Institute of Business and Economic Research
University of California at Berkeley

FISHER CENTER FOR REAL ESTATE AND URBAN ECONOMICS

WORKING PAPER SERIES

WORKING PAPER NO. 97-255

MORTGAGE TERMINATIONS, HETEROGENEITY AND THE EXERCISE OF MORTGAGE OPTIONS

By

Yongheng Deng

John Quigley

Robert VanOrder

WALTER A. HAAS SCHOOL OF BUSINESS

Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options*

Yongheng Deng

Office of Federal Housing Enterprise Oversight

John M. Quigley

University of California, Berkeley

Robert Van Order

Freddie Mac

May 1997

* The authors thank Brian McCall for providing us the Fortran code used in his 1996 Econometrica paper. The paper benefitted from the comments of Jushan Bai, David Card, Glenn Sueyoshi, and two anonymous referees. The views expressed in this research are those of the authors and do not represent the policies or positions of the Office of Federal Housing Enterprise Oversight or Freddie Mac. Quigley's research is supported by the Fisher Center for Real Estate and Urban Economics, University of California, Berkeley.

Abstract

As applied to the behavior of homeowners with mortgages, option theory predicts that mortgage prepayment or default will be exercised if the call or put option is "in the money" by some specific amount. The empirical evidence in this paper indicates that the financial value of the call option is strongly associated with exercise of the prepayment option, and the probability that the put option is in the money is strongly associated with exercise of the default option. Nevertheless, the evidence also shows that some borrowers may not behave as "ruthlessly" as the option theory predicts.

Our analysis: tests the extent to which the option approach can explain default and prepayment behavior; evaluates the practical importance of modeling both options simultaneously; and models the unobserved heterogeneity of borrowers in the home mortgage market. The paper presents a unified model of the competing risks of mortgage termination by prepayment and default. The model considers these two hazards as dependent competing risks and estimates them jointly. It also accounts for the unobserved heterogeneity among borrowers, and estimates the unobserved heterogeneity simultaneously with the parameters and baseline hazards associated with prepayment and default functions.

Our results show that the option model, in its most straightforward version, does a good job of explaining default and prepayment; but it is not enough by itself. The simultaneity of the options is very important empirically in explaining behavior. The results also show that there exists significant heterogeneity among mortgage borrowers. Ignoring this heterogeneity results in serious errors in estimating the prepayment by homeowners.

Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options

1. Introduction

The mortgage market is quite large and is increasing in importance. The outstanding volume of residential mortgages is currently over $ 3 trillion, and volume has doubled in the past decade. In comparison, the stock of outstanding U.S. government debt is currently about $ 5 trillion. Almost half of the stock of mortgages is held in "mortgage-backed securities," and about half of all new mortgages are "securitized." The rise of securitization, the trading of these securities, and the growing use of mortgage-backed securities as collateral for "derivatives" (*e.g.,* collateralized mortgage obligations) has generated a great deal of interest in the economics of mortgage and mortgage-backed securities.

Pricing mortgage contracts is complicated, primarily by the options available to the borrower to default or to prepay. These options are distinct, but not independent. Thus, one cannot calculate accurately the economic value of the default option without considering simultaneously the financial incentive for prepayment. Furthermore, risk preferences and other idiosyncratic differences across borrowers may vary widely. Typically, it is very difficult to measure this heterogeneity explicitly. Appropriately valuing these prepayment and default risks is crucial to the pricing of mortgages and to understanding the economic behavior of homeowners.

The contingent claims models, developed by Black and Scholes (1972), Cox, Ingersoll, and Ross (1985), and others, provide a coherent motivation for borrower behavior, and a number of studies have applied this model to the mortgage market. Hendershott and Van Order (1987) and Kau and Keenan (1995) have surveyed much of the literature related to mortgage pricing.

Virtually all previous studies using option models, however, focus on applying them to explain either prepayment or default behavior, but not both. For instance, in the first application of option models to mortgages, Findley and Capozza (1977) analyzed the prepayment options of holders of adjustable-rate and fixed-rate mortgages. Dunn and McConnell (1981), Buser and Hendershott (1984), and Brennan and Schwartz (1985) used option pricing to price callable mortgages, relying on simulation methods. Green and Shoven (1986), Schwartz and Torous (1989) and Quigley and Van Order (1990) provided empirical estimates of option-based prepayment models.

Cunningham and Hendershott (1984) and Epperson, Kau, Keenan, and Muller III (1985) applied option models to price default risk, modeling default as a put option, and Foster and Van Order (1984) and Quigley and Van Order (1995) estimated default models empirically in an option-based framework. Quercia and Stegman (1992) and Vandell (1993) reviewed many of these default models.

A series of papers by Kau, Keenan, Muller, and Epperson (1992, 1995), Kau and Keenan (1996) and Titman and Torous (1989) provided theoretical models which emphasized the theoretical importance of the jointness of prepayment and default options. A homeowner who exercises the default option today gives up the option to default in the future, but she also gives up the option to prepay the mortgage. Foster and Van Order (1985) estimated simultaneous models of default and prepayment using data on large pools of FHA loans, and Schwartz and Torous (1993) estimated the joint hazard using a Poisson regression approach and aggregate data. Deng, Quigley, and Van Order (1996) and Deng (1997) were the first to analyze residential mortgage prepayment and default behavior using micro data on the joint choices of individuals. However, the competing risks hazard model common to all these studies ignores the heterogeneity among borrowers. Presumably a substantial number of homeowners are less likely to exercise put and call options on mortgages in the fully rational way predicted by finance theory. Accounting for this group is potentially important in understanding market behavior.

In this paper, we present a unified economic model of the competing risks of mortgage termination by prepayment and default. We adopt a proportional hazard framework to analyze these competing risks empirically, using a large sample of individual loans, and we extend the model to analyze unobserved heterogeneity. We test three aspects of homeowner behavior in the mortgage market:

1. The extent to which the option approach can explain the default and prepayment behavior of borrowers with single family mortgages;
2. The importance of modeling both options simultaneously; and
3. The importance of heterogeneity of borrowers in explaining behavior in the market.

We find that:

1. The option model, in its most straightforward version, does a good job of explaining default and prepayment; but it is not enough by itself. Either transactions costs are quite large and vary a great deal across borrowers, or else some people are simply much worse at exercising options;
2. The simultaneity of the options is very important empirically in explaining behavior. In particular, factors that trigger one option are also important in triggering or foregoing exercise of the other; and
3. Unobserved borrower heterogeneity is quite important in accounting for borrower behavior. We allow for heterogeneity by incorporating into the estimation the possibility that there are two sorts of borrowers, those who are more astute ("ruthless players"), and those who are less astute ("woodheads[1]"). We find that heterogeneity is significant. It has important effects on key elasticities explaining behavior, particularly with respect to prepayment.

[1]The term woodhead is due to Ann Dougherty.

The paper is organized as follows: section 2 reviews the application of option models to mortgage terminations. Section 3 discusses the proportional hazard model, specified with competing risks, time-varying covariates and unobserved heterogeneity. Section 4 presents an extensive empirical analysis. Section 5 is a brief conclusion.

2. Mortgage Terminations and Option Pricing

Well-informed borrowers in a perfectly competitive market will exercise financial options when they can thereby increase their wealth. In the absence of either transactions costs or reputation costs which reduce credit ratings, and with assumable mortgages (or no exogenous reasons for residential mobility), default and prepayment are essentially financial decisions which can be separated from real (housing) decisions, and the simplest version of the Miller and Modigliani theory of the irrelevance of financial structure holds.[2] Under these conditions, individuals can increase their wealth by defaulting on a mortgage when the market value of the mortgage equals or exceeds the value of the house. Similarly, by prepaying the mortgage when market value equals or exceeds par, they can increase wealth by refinancing. A necessary condition for exercising an option is that it be "in the money," but that is not sufficient. Exercising either option now means giving up the option to exercise both options later. For instance, a borrower whose house price declines below the mortgage balance may not default immediately, in part because after the price decline the mortgage has a below market rate,[3] but also because by defaulting, the borrower would also lose the option to refinance later on.

While virtually all the recent research on prepayment and default, summarized above, has used option-based models, the underlying theories behind the models differ importantly in the treatment

[2] See Kau, Keenan, Muller, and Epperson (1995) for a recent discussion.

