Model 2 by including the borrower's credit score information. In addition, the put and call values are interacted with borrower credit scores. All specifications also include a rich set of time-constant controls for borrower, loan, and locational determinants of exercise of the default and prepayment options. Overall, the competing risks models are well-specified and control for approximately 20 different characteristics of the loan, the borrower, and the census tract or area in which the property is located.

As evidenced Model 1, estimation results indicate that increases in local unemployment rates negatively affect the exercise of the prepayment option but positively affect exercise of the default option. These results are highly significant across model specifications and are consistent with previous studies based on agency conforming loan data (see for example Deng, Quigley, and Van Order [2000]).

[8] Here we assume that the sampling mechanism is independent of error distribution of the competing risks of FHA loan prepayment and default risks.
[9] We use 100% SDA and PSA curves as our baselines for loan default and prepayment, respectively. All of the models reported in Table III use the same baselines.
[10] For a recent micro-data based analysis of the FHA default experience, see Cotterman (2001).

The estimates from Model 1 suggest that the initial loan-to-value ratio is negatively associated with prepayment risk and positively associated with default risk.[11,12] The estimated LTV coefficients are statistically significant across all model specifications. Higher levels of LTV may reflect in part borrower difficulties in loan re-qualification that diminish the exercise of the prepayment option. As would be expected, Model 1 also reports that prepayment likelihoods vary positively with mortgage expense burdens. An increase in the ratio of housing expense-to-income from below 20% to 20-38% and to in excess of 38% results in statistically significant increases in the likelihood of mortgage prepayment. In contrast, borrower total debt-to–income burdens do not figure significantly in the exercise of prepayment options. In the competing risk model, neither the front-end or back-end mortgage obligation ratios are significant in the exercise of the default option.[13,14]

Model 1 indicates that prepayment likelihoods are elevated among loans subject to interest rate "buy-downs". The estimated coefficient associated with that variable is insignificant in the default equation. In contrast, exercise of the prepayment option is significantly damped among first-time borrowers, single-female borrowers, and borrowers with higher value properties. Compared to married couples, single male borrowers are of significantly higher default risk. As would be expected, shorter-term mortgage loans are characterized by significantly lower prepayment and default risk. Borrowers with larger number of dependents are significantly less likely to exercise the prepayment option but significantly more likely to exercise the default option. Borrowers with greater liquid assets (and hence fewer liquidity constraints)

---

[11] In many prior studies, the ratio of the size of loan to the market value of the property at the time of loan origination is particularly important in predicting default probability, with higher LTVs associated with higher likelihoods of default. See, for example, BCGH (1998).

[12] Note that the model is estimated using a non-random choice-based sample. With the sample selection weights, we obtain consistent estimates for the regression point estimates. However, since the weights vary by individual, the estimated second moments are upward biased, in turn suggesting that the t-statistics on the regression coefficients are downward biased. Accordingly, those estimated coefficients that appear to be borderline significant may indeed be estimated with a high level of statistical significance.

[13] The two "obligation ratios" of housing expense-to-income and total debt payment-to-income are presented as a series of dummy variables indicating specific ranges of these ratios. This approach was adopted because the cut-off values are relevant to FHA loan underwriting guidelines. Therefore we allow for these nonlinearities in our estimation procedure.

[14] Earlier micro-data analyses of default likelihood indicate the importance of increases in the front-end ratio to exercise of the default option. As suggested, those results are not robust to the competing risk specification of mortgage default and prepayment.

are less likely to exercise the default option; however, borrower liquid assets do not significantly affect exercise of the prepayment option. Younger and higher income borrowers are more likely to prepay; however, those factors are not statistically significant in the exercise of the default option. Having accounted for borrower and loan characteristics, findings indicate that census tract level controls are not significant to the exercise of the mortgage options.[15] Further, estimation findings are largely robust to the exclusion of those controls. Research findings also indicate little systematic variations in loan termination propensities across central city, suburban, or rural areas.

The competing risks model also tests for variation in the exercise of default and prepayment options across borrower race and ethnicity. As evidenced in Model 1, Asian borrowers do not appear to be statistically different from white borrowers in their exercise of either the mortgage put or call options. In marked contrast, both Hispanic and black borrowers are characterized by statistically damped prepayment likelihoods. In this case, the reduced exercise of the prepayment option among Hispanic and black borrowers serves to enhance the profitability of those loans among investors in FHA-insured mortgages. In contrast to earlier studies, results of the estimation of the competing risks model do not indicate the presence of statistically elevated default risks among black and Hispanic borrowers.[16]

Our empirical results show damped exercise of the default and prepayment options in more concentrated metropolitan lending markets (as measured by the Herfindahl Index). In Model 1, however, only the association with default likelihood is statistically significant. In general, results are consistent with the hypothesis that in highly concentrated lending markets, lenders may impose more stringent underwriting standards so as to improve loan quality and reduce the risk of default. The Herfindahl-Hirschmann Index of Market Concentration also is interacted with the minority (black and Hispanic)

[15] Estimation of Table III inclusive of census tract controls is contained in Appendix A.1.

[16] This result stands in marked contrast to earlier results indicating statistically elevated default probabilities among black borrowers (see, for example, BCGH [1998]). As well appreciated, however, the FHA data utilized herein derives from a more recent period. Further, earlier results did not derive from a competing risks model of mortgage default and prepayment replete with time-varying controls.

borrower status variables. For the most part, results fail to suggest differential minority borrower exercise of default and prepayment options in more concentrated lending markets.[17]

Model 2 extends Model 1 through the introduction of the option-related time-varying covariates into both the prepayment and default equations. The call and put option controls are similar to those used by Deng, Quigley, and Van Order (1996). Note, however, that the FHA data utilized herein enables a much richer specification of the competing risks than has been previously estimated using conventional loan data (see, for example, Deng, Quigley and Van Order [1996, 2000], and Van Order and Zorn [2001]). The estimates confirm that the call option value is positive and highly significant in the exercise of the prepayment option; similarly, the value of the put option (probability of negative equity) also is positive and highly significant in the exercise of the default option. In other words, declines in mortgage interest rates that bring the call option "into-the-money" will lead to a high volume of prepayment activities, as is observed in the data in the sharp upward movement in mortgage prepayment activity in 1993 and 1998. On the other hand, when the probability of negative equity becomes imminent, the incidence of default increases dramatically. These findings strongly support the predictions of option theory in explaining the exercise of default and prepayment options on the part of mortgage borrowers.

