

Valuing Fixed Rate Mortgage Loans with Default and Prepayment Options

ROBERT M. DUNSKY AND THOMAS S.Y. HO

ROBERT M. DUNSKY
is principal financial engineer & economist at OFHEO in Washington, D.C.
robert.dunsky@ofheo.gov

THOMAS S.Y. HO
is president of Thomas Ho Company, Ltd in New York, NY.
tom.ho@thomasho.com

The mortgage market in the United States is broadly based and vertically deep. From the most general perspective, we can separate the primary mortgage market from the secondary market. The primary mortgage market originates loans to borrowers (the "mortgagors") and sells the loans into the secondary market for securitization. Not all mortgage loans are sold into the secondary market for securitization. In this case, banks or financial institutions in general, are direct investors in these mortgage loans.

The traditional 30 year fixed rate mortgages contain a call option to prepay without penalty and the put option to default. Mortgage lenders are selling embedded American straddles (combined call and put options) to the mortgagors. However, the mortgagors do not exercise these options efficiently; their behavior is heterogeneous and cannot be represented by a typical mortgagor. The average refinancing and defaulting behavior of a portfolio of mortgages has to be described in terms of the age of the mortgage, seasonality, burnout level, underwriting standards, and credit scores in order to value the mortgages. Such risks must be priced into the valuation of the mortgage. Therefore, the option adjusted spread has to be estimated to determine the market pricing of the risks embedded in these mortgage loans. In sum, valuation of mortgage loans must take into consideration three salient features: 1) the

embedded options, 2) inefficiency in exercising the options, and 3) the market price of mortgage loan risks.

Despite the importance of the subject, scant research articles address all three salient features. Chen [1996]; Levin [2005]; Cheyette [1996], and Longstaff [2002] determine the call option by the optimal exercise rule, incorporating a measure of refinancing inefficiency. However, these articles fail to model in the aging, seasonality and other salient features of the mortgagors' heterogeneous behavior. While Schwartz and Torous [1989, 1992, 1993]; Boudoukh et al. [1997], and Hayre [2001] use econometrically estimated prepayment models to capture the factors affecting mortgagors' behavior in prepayment, they ignore defaults. They fail to identify default and prepayment as separate events, thus missing the insights into mortgagors' behavior that would provide a more accurate valuation model, even for mortgage securities with full credit guarantees.

Incorporating default risk in mortgages is challenging. There is voluminous literature on valuing defaultable bonds, yet sparse literature on mortgages. Credit risk literature on bonds uses the structural model and the reduced form models to value the default risk. However, such approaches cannot be adapted to valuing mortgage loans because of the complex interactions of the prepayment and default risks with the mortgagors' heterogeneous behavior. In addition, underwriting

requirements and the borrower's behavior in the case of mortgage loans are strictly different from corporate bonds.

A second genre of mortgage termination research is inferential (and normative) in nature. It attempts to take the perspective of the borrower toward asking the question, why do borrowers default or prepay their mortgage? For example, there is a growing literature in modeling the competing prepayment and default risks. Downing et al. [2002] focuses on the housing index to specify the structural model in default risk. Calhoun and Deng [2002]; Deng, Quigley, and Van Order [2000], and Dunskey and Pennington-Cross [2003] present a method to estimate the mortgagors' behavior with these competing options. However, these articles fail to consider the question of mortgage valuation.

One of this article's contributions is modeling the termination risks from the borrowers' perspective and using that model to value the mortgages. We propose a valuation model of 30 year fixed rate mortgage loans, which uses a multinomial logit model to specify the mortgagors' behavior, and a two factor arbitrage-free interest rate model to specify the interest rate movements. The model also specifies the determinants of the option adjusted spread for mortgage loans using the mortgage origination information.

Another contribution of the article is to provide insights into the question: given the conceptual structure of the econometric model, what factors account for the default event and the prepayment event? At the borrower level of investigation, we cannot overlook the competing risks of the nature of the default and prepayment events. In particular, during each month of the mortgage, the borrower must choose between making the scheduled mortgage payment, prepaying the mortgage (owing to refinance, or move) or failing to make the scheduled payment. The default and prepayment events are mutually exclusive in that the borrower is unable to default on a prepaid loan or prepay on a defaulted loan.

The main empirical results based on our sample data and valuation model are as follows. 1) While the mortgage age, seasonality, burnout factor, and credit score at origination are important factors affecting both the prepayment and default experiences, their impacts are different. The results highlight the competing nature of the two options and the importance of separating the conditional default rate from the conditional prepayment rate. The model can provide reasonable explanatory power for both the conditional prepayment rate and the conditional

default rate. 2) The prepayment model is shown to be reasonable in fitting the historical prepayment and default experiences.

The present valuation model is next analyzed with the following three main results: 1) The option adjusted spreads vary significantly across the loans, suggesting that the market price of mortgage loan risk is not necessarily cohort or pool specific, but rather borrower/loan specific. Our results show that if the recovery ratio is between 90%–95%, the OAS at origination would not tighten or widen with FICO scores above 700. However for lower FICO scores, the option adjusted spread is sensitive to the assumed recovery ratio in the valuation. This result is important to determine the value of a mortgage loan where there is no active securities market to provide the market price. The OAS model can be used instead. 2) The economic cost of guaranteeing the default risk depends primarily on the principal amount borrowed, the loan to value ratio, and the FICO score. The guarantee price elasticities to the original loan size and the FICO scores are -0.44 and -11.89 , respectively. The guarantee model is useful for any financial institutions providing credit risk guarantees and for banks to hedge their credit positions. 3) The duration of a mortgage loan is then studied over a range of recovery rates. Our findings indicate that high OLV loans originated to low credit score borrowers have shorter duration relative to all other mortgages when fully guaranteed. However, for an investor in whole loans with an expected recovery rate of 80%, this result does not hold. This result has broad implications for banks and asset management in managing the interest rate risks of a risky mortgage portfolio.

The article proceeds in four sections. The following section describes the empirical implementation of the model. *Prepayment/Default Model Specifications & Empirical Results* section describes the empirical results of the prepayment/default model. *Valuation of the Mortgages* section provides the results of valuing the mortgage loans. *Conclusion* section offers the implications and conclusions of this research.

EMPIRICAL IMPLEMENTATION

We introduce some new independent variables to the standard prepayment-default modeling literature. The following subsection describes these novel independent variables. Discussion is separated into two groups; factors included into both equations, and factors contained only

in the prepayment equation. Following the discussion of the prepayment-default factors, the estimation methodology and the valuation methodology are discussed. The final subsection reviews the study's data.

Joint Default and Prepayment Factors

Four sets of variables are common to both the default and prepayment equations. The first two variable sets are designed to capture seasoning and seasonality of the borrower's repayment behavior. Seasoning or the baseline function capture borrowers' behavior through time as measured in months. The baseline function is specified as a spline function with twelve knot or inflection points.¹ Seasonality captures the variation in borrowers' behavior throughout the calendar months (January to December). Seasonality is captured with a series of 11 dummy variables with the excluded month as January. As expected, the panel of loan histories is quite uniformly distributed through the calendar.

The next set of variables included in both equations captures the borrower's credit worthiness and loan characteristics at origination. Jointly, credit scores, the original loan to value (OLTV) ratio, and the size of the loan capture a borrower's ability to repay the mortgage. For a given OLTV, large mortgages require greater down payments. OLTV, loan size and credit scores are incorporated into the model in the form of spline functions. The FICO scores are scaled by 100, such that the smallest value is 4.00 and the largest value is 8.50. Unlike the baseline or seasoning function, the spline knots for the loan size and OLTV are chosen at expected kinks in the household demand function for mortgage debt (Dunsky and Follain [2000]) and standard underwriting requirements.

Two additional sets of variables are included only in the prepayment equation: refinance function and yield curve spread. The yield curve spread measures the slope of the yield curve from the two year maturity to the ten year maturity. The refinance function is constructed to capture the borrower sensitivity to changes in market rates toward prepayment. The refinance function is specified as the interaction between a refinance spread and a burnout measure, a feature not modeled in standard literature. The burnout factor captures the difference in the refinancing efficiency between two otherwise identical loans that have gone through different historical interest rate experiences. Note that burnout is most commonly measured in terms of a pool or cohort of mortgages. Our broader inter-

pretation here focuses on the number of missed refinance opportunities over a fixed transaction cost threshold. The refinance function is defined as

$$\text{refinance_function}_{i,t} = \text{refinance_spread}_{i,t} * \text{burn_out}_{i,t}$$

where

$$\text{refinance_spread}_{i,t} = \left(\frac{\text{PMMS}_{i,t=0}}{\text{PMMS}_{i,t}} \right) \quad (1)$$

The refinance spread is the ratio between the Freddie Mac Primary Mortgage Market Survey (PMMS) rate for the i^{th} loan at origination ($t = 0$) to the current period PMMS rate. The PMMS rate is the current mortgage rate at time t .

There are two motivations to specify the ratio in terms of the PMMS. The primary goal of the refinance ratio is to measure the opportunity cost between refinancing at the current market rate and paying the fixed contract rate on the existing loan. It is a common practice to define this ratio in terms of the contract rate relative to the market rate. However, since the borrower can select from a menu of rates and points conditional on their credit score, the former measure contains noise for our purpose. More specifically, original coupons are positively related to the credit risk level and inversely related to the number of points paid at origination. Both the credit spread and points paid are interpreted as noise, since our primary goal is to capture the borrower's response to changes in the market interest rate for their prepayment decision.

The burnout factor is defined in terms of the significantly positive refinance spread cumulated over the age of the mortgage, reflecting missed refinance opportunities. As such,

$$\text{burn_out}_{i,t} = \sum_{t=0}^T \text{MAX} \left(\frac{\text{PMMS}_{i,t=0}}{\text{PMMS}_{i,t}} - 1.10, 0 \right) \quad (2)$$

where 1.10, 0 is assumed as the refinance transaction cost. That is, we assume that a refinance opportunity occurs whenever the original PMMS rate exceeds the prevailing PMMS by 10%.

Given a set of knot points, the *refinance_spread* is specified as a linear spline function. The function should have the qualitative behavior of an S curve: when the refinance ratio is low, there is a constant base refinancing rate.

As the *refinance_spread* increases, the refinancing rate increases as well. But when the *refinance_spread* exceeds a certain level, the refinancing rate should remain stable, at a high level. However, the behavior of this *S* function varies with the burnout level. Our *refinance* function estimates *n* *S* functions as we categorize the sample into *n* buckets by the burnout function. The function is presented in Appendix A.

Default-Prepayment Model and the Valuation Methodology

The Default-Prepayment model is a Multinomial Logit (MNL) model. MNL models are commonly deployed to model competing risk events in mortgage termination literature. One property of the MNL is independence between observations. For example, given a panel or event history of a loan from origination to termination, each monthly observation is treated as though it were independent from the prior observation. A second property of the MNL is that the sum of the three events (prepay, default, no exit) must sum to one. Increases in one risk must be offset by decreases elsewhere, if the risks are truly competing.