[3] This is because mortgage is now riskier, defaulting and buying back the same house would require paying a higher rate or making a larger downpayment.

of transactions costs.[4] For simplicity, we divide these approaches into polar cases. The first case, Model I, assumes no transactions costs (see Titman and Torous, 1989 and Kau et al., 1992), and "ruthless" exercise of both options. The second case, Model II, emphasizes transactions costs, particularly in exercise of the default option. It is assumed that transactions costs are sufficiently high that default requires, not only negative equity, but also a "trigger event" which forces the borrower to leave the house. Model II also entertains the possibility that there are significant transactions costs involved in prepaying, or else that some borrowers are more astute than others at exercising options (see Archer et al., 1996). Finally, Model II also allows the possibility of significant unobserved heterogeneity. Thus the separation between housing and finance decisions is incomplete.

According to Model I, understanding when to exercise either option requires specifying the underlying state variables and the parameters that determine the value of the contract and then deducing the rule for exercise that maximizes borrower wealth. For residential mortgages, the key state variables are interest rates and house values. The value of a mortgage $M(c, r, H, \tau, \Gamma)$ depends upon the coupon rate, c, a vector of relevant interest rates, r, property value, H, the outstanding balance, τ, the remaining time to maturity, Γ, and some other parameters. With continuous time, a standard arbitrage argument is sufficient to derive an equilibrium condition for M (a second order partial differential equation) such that the value of the mortgage equals the risk-adjusted expected present value of its net cash flows.

Assume that house price changes are continuous with an instantaneous mean μ and a standard deviation σ_h. Let d be the imputed rent payout ("dividend") rate. For simplicity, assume there is only one interest rate, the instantaneous short rate r, which determines the yield curve. Let θ be the mean value of the short rate, γ be the rate of convergence for the short rate, σ_r be the volatility

[4] See Kau, Keenan, and Kim (1993) for an explicit discussion of transaction costs.

of the short rate, and ρ be the correlation between interest rate changes and house price changes. Then it has been shown (Kau, Keenan, Muller, and Epperson, 1995) that the value of the mortgage M satisfies

$$\begin{aligned} &\tfrac{1}{2} r\sigma_r^2 \frac{\partial^2 M}{\partial r^2} + \rho\sqrt{r} H \sigma_r \sigma_h \frac{\partial^2 M}{\partial r \partial H} + \tfrac{1}{2} H^2 \sigma_h^2 \frac{\partial^2 M}{\partial H^2} + \gamma(\theta - r)\frac{\partial M}{\partial r} \\ &\quad + (r-d) H \frac{\partial M}{\partial H} + \frac{\partial M}{\partial \tau} - rM = 0. \end{aligned} \tag{2.1}$$

The value of $M(\cdot)$ and the optimal default and prepayment strategy are determined simultaneously. Equation (2.1) is consistent with an infinite number of functions $M(\cdot)$. The appropriate function is determined by choosing the optimal level of r, r^*, and the optimal level of H, H^*, at which to terminate the mortgage through default or prepayment. These are the levels of r and H that minimize M given equation (2.1) (See Kau, Keenan, Muller, and Epperson, 1995); these levels are functions of c, d, τ, Γ and the parameters governing the stochastic processes for r and H. Due to the jointness of the options, there are two pairs of r and H that trigger termination. There are levels of r that trigger default as well as prepayment, and levels of H that trigger prepayment as well as default. For instance, a borrower with positive equity may default at a low enough level of r as a means of prepayment, and a borrower might refinance when equity value has risen because the loan is now safer and would carry a lower interest rate. The probability of default or prepayment is the probability of these levels of r and H occurring, conditional on the information set of actors in the market.

Note that the borrower does not have to solve (2.1) and the boundary conditions in order to know when to exercise either option. All that is necessary is knowledge of market prices. For instance, for a fixed rate mortgage, the prepayment option should be exercised whenever the borrower can refinance the loan for the same remaining term at par at a mortgage rate less than the coupon on the current loan or, alternatively when the market value of the mortgage equals (or exceeds) the mortgage balance. Default should be exercised when the borrower's payments would be lower on a new zero-downpayment loan for the same remaining term, used to purchase the same house. Of course, we on the outside do not observe these market alternatives (and markets are not that complete anyway); this greatly complicates testing the model.

Due to "data limitations," the analyst does not observe the critical levels of house price and the mortgage rate that trigger exercise from the details of the mortgage contract. All that we can hope to observe is the extent to which either option is "in the money." But we cannot even measure directly the extent to which the default option is in the money without data on the course of individual house prices. We can however estimate the probability that the option is in the money, given the initial loan-to-value ratio and the stochastic process for house prices. The analyst can control for the remaining term of the loan, but not for changes in the parameters of the house price or interest rate process. This reality suggests that it is more productive to consider optimal exercise in probabilistic terms and then to test some of the major predictions of Model I: First, the probability of exercise should increase as the option moves further into the money. Second, we should expect that the probability of exercise will accelerate as the option moves further into the money. Third, because exercising one option means giving up the other option, the "in-the-moneyness" of the one should have an effect on exercise of the other. Thus, for example, the probability of prepayment is a function of the extent to which the default option is in the money.

Model I can be extended to address asymmetric information. For instance, we cannot observe directly the parameters governing house price volatilities. This can be a problem if the volatilities vary in a systematic way, for instance if borrowers know more about their own house price volatility than lenders do. Then risky houses might be financed with high loan-to-value (LTV) loans, as borrowers exploit underpriced options.[5] One may control for this by using initial LTV as an explanatory variable in predicting defaults.

[5]See Yezer et al. (1994) for a discussion.

Model I has the great advantage of simplicity.

Model II incorporates transactions costs in a broad sense. It is not simple because transaction costs are complicated and are generally not observable. For instance, different transaction costs across borrowers have been used to explain the observation that the prepayments in mortgage pools tend to be slower than expected and drawn out over time (e.g., see Stanton, 1995, Harding, 1994 and Archer and Ling, 1993). This raises the general question of unobserved differences among mortgage holders. Whether this empirical finding arises from variations in transaction costs or differences in the astuteness of homeowners exercising options, unobserved heterogeneity means that surviving borrowers are systematically different over time. For instance, surviving borrowers may be increasingly less interest-rate sensitive over time if more astute borrowers refinance first.

Transactions costs are more complicated on the default side, particularly if the mortgage is not assumable. A borrower forced to move (e.g., due to divorce or job loss) who cannot have the mortgage assumed has a very short remaining term and may thus default with little negative equity. On the other hand, if there are costs to defaulting, H^* may be lower than model I implies. For these reasons, many researchers (see Quigley and Van Order, 1995 for a discussion) estimate modified option models, which predict that exercise is a function of both "trigger events" like default or divorce and also the extent to which the option is in the money.

We follow Kau and Keenan (1996) who introduce random terminations into the model. These terminations force either a prepayment or a default. If mortgages are not assumable[6] and there are no (e.g., reputation) costs to default, a random termination will lead to default if the house is worth less than the mortgage balance, and prepayment otherwise. Note that in Model II, it is the par value of the mortgage that is relevant for default. In contrast, in Model I a borrower is less

[6] The empirical analysis below is based on mortgages which were nominally not assumable, but some states forbade exercise of due on sale clauses during the observation period, and in any event due on sales clauses were typically not enforced.

likely to default when interest rates increase due to the value of the low-rate mortgage. According to model II, a borrower who is forced to leave the house does not have the option to keep the mortgage alive. As is the case with prepayments, transactions costs matter, especially if they vary across borrowers or if there are unobservable differences in astuteness among borrowers.

Estimates of default and prepayment are reported below in four stages. First, we estimate proportional hazard models which use as explanatory variables only the extent to which the options are "in the money," in order to test the predictions of Model I. Second, we add variables that are proxies for trigger events (e.g., unemployment and divorce) and information asymmetry (e.g., the original loan-to-value ratio). Third, we allow for unobserved heterogeneity and estimate the nonparametric distribution of the unobserved heterogeneity simultaneously with the competing risks of prepayment and default functions. Finally, we assume that the probability that an individual borrower belongs to the "woodheads" group also affects the estimation of the competing risks of prepayment and default. In order to correct for this selection bias, we estimate the competing risks model using a selectivity-corrected maximum likelihood approach.

3. A Competing Risks Model of Mortgage Termination with Unobserved Heterogeneity

The proportional hazard model introduced by Cox (Cox and Oakes, 1984) provides a convenient framework for considering the exercise of options empirically and the importance of other trigger events in mortgage terminations. We start with a brief review of the hazard model. Then we present the competing risks model used to analyze mortgage terminations by prepayment and default.