Model 2 further suggests that a higher probability of negative equity significantly reduces the risk of mortgage prepayment. Such an outcome is indeed plausible, in that households with poor equity positions may be less willing to exercise the refinance option owing to equity values that may be insufficient to refinance the remaining loan balance. On the other hand, the value of the call option exerts a significant positive influence on default propensities. This may be explained by the fact that when market rates drop, the value of call option increases, as does the market value of the mortgage. Relative to the market value of the outstanding balance of the loan, the underlying collateral (the house) is less

[17]Among black and Hispanic borrowers, lending market concentration is statistically insignificant in determination of default likelihood. Results similarly failed to indicate a statistically different prepayment risk between blacks and whites in a more concentrated lending market.

valuable to the borrower so as to encourage borrower exercise of the default option. These findings are consistent with Deng, Quigley, and Van Order (2000).

For the most part, the remaining estimated coefficients of Model 2 are robust to the inclusion of the call and put option values. Note, however, that the market concentration measure becomes negative and statistically significant in its effect on prepayment likelihood, suggesting that a reduced number of market competitors results in diminished borrower exercise of the prepayment option.[18]

Model 3 extends Model 2 through the introduction of borrower credit scores into both the default and prepayment equations. The credit scores are entered in a nonlinear fashion roughly consistent with loan underwriting policy. As evidenced in Table III, the credit score terms are statistically significant in the default equation. As would be expected, relative to the excluded highly credit qualified borrowers (credit score > 740), lower score borrowers are more likely to exercise the default option. On average, the default likelihood of the middle qualified group (credit score in the 620-680 range) is almost twice as high as the more highly qualified group (score in the 680-740 range), whereas the default risks associated with the least qualified group (credit score below 620) are about 2.4 times higher than the group with score between 680-740.[19]

The credit score variables also are interacted with the time-varying estimates of the call and put options. As evidenced in Model 3 results, the interactive credit score and call option terms are positive and highly significant in the loan prepayment equation. Further, the estimated coefficients indicate more ruthless exercise of the call option among the most credit worthy borrowers.[20] Among borrowers with credit scores in excess of 740, for example, the influence of the call option value on prepayment propensities is about one-third higher than that of borrowers with credit scores below 620. The estimated interactions between credit scores and call option proxy also underscore the relatively damped

[18] The market concentration measure also exerts and negative and significant influence on the exercise of the put option.

[19] The default likelihood of the middle qualified group (credit score in the 620-680 range) relative to that of the more highly qualified group (score in the 680-740 range) is 1.347/0.702=1.9, whereas the default risks associated with the least qualified group (credit score below 620) relative to the group with a score between 680-740 is 1.672/0.702=2.38.

prepayment propensities of less credit worthy borrowers, even as that prepayment is "in the money". Similarly, the estimated coefficients on the interactive put option and credit score terms also are positive and highly significant, suggesting a U-shaped relation with elevated propensities to default among both relatively low and high credit score borrowers. The estimated coefficients of the interactive put option and credit score variables also are negative and highly significant in the loan prepayment equation.

Table IV reports on the *unadjusted* cumulative probability of prepayment and default by various covariates and at the end of post-origination years one, three, and five. The unadjusted probabilities derive from the full sample of FHA loans. Overall, the data indicate very substantial upward movement in prepayment probabilities over the five years subsequent to mortgage origination; default propensities similarly are shown to move up perceptibly over that period. The top panel reports on the cumulative probabilities of prepayment and default by borrower liquid assets. As would be expected, the data indicate substantially higher prepayment probabilities and similarly damped default probabilities among those borrowers with liquid assets in excess of median levels. Among other borrower and loan characteristics, elevated prepayment propensities are observed among loans with LTVs below 95 percent and housing expense-to-income ratios of 20-38 percent, and among repeat buyers. Those same borrower and loan categories are associated with relatively damped five-year cumulative default probabilities.

Table V *simulates* the cumulative probabilities of prepayment and default by those borrower and loan characteristics identified in Table IV. As in Table IV, those probabilities are computed for one, three, and five years post loan origination. The simulations are based on a ten percent random sample of loans originated in June 1992. The baseline borrower is assumed to be a white household purchasing an existing suburban home with a 30-year fixed rate mortgage. The values of the other time-invariant control variables are set at their sample means, whereas time-varying covariates are set at their sample mean in each period.[21] Those covariates that are the focus of model simulation are specified in the table.

[20] These findings are consistent with Bennett et al (2001).
[21] Among time-invariant controls, for example, the simulation assumes two dependents per household. Further, the average loan-to-value ratio is set equal to 94 percent, whereas the log values of property value, household liquid

For that simulated mortgage borrower, higher levels of credit risk serve both to elevate default likelihoods and to damp prepayment propensities. For example, as shown in Table V, at 5 years post loan origination, borrowers with high LTVs (LTVs ≥ 95%) are characterized by 1.4 times the default risk (3.53/2.48=1.4) of borrowers with lower LTVs. Also evident, however, are the substantially lower prepayment propensities of those high LTV borrowers; at 5 years post loan origination, the prepayment likelihoods of high LTV borrowers were 20 percent below those of lower LTV loans. A similar outcome is evidenced, for example, in the simulation of default and prepayment propensities among more or less credit worthy borrowers. At 5 years post loan origination, borrowers with lower credit scores (credit scores < 680) are characterized by almost 3 times the default risk (6.01/2.19= 2.74) of borrowers with higher credit scores. Those same lower credit score borrowers are characterized by damped prepayment risk relative to their higher credit score counterparts.[22]

The bottom rows of Table V provide simulations of default and prepayment propensities among more fully specified high- and low- credit and prepayment risk borrowers. The precise specification of those borrower profiles is articulated in notes to Table 5. In general, higher prepayment risk borrowers are

---

assets, and family income are set to 11.13, 8.54, and 8.00, respectively. These simulations further assume that the borrowers are married, first-time buyers and that the loan interest rate is not subject to buy-down.