MNL treats the dependent variables as a polytomous qualitative choice variable. Clapp et al. [2000, 2001] show that the bivariate logit (MNL with two outcomes) provides a convenient method for dealing with prepayment risk and default risk. The information for each loan is restructured to include one observation for each time period during which that loan is active.

Specifically, the monthly conditional prepayment rate (CPR) and the monthly conditional default rate (CDR) are defined as $pp_{i,t}$ and $dft_{i,t}$ respectively for the i^{th} loan in the t^{th} month.

$$pp_{i,t} = \left(\frac{\exp(x_{i,t}'\beta_{pp})}{1 + \exp(x_{i,t}'\beta_{pp}) + \exp(x_{i,t}'\beta_{dft})} \right) \quad (3)$$

and

$$dft_{i,t} = \left(\frac{\exp(x_{i,t}'\beta_{dft})}{1 + \exp(x_{i,t}'\beta_{pp}) + \exp(x_{i,t}'\beta_{dft})} \right) \quad (4)$$

and $x_{i,t}$ are the variables and the vectors of estimated parameters (β_{pp} and β_{dft}) are described in the empirical section.

The log likelihood function is given by

$$\ln L(\beta_{pp}, \beta_{dft}) = \sum_i \sum_t \alpha \ln pp_{i,t} + (1 - \alpha) \ln dft_{i,t} \quad (5)$$

where $\alpha = [0,1]$, 1 is an indicator variable depending whether the event is default or prepay.

The MNL is estimated using maximum likelihood by treating restructured discrete time information for each loan taken from identical and independent distributions. This estimation assumes irrelevant alternatives, where the elimination of one of the choices should not change the ratios of probabilities for the remaining choices. The estimation further assumes that the borrower's prior choices at any point in time are independent of those at any other point in time.

Equation (5), $\ln L(\beta_{pp}, \beta_{dft})$ is estimated from the loan level panel data, which we discuss in the following section. Given a set of scenarios of yield curve evolutions generated from an arbitrage-free interest rate model, the prepayment-default model provides a set of cash flows of a mortgage loan corresponding to the interest rate scenarios. These cash flows are used to determine the mortgage loan value. We now proceed to describe the valuation procedure in more detail.

In this article, we use the two factor arbitrage-free Generalized Ho-Lee Model [2005]. The model is provided in Appendix B for completeness of the article. This model is chosen for a number of reasons: it is arbitrage-free, therefore the valuation is consistent with the market observed yield curve and the swaption prices that determine implied volatilities of the interest rates. The interest rate model is calibrated to the market treasury curve and the swaption volatility surface. Calibration to the swaption prices uses over 90 prices for each sample date. Being a two factor model, the model can generate a broad range of yield curve shapes, unlike the one factor model. The model uses the implied volatility functions that are based on implied correlations of rates, resulting in a more parsimonious specification of the model (see Ho and Mudavanhu [2006]), providing accurate pricing at the same time. The mix normal and lognormal interest rate model truncates extremely high and low interest rates and describes the mean reversion behavior of the interest rate movements.

Given the two factor interest rate model, we now generate random interest rate paths in the two dimension lattice, with monthly step size. At each node point of the

lattice, we calculate the one month rate, two year, and 10 year rates from the arbitrage-free interest rate model.

We use the two year and 10 year rates, represented by $r_2(t)$ and $r_{10}(t)$ respectively, to determine the PMMS(t) rate based on the empirical model, estimated using the historical experience.

$$\text{PMMS}(t) = 2.638 + 0.0706 r_2(t) + 0.7576 r_{10}(t) + \epsilon(t) \quad (6)$$

PMMS(t) is used in the *refinance_spread* and *burn_out* factors defined in Equations (1) and (2). The generated ten and two year note rates are also used to construct the yield curve spread variable. The slope of the yield curve is defined as the 10 year rate net of the two year rate. Given these inputs to the prepayment-default model, we can now determine CPR(t) and CDR(t) at each node point. The prepayment-default model provides the probability of each event. The expected cash flow for a particular interest scenario is calculated as follows.

For a particular i^{th} loan with maturity T (in months) and monthly coupon rate c , let $P(t)$ be the remaining principal amount and $Y(t)$ the scheduled payment of the loan at time $t = 0, 1, \dots, T$, where T is the stated maturity date for \$1 principal. For clarity, we suppress the index i for each loan. In specifying the model, we are referring the model for a representative loan. Since the mortgage loans are fixed rate amortization bonds, the scheduled payments are determined as below.

$$Y(t) = 1/(1/(1+c) + 1/(1+c)^2 + \dots + 1/(1+c)^{T-t}) \quad (7)$$

The scheduled principal paydown at time t is given by:

$$\text{Scheduled principal paydown} = (Y(t) - c) P(t) \quad (8A)$$

$$\text{Unscheduled principal paydown} = (\text{CPR}(t) + \text{CDR}(t)) P(t) \quad (8B)$$

The total change in principal over one period is:

$$P(t) - P(t+1) = (Y(t) - c)P(t) + (\text{CPR}(t) + \text{CDR}(t)) P(t) \quad (9)$$

In simplifying, we can determine the remaining principal recursively by:

$$P(t+1) = (1 + c - Y(t) - \text{CPR}(t) - \text{CDR}(t)) P(t) \quad (10)$$

The paydown of the principal of a mortgage portfolio is not the same as the principal payments to the lenders (the banks). This is because at default, the lenders receive only a portion of the promised principal. The proportion of the received amount to the promised principal is referred to as the recovery ratio, ρ , where $0 < \rho < 1$. Our model does not seek to model the recovery ratio; therefore we simulate the valuation using a range of recovery ratios. For the guaranteed pool value, the recovery ratio is 1, as the investors cannot distinguish between prepayments and defaults.

Given these assumptions, the cash flow to the bank conditional on neither unscheduled prepayment nor default would be $P(t) Y(t)$. Conditional on default, the payment is $\text{CDR}(t) \rho P(t)$, and conditional on an unscheduled prepayment, the payment is $\text{CPR}(t) P(t)$. Therefore, we have the following cashflow model:

$$E(X(t) | \Omega(t)) = [(1 - \text{CPR}(t) - \text{CDR}(t))Y(t) + (\text{CPR}(t) + \text{CDR}(t) \rho)] P(t) \quad (11)$$

where $X(t)$ is the cashflow subject to the uncertain prepayment and default risks. $\Omega(t)$ is the information set available at time t , and ρ is the recovery ratio at default.

But the CPR(t) and CDR(t) depend on the prevailing yield curve level and shape. Therefore we need an interest rate model to generate the interest rate scenarios. Using Equations (7)–(8), we can specify the cash flows of a mortgage loan each month along the interest rate scenario. The present value of this cash flow is then determined using the vector of one month rates discounting the cash flows. The value is called the pathwise value for an interest rate scenario. Interest rate scenarios are determined randomly based on the generalized Ho-Lee model from the average of the 1000 pathwise values. This is the fair value of a mortgage loan, before introducing the option adjusted spread, which we discuss in a later section.

Data

The loan level data are from LoanPerformance, a firm that remarkets mortgage servicing data. The loan sample is restricted to the fixed rate fully amortizing 30 year maturity loan product. All of the loans are backed by owner occupied single family residences. The original loan balance spans both the conforming market and the Jumbo market, and credit quality is broad.

Exhibit 1 summarizes the loan level data at origination. Overall the sample contains 531,929 loans which were originated between 1995 and 2006. Across the loan cohorts, the average original loan size varies between \$240,000 and \$365,000, and the variation is partly due to the mix of jumbo/non-jumbo loans in the cohort. Interestingly, the average original LTV across the cohorts is fairly constant at 75%. The contract rates are not annual percentage rates (APR) and must be viewed carefully. In particular, the contract rate on any particular loan may have been "bought" down through the purchase of points at origination. During model estimation, we would ideally include both points paid and contract rate, yet in the absence of APR we exclude both for estimation to avoid measurement error associated with the contract rate. This issue is reviewed in the section focusing on the construction of the refinance spread variable. Note that the contract rate is used during our valuation stage. The average cohort credit scores are tightly distributed between 685 and 712; however, within any particular cohort, the credit score ranges between 400 and the low 800s. To this extent, the sample covers a wide spectrum of borrower credit quality.

There are more than 15 million loan-month observations. In terms of original unpaid principal balance, the

sum of the original loan balances is approximately \$150 billion. Over the reporting period, we observe 410,399 prepayments and 21,982 defaults. The remaining observations represent scheduled payments. Exhibit 2 offers a closer view of the cohort performance. The exhibit identifies peak defaults at 9% for the year 2000 loan cohort. The refinance boom in 1998 is seen with the great volume of originations in our sample. Interestingly, 95% of these 1998 originated loans prepaid by July 2006.

Prepaid loans are identified in the transaction data as loans with a zero unpaid principal balance prior to maturity, and which have never been in the state of 90 days delinquent. The second criterion is required to avoid mislabeling "work outs" as prepayments when they actually should be classified as defaults. Also critical to the data design and estimation effort is the definition of the date of prepayment and default. For prepayments, the date in which the balance transitions to zero is the date of the prepayment. In the case of defaults, the date of the last completed full payment is labeled the default date.

The market data is based on June 1, 2006 data. The Treasury curve is based on Federal Reserve Bank data. Swap prices used to determine the implied volatility of the yield curve movement are obtained from Bloomberg Financial Services.