Let $T \in R^{+}$ be a continuous random variable which measures the duration of stay, *i.e.*, the length of time since a mortgage was originated. If each individual enters the state at the same calendar time (*i.e.*, all take out mortgages on the same day), then there is no difference between duration and calendar time. In general, however duration is not the same as calendar time.

Define

$$S(t) = \Pr(T \geq t) \tag{3.1}$$

as the survivor function. The probability density function of the random variable t is:

$$\begin{aligned} f(t) &= \lim_{\triangle t \to 0^+} \frac{\Pr(t \leq T < t + \triangle t)}{\triangle t} \\ &= \frac{-dS(t)}{dt}. \end{aligned} \tag{3.2}$$

Define a hazard function that specifies the instantaneous rate of failure at $T = t$ conditional upon survival to time T such that

$$\begin{aligned} h(t) &= \lim_{\triangle t \to 0^+} \frac{\Pr(t \leq T < t + \triangle t | T \geq t)}{\triangle t} \\ &= \frac{f(t)}{S(t)} \\ &= \frac{-d \ln S(t)}{dt}. \end{aligned} \tag{3.3}$$

The relationships among the density function $f(t)$, the survivor function $S(t)$, and the hazard function $h(t)$ are thus:

$$S(t) = \exp\left(-\int_0^t h(k)\, dk\right), \tag{3.4}$$

$$f(t) = h(t) \exp\left(-\int_0^t h(k)\, dk\right). \tag{3.5}$$

Since mortgage termination data are collected in a discrete pattern, a grouped discrete-time approach is applied in the empirical part of this paper.

Han and Hausman (1990), Sueyoshi (1992) and McCall (1996) (HHSM, for short) suggested a maximum likelihood estimation approach for the proportional hazard model with grouped duration data. The HHSM approach estimates the competing risks simultaneously, accounts for the fact that risks may be correlated, also that covariates may be time-varying. This implies no restriction on the baseline hazard function (described below). Following HHSM, the competing risks model for mortgage prepayment and default can be derived:

Let T_p and T_d be the discrete random variables representing the duration of a mortgage until it is terminated by the mortgage holder in the form of prepayment or default, respectively. The joint survivor function conditional on η_p, η_d, r, H, Y and X can be expressed in the following form:

$$\begin{aligned} S\left(t_p, t_d \mid r, H, Y, X, \eta_p, \eta_d\right) &= \exp\left(-\eta_p \sum_{k=1}^{t_p} \exp\left(\gamma_{pk} + g_{pk}\left(r, H, Y\right) + \beta_p' X\right)\right. \\ &\qquad \left. -\eta_d \sum_{k=1}^{t_d} \exp\left(\gamma_{dk} + g_{dk}\left(r, H, Y\right) + \beta_d' X\right)\right), \end{aligned} \tag{3.6}$$

where $g_{jk}\left(r, H, Y\right)$ are time-varying functions of options-related variables;[7] r and H are the relevant interest rates and property values, respectively, as discussed in the previous section; Y is a vector of other variables which will be used, together with r and H, to estimate the market values of the options empirically; X is a vector of other non-option-related variables, which may include indicators reflecting a borrower's credit risk or financial strength, as well as other trigger events, such as unemployment and divorce. To simplify the notation, we suppress the time-varying subscripts for r, H, Y and X. γ_{jk} are parameters of the baseline function which may be estimated non-parametrically, following Han and Hausman (1990)

$$\gamma_{jk} = \log\left[\int_{k-1}^{k} h_{0j}\left(t\right) dt\right], \qquad j = p, d. \tag{3.7}$$

η_p and η_d are unobserved heterogeneities associated with the hazard functions for prepayment and default respectively.

[7]The details of the function $g_{jk}\left(r, H, Y\right)$ are specified in the next section.

We allow for the possibility that the population of mortgage borrowers consists of two distinct groups: "woodheads," who may not terminate mortgages when their options are "in the money," and "ruthless players," who behave just as ruthlessly as the option theory predicts. The joint distribution of the unobservables $\left(\eta_p, \eta_d\right)$ is modeled by assuming that these distinct, but unobserved types of individuals, $\tau = 1, 2$, (an individual in group τ is characterized by the doublet of location parameters $\left(\eta_{p\tau}, \eta_{d\tau}\right)$), occur in the population with relative frequency p_τ, $\tau = 1, 2$.

Due to the nature of the competing risks between prepayment and default, only the duration associated with the type which terminates first is observed, *i.e.* $t = \min\left(t_p, t_d\right)$. Define $F_p\left(k \mid \eta_p, \eta_d\right)$ as the probability of mortgage termination by prepayment in period k, $F_d\left(k \mid \eta_d, \eta_d\right)$ as the probability of mortgage termination by default in period k, $F_m\left(k \mid \eta_p, \eta_d\right)$ as the probability of mortgage termination in period k but information on the cause of the termination is missing, and $F_c\left(k \mid \eta_p, \eta_d\right)$ as the probability that mortgage duration data are censored in period k due to the ending of the data collecting period.

Following McCall (1996), these probabilities can be expressed as:

$$\begin{aligned} F_p\left(k \mid \eta_p, \eta_d\right) &= S\left(k, k \mid \eta_p, \eta_d\right) - S\left(k+1, k \mid \eta_p, \eta_d\right) - \frac{1}{2} S\left(k, k \mid \eta_p, \eta_d\right) \\ &\quad + S\left(k+1, k+1 \mid \eta_p, \eta_d\right) - S\left(k, k+1 \mid \eta_p, \eta_d\right) - S\left(k+1, k \mid \eta_p, \eta_d\right), \end{aligned} \tag{3.8}$$

$$\begin{aligned} F_d\left(k \mid \eta_p, \eta_d\right) &= S\left(k, k \mid \eta_p, \eta_d\right) - S\left(k, k+1 \mid \eta_p, \eta_d\right) - \frac{1}{2} S\left(k, k \mid \eta_p, \eta_d\right) \\ &\quad + S\left(k+1, k+1 \mid \eta_p, \eta_d\right) - S\left(k, k+1 \mid \eta_p, \eta_d\right) - S\left(k+1, k \mid \eta_p, \eta_d\right), \end{aligned} \tag{3.9}$$

$$F_m\left(k \mid \eta_p, \eta_d\right) = S\left(k, k \mid \eta_p, \eta_d\right) - S\left(k+1, k+1 \mid \eta_p, \eta_d\right), \tag{3.10}$$

and

$$F_c\left(k \mid \eta_p, \eta_d\right) = S\left(k, k \mid \eta_p, \eta_d\right), \tag{3.11}$$

where the dependence of these functions on r, H, Y and X has been omitted for notational simplicity.

The unconditional probability is given by

$$F_j(k) = \sum_{\tau=1}^{2} p_\tau F_j\left(k \mid \eta_{p\tau}, \eta_{d\tau}\right), \qquad j = p, d, m, c. \tag{3.12}$$

The log likelihood function of the competing risks model is given by

$$\log L = \sum_{i=1}^{N} \delta_{pi} \log\left(F_p\left(K_i\right)\right) + \delta_{di} \log\left(F_d\left(K_i\right)\right) + \delta_{mi} \log\left(F_m\left(K_i\right)\right) + \delta_{ci} \log\left(F_c\left(K_i\right)\right), \tag{3.13}$$

where δ_{ji}, $j = p, d, m, c$ are indicator variables which take value one if the ith loan is terminated by prepayment, default, unknown type, or censoring, respectively, and take a value of zero otherwise.

Importantly, the probability that an individual borrower belongs to either one of the groups may also affect the estimation of the risks model. Define the probability that an individual is a "woodhead" as

$$P\left(WH = 1 \mid \eta_w\right) = 1 - \exp\left(-\eta_w \exp\left(\beta_w' Z\right)\right), \tag{3.14}$$

where Z is a vector of regressors, not including a constant term, and η_w is an unmeasured random variable affecting "woodhead" behavior — which is possibly correlated with η_p and η_d but is distributed independently of Z.