[22]Other simulations suggest that by the end of year five post-origination, younger borrowers (age of household head is less than 25 years old) are characterized by 1.3 times the prepayment risks (55.98/42.4=1.3) of older households (age of household head greater than 45 years old). While the simulated risks of loan default similarly move up over the five-year period post origination, the differences between age groups is slight. Findings further suggest that the cumulative 5-year risk of prepayment is relatively higher among married couples (52%) than single females (46%). In marked contrast, the 5-year cumulative probability of default among single males is about 1-1/2 times that of single females. We further find little quantitative variation in the cumulative probabilities of default across first-time buyer and loan rate buydown status.We also simulated the cumulative probability of prepayment and default by borrower race and credit score. As would be expected, the 5-year cumulative probability of prepayment rises substantially with borrower credit worthiness (as reflected in borrower credit score); among white borrowers, that probability is 22 percent higher among borrowers with scores in excess of 740 than among those with scores below 620 (55%/45%). A similar pattern is detected among black and Latino borrowers, albeit at substantially lower cumulative probabilities. Among white borrowers, for example, the 5-year cumulative probability of prepayment of 55% is more than 2-1/2 times in excess of the 20% rate estimated for similarly credit worthy blacks. Indeed, computation of cumulative prepayment rates by race and credit worthiness illustrates the strikingly lower prepayment propensities of black borrowers, relative to their white, Latino, and Asian counterparts. Likewise, cumulative default rates among black borrowers are estimated to be substantially in excess of those for other racial groups. At almost 7%, the 5-year cumulative default rate of highly credit worthy black borrowers (credit scores in excess of 720) is almost 6 times greater than similarly qualified white borrowers. Results of these analyses are available from the authors upon request.

those with higher levels of liquid assets, elevated credit scores, and more conservatively underwritten mortgages (as regards loan-to-value and payments-to-income ratios). With some limited nuance, higher credit risk borrowers are the opposite.

As is evidenced in Table V, loan performance behavior differs markedly over these borrower risk profiles. For example, by end of year 5 post loan origination, the simulated prepayment propensity of the higher prepayment risk borrower is fully double that of the lower prepayment risk borrower. Further, higher prepayment risk borrowers are characterized by only 1/6th the default risk of their lower prepayment counterparts. Those same results are roughly replicated among higher- and lower-credit risk borrower profiles. As is evidenced in the table, high default risk borrowers prepay at about one-half the rate of low default risk borrowers; similarly, the default propensities of the higher credit risk group are in excess of 5 times that of the lower credit risk group. Results then provide clear evidence of the damped prepayment propensities of the higher credit risk group.

The right-hand columns of Table V provide an assessment of total termination risk of FHA-insured mortgage loans. Those risks are defined as the sum of the default and prepayment propensities at the end of years 1, 3, and 5. Total loan terminations (from all sources) are relevant to the profitability of investment in FHA-insured mortgages. Typically, those loans not only are FHA-insured, but if pooled and sold also often are backed by a Ginnie Mae guarantee of timely repayment of principal and interest in the event of borrower default. Accordingly, from the perspective of the FHA-insured loan investor, a loan termination via default is equivalent to that which derives from prepayment. Clearly, borrower groups with lower total loan termination risks represent more profitable loan investment opportunities, relative to those groups with higher total termination propensities.

As is evident in Table V, total loan termination risk is substantially elevated among both high prepayment risk and low default risk borrowers. Further, total termination risk among low default risk borrowers is about 50 percent in excess of that of high default risk borrowers. A similar termination risk differential is indicated between high- and low-prepayment risk groups. As is further apparent, the substantially elevated default probabilities among the high default risk group are more than offset by the damped prepayment propensities of those higher credit-risk borrowers, resulting in significantly lower loan termination propensities overall. Indeed, among high default risk borrowers, loan termination probabilities via prepayment at end of year 5 post origination are about 4 times that of loan termination propensities from default. Similarly, the very substantial default probabilities of low prepayment risk borrowers, when added to the damped prepayment propensities of that group, result in overall loan termination risks that remain at only two-thirds of that of high prepayment risk borrowers. Clearly those loans originated among high default risk borrowers are relatively more profitable to the investor, given their substantially depressed overall termination propensities.

V. Summary and Conclusion

This paper applies micro-data from the FHA to estimate an option-based hazard model of the competing risks of mortgage default and prepayment. The empirical model derives from option theory and predicts that well-informed mortgage borrowers in competitive markets will exercise the default or prepayment options in order to increase their wealth. The analysis is well specified and includes proxies for mortgage put and call options, borrower credit worthiness, lending market concentration, and numerous other time-constant and time-varying borrower, loan, and locational characteristics. The estimated competing risks of prepayment and default are then used to simulate the potential risks to FHA mortgage lending as derive from various borrower characteristics, notably including borrower age, marital status, first time buyer status, credit worthiness, and race/ethnicity.

Results of the analysis strongly support the predictions of option theory in explaining the exercise of default and prepayment options among FHA mortgage borrowers. The estimates confirm that the call option value is positive and highly significant in the exercise of the prepayment option; similarly, the value of the put option (probability of negative equity) also is positive and highly significant in the exercise of the default option. Results further suggest that a higher probability of negative equity reduces the risk of mortgage prepayment. Such an outcome is indeed plausible, in that households with poor equity positions may be less willing to exercise the refinance option owing to equity values that may be insufficient to refinance the remaining loan balance.

Results further point to the importance of other borrower, loan, and market characteristics in the estimation of mortgage termination risks. For example, findings indicate reduced consumer refinance propensity in more concentrated and less competitive loan markets. Among FHA borrowers, the initial loan-to-value ratio is negatively associated with prepayment propensity and positively associated with default propensity. As would be expected, higher credit score borrowers are less likely to exercise the default option, whereas lower credit score borrowers are less likely to prepay. In that regard, the 5-year cumulative probability of prepayment is 22 percent higher among borrowers with scores in excess of 740 than among those with scores below 620. Borrowers with a larger number of dependents as well as older and lower income borrowers are less likely to exercise the prepayment option. Relative to white borrowers, estimates suggest that black and Hispanic borrowers are statistically less likely to exercise the prepayment option. Among white borrowers, for example, the 5-year cumulative probability of prepayment of 55% is more than 2-1/2 times in excess of the 20% rate estimated for similarly credit worthy blacks. Indeed, computation of cumulative prepayment rates by race and credit worthiness illustrates the strikingly lower prepayment propensities of black borrowers, relative to their white, Latino, and Asian counterparts.