EXHIBIT 1

Loan Characteristics at Origination by Cohort Year

		Loans	Mean	Std. Dev.	Min	Max
1995 Cohort	Loan Amount	4897	\$ 241,461	140809	\$ 10,650	\$ 1,000,000
	Contract Rate	4897	8.41	0.87	5.50	16.25
	Credit Score	4897	693	73	444	823
	Original LTV	4897	78	14	0	119.41
1996 Cohort	Loan Amount	16657	\$ 241,632	142968	\$ 10,000	\$ 1,600,000
	Contract Rate	16657	8.49	0.93	4.75	15.76
	Credit Score	16657	696	64	406	829
	Original LTV	16657	76	13	0	216
1997 Cohort	Loan Amount	44123	\$ 243,585	135439	\$ 11,700	\$ 1,500,000
	Contract Rate	44123	8.21	0.80	5.88	18.46
	Credit Score	44123	711	58	440	832
	Original LTV	44123	76	12	0	110
1998 Cohort	Loan Amount	103258	\$ 270,949	147803	\$ 8,800	\$ 2,000,000
	Contract Rate	103258	7.75	1.09	5.00	18.25
	Credit Score	103258	710	62	400	845
	Original LTV	103258	74	12	0	120
1999 Cohort	Loan Amount	66992	\$ 265,811	155921	\$ 10,000	\$ 2,000,000
	Contract Rate	66992	7.97	1.43	5.00	17.55
	Credit Score	66992	699	68	400	827
	Original LTV	66992	75	13	0	151.26
2000 Cohort	Loan Amount	46806	\$ 240,382	166451	\$ 10,000	\$ 3,770,000
	Contract Rate	46806	9.27	1.46	5.25	18.00
	Credit Score	46806	685	74	400	829
	Original LTV	46806	75	18	0	180.29
2001 Cohort	Loan Amount	66305	\$ 335,845	174354	\$ 10,000	\$ 2,105,000
	Contract Rate	66305	7.87	1.08	5.00	15.30
	Credit Score	66305	710	60	424	832
	Original LTV	66305	75	14	6	143.66
2002 Cohort	Loan Amount	50927	\$ 309,970	185943	\$ 14,450	\$ 1,600,000
	Contract Rate	50927	7.33	1.04	4.70	14.63
	Credit Score	50927	712	59	436	848
	Original LTV	50927	75	16	7.1	120.88
2003 Cohort	Loan Amount	55086	\$ 281,427	188239	\$ 15,000	\$ 1,800,000
	Contract Rate	55086	6.60	1.02	3.88	12.38
	Credit Score	55086	707	59	438	837
	Original LTV	55086	76	17	6	121.3
2004 Cohort	Loan Amount	27694	\$ 244,386	182180	\$ 14,725	\$ 1,800,000
	Contract Rate	27694	6.68	1.01	4.75	12.88
	Credit Score	27694	703	60	425	827
	Original LTV	27694	79	16	9.62	116.42
2005 Cohort	Loan Amount	30279	\$ 321,684	231922	\$ 15,000	\$ 3,000,000
	Contract Rate	30279	6.49	0.95	1.00	13.05
	Credit Score	30279	707	62	477	847
	Original LTV	30279	75	15	7.9	126
2006 Cohort	Loan Amount	18905	\$ 365,847	260079	\$ 13,600	\$ 3,500,000
	Contract Rate	18905	6.92	0.86	4.75	12.38
	Credit Score	18905	709	59	482	821
	Original LTV	18905	74	14	8	107

EXHIBIT 2

Termination Events by Cohort

	\$Volume Cohort	Prepaid (millions)	Defaulted (loans)	Total Terminations	Current @ 7/2006
1995	\$ 1,182	4,455	271	4,726	171
1996	\$ 4,025	15,589	829	16,418	239
1997	\$ 1,075	42,605	1,223	43,828	295
1998	\$ 27,978	98,685	3,618	102,303	955
1999	\$ 17,807	61,787	3,969	65,756	1,239
2000	\$ 11,251	41,672	4,294	45,966	840
2001	\$ 22,268	61,985	2,590	64,575	1,730
2002	\$ 15,786	43,089	1,991	45,080	5,847
2003	\$ 15,503	27,848	1,902	29,750	26,336
2004	\$ 6,768	9,780	886	10,666	17,028
2005	\$ 9,740	3,536	381	3,917	26,362
2006	\$ 6,916	368	28	396	18,509

PREPAYMENT/DEFAULT MODEL SPECIFICATIONS & EMPIRICAL RESULTS

The default and prepayment model parameters are reported in Exhibit 7. Exhibit 7 includes four alternative default and prepayment model specifications: 1) the age spline only; 2) the age spline plus cohort and transaction month variables; 3) the full model less credit scores and unpaid principal balance spline; and 4) the full model. Across the four models marginal improvements are seen, particularly within the default equation. The parameter estimates for the default equation tend to settle down when the credit scores and original balances are included in the model. The following subsections consider each group of parameter estimates in turn.

Seasoning Effects

"Seasoning" with regard to the loans describes the prepayment and default rate as a function of the age (in months) of the loans. The top panel in Exhibit 7 lists the age spline parameter estimates with standard errors reported under each parameter estimate. Across the four reported models, it is noticeable that the parameter estimates for model 1 are dissimilar to the other models. In particular, notice the different magnitude and direction of several of the parameters. For example, the first two age parameters are four times larger than as well as the opposite sign from the other models. This suggests that modeling defaults and

prepayments purely based on historical outcomes in the absence of controlling for borrower, loan and economic characteristics could lead to mispricing. The inclusion of these control variables in the other models appears to stabilize the seasoning parameters.

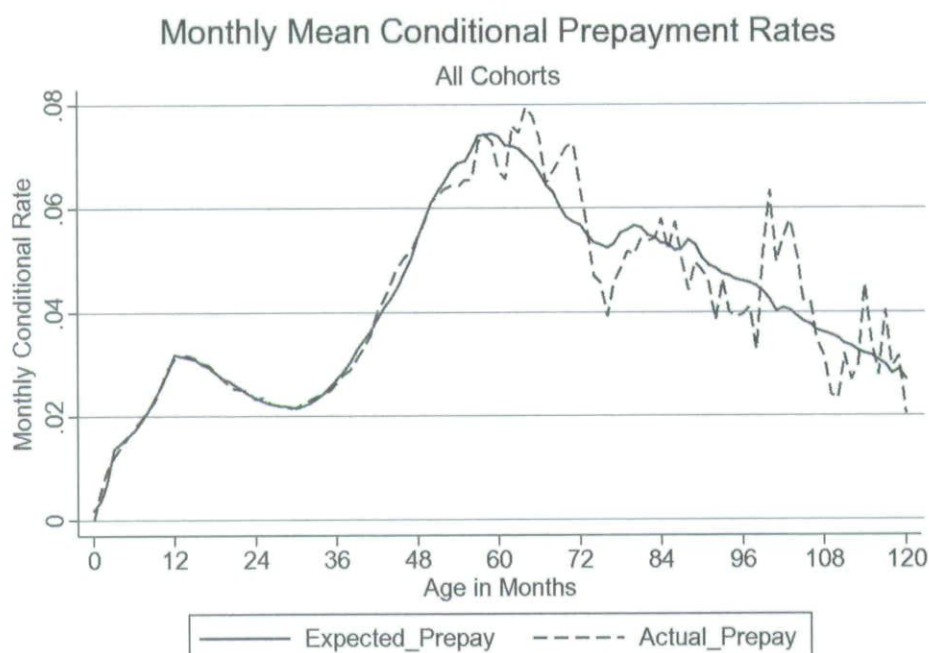
All of the prepayment seasoning parameters are statistically different than zero. However, the default seasoning parameters lack the same level of precision as the prepayment model. The later months in the times series are supported by fewer and fewer observations. As such, fitting the seasoning parameters in the later periods is challenging. Otherwise, the spline parameters are rather robust to the alternative specifications considered in Exhibit 7.

Interpreting the parameter estimates in Exhibit 7 is not straightforward. The spline parameters are cumulative in design and the Multinomial Logit structure requires the transformation depicted in equation 3 for interpretation. To this extent, Exhibit 3 plots the expected monthly conditional prepayment rates (equation 3) over loan age in months against the actual prepayment rates.

Exhibit 3 demonstrates the ability of the age spline function to capture the nonlinearity of the aging prepayment process.² Although the model utilizes 12 spline knots, there are three distinct segments depicted in the prepayment plot; a ramping up of prepayments during the first 13 months, a gradual decline in prepayments over the next year, and a final ramping up. Although the chart spans only the first 120 months of the contracted term, the model can forecast out to the full maturity, 360 months.

EXHIBIT 3

Conditional Prepayment Rates



The second way to view the model accuracy is to plot actual and expected rates across multiple cohorts through calendar time. Exhibit 4 demonstrates in sample robustness across all loan cohorts and through time.

Exhibit 5 plots the expected conditional default rates against the actual default rates. In comparison to the conditional prepayment rates, the default rates are ten times lower than the monthly prepayment rates. For example, at the 20th month the expected conditional monthly default rate is about 0.25%, while the conditional monthly prepayment rate is approximately 2.5%.

Exhibit 5 also demonstrates the impact of the thinning loan sample in the later months. Although one loan lived 139 months, the majority of the loan sample terminates before the 100th month. For example, at month 100 there are only 1,936 loans at risk of terminating, and these loans must have been originated prior to March of 1998. Beyond month 75 the absence of loans in the sample leads to large swings in the actual default rate. And after the 90th month, fewer than 10 loans default per-month in the sample, and in Exhibit 6, where the actual default rate touches zero, a single loan defaulted. Although the forecasting power of the default model appears lacking in the

later years, Exhibit 6 suggests that model is rather reasonable. Exhibit 6 plots expected and actual default rates for all loans in the sample over calendar.

Seasonality

The estimates of the seasonality coefficients are presented in Exhibit 7. January is the excluded month; as such, the parameter estimates in the model are relative to January. In the full model, both the conditional default rate and the conditional prepayment rate are related to the calendar months, but in a different way. Statistically, the seasonality parameters in the prepayment function are significantly different than zero. The parameter for September is relatively smaller than all of the other months—suggesting fewer prepayments in September. Most of the seasonality parameters are significantly different than zero.

The model suggests that loans are more likely to default (relative to January) during the month of March and least likely to default in May. Meanwhile, September is the slowest month of prepayment, according to the parameter estimates.

EXHIBIT 4

Cohort Conditional Prepayment Rates

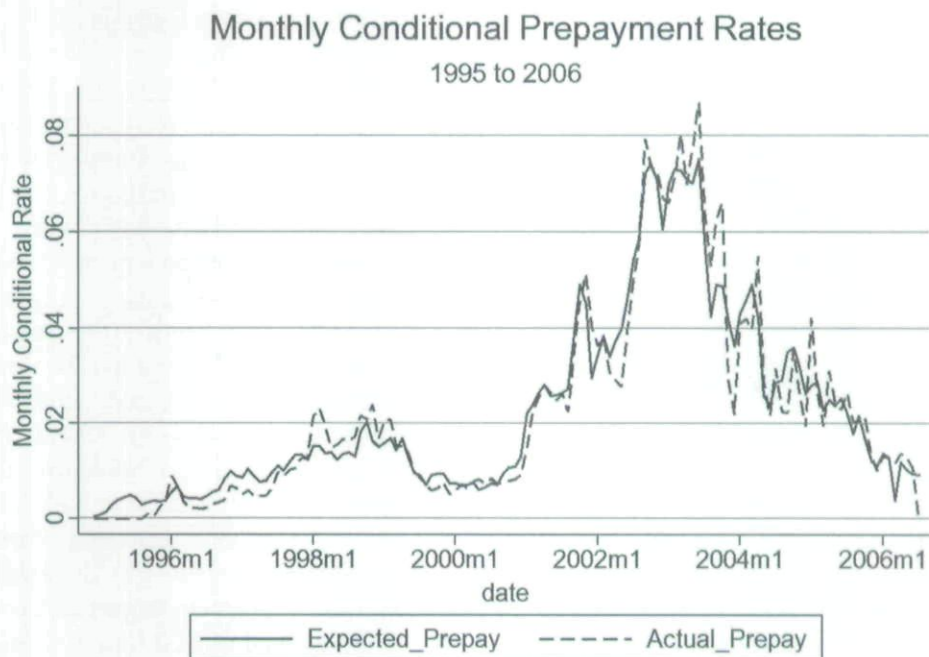
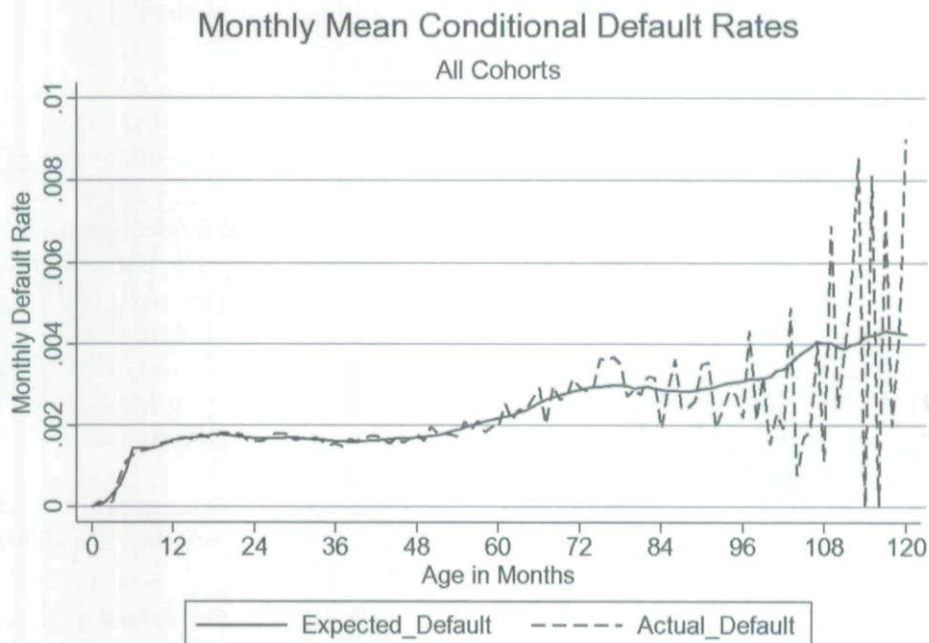