Let

$$G_j\left(k, WH = 1 \mid \eta_w, \eta_p, \eta_d\right) = F_j\left(k \mid \eta_p, \eta_d\right) P\left(WH = 1 \mid \eta_w\right), \qquad j = p, d, m, c, \tag{3.15}$$

and

$$G_j\left(k, WH=0 \mid \eta_w, \eta_p, \eta_d\right) = F_j\left(k \mid \eta_p, \eta_d\right)\left(1 - P\left(WH=1 \mid \eta_w\right)\right), \qquad j = p, d, m, c. \quad (3.16)$$

Again, the joint distribution of the unobservables $\left(\eta_w, \eta_p, \eta_d\right)$ is modeled by assuming that there are two distinct but unobserved types of individuals in the population, where individual τ is characterized by the triplet of location parameters $\left(\eta_{w\tau}, \eta_{p\tau}, \eta_{d\tau}\right)$ and occurs in the population with relative frequency p_τ, $\tau = 1, 2$. The unconditional probabilities for the "woodheads" and the "ruthless players," respectively, are

$$G_j\left(k, WH=1\right) = \sum_{\tau=1}^{2} p_\tau G_j\left(k, WH=1 \mid \eta_{w\tau}, \eta_{p\tau}, \eta_{d\tau}\right), \qquad j = p, d, m, c. \qquad (3.17)$$

and

$$G_j\left(k, WH=0\right) = \sum_{\tau=1}^{2} p_\tau G_j\left(k, WH=0 \mid \eta_{w\tau}, \eta_{p\tau}, \eta_{d\tau}\right), \qquad j = p, d, m, c. \qquad (3.18)$$

Finally, the log likelihood function of the selectivity-corrected competing risks model is given by

$$\begin{aligned}
\log L \;=\; & \sum_{i=1}^{N} WH_i \delta_{pi} \log\left(G_p\left(K_i, WH=1\right)\right) + WH_i \delta_{di} \log\left(G_d\left(K_i, WH=1\right)\right) \qquad (3.19)\\
& + WH_i \delta_{mi} \log\left(G_m\left(K_i, WH=1\right)\right) + WH_i \delta_{ci} \log\left(G_c\left(K_i, WH=1\right)\right)\\
& + \left(1 - WH_i\right) \delta_{pi} \log\left(G_p\left(K_i, WH=0\right)\right) + \left(1 - WH_i\right) \delta_{di} \log\left(G_d\left(K_i, WH=0\right)\right)\\
& + \left(1 - WH_i\right) \delta_{mi} \log\left(G_m\left(K_i, WH=0\right)\right) + \left(1 - WH_i\right) \delta_{ci} \log\left(G_c\left(K_i, WH=0\right)\right),
\end{aligned}$$

where WH_i is an indicator variable which takes value of one if the ith individual has experienced an "in-the-money" option but "chose" not to exercise the option.

4. The Empirical Analysis

The empirical analysis is based upon individual mortgage history data maintained by the Federal Home Loan Mortgage Corporation (Freddie Mac). The data base contains 1,489,372 observations on single family mortgage loans issued between 1976 to 1983 and purchased by Freddie Mac. All are fixed-rate, level-payment, fully-amortized loans, most of them with thirty-year terms. The mortgage history period ends in the first quarter of 1992. For each mortgage loan, the available information includes the year and month of origination and termination (if it has been closed), indicators of prepayment or default, the purchase price of the property, the original loan amount, the initial loan-to-value ratio, the mortgage contract interest rate, the monthly principal and interest payment, the state, the region and the major metropolitan area in which the property is located. For the mortgage default and prepayment model, censored observations include all matured loans as well as the loans active at the end of the period.

The analysis is confined to mortgage loans issued for owner occupancy, and includes only those loans which were either closed or still active[8] at the first quarter of 1992. The analysis is confined to loans issued in 30 major metropolitan areas (MSAs) — a total of 447,042 observations. Loans are observed in each quarter from the quarter of origination through the quarter of termination, maturation, or through 1992:I for active loans.

The key variables are those measuring the extent to which the put and call options are in the money. To value the call option, the current mortgage interest rate and the initial contract terms are sufficient. We compute a variable "$POPTION$" measuring the ratio of the present discounted value of the unpaid mortgage balance at the current quarterly mortgage interest rate[9] relative to

[8] It excludes those observations which were in delinquency or foreclosure at the time data were collected.

[9] The rate used is the average interest rate charged by lenders on new first mortgages reported by Freddie Mac's quarterly market survey. This mortgage interest rate varies by quarter across five major US regions.

the value discounted at the contract interest rate.[10]

To value the put option analogously, we would measure the market value of each house quarterly and to compute homeowner equity quarterly. Obviously, we do not observe the course of price variation for individual houses in the sample. We do, however, have access to a large sample of repeat (or paired) sales of single family houses in these 30 metropolitan areas (MSAs). This information is sufficient to estimate a weighted repeat sales house price index (WRS) separately for each of the 30 MSAs. The WRS index provides estimates of the course of house prices in each metropolitan area. It also provides an estimate of the variance in price for each house in the sample, by metropolitan area and elapsed time since purchase.[11]

---

[10]Specifically, POPTION for the lth loan observation is defined as

$$\begin{aligned} poption_l &= \frac{\sum_{t=1}^{term_l-\tau_i} \frac{mopipmt_l \times 3}{(1+mktrate_{\omega_l,\kappa_l+\tau_i}/400)^t} - \sum_{t=1}^{term_l-\tau_i} \frac{mopipmt_l \times 3}{(1+noterate_l/400)^t}}{\sum_{t=1}^{term_l-\tau_i} \frac{mopipmt_l \times 3}{(1+mktrate_{\omega_l,\kappa_l+\tau_i}/400)^t}} \\ &= 1 - \frac{mktrate_{\omega_l,\kappa_l+\tau_i} \times \left(1-\left(\frac{1}{1+noterate_l/400}\right)^{term_l-\tau_i}\right)}{noterate_l \times \left(1-\left(\frac{1}{1+mktrate_{\omega_l,\kappa_l+\tau_i}/400}\right)^{term_l-\tau_i}\right)}, \end{aligned} \tag{4.1}$$

where τ_i is loan age measured in quarters, ω_l is a vector of indices for geographical location, κ_l is loan origination time, $mopipmt_l$ is monthly principal and interest payment, $noterate_l$ is mortgage note rate, $mktrate_{\omega_l,\kappa_l+\tau_i}$, is the current local market interest rate, and $term_l$ is mortgage loan term calculated by

$$term_l = \frac{\log\left(\frac{mopipmt_l}{origamt_l \times (noterate_l/1200) + mopipmt_l}\right)}{\log(1+noterate_l/1200) \times 3}, \tag{4.2}$$

where $origamt_l$ is original loan amount.

[11]Housing price indices and their volatilities are estimated according to the three stage procedure suggested by Case and Shiller (1987)and modified by Quigley and Van Order (1995). The model assumes that log price for ith house at time t is given by

$$P_{it} = I_t + H_{it} + N_{it} \tag{4.3}$$

where I_t is the logarithm of the regional housing price level, H_{it} is a Gaussian random walk, such that,

$$\begin{aligned} E\left[H_{i,t+\tau} - H_{it}\right] &= 0, \\ E\left[H_{i,t+\tau} - H_{it}\right]^2 &= \tau\sigma_{\eta_1}^2, + \tau^2\sigma_{\eta_2}^2; \end{aligned}$$

and N_{it} is white noise, such that,

$$\begin{aligned} E\left[N_{it}\right] &= 0, \\ E\left[N_{it}\right]^2 &= \tfrac{1}{2}\sigma_\nu^2. \end{aligned}$$

The model is estimated on paired sales of owner occupied housing. In the first stage, the log price of the second

Estimates of the mean and variance of individual house prices, together with the unpaid mortgage balance (computed from the contract terms), permit us to estimate the distribution of homeowner equity quarterly for each observation. In particular "*PNEQ*" is the probability that equity is negative, *i.e.*, the probability that the put option is in the money.[12]

As proxies for other "trigger events," we include measures of the quarterly unemployment rate and the annual divorce rate by state.[13]

---

sale minus the log price of the first sale is regressed on a set of dummy variables, one for each time period in the sample except the first period. The dummy variables have values of zero in every quarter except the quarter in which the sales occurred. For the quarter of the first sale, the dummy is −1, and for the quarter of the second sale, the dummy is +1. (This follows Bailey, Muth, and Nourse (1963)exactly.)

In the second stage, the squared residuals $\left(e^2\right)$ from each observation in the first stage are regressed upon τ and τ^2

$$e^2 = A + B\tau + C\tau^2, \tag{4.4}$$

where τ is the interval between the first and second sale. The coefficients A, B, and C are estimates of σ_ν^2, $\sigma_{\eta_1}^2$, and $\sigma_{\eta_2}^2$ respectively.