Overall, results indicate the appropriateness of the competing risk specification and indicate the importance of slower prepayment speeds among higher risk borrowers. As is evidenced, the substantially elevated default probabilities of higher credit risk groups are more than offset by the damped prepayment propensities of those borrowers, resulting in significantly lower loan termination propensities overall. Indeed, among high default risk borrowers, at 5 years post loan origination, loan termination probabilities via prepayment are about 4 times those emanating from loan default. For the investor in FHA-insured mortgage pools, the estimated 5-year cumulative probability of mortgage termination among high default risk borrowers (at 37%) remains well below that value (at 60%) for low-default risk borrowers. Recognition of this advantage among investors should facilitate their greater willingness to acquire such loans and at more competitive pricing.

References

Archer, Wayne R., David C. Ling, and Gary A. McGill, (1996). "The Effect of Income and Collateral Constraints on Residential Mortgage Terminations," *Regional Science and Urban Economics*, 26, 235-261.

Archer, Wayne R., David C. Ling, and Gary A. McGill, (2001). "Prepayment Risks and Lower Income Mortgage Borrowers," University of Florida Working Paper.

Avery, R., R. Bostic, P. Calem, and G. Canner, (1996). "Credit Risk, Credit Scoring, and the Performance of Home Mortgages," *Federal Reserve Bulletin*, 82 (7), 621-648.

Bennett, P., R. Peach, and S. Peristiani, (2001) "Structural Change in the Mortgage Market and the Propensity to Refinance," *Journal of Money Credit and Banking*, 33 (4): 955-975.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan, (1994). "Race, Redlining, and Residential Mortgage Loan Performance," *Journal of Real Estate Finance and Economics*, 9, 263-294.

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy H. Hannan, (1998). "Discrimination, Competition, and Loan Performance in FHA Mortgage Lending," *Review of Economics and Statistics*, 241-250.

Bradley, Michael, Stuart Gabriel, and Mark Wohar, (1995). "The Thrift Crisis, Mortgage Credit Intermediation, and Housing Activity." *Journal of Money, Credit, and Banking*.

Canner, Glenn B., Wayne Passmore, and Delores S. Smith, (1994). "Residential Lending to Low-Income and Minority Families: Evidence from the 1992 HMDA Data," *Federal Reserve Bulletin* 80.

Cotterman, Robert F., (2001) "*Neighborhood Effects in Mortgage Default Risk*," U.S. Department of Housing and Urban Development, Office of Policy Development and Research

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

Deng, Yongheng, and John M. Quigley, (2001). "Woodhead Behavior and the Pricing of Residential Mortgages," *Lusk Center for Real Estate Working Paper Series*, University of Southern California.

Deng, Yongheng, John M. Quigley, and Robert. Van Order, (2000). "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options," *Econometrica*, 68 (2), 275-307.

Fabozzi, Frank J., editor (1195) *Handbook of Mortgage Backed Securities*, fourth edition, New York: McGraw-Hill.

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

Han, A. K., and J. A. Hausman (1990), "Flexible Parametric Estimation of Durationduration and Competing Risk Models," *Journal of Applied Econometrics*, 5(1), 1-28.

Kain, John F., and John M. Quigley, (1972). “Housing Market Discrimination, Homeownership, and Savings Behavior,” *American Economic Review* 62, 263-277.

Kau, James, and Donald Keenan (1995), “An Overview of Option-Theoretic Pricing of Mortgages”, *Journal of Housing Research*, 6(2) , 217-244.

Munnell, Alicia H., Lynn E. Browne, James McEneaney, and Geoffrey M. B. Tootell, (1996). “Mortgage Lending in Boston: Interpreting HMDA Data,” *American Economic Review*, 86, 25-53.

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

Quigley, John M., (1987). “Interest Rate Variations, Mortgage Prepayments and Household Mobility,” *Review of Economics and Statistics*, 69, 636-643.

Schwartz, E.S. and W. N. Torous (1993), “Mortgage Prepayment and Default Decisions: A Poisson Regression Approach”, *Journal of the American Real Estate and Urban Economics Association*, 21(4), 431-449.

Van Order, Robert, and Peter Zorn, (2001). “The Performance of Low Income and Minority Mortgages: A Tale of Two Options,” Freddie Mac Working Paper.

TABLE I. MEANS AND VARIANCES OF TIME-CONSTANT VARIABLES

Variables	Means (Variances)	Variables	Means (Variances)
Credit Scores < 620 (categorical variable)	0.1874 (0.158)	Mortgage Term < 30 Years (categorical variable)	0.0380 (0.038)
Credit Scores 620~680 (categorical variable)	0.3098 (0.224)	Central City Location (categorical variable)	0.4405 (0.262)
Credit Scores 680~740 (categorical variable)	0.3161 (0.228)	Rural (categorical variable)	0.0662 (0.064)
Black (categorical variable)	0.1240 (0.117)	First Time Home Buyer (categorical variable)	0.6711 (0.244)
Asian (categorical variable)	0.0168 (0.019)	New House (categorical variable)	0.0789 (0.077)
Hispanic (categorical variable)	0.1214 (0.130)	Unmarried Co-borrower (categorical variable)	0.1114 (0.103)
Others (categorical variable)	0.0205 (0.023)	Single Male (categorical variable)	0.1936 (0.162)
Loan-to-Value Ratio	0.9383 (0.034)	Single Female (categorical variable)	0.2022 (0.168)
Housing Exp. to Income Ratio 20~38% (categorical variable)	0.6447 (0.254)	Number of Dependents	0.7692 (1.288)
Housing Exp. to Income Ratio > 38% (categorical variable)	0.0103 (0.011)	Log Value of Liquid Assets	8.4707 (4.966)
Debt to Income Ratio 20~41% (categorical variable)	0.8041 (0.187)	Borrower Age < 25 (categorical variable)	0.1101 (0.101)
Debt to Income Ratio 41~53% (categorical variable)	0.1577 (0.137)	Borrower Age 25~35 (categorical variable)	0.4977 (0.267)
Debt to Income Ratio > 53% (categorical variable)	0.0084 (0.009)	Borrower Age 35~45 (categorical variable)	0.2614 (0.204)
Buydown (categorical variable)	0.0242 (0.025)	Log Value of Household Income	8.0379 (2.340)
Log of Property Appraisal Value	11.2102 (4.423)		
Number of Observations		12,021	

Note: Variances are in parentheses.