EXHIBIT 5

Conditional Default Rates



Cohort Effect

Both default and prepayment experiences may be affected by factors not considered in the model. For example, such factors may include macroeconomics variables like consumer income and confidence. In addition, local economic conditions driving housing prices are rarely directly observed. These factors are known to be time varying and known to impact the default and prepayment experience. In order to capture these factors without an explicit model of them, we use the cohort indicators defined by the origination year. In addition, omitted structural changes in the primary mortgage are captured by the cohort variables. Statistically, all of the cohort parameters are significantly different than zero. Refer to Exhibit 7 for the full set of parameter estimates.

FICO Effect

The credit scores deployed in the model are FICO scores reported at loan origination. FICO scores enter into the model in the form of a spline function with three

knot points. Exhibit 8 demonstrates the sensitivity of the conditional prepayment rate to different credit scores. As suggested by the model, higher credit scores lead to greater prepayment speeds. The exhibit also suggests that the difference in prepayment speed increases as the loan ages. For example, while a high credit score mortgagor (who has a high credit rating) tends to prepay faster, the prepay speeds also increases with age. To this extent, the plane covering months zero to 25 appears rather similar across the FICO scores. However, following the first 25 months, the conditional prepayment rates grow faster with increasing credit scores.

In the default model, the credit score spline is a very explanatory set of parameters. The conditional default rate increases rapidly as the credit score declines below 650. In Exhibit 8, the conditional default rate for a loan with an 800 FICO appears negligible over the first 100 months. However, the identical loan with a FICO of 600 faces a much steeper conditional default function. Another interesting finding suggested by the exhibit is that loans with lower credit scores face higher expected conditional default rates earlier in their lives relative to higher score borrowers.

EXHIBIT 6

Conditional Default Rates Over Calendar Time

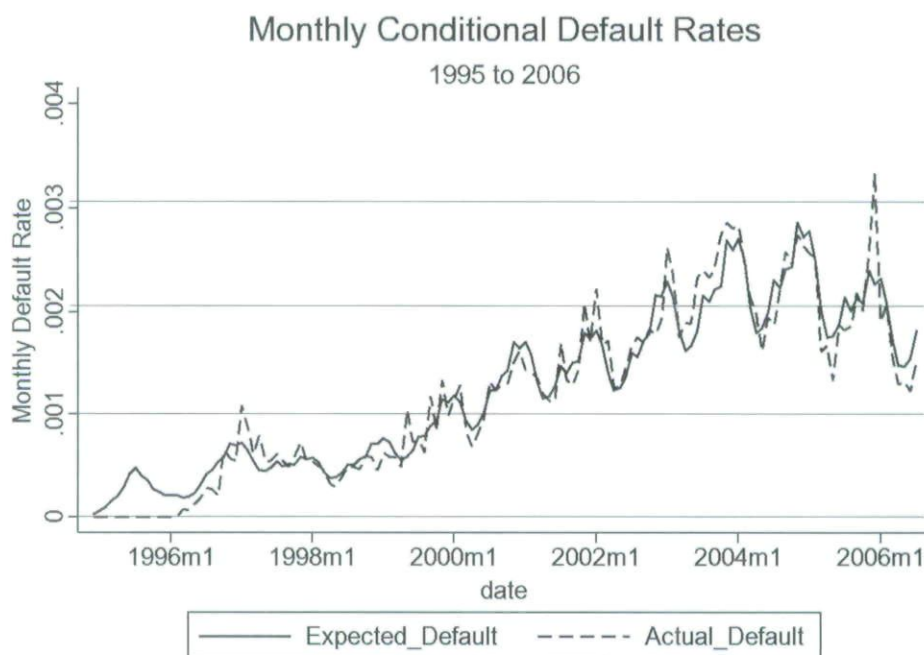


EXHIBIT 7

Multinomial Logit Model Estimates of Default and Prepayment (Standard Errors under Estimates)

	Model (1)		Model (2)		Model (3)		Model (4)	
	Age Spline only		Age Spline, Cohort & Transaction Month		Full Model Less UPB & Refi/Burnout		Full Model	
Log Likelihood	-2641381.9		-1.95E+06		-1894401.8		-1860570.9	
LR ChiSquared	8.69E+06		3.01E+07		3.02E+07		3.03E+07	
ChiSquared d.o.f	24		70		83		120	
Pseud R2	0.8445		0.8853		0.8885		0.8905	
Seasoning Spline Variables	Prepay	Default	Prepay	Default	Prepay	Default	Prepay	Default
age1 (0 < Months < 3)	-2.486	-5.705	0.641	0.876	0.630	0.870	0.630	1.215
	0.004	0.024	0.008	0.064	0.008	0.064	0.008	0.060
age2 (3 < Months < 6)	1.261	3.640	0.080	0.504	0.070	0.499	0.070	0.531
	0.005	0.026	0.005	0.022	0.005	0.023	0.004	0.015
age3 (6 < Months < 9)	-0.045	-0.134	0.104	0.005	0.101	0.002	0.106	-0.003
	0.004	0.015	0.004	0.015	0.004	0.015	0.004	0.010
age4 (9 < Months < 12)	0.141	0.077	0.117	0.057	0.114	0.055	0.114	0.037
	0.003	0.013	0.003	0.013	0.003	0.013	0.003	0.009
age5 (12 < Months < 18)	-0.022	0.011	-0.007	0.027	0.000	0.018	-0.015	0.014
	0.002	0.006	0.002	0.006	0.002	0.006	0.001	0.004
age6 (18 < Months < 24)	-0.034	-0.009	-0.023	0.005	-0.016	-0.001	0.002	-0.006
	0.002	0.006	0.002	0.006	0.002	0.006	0.001	0.004
age7 (24 < Months < 30)	-0.020	0.000	-0.012	0.013	-0.021	0.009	-0.008	0.004
	0.002	0.006	0.002	0.006	0.002	0.006	0.002	0.004
age8 (30 < Months < 40)	0.049	-0.005	0.053	0.007	0.022	0.000	0.028	-0.003
	0.001	0.004	0.001	0.004	0.001	0.004	0.001	0.003
age9 (40 < Months < 50)	0.064	0.011	0.072	0.024	0.048	0.004	0.017	-0.003
	0.001	0.005	0.001	0.005	0.001	0.005	0.001	0.003
age10 (50 < Months < 60)	0.019	0.030	0.027	0.028	0.019	0.000	0.018	-0.008
	0.001	0.005	0.001	0.006	0.001	0.006	0.001	0.004
age11 (60 < Months < 90)	-0.017	0.010	-0.008	0.022	-0.002	0.004	0.011	0.004
	0.001	0.003	0.001	0.003	0.001	0.003	0.001	0.002
age12 (90 < Months)	-0.015	0.007	-0.004	0.014	0.014	0.014	-0.002	0.016
	0.002	0.006	0.002	0.006	0.002	0.006	0.002	0.005
Transaction Month								
February			0.165	0.035	0.138	0.037	0.107	0.013
			0.008	0.031	0.008	0.031	0.007	0.021
March			0.203	-0.061	0.234	-0.058	0.165	-0.104
			0.008	0.031	0.008	0.031	0.007	0.021
April			0.122	-0.267	0.237	-0.262	0.168	-0.333
			0.008	0.033	0.008	0.033	0.007	0.023
May			0.147	-0.403	0.196	-0.396	0.218	-0.452
			0.008	0.034	0.008	0.034	0.007	0.023
June			0.044	-0.395	0.096	-0.389	0.138	-0.427
			0.008	0.034	0.008	0.034	0.008	0.023
July			0.067	-0.321	0.144	-0.312	0.130	-0.354
			0.008	0.033	0.008	0.033	0.008	0.023
August			0.047	-0.178	0.084	-0.166	0.110	-0.245
			0.008	0.032	0.008	0.032	0.008	0.022
September			0.034	-0.202	0.031	-0.202	0.043	-0.208
			0.008	0.033	0.008	0.033	0.008	0.022
October			0.268	-0.117	0.279	-0.116	0.188	-0.158
			0.008	0.032	0.008	0.032	0.007	0.022
November			0.328	-0.107	0.301	-0.109	0.216	-0.142
			0.008	0.032	0.008	0.032	0.007	0.022
December			0.222	0.054	0.223	0.050	0.173	0.014
			0.008	0.031	0.008	0.031	0.007	0.021
Cohort Year Indicators								
Cohort Year 1995			-7.50	-11.63	-11.77	-5.41	-13.66	-7.25
			0.03	0.19	0.07	0.22	0.17	0.20
Cohort Year 1996			-7.18	-11.36	-11.52	-4.77	-13.42	-6.60
			0.02	0.18	0.07	0.21	0.17	0.20
Cohort Year 1997			-6.76	-11.53	-11.24	-4.73	-13.18	-6.54
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 1998			-6.82	-11.38	-11.46	-4.53	-13.10	-6.27
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 1999			-6.66	-10.72	-11.39	-4.16	-13.14	-5.99
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2000			-5.94	-9.51	-10.73	-3.48	-12.83	-5.52
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2001			-5.59	-10.08	-10.82	-3.72	-12.71	-5.60
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2002			-5.78	-10.19	-11.02	-3.71	-12.90	-5.68
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2003			-6.58	-10.69	-11.60	-4.04	-12.97	-6.01
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2004			-6.68	-10.45	-11.29	-3.90	-12.85	-6.02
			0.02	0.18	0.07	0.21	0.17	0.19
Cohort Year 2005			-7.06	-10.55	-11.32	-3.97	-12.81	-6.02
			0.03	0.18	0.07	0.21	0.17	0.20
Cohort Year 2006			-7.02	-10.43	-11.28	-3.90	-12.97	-6.04
			0.0560	0.2540	0.0873	0.2780	0.1768	0.2340