In the third stage, the stage one regression is reestimated by GLS with weights $\sqrt{A + B\tau + C\tau^2}$.

The estimated log price level difference $\left(\widehat{I}_{t+\tau} - \widehat{I}_t\right)$ is normally distributed with mean $(I_{t+\tau} - I_t)$, and variance $\left(\tau\sigma_{\eta_1}^2 + \tau^2\sigma_{\eta_2}^2 + \sigma_\nu^2\right)$. Denote $msa_\tau = \exp\left(\widehat{I}_\tau\right)$ as the estimated regional housing price index; then $\log\left(\frac{msa_{\kappa+\tau}}{msa_\kappa}\right)$ is normally distributed with mean $(I_{\kappa+\tau} - I_\kappa)$ and variance $\left(\tau\sigma_{\eta_1}^2 + \tau^2\sigma_{\eta_2}^2 + \sigma_\nu^2\right)$.

Means and Variances are estimated for each of 30 major MSA regions using samples of paired sales. There are about four million paired sales in the Freddie Mac data base.

[12]Specifically, the mean value of equity ratio for the lth loan observation is defined as:

$$\begin{aligned} eqr_l &= \frac{mktvalue_l - pdvupb_l}{mktvalue_l} \\ &= \frac{purprice_l \times \dfrac{msa_{\omega_l,\kappa_l+\tau_i}}{msa_{\omega_l,\kappa_l}} - \displaystyle\sum_{t=1}^{term_l-\tau_i} \frac{mopipmt_l \times 3}{(1 + mktrate_l/400)^t}}{purprice_l \times \dfrac{msa_{\omega_l,\kappa_l+\tau_i}}{msa_{\omega_l,\kappa_l}}} \\ &= 1 - \frac{(LTV/100) \times \left(1 - \left(\dfrac{1}{1 + mktrate_l/400}\right)^{term_l-\tau_i}\right)}{\left(\dfrac{msa_{\omega_l,\kappa_l+\tau_i}}{msa_{\omega_l,\kappa_l}}\right) \times \left(1 - \left(\dfrac{1}{1 + mktrate_l/400}\right)^{term_l}\right)}, \end{aligned} \tag{4.5}$$

where *purprice*$_l$ is the purchasing price of the house at the time of loan initiation, and *pdvupb*$_l$ is the present discounted value of the remaining loan balance evaluated at the current market mortgage rate.

The probability of negative equity, *pneq*, is thus

$$pneq_l = \Phi\left(\frac{\log\left(pdvupb_l\right) - \log\left(mktvalue_l\right)}{\sqrt{e^2_{\omega_l,\kappa_l+\tau_i}}}\right), \tag{4.6}$$

where $pdvupb_l$ and $mktvalue_l$ are defined above, $\Phi(\cdot)$ is cumulative standard normal distribution function, and $e^2_{\omega_l,\kappa_l+\tau_i}$ is the variance of the WRS index defined in footnote 10.

[13]Unemployment and divorce rates are measured at the state level. State unemployment data are reported in various issues of: US Department of Labor, "*Employment and Unemployment in States and Local Areas (Monthly)*" and in the "*Monthly Labor Review*". State divorce data are reported in various issues of U.S. National Center

Figure 1 summarizes the raw data used in the empirical analysis described below. Panel A of Figure 1 displays the average conditional prepayment rate, separately by the loan-to-value ratio at origination (LTV), as a function of duration. Conditional prepayment rates are slightly higher for higher LTV loans. Rates increase substantially after the first fifteen quarters. Panel C of Figure 1 displays raw conditional default rates by LTV. Note again that default rates increase substantially after about fifteen quarters. Note also that the default rates vary substantially by initial LTV. Default rates for 90 percent LTV loans are four or five times higher than default rates for 80 to 90 percent LTV loans. The default rates for these latter loans are, in turn, about twice as high as for those with LTV below 80 percent.

[Figure 1 about here.]

Finally, note that conditional default rates are quite low. Even for the riskiest class of loans, conditional default rates are no higher than two in a thousand in any quarter. Residential mortgages are relatively safe investments, and the period as a whole was one of generally rising house prices (keeping the put option out of the money).

Table 1 presents the means and standard deviations of the explanatory variables measured at origination and termination of the mortgage loans. The mean value of the prepayment option, "*POPTION*," is deeply "out of the money" when mortgages were originated, but is much less so when the mortgages were terminated. The mean value of the probability of negative equity, "*PNEQ*," is quite low when mortgages were originated. The probability remains about the same for those mortgages that were terminated by prepayment, but the probability is much larger for those loans that were terminated by default. Finally, the defaulted mortgages are associated with higher average unemployment rates and divorce rates but the prepaid mortgages are associated with relatively low rates of unemployment or divorce.

---

for Health Statistics, *"Vital Statistics of the United States, Volume III, Marriage and Divorce"*, and in *"Statistical Abstract of the U.S."*.

[Table 1 about here.]

Table 2 presents maximum likelihood estimates of the parameters of models of competing risks of mortgage prepayment and default. Estimates in this table assume prepayment and default risks are interdependent. However, the models do not address unobserved heterogeneity.

[Table 2 about here.]

Model 1 in table 2 tests the "ruthless" model, *i.e.,* Model I, as described in section 2. The model includes only measures of the financial value of the prepayment and default options. The results provide very strong support for the option theory in that the prepayment hazard increases when the call option is in the money. Similarly, a higher probability of negative equity increases the default hazard and reduces the prepayment hazard. The results also indicate that the estimated second order effect of the prepayment option value is significant and positive, suggesting that there is a pivot point in the option-driven prepayment risks. After the interest rate drops below the critical point r^* as discussed in section 2, the prepayment speed increases substantially.

Model 2 in table 2 extends the "ruthless" model by adding the trigger event variables, such as original LTV category, unemployment and divorce. The results show that financial motivation is still of paramount importance governing the prepayment and default behavior. In addition, the results suggest that borrowers' willingness to exercise financial options may be triggered or hindered by other events. For example, it suggests that higher default risks are associated with higher original LTV's. This is consistent with the argument in section 2 (see Yezer et al., 1994) that there is asymmetric information and riskier borrowers chose high LTV loans. The prepayment risk increases slightly as original LTV increases, except for the highest LTV category. For loans with original LTV over 90 percent, the prepayment risk is reduced. This is probably attributable to the liquidity constraint of the higher LTV borrowers. The estimates also show that unemployment and divorce are positive and highly significant in the default function, reflecting liquidity constraints and the effect of trigger events upon the exercise of put options.

Table 3 reports the maximum likelihood estimates of the interdependent competing risks of mortgage prepayment and default with unobserved heterogeneity as specified in section 3. We assume that there are two populations among borrowers. Each borrower may either belong to the high risk group — the "ruthless" players — or the low risk group — the "woodheads." We do not observe directly the group to which an individual borrower belongs. Since unobserved heterogeneity may be correlated with the errors in the competing risks of prepayment and default hazard functions, we estimate the distribution of the unobserved heterogeneity jointly with the competing risks of prepayment and default hazard functions.

[Table 3 about here.]

Model 3 reestimates the "ruthless" model reported in table 2. The estimates in general still provide strong support for the prediction of option theory: the prepayment hazard increases when the call option is in the money; similarly a higher probability of negative equity increases the default hazard and reduces the prepayment hazard. However, the marginal effect of the prepayment option, $POPTION$, reported in table 3 increases substantially — by about 20 percent compared to that reported in table 2. This suggests that estimating the prepayment risk without accounting for heterogeneity leads to a substantial underestimate of option-driven prepayment. The estimates also show that there is a substantial and "statistically significant" difference between the two groups in exercising the prepayment option. The "ruthless" players are about 4.72 times riskier than "woodheads" in terms of prepayment risks. However, there is almost no difference between the two groups in terms of default risks. The estimates also suggest that about three quarters of the borrowers are "ruthless" players when it comes to exercising mortgage options (i.e., $1/[1+0.343]$).