TABLE II. MEANS AND VARIANCES OF TIME-CONSTANT VARIABLES AT ORIGINATION AND TERMINATION

Variables	At Origination				At Termination	
	All Loans	Prepaid	Defaulted	Other*	Prepaid	Defaulted
Call Option (fraction of contract Value)	-0.0120 (0.007)	0.0001 (0.007)	-0.0022 (0.005)	-0.0265 (0.008)	0.0750 (0.007)	0.0352 (0.005)
Put Option (probability of negative equity)	0.2732 (0.057)	0.2694 (0.051)	0.2160 (0.062)	0.2869 (0.061)	0.0407 (0.013)	0.1011 (0.034)
SMSA Unemployment Rate (percent)	5.6901 (7.498)	5.7902 (5.526)	4.2433 (12.690)	5.8291 (8.333)	4.2383 (4.120)	3.4211 (8.283)
Log Value of Herfindahl-Hirschmann Index	5.8737 (1.500)	6.1111 (0.666)	3.7181 (6.669)	5.9866 (0.649)	5.9782 (0.494)	3.6548 (6.556)
Log Value of HH Index For Blacks	0.7038 (3.803)	0.4738 (2.714)	0.8284 (4.190)	0.9277 (4.789)	0.4672 (2.631)	0.8184 (4.113)
Log Value of HH Index For Hispanics	0.6968 (4.316)	0.4967 (3.092)	0.6857 (3.812)	0.9118 (5.617)	0.4875 (2.961)	0.6760 (3.727)
Number of Observations	12,021	5,730	913	5,378	5,730	913

Note: Variances are in parentheses.
*Other includes those outstanding at the end of the observation period.

TABLE III. MAXIMUM LIKELIHOOD ESTIMATES FOR COMPETING RISKS OF FHA MORTGAGE PREPAYMENT AND DEFAULT

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Call Option (fraction of Contract value)			5.087 (29.37)	2.278 (4.04)		
Put Option (probability of negative equity)			-2.307 (17.44)	1.746 (7.13)		
Interaction of Call Option Credit Scores < 620					4.554 (10.80)	2.450 (2.47)
Interaction of Call Option Credit Scores 620~680					4.292 (13.82)	1.690 (1.88)
Interaction of Call Option Credit Scores 680~740					5.676 (19.41)	1.777 (1.66)
Interaction of Call Option Credit Scores >740					6.002 (16.75)	2.386 (1.36)
Interaction of Put Option Credit Scores < 620					-3.462 (8.57)	2.109 (4.83)
Interaction of Put Option Credit Scores 620~680					-3.088 (11.57)	1.618 (4.41)
Interaction of Put Option Credit Scores 680~740					-2.302 (11.24)	2.007 (4.94)
Interaction of Put Option Credit Scores >740					-1.405 (6.59)	2.112 (3.08)
SMSA Unemployment Rate (percent)	-0.186 (25.07)	0.111 (6.16)	-0.136 (17.20)	0.058 (2.79)	-0.135 (17.07)	0.057 (2.75)
Log Value of Herfindahl Hirschmann Index	-0.048 (1.64)	-0.493 (4.20)	-0.099 (3.42)	-0.462 (3.90)	-0.110 (3.79)	-0.458 (3.82)
Interaction of Black and Log of H-H Index	0.185 (1.61)	-0.187 (0.74)	0.150 (1.32)	-0.216 (0.86)	0.147 (1.29)	-0.198 (0.77)
Interaction of Hispanic and Log of H-H Index	0.189 (1.68)	-0.244 (0.84)	0.306 (2.78)	-0.098 (0.33)	0.324 (2.92)	-0.118 (0.40)
Credit Scores < 620 (dummy)					-0.059 (0.89)	1.672 (5.75)
Credit Scores 620~680 (dummy)					0.019 (0.36)	1.347 (4.82)
Credit Scores 680~740 (dummy)					-0.012 (0.23)	0.702 (2.41)

TABLE III. (continued)

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Black (dummy)	-1.493 (2.30)	1.732 (1.23)	-1.395 (2.17)	1.895 (1.34)	-1.336 (2.08)	1.530 (1.07)
Asian (dummy)	-0.051 (0.51)	-0.057 (0.15)	0.010 (0.10)	-0.153 (0.38)	0.017 (0.17)	-0.160 (0.37)
Hispanic (dummy)	-1.292 (2.04)	1.663 (1.02)	-2.042 (3.28)	0.795 (0.48)	-2.129 (3.40)	0.819 (0.49)
Others (dummy)	-0.212 (2.01)	0.205 (0.74)	-0.321 (3.09)	0.091 (0.33)	-0.278 (2.68)	-0.098 (0.35)
Loan-to-Value Ratio	-2.095 (11.05)	1.630 (2.03)	-2.067 (12.00)	2.109 (2.58)	-2.106 (12.17)	1.921 (2.30)
Housing Exp. to Income 20~38% (dummy)	0.296 (7.00)	0.207 (1.43)	0.353 (8.20)	-0.167 (1.08)	0.359 (8.30)	-0.152 (0.96)
Housing Exp. to Income > 38% (dummy)	0.448 (2.92)	0.085 (0.17)	0.179 (1.16)	-0.028 (0.05)	0.172 (1.11)	0.160 (0.31)
Debt to Income Ratio 20~41% (dummy)	0.079 (0.98)	-0.275 (0.94)	-0.030 (0.38)	-0.277 (0.94)	-0.022 (0.27)	-0.418 (1.41)
Debt to Income Ratio 41~53% (dummy)	0.168 (1.90)	-0.051 (0.17)	0.036 (0.42)	(4.04) 1.746	0.064 (0.73)	-0.292 (0.94)
Debt to Income Ratio > 53% (dummy)	0.233 (1.32)	-0.556 (0.84)	-0.031 (0.19)	-0.065 (0.21)	-0.012 (0.07)	-0.898 (1.35)
Buydown (dummy)	0.293 (3.43)	0.076 (0.28)	0.226 (2.75)	-0.627 (0.94)	0.217 (2.67)	-0.012 (0.04)
Log Value of Property Appraisal Value	-0.399 (7.50)	0.059 (0.33)	-0.151 (2.86)	-0.060 (0.22)	-0.156 (2.94)	0.221 (1.13)
Mortgage Term < 30 Year (dummy)	-0.316 (4.36)	-1.131 (2.63)	-0.087 (1.19)	0.211 (1.11)	-0.099 (1.36)	-0.860 (1.95)
Central City Location (dummy)	0.028 (0.97)	-0.154 (1.70)	0.031 (1.10)	-0.909 (2.10)	0.033 (1.17)	-0.142 (1.53)
Rural (dummy)	0.007 (0.11)	-0.255 (1.28)	0.014 (0.25)	-0.156 (1.71)	0.019 (0.32)	-0.242 (1.20)
First Time Home Buyer (dummy)	-0.195 (6.42)	0.133 (1.34)	-0.207 (6.99)	-0.293 (1.48)	-0.196 (6.61)	0.063 (0.62)
New House (dummy)	-0.073 (1.40)	-0.099 (0.59)	0.018 (0.36)	0.128 (1.28)	0.014 (0.27)	-0.029 (0.17)