EXHIBIT 7 (Continued)

	Age Spline, Cohort & Transaction				Full Model Less			
	Age Spline only		Month		UPB & Refi/Burnout		Full Model	
Log Likelihood	-2641381.9		-1.95E+06		-1894401.8		605858	
LR ChiSquared	8.69E+06		3.01E+07		3.02E+07		1.01E+07	
ChiSquared d.o.f	24		70		83		101	
Pseud R2	0.8445		0.8853		0.8885		0.8931	
Original LTV Spline	Prepay	Default	Prepay	Default	Prepay	Default	Prepay	Default
0 < OLTV < 70					0.0036	0.0011	-0.0008	0.0081
					0.0002	0.0009	0.0002	0.0008
70 < OLTV < 81					-0.0019	0.0193	-0.0024	0.0160
					0.0005	0.0021	0.0004	0.0015
81 < OLTV					-0.0070	0.0321	-0.0053	0.0230
					0.0003	0.0012	0.0003	0.0008
Credit Score Spline								
450 < Fico < 635					0.614	-0.939	0.198	-0.727
					0.011	0.017	0.007	0.012
635 < Fico < 695					0.426	-2.217	0.209	-1.755
					0.011	0.043	0.009	0.034
695 < Fico < 850					0.060	-1.941	-0.012	-1.742
					0.006	0.057	0.006	0.048
CMT2 to CMT10 Slope								
Yield Curve Slope					0.538		0.216	
					0.002		0.002	
Oringal UPB Spline								
Original UPB Spline (50k < OUPB)							0.0117	-0.0100
							0.0006	0.0009
Original UPB Spline (50k,150k)							0.0067	-0.0026
							0.0001	0.0002
Original UPB Spline (150k,250k)							0.0023	-0.0052
							0.0001	0.0002
Original UPB Spline (250k>OUPB)							-0.0001	-0.0008
							0.0000	0.0001
Refinance-Burnout Model for the Full Model								
(Standard Errors Under Estimates)								
	yt_i_0	yt_i_1	yt_i_2	yt_i_3	yt_i_4	yt_i_5		
yt_1_j	3.535	5.195	5.553	4.765	0.789	0.000		
	0.181	0.058	0.152	0.208	0.539	0.000		
yt_2_j	2.676	12.114	0.674	3.493	-0.176	4.767		
	0.198	0.641	0.755	0.416	0.297	0.260		
yt_3_j	2.715	9.319	8.952	1.431	0.550	2.965		
	0.223	0.974	1.022	0.603	0.433	0.168		
yt_4_j	2.835	5.949	14.811	5.636	-2.954	1.873		
	0.295	1.677	1.673	0.997	0.516	0.175		
yt_5_j	4.514	-5.485	15.487	4.313	2.383	-0.148		
	0.357	2.070	1.304	0.824	0.479	0.061		

Notes: All estimation runs utilized 15459583 observations.

Size Effect: Hot and Cold Money

The coefficients of the linear spline model of the original loan size effect are presented in Exhibit 7. Similar to the age spline used in the model, this spline is cumulative in design, such that large loans utilize more spline

knots than small loans in the default and prepayment calculations. In both the default and prepayment equations, all of the spline parameters are statistically significant from zero. Exhibit 10 suggests that the prepayment rate increases with the original loan size, though the rate is highest around \$250,000. This result suggests that the larger the

EXHIBIT 8

Conditional Prepayment Rate and Credit Scores

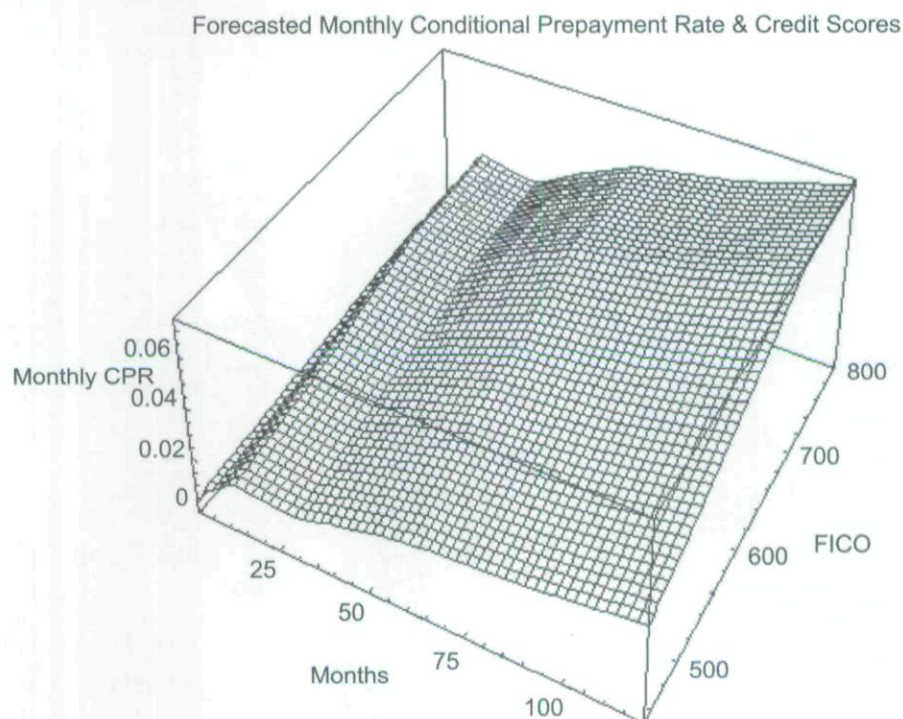


EXHIBIT 9

Conditional Default Rates and Credit Scores

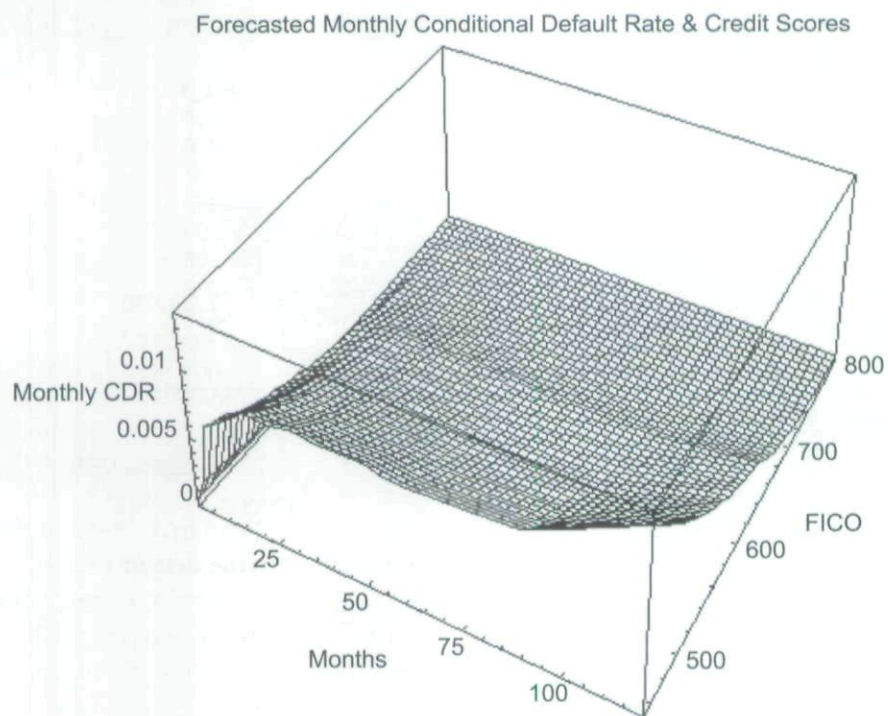
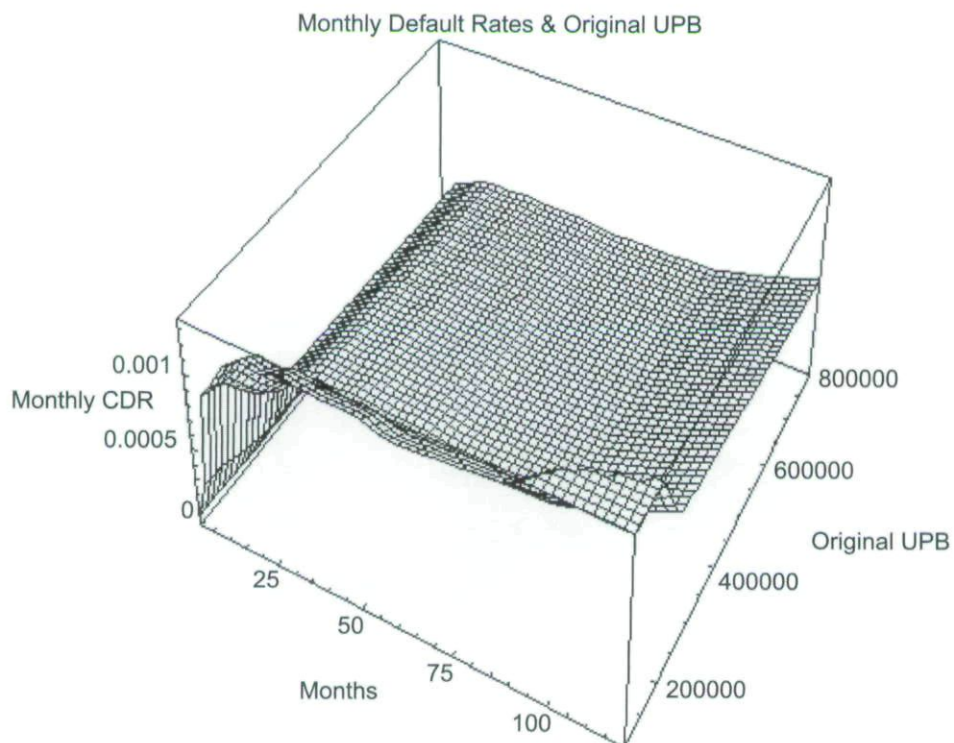
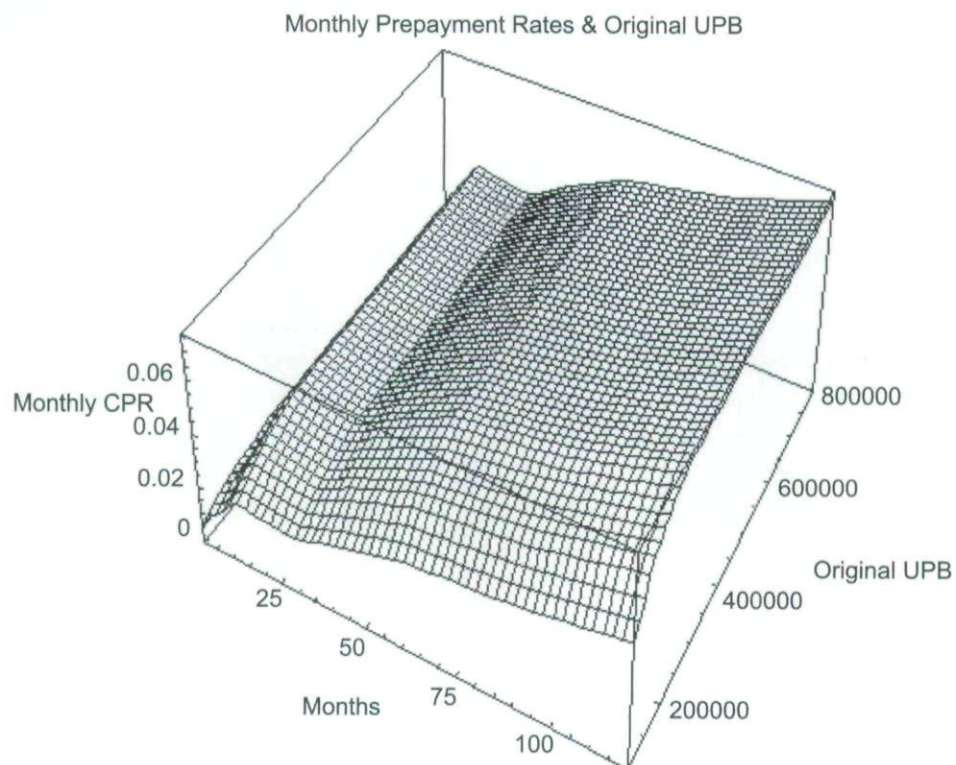