Model 4 reestimates model 2 reported in table 1. In general, the importance of the option values reported in model 3 is confirmed. In addition, the unemployment is negative and highly significant in the prepayment function — indicating that liquidity constraints (which make refinancing more difficult for unemployed and divorced households) keep them from exercising in-the-money call options. It seems clear that the original LTV, unemployment and divorce may trigger or hinder the borrower's willingness to exercise the options. These findings are analogous to those noted in model 2 in table 2. Note that, in model 4, after including trigger event variables explicitly in the prepayment and default hazard function, the estimated heterogeneity becomes less significant relative to its importance in model 3. Nonetheless, we still find that unobserved heterogeneity is statistically significant in exercising the prepayment option, and "ruthless" players are about 4.58 times riskier than "woodheads" in terms of prepayment risks. The results also show that by adding trigger event variables, model 4 has a much better fit than the "ruthless" model, *i.e.*, model 3.

Table 4 reports the selectivity corrected maximum likelihood estimates for the interdependent competing risks of mortgage prepayment and default with unobserved heterogeneity. We assume that the probability that an individual borrower belongs to the "woodheads" group may also affect the estimation of the competing risks of prepayment and default. In order to account for this possible selection bias, we integrate the "woodheads" probability function into the maximum likelihood estimation of the interdependent competing risks of prepayment and default.

[Table 4 about here.]

Table 4 indicates clearly that borrowers with higher original LTV and higher initial payment-to-income ratios are more likely to behave as "woodheads" in exercising mortgage options. The general findings reported from table 3 still persist. In addition, most of the key parameters are estimated more precisely when compared to those obtained from the estimation without a correction for selectivity bias. Both models in table 4 also report that "ruthless" players are about 5.7 times more risky than "woodheads" in terms of prepayment, but the difference between the two groups in terms of default risks is much smaller.

Table 5 reports the results of estimating the mortgage prepayment risk and default risk functions independently. It is clear that, for this sample at least, ignoring the interdependence between prepayment risk and default risk has a substantial effect upon the accuracy of the estimation for the default function.

[Table 5 about here.]

Tables 6 and 7 report a comparison of predicted cumulative default and prepayment risks at the end of the first 10 years of the mortgage contract term. The cumulative default risks were simulated for mortgage loans originated in 1979 in the Detroit metropolitan area. The cumulative prepayment risks were simulated for mortgage loans originated in 1979 in the Los Angeles metropolitan area. These two tables also report empirical default and prepayment risks observed from the raw data in the same period and metropolitan area.

[Table 6 about here.]

[Table 7 about here.]

The simulation results confirm the results of the statistical analysis: forecasts which ignore the interdependence between default and prepayment and which estimate these two risks separately lead to serious errors in estimating the default risk. Moreover, forecasts which ignore unobserved heterogeneity lead to serious errors in estimating prepayment risk. The error structure "matters" for forecasting and simulation purposes.

5. Conclusion

This paper has presented a unified model of the competing risks of mortgage termination by prepayment and default. The model considers these two hazards as dependent competing risks and estimates them jointly. The model also accounts for the unobserved heterogeneity among borrowers, and estimates the unobserved heterogeneity simultaneously with the parameters and baseline hazards associated with prepayment and default functions.

The substantive results of the analysis provide powerful support for the contingent claims model which predicts the exercise of financial options. The financial value of the call option is strongly associated with exercise of the prepayment option, and the probability that the put option is in the money is strongly associated with exercise of the default option. The results also provide strong support for the interdependence of the decisions to prepay and to default on mortgage obligations.

The results also show that there exists significant heterogeneity among mortgage borrowers. The results indicate that ignoring heterogeneity among mortgage borrowers leads to serious errors in estimating the prepayment risk. Moreover, forecasts which ignore the interdependence between default and prepayment risks and which estimate these two risks separately lead to serious errors in estimating the default risk.

Further, the results suggest that, holding other things constant, those who have chosen high initial LTV loans are more likely to exercise options in the mortgage market — prepayment as well as default. It appears that the initial LTV ratio, known at the time mortgages are issued, may well reflect investor preferences for risk in the market for mortgages on owner-occupied housing.

Office of Federal Housing Enterprise Oversight, 1700 G Street, NW, Washington DC 20552, U.S.A.

Department of Economics and Graduate School of Public Policy, University of California, Berkeley, CA 94720, U.S.A.

Freddie Mac, 8200 Jones Branch Drive, McLean, VA 22102, U.S.A.

References

Archer, W. R. and D. C. Ling. (1993). "Pricing Mortgage-backed Securities: Integrating Optimal Call and Empirical Models of Prepayment," *Journal of The American Real Estate and Urban Economics Association 21*(4), 373–404.

Archer, W. R., D. C. Ling, and G. A. McGill. (1996). "The Effect of Income and Collateral Constraints on Residential Mortgage Terminations," *Regional Science and Urban Economics 26*(3-4), 235–261.

Bailey, M. J., R. F. Muth, and H. O. Nourse. (1963). "A Regression Method for Real Estate Price Index Construction," *Journal of The American Statistical Association 58*, 933–942.

Black, F. and M. S. Scholes. (1972). "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy 81*, 637–654.

Brennan, M. J. and E. S. Schwartz. (1985). "Determinants of GNMA Mortgage Prices," *Journal of The American Real Estate and Urban Economics Association 13*, 209–228.

Buser, S. A. and P. H. Hendershott. (1984). "Pricing Default Free Mortgages," *Housing Finance Review 3*, 405–429.

Case, K. E. and R. J. Shiller. (1987). "Prices of Single-Family Homes Since 1970: New Indexes for Four Cities," *New England Economic Review*, 45–56.

Cox, D. R. and D. Oakes. (1984). *Analysis of Survival Data.* Monographs on Statistics and Applied Probability. London: Chapman and Hall.

Cox, J. C., J. E. Ingersoll, and S. A. Ross. (1985). "A Theory of the Term Structure of Interest Rates," *Econometrica 53*, 385–467.

Cunningham, D. F. and P. H. Hendershott. (1984). "Pricing FHA Mortgage Default Insurance," *Housing Finance Review 13*, 373–392.

Deng, Y. (1997). "Mortgage Termination: An Empirical Hazard Model with Stochastic Term Structure," *The Journal of Real Estate Finance and Economics 14*(3), 309–331.

Deng, Y., J. M. Quigley, and R. Van Order. (1996). "Mortgage Default and Low Downpayment Loans: The Cost of Public Subsity," *Regional Science and Urban Economics 26*(3-4), 263–285.

Dunn, K. B. and J. J. McConnell. (1981). "Valuation of Mortgage-Backed Securities," *The Journal of Finance 36*, 599–617.

Epperson, J. F., J. B. Kau, D. C. Keenan, and W. J. Muller III. (1985). "Pricing Default Risk in Mortgages," *Journal of The American Real Estate and Urban Economics Association 13*, 261–272.

Findley, M. C. and D. R. Capozza. (1977). "The Variable Rate Mortgage: An Option Theory Perspective," *Journal of Money, Credit and Banking 9*(2), 356–364.

Foster, C. and R. Van Order. (1984). "An Option-Based Model of Mortgage Default," *Housing Finance Review 3*(4), 351–372.

Foster, C. and R. Van Order. (1985). "FHA Terminations: A Prelude to Rational Mortgage Pricing," *Journal of The American Real Estate and Urban Economics Association 13*(3), 273–291.

Green, J. and J. B. Shoven. (1986). "The Effect of Interest Rates on Mortgage Prepayments," *Journal of Money, Credit and Banking 18*, 41–50.

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

Harding, J. P. (1994). "Rational Mortgage Valuation Using Optimal Intertemporal Refinancing Strategies and Heterogeneous Borrowers,". Ph.D. Dissertation, Haas School of Business, University of California at Berkeley.

Hendershott, P. H. and R. Van Order. (1987). "Pricing Mortgages: An Interpretation of the Models and Results," *Journal of Financial Services Research 1*, 77–111.

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

Kau, J. B. and D. C. Keenan. (1996). "Patterns of Rational Default,". Working Paper, The University of Georgia, Athens, GA.

Kau, J. B., D. C. Keenan, and T. Kim. (1993). "Transaction Costs, Suboptimal Termination and Default Probabilities," *Journal of The American Real Estate and Urban Economics Association 21*(3), 247–263.

Kau, J. B., D. C. Keenan, W. J. Muller, and J. F. Epperson. (1992). "A Generalized Valuation Model for Fixed-Rate Residential Mortgages," *Journal of Money, Credit and Banking 24*(3), 279–299.

Kau, J. B., D. C. Keenan, W. J. Muller, and J. F. Epperson. (1995). "The Valuation at Origination of Fixed Rate Mortgages with Default and Prepayment," *The Journal of Real Estate Finance and Economics 11*(1), 5–36.