TABLE III. (continued)

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Unmarried Co-borrower (dummy)	-0.035 (0.73)	-0.107 (0.65)	-0.032 (0.70)	-0.062 (0.36)	-0.035 (0.76)	-0.076 (0.45)
Single Male (dummy)	-0.038 (0.96)	0.293 (2.54)	-0.079 (2.08)	-0.112 (0.67)	-0.074 (1.92)	0.275 (2.34)
Single Female (dummy)	-0.150 (3.60)	-0.213 (1.63)	-0.185 (4.60)	0.316 (2.72)	-0.182 (4.51)	-0.210 (1.56)
Number of Dependents	-0.065 (4.57)	0.131 (3.48)	-0.083 (5.96)	-0.182 (1.38)	-0.072 (5.21)	0.078 (1.99)
Log Value of Liquid Assets	0.013 (1.44)	-0.092 (3.22)	0.024 (2.61)	0.122 (3.26)	0.018 (2.02)	-0.067 (2.22)
Borrower Age < 25 (dummy)	0.367 (6.07)	0.227 (1.36)	0.392 (6.65)	-0.094 (3.27)	0.399 (6.78)	0.187 (1.10)
Borrower Age 25~35 (dummy)	0.236 (5.13)	-0.194 (1.44)	0.278 (6.15)	0.289 (1.71)	0.279 (6.16)	-0.203 (1.45)
Borrower Age 35~45 (dummy)	0.063 (1.28)	-0.181 (1.26)	0.071 (1.45)	-0.146 (1.06)	0.070 (1.43)	-0.169 (1.15)
Log Value of Household Income	0.636 (9.16)	-0.330 (1.45)	0.276 (3.97)	-0.122 (0.84)	0.297 (4.25)	-0.716 (2.95)
Log Likelihood	-36,473		-35,865		-35,757	

Note: T-ratios are in parentheses. All models are estimated by ML approach. Prepayment and default functions are considered as correlated competing risks and they are estimated jointly. Baseline functions for prepayment and default are specified as 100% PSA curve and 100% SDA curve, respectively. PSA curve stands for the Public Securities Association Prepayment Curve; and SDA curve stands for Bond Market Association (formerly the Public Securities Association) Standard Default Assumption curve. PSA curve and SDA curve are widely used prepayment and default benchmarks in the mortgage industry.

TABLE IV. UNADJUSTED CUMULATIVE PROBABILITY OF PREPAYMENT AND DEFAULT BY VARIOUS COVARIATES AT THE END OF ONE-, THREE-, AND FIVE-YEAR

	Prepayment			Default		
	End of Year1	End of Year 3	End of Year5	End of Year1	End of Year 3	End of Year5
By Liquid Asset						
Liquid Asset ≥ Median	2.36%	23.50%	40.13%	0.48%	2.31%	3.47%
Liquid Asset < Median	2.00%	18.49%	33.44%	0.75%	3.84%	5.53%
By LTV						
LTV < 95%	2.67%	21.56%	37.67%	0.58%	2.75%	4.01%
LTV ≥ 95%	1.87%	20.63%	36.22%	0.63%	3.28%	4.82%
By Buyers' Type						
Non First Time Buyer	3.26%	24.74%	41.18%	0.56%	2.56%	3.83%
First Time Buyer	1.65%	19.16%	34.63%	0.64%	3.33%	4.83%
By Housing Expense Ratio						
20% < HEI ≤ 38%	2.48%	23.05%	39.49%	0.21%	1.32%	2.31%
Otherwise	1.89%	18.92%	34.05%	1.01%	4.85%	6.72%
By Credit Score						
Credit Score ≥ 680	2.32%	21.42%	36.74%	0.46%	2.29%	3.39%
Credit Score < 680	2.11%	20.76%	36.80%	0.69%	3.50%	5.12%

Note: Unadjusted probabilities are calculated based on cleaned full sample.