EXHIBIT 10

The Size Effect on the CPR and CDR



outstanding amount, the *hotter* is the money. Conversely, for the default model, the default rate decreases with the increase in the loan size. Exhibit 10 depicts this behavior.

This is the last effect that applies to both prepayments and defaults. The remaining two effects are related only to prepayments.

Yield Curve Slope

The slope of the yield curve is defined as the 10 year Treasury rate (in percent) net of the 2 year rate (in percent). The slope is used for the conditional prepayment rate model and not for the conditional default rate model, as the default tends not to be dependent on the shape of the yield curve. We expect that the steeper yield curve could lead to higher prepayments out of the 30yr FRM. For example, the relative cost to short termed mortgage products could tilt borrowers toward short termed maturities.

The estimates suggest that a steeper yield curve generates faster prepayment speeds. Interestingly, an inverted yield curve, where the spread is negative, leads to relatively slower prepayment speeds.

Refinance Ratio and the Burnout Effect

The refinance-burnout function is depicted in Exhibit 11. The exhibit plots the plane(s) between refinance ratio, burnout function and the monthly conditional prepayment rate. As the refinance ratio increases from unity, the mortgage becomes more in the money to refinance. The burnout function captures the number

forgone refinance basis points above a threshold. Interestingly, some borrowers tend to prepay less at burnout levels less than 0.25. This suggests less than ruthless refinancing. For example, when burnout equals zero, the prepayment rates for refinance ratios greater than 1.2 are less than those with burnout greater than 0.25. The refinancing behavior is consistent with the shape of the S curve at different levels of the burnout. That is, the results are consistent with the standard result that the refinance rate increases with a fall in the current mortgage rate.

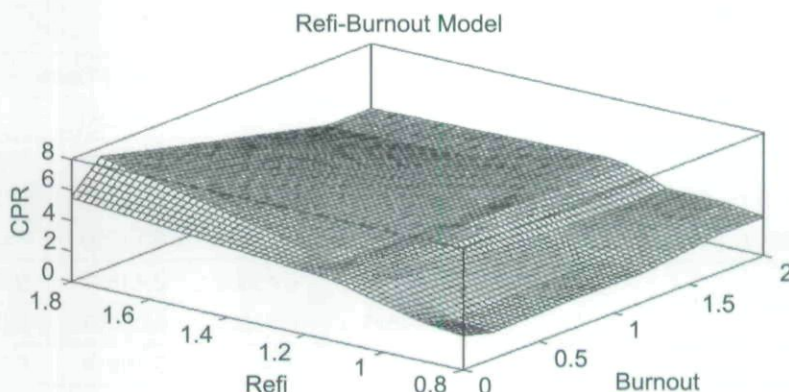
Higher refinancing is observed in light of burnout, though is not the case when the refinance ratio is less than 1.10. There are fewer observations for the burnout factor higher than 1.7, as such, the lack of data may lead to refinancing rate estimates higher than expected.

Default and Prepayment Model Summary

The results suggest that the interactions between default and prepayment risks are complex. Specifically, the very low and very high FICO scores slow the prepayment rate, yet higher FICO scores always reduce the default rate. Both the seasonality and seasoning have impacts on both the default and the prepayment rates, but the functional forms with respect to time are different. Meanwhile, the loan size is nonlinearly related to both default and prepayment rates. Increasing loan size reduces default rates at all levels yet increases prepayments until the loan is very large. These results suggest that using a simple reduced form default model to determine the "mortality" of a mortgage loan is inappropriate, because such a model assumes that we can isolate the "event of a death of a

EXHIBIT 11

Refinance Ratio and Burnout



mortgage" from other effects, something that our empirical results fail to support.

VALUATION OF THE MORTGAGES

This section reviews the calculation of the option adjusted spread given the estimated default and prepayment equations and applies the OAS to further investigate both the credit and interest rate behavior of a sample of loans. More specifically, we first discuss the valuation methodology and we then report the OAS values across alternative recovery rate assumptions. The last two sections of the article consider direct applications of our model, namely, estimating the cost of guaranteeing mortgage loans and measuring the credit impact on duration.

Determining the OAS

We now proceed to use the estimated prepayment-default model to determine the valuation of a mortgage loan at origination. The loan data is based on originations during the month of July in 2006. In order to control various factors in determining the contract rate, we selected 240 loans that have a contract rate of 6.285%. In our research, we have repeated the tests without selecting a sample of a specific interest rate. The main results do not change.

At origination, the bank pays the par amount for the mortgage. When a mortgage prepays, we assume that the bank receives the remaining principal. However, when a mortgage defaults, the bank receives the remaining principal times a constant recovery ratio. A constant spread or option adjusted spread (OAS) is added to the one month

Treasury rate along a scenario. These are the monthly discount rates to determine the pathwise value of the mortgage loan. Following the standard methodology in determining the option adjusted spread (OAS) of a mortgage loan, we determine the OAS to be the constant spread such that the mean of the pathwise values of the mortgage loan equals par. The OAS can be interpreted as the gross profit of funding a mortgage loan. It is the interest income net of the combined prepayment and default options sold to the mortgagors.

All the independent variables in the default-prepayment equations are observable at the issuance date for each mortgage loan, except for the cohort variable. The coefficient of the cohort variable is assumed to be the average of the past estimates. We use the Treasury yield curve observed on the origination date to value each mortgage. While the term structure of volatilities does affect the embedded options in the mortgages, it does not affect the value of the mortgage value significantly.

OAS Analysis

The OAS is estimated over a range of recovery ratios: 80%, 85%, 90%, 95% and 100%. The results are presented in Exhibit 12. The average size of the principal amount of the sample is \$311,878, and the average FICO score is 715. The variation of the principal amount is quite significant, with a standard deviation of \$228,665, or a coefficient of variation of 73%. The average OAS, assuming a 100% recovery ratio, the case for mortgage guaranteed loans, is 1%. This result is consistent with that obtained by Longstaff [2002], where the spread is estimated for mortgage-backed securities. However, when the principal

EXHIBIT 12

OAS Analysis of the Mortgage Loans Interest Rate 6.285

	Principal	FICO	LTV	-----recovery ratio-----				
				80	85	90	95	100
Mean	\$311,878	715.621	73.9163	0.00713	0.00809	0.00905	0.01000	0.01094
Std Dev	\$228,665	44.5498	12.8826	0.00238	0.00168	0.00100	0.00035	0.00042
Var Coef	0.73319	0.06225	0.17429	0.33342	0.20796	0.11050	0.03505	0.03799
Correl (Principal)				0.41329	0.43652	0.48893	0.69935	0.00782
Correl (FICO)				0.84677	0.83627	0.80802	0.61151	-0.89904
Correl (LTV)				-0.16427	-0.16599	-0.16914	-0.17046	0.11742

is not guaranteed, the spread would be tightened, since part of the principal would be deducted as a loss in default. Specifically, the result shows that the OAS tightens by approximately by 38 basis points when the recovery ratio lowers to 80%.

The variations of the OAS across the mortgages are significant. The loans' OAS for the range of recovery ratios are depicted in Exhibit 13.

Exhibit 13 shows that, as expected, the variations of the OAS across different recovery ratios are minimal when the FICO scores are high, yet this is not the case when, in light of greater default risk, the FICO scores are lower (on the left hand side). However, the variation of OAS is high across the loans. Part of the variation is explained by the variation of the FICO scores and loan size, referring to Exhibit 12 where we report the significant correlations of the FICO scores and the size to the OAS of the loans.

As expected, for low recovery ratios the correlation between the FICO scores and the OAS is positive (loans with lower FICO scores have a lower option adjusted spread). The result shows that when the recovery ratio is between 90% and 95%, the correlation is near zero, and that the default adjusted spread (or profitability) is not biased for or against the default risk of the loans.

Consider the case of 100% recovery ratio: with the presence of transaction costs, the OAS tends to be higher for smaller principal mortgage loans. For this reason, the correlation of the variation of the principal size and the OAS is negative (-39%). However, when the recovery ratio decreases, and since smaller mortgage loans tend to have higher default rates, the OAS tends to tighten. Neither the

variations of the mortgage loan size nor the FICO scores can explain the variations of the OAS when the recovery ratio is 90%. This result suggests that the loan origination profitability varies significantly depending on the servicer's ability (or bank's ability) to recover defaulted loans.

The Cost of Guarantee

Our results can be extended to study the appropriate charges in guaranteeing default losses. If the mortgage loan is guaranteed, then the investor would receive the par amount in the event of default. Hence, the investor can assume that the recovery ratio is 100%. The widening of the OAS with the guarantee is therefore the expected cost to the guarantor. In particular, given a recovery ratio, R , we define the guaranteed premium P to be

$$P = \text{OAS (recovery ratio} = 100\%) - \text{OAS (recovery ratio} = R) \quad (12)$$

The summary results are presented Exhibit 14.

The results show the average premium for each recovery ratio assumed ranges from 10 basis points to 38 basis points. However, the guaranteed premium should be dependent on the FICO score and the principal size, for a given recovery ratio. The model of such a relationship is presented below.

$$\ln P = \alpha + \beta_{OUPB} \ln OUPB + \beta_{FICO} \ln FICO + \beta_{OLTV} \ln OLTV \quad (13)$$

EXHIBIT 13

Option Adjusted Spreads and the FICO Relationship

	OAS (basis points)				
	Recovery Ratio				
FICO	80	85	90	95	100
550	-38	6	49	91	132
600	5	36	66	97	127
650	45	64	83	102	121
700	82	90	98	106	114
750	97	101	104	108	111
800	104	105	107	108	110
850	107	107	108	108	109

EXHIBIT 14

Guaranteed Premiums & Recovery Ratios

	Descriptive Stats on the Guaranteed Premium			
Recovery Ratio	80	85	90	95
Mean Premium	0.381%	0.285%	0.189%	0.094%
St-Dev	0.275%	0.205%	0.136%	0.068%

Where β_{OUPB} and β_{FICO} are the elasticities of the guaranteed premium to the principal size and FICO score respectively.