McCall, B. P. (1996). "Unemployment Insurance Rules, Joblessness, and Part-Time Work," *Econometrica 64*(3), 647–682.

Quercia, R. G. and M. A. Stegman. (1992). "Residential Mortgage Default: A Review of the Literature," *Journal of Housing Research 3*, 341–379.

Quigley, J. M. and R. Van Order. (1990). "Efficiency in the Mortgage Market: The Borrower's Perspective," *Journal of The American Real Estate and Urban Economics Association 18*(3), 237–252.

Quigley, J. M. and R. Van Order. (1995). "Explicit Tests of Contingent Claims Models of Mortgage Default," *The Journal of Real Estate Finance and Economics 11*(2), 99–117.

Schwartz, E. S. and W. N. Torous. (1989). "Prepayment and the Valuation of Mortgage-Backed Securities," *The Journal of Finance 44*(2), 375–392.

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.

Stanton, R. (1995). "Rational Prepayment and the Valuation of Mortgage-Backed Securities," *The Review of Financial Studies 8*(3), 677–708.

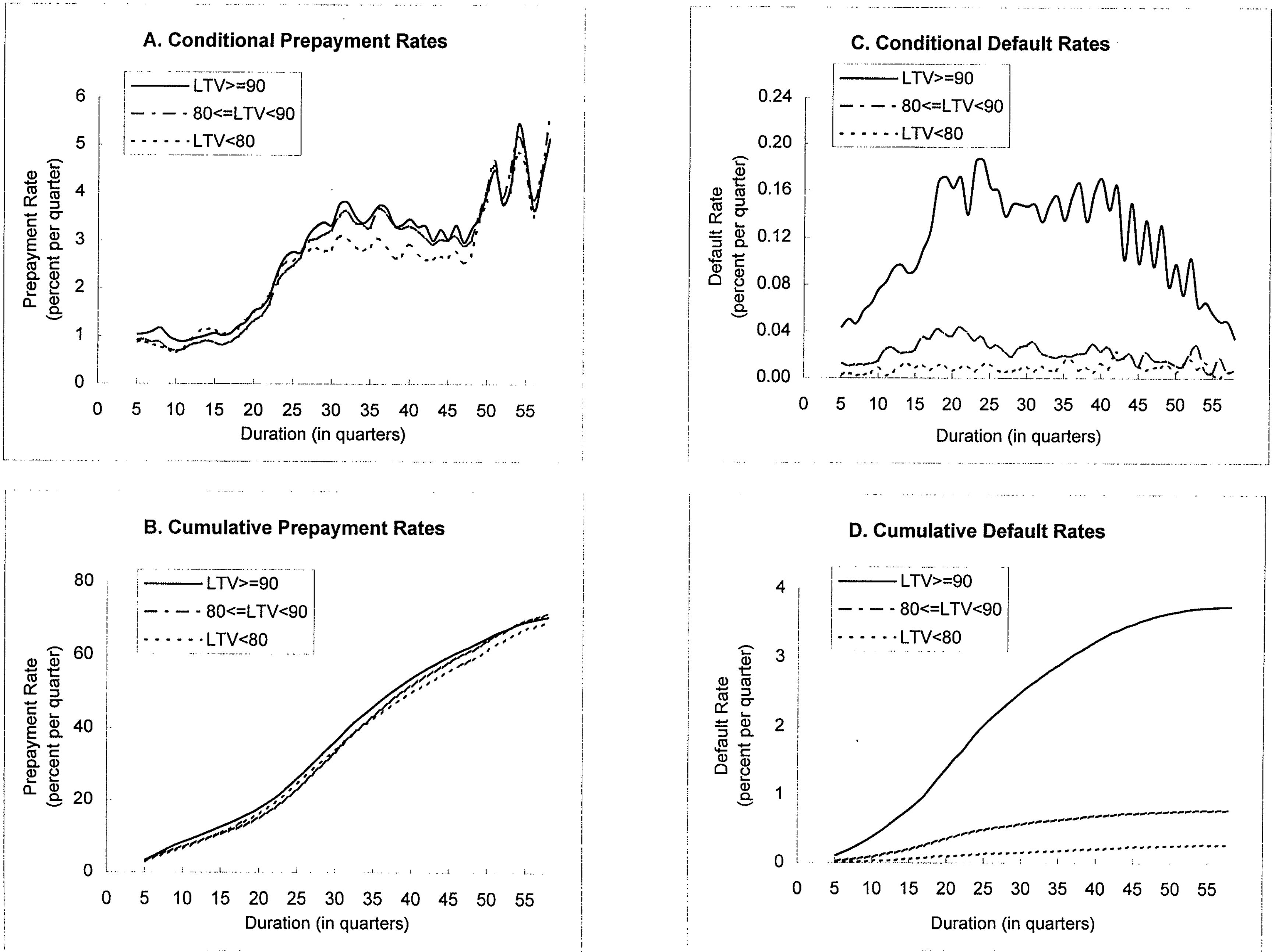
Sueyoshi, G. (1992). "Semiparametric Proportional Hazards Estimation of Competing Risks Models with Time-varying Covariates," *Journal of Econometrics 51*(1-2), 1011–1029.

Titman, S. and W. N. Torous. (1989). "Valuing Commercial Mortgages: An Empirical Investigation of the Contingent Claims Approach to Pricing Risky Det," *The Journal of Finance 44*, 345–373.

Vandell, K. D. (1993). "Handing Over the Keys: A Perspective on Mortgage Default Research," *Journal of The American Real Estate and Urban Economics Association 21*(3), 211–246.

Yezer, A. M., R. F. Phillips, and R. P. Trost. (1994). "Bias in Estimates of Discrimination and Default in Mortgage Lending: The Effects of Simultaneity and Self-Selection," *The Journal of Real Estate Finance and Economics 9*(3), 179–215.

Figure 1



A. Conditional Prepayment Rates
LTV>=90
80<=LTV<90
LTV<80
Prepayment Rate (percent per quarter)
Duration (in quarters)
C. Conditional Default Rates
Default Rate (percent per quarter)
B. Cumulative Prepayment Rates
D. Cumulative Default Rates

**Table 1. Descriptive Statistics on the Mortgage Loans
Mean Values for 22,294 Observations**

Variable	At Origination	At Termination by Prepayment	At Termination by Default
Call Option (fraction of contract value)	-0.0529 (0.0067)	-0.0265 (0.0329)	0.0321 (0.0273)
Put Option (probability of negative equity)	0.0072 (0.0005)	0.0067 (0.0013)	0.0918 (0.0283)
Squared Term of Call Option	0.0095 (0.0005)	0.0336 (0.0035)	0.0283 (0.0021)
Squared Term of Put Option	0.0006 (0.0000)	0.0014 (0.0005)	0.0367 (0.0162)
Initial Loan-To-Value Ratio (LTV)	0.7657 (0.0240)	-	-
Initial Payment-To-Income Ratio	0.1807 (0.0068)	-	-
State Unemployment Rate (percent)	6.8605 (2.4960)	6.6796 (2.8578)	7.6085 (3.5006)
State Divorce Rate (percent)	5.4486 (0.8015)	4.8635 (0.5585)	5.2587 (0.3650)
No. of Observations	22,294	16,402	363

Note: Standard deviations are in parentheses.

Table 2. Maximum Likelihood Estimates for Competing Risks of Mortgage Prepayment and Default without Heterogeneity

	Model 1		Model 2	
	Prepayment	Default	Prepayment	Default
Call Option (fraction of contract value)	4.795 (124.84)	6.283 (20.10)	4.837 (118.61)	6.768 (19.78)
Put Option (probability of negative equity)	-3.607 (8.76)	15.286 (19.79)	-3.495 (7.79)	8.662 (9.19)
Squared Term of Call Option	2.663 (20.76)	-0.359 (0.32)	2.695 (20.49)	0.236 (0.20)
Squared Term of Put Option	3.476 (5.45)	-16.343 (14.76)	3.373 (5.00)	-9.199 (7.42)
0.6<LTV≤0.75			0.035 (1.39)	1.384 (2.16)
0.75<LTV≤0.8			0.060 (2.69)	2.231 (3.75)
0.8<LTV≤0.9			0.077 (3.19)	3.146 (5.36)
LTV>0.9			-0.056 (1.89)	3.518 (5.98)
State Unemployment Rate (percent)			-0.007 (1.37)	0.097 (3.00)
State Divorce Rate (percent)			0.032 (3.10)	0.415 (4.59)
LOC	1.330 (111.42)	2.517 (13.28)	0.803 (18.45)	0.059 (1.25)
Log Likelihood	-74,981		-74,813	

Note: T-ratios are in parentheses. All models are estimated by ML approach with flexible baseline hazard function. Prepayment and default functions are considered as correlated competing risks and they are estimated jointly. Restrictions of homogeneous error terms were imposed during the maximum likelihood estimation. LOC is the location parameter of the error term.