TABLE V. PREDICTED CUMULATIVE PROBABILITY OF PREPAYMENT AND DEFAULT RISKS BY VARIOUS COVARIATES AT THE END OF ONE-, THREE-, AND FIVE-YEAR

	Prepayment			Default			Total Termination		
	End of Year1	End of Year3	End of Year 5	End of Year1	End of Year 3	End of Year5	End of Year1	End of Year 3	End of Year5
By Liquid Asset									
Liquid Asset ≥ Median	4.28%	20.70%	40.38%	0.55%	2.07%	3.30%	4.83%	22.78%	43.68%
Liquid Asset < Median	4.20%	20.36%	39.77%	0.59%	2.22%	3.54%	4.79%	22.58%	43.30%
By LTV									
LTV < 95%	5.64%	26.54%	49.80%	0.44%	1.60%	2.48%	6.07%	28.15%	52.28%
LTV ≥ 95%	4.14%	20.06%	39.27%	0.59%	2.21%	3.53%	4.72%	22.27%	42.80%
By Buyers' Type									
Non First Time Buyer	4.79%	22.92%	44.01%	0.55%	2.05%	3.22%	5.34%	24.96%	47.23%
First Time Buyer	3.95%	19.24%	37.87%	0.59%	2.22%	3.57%	4.54%	21.46%	41.44%
By Housing Exp/Inc Ratio									
20%<HEI≤38%	4.76%	22.82%	43.86%	0.54%	2.02%	3.18%	5.31%	24.85%	47.05%
Otherwise	3.36%	16.56%	33.17%	0.64%	2.44%	3.97%	4.00%	19.00%	37.14%
By Credit Score									
Credit Score ≥ 680	4.81%	22.58%	43.27%	0.37%	1.39%	2.19%	5.19%	23.97%	45.46%
Credit Score < 680	3.99%	19.89%	39.03%	0.99%	3.77%	6.01%	4.98%	23.66%	45.03%
By Prepayment Risk									
High Prepayment Risk	8.21%	35.97%	62.84%	0.25%	0.86%	1.26%	8.45%	36.82%	64.10%
Low Prepayment Risk	2.86%	14.59%	29.60%	1.18%	4.60%	7.51%	4.05%	19.19%	37.11%
By Default Risk									
High Default Risk	3.07%	15.58%	31.40%	1.16%	4.48%	7.27%	4.23%	20.05%	38.67%
Low Default Risk	7.27%	32.48%	58.20%	0.26%	0.91%	1.37%	7.53%	33.39%	59.57%

Note: The calculation of probability for each risk group is based on 10% random sample of the mortgage pools originated in June 1992. For each group, probabilities are evaluated at the mean value of each covariate (time-varying means are calculated for time-varying covariates) except for those specified in each risk category such that

(1) Liquid asset is evaluated at 75% and 25% quartiles of the sample for higher and lower liquid asset groups, respectively.

(2) LTV is set to 95% , and 80% for high LTV category, and low LTV category, respectively.

(3) Weighted average of borrowers with credit score between 680 and 740, and credit score above 740 is used to define high credit score group; and weighted average of borrowers with credit score between 620 and 680, and credit score below 620 is used to define low credit score group.
(4) The high prepayment risk group consists of borrowers with liquid asset above sample median, LTV under 95%, not first-buyer, housing expense to income ratio between 20% and 38%, and credit score above 680.
(5) The low prepayment risk group consists of borrowers with liquid asset below sample median, LTV above 95%, first-buyer, housing expense to income ratio greater than 38% , and credit score under 680.
(6) The high default risk group consists of borrowers with liquid asset below sample median, LTV above 95%, housing expense to income ratio above 38%, credit score under 680.
(7) The low default risk group consists of borrowers with liquid asset above sample median, LTV under 95%, housing expense to income ratio between 20% and 38%, credit score above 680.
(8) The cumulative probabilities of total termination at the end of year 1, 3, and 5 are the sum of predicted cumulative prepayment and default rates at the end of year 1, 3, and 5, respectively.

APPENDIX A.1. MAXIMUM LIKELIHOOD ESTIMATES FOR COMPETING RISKS OF FHA MORTGAGE PREPAYMENT AND DEFAULT WITH CENSUS TRACT CONTROLS

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Call Option (fraction of Contract value)			5.099 (29.44)	2.291 (4.05)		
Put Option (probability of negative equity)			-2.313 (17.47)	1.751 (7.14)		
Interaction of Call Option Credit Scores < 620					4.560 (10.81)	2.473 (2.48)
Interaction of Call Option Credit Scores 620~680					4.305 (13.87)	1.687 (1.86)
Interaction of Call Option Credit Scores 680~740					5.686 (19.45)	1.765 (1.64)
Interaction of Call Option Credit Scores >740					6.018 (16.78)	2.491 (1.41)
Interaction of Put Option Credit Scores < 620					-3.469 (8.58)	2.126 (4.90)
Interaction of Put Option Credit Scores 620~680					-3.090 (11.56)	1.617 (4.40)
Interaction of Put Option Credit Scores 680~740					-2.303 (11.23)	1.989 (4.91)
Interaction of Put Option Credit Scores >740					-1.416 (6.63)	2.164 (3.12)
SMSA Unemployment Rate (percent)	-0.185 (25.02)	0.111 (6.16)	-0.136 (17.13)	0.058 (2.77)	-0.135 (17.00)	0.057 (2.73)
Log Value of Herfindahl Hirschmann Index	-0.045 (1.54)	-0.487 (4.11)	-0.095 (3.28)	-0.456 (3.82)	-0.106 (3.65)	-0.448 (3.72)
Interaction of Black and Log of H-H Index	0.182 (1.58)	-0.197 (0.78)	0.146 (1.28)	-0.224 (0.89)	0.143 (1.25)	-0.211 (0.82)
Interaction of Hispanic and Log of H-H Index	0.186 (1.65)	-0.249 (0.86)	0.301 (2.72)	-0.101 (0.34)	0.319 (2.86)	-0.127 (0.42)
Credit Scores < 620 (dummy)					-0.056 (0.85)	1.687 (5.73)
Credit Scores 620~680 (dummy)					0.020 (0.38)	1.359 (4.82)
Credit Scores 680~740 (dummy)					-0.011 (0.22)	0.719 (2.45)

APPENDIX A.1. (continued)