Exhibit 15 suggests that the model of the guaranteed premium has significant explanatory power, with adjusted R square in excess of 97% in all cases. Further, the coefficients have *t* statistics which are significantly different than zero and are stable over the range of recovery ratios. The elasticities are -0.44, -11.89 and 0.48 for the original loan amount, FICO score, and LTV respectively. This result shows that the principal amount is less important in determining the guaranteed premium, and a 1% increase in the FICO score would lead to 11.89% decrease in the guaranteed premium for the given loan sample.

Impact of the Credit Risk on Duration

The credit risk has confounding effects on both the timing and size of mortgage cash flows. For example, for borrowers with very low credit scores (high credit risk), the results show that these mortgagors tend to prepay more slowly than more creditworthy borrowers. On the other hand, the same borrower is more likely to default. The net risk to investors of receiving their principal back too soon is analytically indeterminant yet empirically measurable. The duration of a mortgage, as defined to be the price sensitivity to a parallel shift of the yield curve, intersects these two opposing factors of the mortgagor's put option on the credit risk and call option on the interest rate risk into account.

EXHIBIT 15

Guaranteed Premium Model

Coefficient	ratio 80	ratio 85	ratio 90	ratio 95
Intercept	75.8095	75.3908	74.8573	74.039
Principal(OUPB)	-0.4459	-0.4455	-0.445	-0.4445
FICO	-11.8947	-11.8757	-11.8572	-11.8391
OLTV	0.4725	0.4720	0.4716	0.4711
ln(intercept)	4.3282	4.3227	4.3156	4.3046
t statistics				
Intercept	92.259	91.780	91.162	90.198
Principal(OUPB)	-41.193	-41.163	-41.134	-41.106
FICO	-99.064	-98.939	-98.819	-98.703
OLTV	14.971	14.961	14.952	14.943
Adj R Square	0.96409	0.96405	0.96402	0.96398

EXHIBIT 16

Duration, FICO and OLV with 100% Recovery Rates

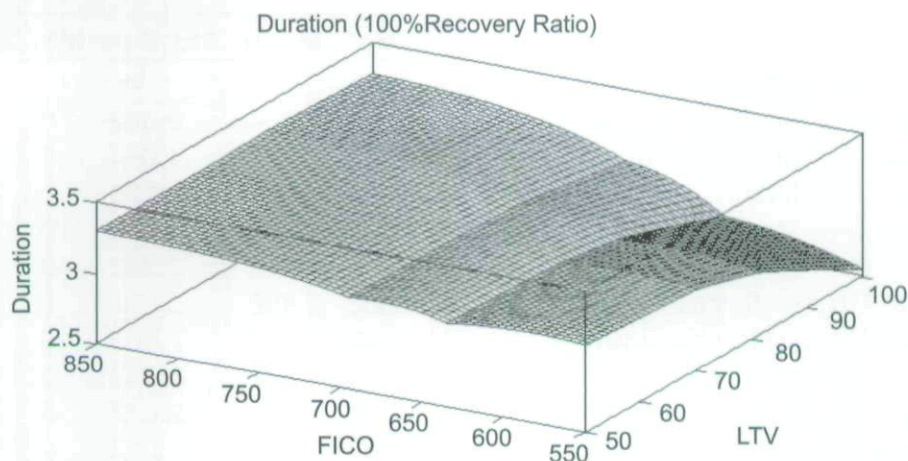


Exhibit 16 depicts the duration as a function of the FICO Score and the original LTV for fully guaranteed loans (i.e., 100% recovery ratio), and it shows that the default effect dominates for borrowers with low FICO scores and high original LTVs. The exhibit suggests that higher FICO scores and lower LTVs lead to longer duration. However, such behavior is changed when the loan is not guaranteed or is assumed to have a recovery rate of less than 100%

Exhibit 17 offers a more direct view of the impact of alternative recovery rate assumptions on duration across

the credit spectrum. This exhibit is the point by point difference between the duration assuming 100% recovery rate, and that assuming 80% recovery rate. The interpretation of Exhibit 17 is straightforward: areas with zero values indicate no difference between duration calculated at 100% and 80% recovery rates. Positive points on the graphic indicate increased duration when we decrease the recovery rate by 20 percentage points. For loans with high credit scores, we see the difference between the recovery rate assumptions as negligible. Yet as we move down the FICO score axis the duration increases for all values of original

EXHIBIT 17

Net Effect of Alternative Recovery Rates on Duration

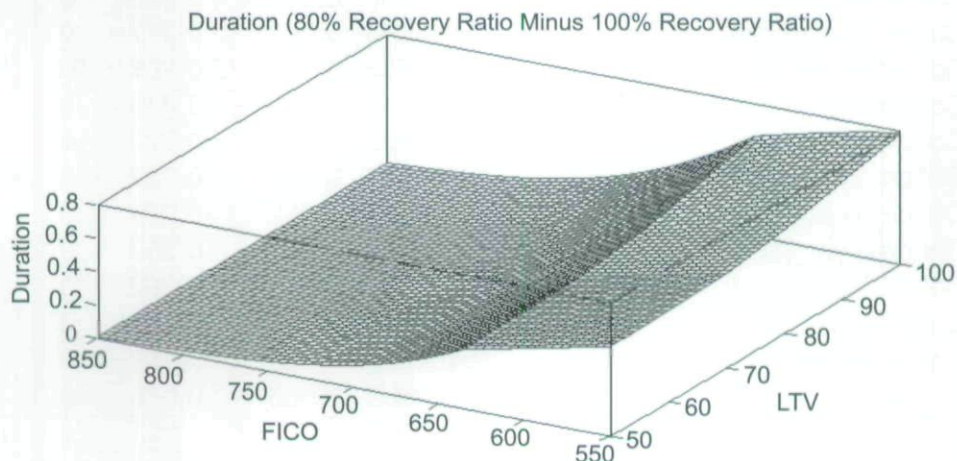


EXHIBIT 18

Sample Descriptive Statistics (Available on Request)

	Obs	Mean	Std. Dev	Min	Max
Seasoning Spline Variables					
age1 [0 < Months < 3)	15459583	2.794	0.662	0	3
age2 (3=< Months < 6)	15459583	2.493	1.064	0	3
age3 (6=< Months <9)	15459583	2.210	1.275	0	3
age4 (9=< Months < 12)	15459583	1.948	1.392	0	3
age5 (12 =< Months < 18)	15459583	3.213	2.852	0	6
age6 (18 =< Months < 24)	15459583	2.458	2.837	0	6
age7 (24 =< Months < 30)	15459583	1.842	2.667	0	6
age8 (30 =< Months < 40)	15459583	1.987	3.737	0	10
age9 (40 =< Months < 50)	15459583	1.036	2.843	0	10
age10 (50 =< Months < 60)	15459583	0.491	2.013	0	10
age11 (60 =< Months < 90)	15459583	0.390	2.745	0	30
age12 (90 =< Months)	15459583	0.037	0.863	0	49
Transaction Month					
February	15459583	0.083	0.276	0	1
March	15459583	0.083	0.276	0	1
April	15459583	0.084	0.277	0	1
May	15459583	0.084	0.278	0	1
June	15459583	0.085	0.279	0	1
July	15459583	0.086	0.280	0	1
August	15459583	0.086	0.281	0	1
September	15459583	0.081	0.272	0	1
October	15459583	0.082	0.274	0	1
November	15459583	0.082	0.274	0	1
December	15459583	0.082	0.275	0	1
Cohort Year (year of first PMT)					
Cohort Year 1995	15459583	0.018	0.132	0	1
Cohort Year 1996	15459583	0.050	0.217	0	1
Cohort Year 1997	15459583	0.106	0.308	0	1
Cohort Year 1998	15459583	0.279	0.449	0	1
Cohort Year 1999	15459583	0.167	0.373	0	1
Cohort Year 2000	15459583	0.068	0.252	0	1
Cohort Year 2001	15459583	0.082	0.274	0	1
Cohort Year 2002	15459583	0.067	0.250	0	1
Cohort Year 2003	15459583	0.094	0.292	0	1
Cohort Year 2004	15459583	0.039	0.193	0	1
Cohort Year 2005	15459583	0.026	0.159	0	1
Cohort Year 2006	15459583	0.006	0.076	0	1

EXHIBIT 18 (Continued)

Variable	Obs	Mean	Std. Dev	Min	Max
Yield Curve Slope	15459583	0.854	1.012	-4.02	2.59
Credit Score Spline					
Fico Score/100	15459583	7.028	0.650	4	8.48
450 < Fico < 635	15459583	6.277	0.236	4	6.35
635 < Fico < 695	15459583	0.448	0.231	0	0.6
695 < Fico < 850	15459583	0.303	0.333	0	1.529999
Original LTV Spline					
Original UPB Spline (50k < OUPB)	15459583	49.469	3.030	8.8	50
Original UPB Spline (50k,150k)	15459583	80.928	32.807	0	100
Original UPB Spline (150k,250k)	15459583	61.090	46.358	0	100
Original UPB Spline (250k>OUPB)	15459583	69.977	116.085	0	3520
Refinance Burnout Spline					
yt_1_0	15459583	0.753	0.325	0	0.9
yt_2_0	15459583	0.055	0.215	0	0.9
yt_3_0	15459583	0.028	0.155	0	0.9
yt_4_0	15459583	0.019	0.128	0	0.9
yt_5_0	15459583	0.040	0.186	0	0.9
yt_1_1	15459583	0.072	0.060	0	0.15
yt_2_1	15459583	0.009	0.035	0	0.15
yt_3_1	15459583	0.004	0.025	0	0.15
yt_4_1	15459583	0.003	0.021	0	0.15
yt_5_1	15459583	0.007	0.031	0	0.15
yt_1_2	15459583	0.006	0.015	0	0.05
yt_2_2	15459583	0.003	0.011	0	0.05
yt_3_2	15459583	0.001	0.008	0	0.05
yt_4_2	15459583	0.001	0.007	0	0.05
yt_5_2	15459583	0.002	0.010	0	0.05
yt_1_3	15459583	0.002	0.009	0	0.05
yt_2_3	15459583	0.002	0.010	0	0.05
yt_3_3	15459583	0.001	0.007	0	0.05
yt_4_3	15459583	0.001	0.006	0	0.05
yt_5_3	15459583	0.002	0.010	0	0.05
yt_1_4	15459583	0.000	0.002	0	0.05
yt_2_4	15459583	0.001	0.007	0	0.05
yt_3_4	15459583	0.001	0.006	0	0.05
yt_4_4	15459583	0.001	0.005	0	0.05
yt_5_4	15459583	0.002	0.009	0	0.05
yt_1_5	15459583	0.000	0.000	0	0.0153237
yt_2_5	15459583	0.000	0.004	0	0.1021033
yt_3_5	15459583	0.001	0.008	0	0.1537285
yt_4_5	15459583	0.001	0.007	0	0.1938814
yt_5_5	15459583	0.004	0.027	0	0.5590822

LTV. The recovery rate assumptions are most important when working with low credit quality loans. As expected, the greatest difference in duration is the combination of high original LTV loans and low credit score borrowers. These results show the importance of taking the credit risk into consideration even for the guaranteed loans, because the confounding effects of the put and call options in the mortgage loans can have conflicting impact on the loan duration.