Table 3. Maximum Likelihood Estimates for Competing Risks of Mortgage Prepayment and Default with Unobserved Heterogeneity

	Model 3		Model 4	
	Prepayment	Default	Prepayment	Default
Call Option (fraction of contract value)	5.810 (89.90)	6.324 (16.39)	5.779 (88.51)	6.750 (17.06)
Put Option (probability of negative equity)	-4.700 (10.06)	15.271 (19.81)	-4.312 (8.53)	8.669 (9.24)
Squared Term of Call Option	4.074 (23.95)	-0.270 (0.23)	4.133 (23.96)	0.193 (0.15)
Squared Term of Put Option	4.656 (6.83)	-16.327 (14.70)	4.268 (5.94)	-9.206 (7.44)
0.6<LTV≤0.75			0.019 (0.60)	1.385 (2.15)
0.75<LTV≤0.8			0.060 (2.06)	2.230 (3.73)
0.8<LTV≤0.9			0.078 (2.46)	3.146 (5.33)
LTV>0.9			-0.011 (0.29)	3.517 (5.94)
State Unemployment Rate (percent)			-0.029 (5.49)	0.097 (2.89)
State Divorce Rate (percent)			-0.005 (0.43)	0.416 (4.55)
LOC_RL	1.972 (37.60)	2.577 (7.43)	1.696 (12.95)	0.058 (1.15)
LOC_WH	0.417 (15.22)	2.403 (4.21)	0.370 (11.23)	0.060 (1.25)
MASS_WH	0.335 (10.28)		0.379 (10.26)	
Log Likelihood	-74,708		-74,560	

Note: T-ratios are in parentheses. All models are estimated by ML approach with flexible baseline hazard function. Prepayment and default functions are considered as correlated competing risks and they are estimated jointly. A bivariate distribution of unobserved heterogeneous error terms is also estimated simultaneously with the competing risks hazard functions. LOC_RL and LOC_WH are the location parameters of the error distribution associated with the ruthless players and the woodheads, respectively. MASS_RL and MASS_WH are the mass points associated with LOC_RL and LOC_WH, respectively. MASS_RL is normalized to 1.0 during the estimation.

Table 4. Selectivity Corrected Maximum Likelihood Estimates for Competing Risks of Mortgage Prepayment and Default with Unobserved Heterogeneity

	Model 5			Model 6		
	Prepayment	Default	Woodhead	Prepayment	Default	Woodhead
Call Option (fraction of contract value)	5.772 (136.38)	6.432 (19.52)		5.783 (129.46)	6.812 (19.00)	
Put Option (probability of negative equity)	-4.180 (9.40)	15.258 (19.69)		-3.518 (7.34)	8.664 (9.16)	
Squared Term of Call Option	2.727 (21.53)	-0.455 (0.41)		2.814 (21.66)	0.207 (0.18)	
Squared Term of Put Option	4.060 (5.96)	-16.325 (14.68)		3.374 (4.65)	-9.204 (7.41)	
0.6<LTV≤0.75			-0.114 (3.03)	0.055 (2.07)	1.385 (2.16)	-0.125 (3.26)
0.75<LTV≤0.8			0.044 (1.29)	0.076 (3.16)	2.231 (3.74)	0.028 (0.81)
0.8<LTV≤0.9			0.652 (15.85)	0.106 (4.09)	3.148 (5.36)	0.621 (14.98)
LTV>0.9			10.170 (1.55)	-0.112 (3.45)	3.510 (5.96)	10.352 (1.33)
Initial Payment-To-Income Ratio			4.162 (18.63)			4.143 (18.58)
State Unemployment Rate (percent)				-0.017 (3.33)	0.096 (2.99)	
State Divorce Rate (percent)				-0.001 (0.10)	0.414 (4.54)	
LOC_RL	6.428 (27.10)	4.098 (3.28)	0.000 (0.15)	4.749 (13.72)	0.072 (1.12)	0.000 (0.13)
LOC_WH	1.117 (93.43)	2.441 (12.50)	0.769 (23.62)	0.830 (17.28)	0.059 (1.25)	0.783 (23.48)
MASS_WH		5.546 (33.19)			5.458 (32.98)	
Log Likelihood		-85,909			-85,753	

Note: T-ratios are in parentheses. All models are estimated by ML approach with flexible baseline hazard function. Prepayment and default functions are considered as correlated competing risks and they are estimated jointly. A trivariate distribution of unobserved heterogeneous error term is also estimated simultaneously with the competing risks hazard functions. LOC_RL and LOC_WH are the location parameters of the error distribution associated with the ruthless players and the woodheads, respectively. MASS_RL and MASS_WH are the mass points associated with LOC_RL and LOC_WH, respectively. MASS_RL is normalized to 1.0 during the estimation.

Table 5. Maximum Likelihood Estimates for Independent Risks of Mortgage Prepayment and Default with Unobserved Heterogeneity

	Model 7		Model 8	
	Prepayment	Default	Prepayment	Default
Call Option (fraction of contract value)	5.869 (90.68)	8.348 (21.07)	5.843 (89.48)	8.520 (21.34)
Put Option (probability of negative equity)	-3.633 (7.89)	15.733 (17.94)	-3.439 (6.85)	8.192 (7.84)
Squared Term of Call Option	4.176 (24.52)	4.242 (3.20)	4.233 (24.53)	3.370 (2.57)
Squared Term of Put Option	3.819 (5.60)	-16.780 (12.75)	3.640 (5.05)	-9.049 (6.24)
0.6<LTV≤0.75			0.021 (0.66)	1.327 (2.07)
0.75<LTV≤0.8			0.064 (2.22)	2.372 (3.98)
0.8<LTV≤0.9			0.088 (2.78)	3.370 (5.74)
LTV>0.9			0.021 (0.54)	3.333 (5.66)
State Unemployment Rate (percent)			-0.027 (5.14)	0.166 (5.03)
State Divorce Rate (percent)			0.004 (0.31)	0.389 (4.23)
LOC_RL	1.976 (38.70)	10.866 (7.05)	1.581 (13.14)	0.144 (1.23)
LOC_WH	0.415 (15.66)	0.507 (1.97)	0.341 (11.31)	0.011 (0.99)
MASS_WH	0.328 (10.61)	0.629 (3.41)	0.359 (10.52)	0.344 (2.30)
Log Likelihood	-80,994	-417,754	-80,961	-417,644

Note: T-ratios are in parentheses. All models are estimated by ML approach with flexible baseline hazard function. Prepayment and default functions are estimated separately. A bivariate distribution of unobserved heterogeneous error terms is also estimated simultaneously with the prepayment or default hazard functions. LOC_RL and LOC_WH are the location parameters of the error distribution associated with the ruthless players and the woodheads, respectively. MASS_RL and MASS_WH are the mass points associated with LOC_RL and LOC_WH, respectively. MASS_RL is normalized to 1.0 during the estimation.

Table 6. Predicted Cumulative Default Risks in 10 Years

Model	Ruthless Players	Woodheads	Weighted Avg.
Model 2	-	-	5.57 %
Model 4	5.54 %	6.36 %	5.74 %
Model 6	5.01 %	5.67 %	5.53 %
Model 8	11.61 %	1.01 %	9.23 %
Observed Default Risk			5.81 %

Note: The predicted and observed cumulative default risks are estimated for mortgage loans originated in 1979 in the Detroit metropolitan area.

Table 7. Predicted Cumulative Prepayment Risks in 10 Years

Model	Ruthless Players	Woodheads	Weighted Avg.
Model 2	-	-	58.66 %
Model 4	73.72 %	24.66 %	64.69 %
Model 6	98.88 %	51.61 %	72.06 %
Model 8	71.50 %	23.32 %	62.87 %
Observed Prepayment Risk			66.66 %

Note: The predicted and observed cumulative prepayment risks are estimated for mortgage loans originated in 1979 in the Los Angeles metropolitan area.