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Black (dummy)	-1.478 (2.28)	1.792 (1.27)	-1.374 (2.14)	1.947 (1.38)	-1.315 (2.04)	1.612 (1.13)
Asian (dummy)	-0.053 (0.54)	-0.050 (0.13)	0.006 (0.06)	-0.137 (0.35)	0.014 (0.14)	-0.136 (0.32)
Hispanic (dummy)	-1.276 (2.01)	1.691 (1.03)	-2.012 (3.23)	0.817 (0.49)	-2.099 (3.34)	0.872 (0.52)
Others (dummy)	-0.215 (2.04)	0.206 (0.74)	-0.324 (3.12)	0.092 (0.33)	-0.281 (2.70)	-0.099 (0.36)
Loan-to-Value Ratio	-2.085 (10.94)	1.654 (2.05)	-2.058 (11.83)	2.140 (2.60)	-2.097 (11.99)	1.962 (2.33)
Housing Exp. to Income 20~38% (dummy)	0.295 (6.97)	0.205 (1.41)	0.350 (8.12)	-0.170 (1.10)	0.356 (8.23)	-0.159 (1.00)
Housing Exp. to Income > 38% (dummy)	0.447 (2.91)	0.069 (0.13)	0.177 (1.14)	-0.046 (0.09)	0.170 (1.09)	0.145 (0.28)
Debt to Income Ratio 20~41% (dummy)	0.083 (1.03)	-0.288 (0.98)	-0.023 (0.29)	-0.287 (0.97)	-0.014 (0.18)	-0.425 (1.44)
Debt to Income Ratio 41~53% (dummy)	0.172 (1.94)	-0.066 (0.21)	0.042 (0.49)	-0.078 (0.25)	0.070 (0.80)	-0.299 (0.96)
Debt to Income Ratio > 53% (dummy)	0.244 (1.39)	-0.554 (0.84)	-0.016 (0.09)	-0.624 (0.94)	0.004 (0.02)	-0.894 (1.34)
Buydown (dummy)	0.294 (3.42)	0.074 (0.27)	0.228 (2.74)	-0.064 (0.23)	0.219 (2.66)	-0.017 (0.06)
Log Value of Property Appraisal Value	-0.390 (7.26)	0.066 (0.36)	-0.136 (2.55)	0.219 (1.14)	-0.141 (2.64)	0.234 (1.19)
Mortgage Term < 30 Year (dummy)	-0.315 (4.34)	-1.124 (2.60)	-0.085 (1.16)	-0.902 (2.07)	-0.097 (1.32)	-0.853 (1.92)
Central City Location (dummy)	0.028 (0.93)	-0.143 (1.50)	0.032 (1.07)	-0.144 (1.50)	0.034 (1.16)	-0.123 (1.26)
Rural (dummy)	-0.002 (0.03)	-0.206 (1.02)	0.005 (0.08)	-0.237 (1.19)	0.010 (0.16)	-0.179 (0.88)
First Time Home Buyer (dummy)	-0.196 (6.45)	0.135 (1.35)	-0.208 (7.02)	0.130 (1.30)	-0.197 (6.64)	0.063 (0.63)
New House (dummy)	-0.074 (1.43)	-0.100 (0.59)	0.018 (0.35)	-0.064 (0.37)	0.014 (0.26)	-0.031 (0.17)

APPENDIX A.1. (continued)

	Model 1		Model 2		Model 3	
	Prepay	Default	Prepay	Default	Prepay	Default
Unmarried Co-borrower (dummy)	-0.034 (0.73)	-0.104 (0.63)	-0.031 (0.68)	-0.110 (0.66)	-0.034 (0.75)	-0.073 (0.43)
Single Male (dummy)	-0.037 (0.93)	0.300 (2.58)	-0.078 (2.03)	0.322 (2.76)	-0.072 (1.87)	0.284 (2.40)
Single Female (dummy)	-0.149 (3.57)	-0.209 (1.58)	-0.183 (4.52)	-0.178 (1.34)	-0.180 (4.44)	-0.203 (1.50)
Number of Dependents	-0.064 (4.55)	0.130 (3.46)	-0.082 (5.92)	0.122 (3.26)	-0.072 (5.18)	0.078 (1.97)
Log Value of Liquid Assets	0.013 (1.39)	-0.093 (3.22)	0.024 (2.57)	-0.094 (3.26)	0.018 (2.00)	-0.067 (2.20)
Borrower Age < 25 (dummy)	0.369 (6.10)	0.234 (1.39)	0.394 (6.67)	0.296 (1.74)	0.400 (6.80)	0.195 (1.13)
Borrower Age 25~35 (dummy)	0.235 (5.11)	-0.188 (1.38)	0.276 (6.10)	-0.138 (0.99)	0.277 (6.12)	-0.196 (1.38)
Borrower Age 35~45 (dummy)	0.064 (1.28)	-0.178 (1.23)	0.070 (1.43)	-0.120 (0.82)	0.069 (1.41)	-0.169 (1.14)
Log Value of Household Income	0.635 (9.14)	-0.329 (1.44)	0.273 (3.92)	-0.599 (2.50)	0.295 (4.21)	-0.721 (2.95)
Percentage of Black in Census Tract Population	0.035 (0.44)	-0.022 (0.08)	0.009 (0.12)	-0.018 (0.07)	-0.002 (0.02)	-0.000 (0.00)
Percentage of Asian in Census Tract Population	-0.415 (1.13)	1.114 (1.22)	-0.499 (1.32)	1.200 (1.28)	-0.441 (1.16)	1.127 (1.14)
Percentage of Hispanic in Census Tract Population	0.017 (0.16)	0.075 (0.21)	0.057 (0.57)	0.055 (0.15)	0.055 (0.55)	0.062 (0.17)
Percentage of Others in Census Tract Population	0.757 (0.72)	-5.596 (0.94)	0.922 (0.85)	-5.693 (0.96)	0.911 (0.86)	-6.232 (1.02)
Census Tract to MSA Median Income Ratio	-0.100 (1.59)	-0.116 (0.58)	-0.141 (2.28)	-0.119 (0.59)	-0.141 (2.28)	-0.174 (0.84)
Census Tract Rental Ratio	-0.088 (0.91)	-0.038 (0.13)	-0.114 (1.20)	-0.059 (0.20)	-0.110 (1.16)	-0.137 (0.45)
Log Likelihood	-36,467		-35,857		-35,749	

Note: T-ratios are in parentheses. All models are estimated by ML approach. Prepayment and default functions are considered as correlated competing risks and they are estimated jointly. Baseline functions for prepayment and default are specified as 100% PSA curve and 100% SDA curve, respectively. PSA curve stands for the Public Securities Association Prepayment Curve; and SDA curve stands for Bond Market Association (formerly the Public Securities Association) Standard Default Assumption curve. PSA curve and SDA curve are widely used prepayment and default benchmarks in the mortgage industry.