CONCLUSION

Mortgage loan valuation inclusive of both the credit and interest rate risk is critically important to all market participants: investment bankers, investors, originators, servicers and regulators. In addition to the obvious need for accurate valuation models, changes in accounting rules provide some of the impetus for enhancing the state of mortgage valuation models. The introduction of fair value accounting, fair disclosure rule, IAS 39/FASB 133 (fair value hedge accounting), FASB 140 (fair valuation of mortgage servicing) and FASB 115 (fair valuation of assets) are some examples. Increased use of risk management techniques and disclosures is another factor. Examples include the Basel II Pillar I and Pillar II for capital requirements and the use of Value-at-Risk (VaR) to identify the potential risks of the portfolio result in banks determining the economic value of the mortgage loans. Beyond satisfying regulatory needs, the best practice of asset liability management and profitability measure of a bank often requires use of fair valuation of these mortgage loans. Therefore, banks find broad applications of such valuation models in their financial disclosures, profitability and risk management.

Valuation of mortgage loans is also important for non-bank institutions. Investors in whole loans have to

evaluate the underlying loan value. Government sponsored enterprises and other credit guarantors of mortgage loans (such as, FHA) need similar models to determine the cost of the guarantee of the credit risks net of fee income. The mortgage insurance business requires consideration of default probabilities with respect to the costs and prepayment likelihoods in regard to the stream of fee income. Mortgage servicers use the fair valuation of the mortgage loans to hedge their servicing fees. Finally, investment banking uses these models to design financial innovations (for example credit derivatives, structured products) and hedging strategies.

This article provides a model of the valuation of mortgage loans on the balance sheets of banks. Many banks have originated and held or bought mortgage loans from external sources. These constitute a significant part of their balance sheet. Many of the mortgages are neither securitized nor do they pass the default risk to the mortgage guarantors.

This article estimates the prepayment-default model and presents a valuation model of a mortgage loan. The result identifies the competing effects of default and prepayment on the valuation of a mortgage loan. Further, the valuation model shows that the OAS of the mortgage loan is sensitive to the assumed recovery ratio. The article presents a model of the premium in guaranteeing a mortgage loan and studies the impact of credit risk on the mortgage duration.

The methodology presented in the article can be extended to value other mortgage loans, including 15 year fixed rate, balloon, and interest-only mortgages. The valuation model can have a broad range of analytical applications including determining the key rate durations of the mortgages for hedging, and it can be applied in valuing whole loans, mortgage servicing fees, and the cost of default guarantees.

APPENDIX A

Refinance-Burnout Function

Let $\delta(\text{refinance_spread} > k)$ be the delta function which equals 1 if the condition is satisfied, otherwise zero. Let burn_out_j ($j = 1, \dots, 5$) be defined below:

$$\text{burn_out}_{k,i,t} = \delta(L(k) < \text{burn_out}_{i,t} < L(k+1)) \text{ for } k = 1, \dots, 5$$

k	1	2	3	4	5
L(k)	0	0.2	0.7	1.2	1.7

We now define the independent variables to be:

$$\begin{aligned} \text{yt_j_h} = & (\delta(L^*(h) < \text{refinance_spread}_{i,t} < L^*(h+1)) * (\text{refinance_spread}_{i,t} - L^*(h)) \\ & + \delta(\text{refinance_spread}_{i,t} > L^*(h+1)) * (L^*(h+1) - L^*(h)) * \text{burn_out}_{j,i,t} \end{aligned}$$

$$\text{yt_j_5} = \delta(1.2 \leq \text{refinance_spread}_{i,t}) * (\text{refinance_spread}_{i,t} - 1.2) * \text{burn_out}_{j,i,t}$$

$$j = 1, \dots, 5 \quad h = 0, \dots, 5$$

h	0	1	2	3	4	5
L*(h)	0	0.9	1.05	1.10	1.15	1.2

APPENDIX B

The Two Factor Generalized Ho-Lee Model

Let $P_{i,j}^n(T)$ be the price of a T year bond at time n , at state (i, j) . Then the bond price is specified by combining two one-factor models. Specifically, we have

$$P_{i,j}^n(T) = \frac{P(n+T)}{P(n)} \prod_{k=1}^n \left(\frac{1 + \delta_{0,1}^{k-1}(n-k)}{1 + \delta_{0,1}^{k-1}(n-k+T)} \right) \prod_{k=1}^n \left(\frac{1 + \delta_{0,2}^{k-1}(n-k)}{1 + \delta_{0,2}^{k-1}(n-k+T)} \right) \prod_{k=0}^{i-1} \delta_{k,1}^{n-1}(T) \prod_{k=0}^{j-1} \delta_{k,2}^{n-1}(T) \quad (\text{A1})$$

where

$$\begin{aligned} \delta_{i,1}^n(T) &= \delta_{i,1}^n \delta_{i,1}^{n+1}(T-1) \left(\frac{1 + \delta_{i+1,1}^{n+1}(T-1)}{1 + \delta_{i,1}^{n+1}(T-1)} \right) \\ \delta_{i,2}^n(T) &= \delta_{i,2}^n \delta_{i,2}^{n+1}(T-1) \left(\frac{1 + \delta_{i+1,2}^{n+1}(T-1)}{1 + \delta_{i,2}^{n+1}(T-1)} \right) \end{aligned} \quad (\text{A2})$$

and the one period forward volatilities are given by definition,

$$\begin{aligned} \delta_{i,1}^m(1) &= \delta_{i,1}^m = \exp \left(-2 \cdot \sigma_1(m) \min(R_{i,1}^m, R) \Delta t^{3/2} \right) \\ \delta_{i,2}^m(1) &= \delta_{i,2}^m = \exp \left(-2 \cdot \sigma_2(m) \min(R_{i,2}^m, R) \Delta t^{3/2} \right) \end{aligned} \quad (\text{A3})$$

where the functions $\sigma_j(n) = (a + bn) \exp(-cn) + d$ are specified by the parameters a, b, c , and d , which can be obtained from the calibration to the market price of swaption.

Using the direct extension, we can specify the one period rates for the two-factor model for any future period m and state i , and $R_{i,1}^m$ and $R_{i,2}^m$ are defined by

$$\begin{aligned} R_{i,1}^m \Delta t &= -\log \left(\frac{P(m+1)}{P(m)} \right) - \sum_{k=1}^m \log \left(\frac{1 + \delta_{0,1}^{k-1}(m-k)}{1 + \delta_{0,1}^{k-1}(m-k+1)} \right) - \sum_{k=0}^{i-1} \log(\delta_{k,1}^{m-1}(1)) \\ R_{i,2}^m \Delta t &= -\log \left(\frac{P(m+1)}{P(m)} \right) - \sum_{k=1}^m \log \left(\frac{1 + \delta_{0,2}^{k-1}(m-k)}{1 + \delta_{0,2}^{k-1}(m-k+1)} \right) - \sum_{k=0}^{i-1} \log(\delta_{k,2}^{m-1}(1)) \end{aligned} \quad (\text{A4})$$

ENDNOTES

The authors thank colleagues at the Office of Federal Housing Enterprise Oversight, and Thomas Ho Company, for their comments and assistance. In particular we thank Samantha Roberts, Lori Bowes, Shi Qiuyun and Shu Hedong. We also thank Jonathan Jones, OTS, for his suggestions. Any errors are our own.

The views expressed in this research are those of the authors and do not represent policies or positions of the Office of Federal Housing Enterprise Oversight, Office of Thrift Supervision or other officers, agencies, or instrumentalities of the United States Government.

¹The baseline function or loan seasoning is specified as a spline of the number of months that the loan is making scheduled payments. Relative to the alternative baseline functions (e.g., flexibly parametric or functions of age and age squared), we believe that the spline specification is a nice balance between parameter parsimony yet still offers flexibly to the model. Twelve spline knots are chosen to capture this seasoning behavior. The estimated parameters capture the slope between the respective knot points. The knot points are chosen non parametrically such that more spline knots exist in regions of greater support.

²The expected conditional prepayment rate is constructed within sample for the 15 million loan months. The exhibit reports the mean expected conditional prepayment rate calculated at each month using model (4). Values of the independent variables are the same as those used in estimation, as such this is an "in sample" forecast. The actual prepayment rate is simply defined as the ratio of the number of prepayments to the number of at risk loans in each month. The loan sample is described in Exhibit 1 and the appendix exhibits.

REFERENCES

- Boudoukh, J., M. Richardson, R. Stanton, and R.F. Whitelaw. "Pricing Mortgage-Backed Securities in a Multifactor Interest Rate Environment: A Multivariate Density Estimation Approach." *Review of Financial Studies*, Vol. 10, No. 2 (1997), pp. 405-446.
- Chen R.R., T.L.T. Yang. "The Relevance of Interest Rate Processes in Pricing Mortgage Backed Securities." *Journal of Housing Research*, Vol. 6, Issue 2 (1995).
- Calhoun, C.A. and Y. Deng. "A Dynamic Analysis of Fixed- and Adjustable-Rate Mortgage Terminations." *Journal of Real Estate Finance and Economics*, Vol. 24, No. 1-2 (2002), pp. 9-33.
- Clapp, J.M., Y. Deng, and X. An. "Unobserved Heterogeneity in Models of Competing Mortgage Termination." August 22, 2005. Available at SSRN: <http://ssrn.com/abstract=512624>
- Deng, Y., J.M. Quigley and R. Van Order. "Mortgage Termination, Heterogeneity, and the Exercise of Mortgage Options." *Econometrica*, Vol. 68, No. 2 (2000), pp. 275-307.
- Dunsky, R. and A. Pennington-Cross. "Estimation, Deployment and Back Testing of Default and Prepayment Equations." Working Paper, 2004.
- Dunsky, R. and J.R. Follain. "Tax Induced Portfolio Reshuffling: The Case of the Mortgage Interest Deduction." *Real Estate Economics*, Vol. 28 (December 2000), pp. 683-718.
- Downing, C., R. Stanton, and N. Wallace. "An Empirical Test of a Two-Factor Mortgage Valuation Model: How Much Does the Housing Pricing Matter?" Working Paper, University of California, Berkeley, 2002.
- Follain, J.R. "Mortgage Choice." *American Real Estate and Urban Economics Association Journal*, Vol. 18, No. 2 (1990), pp. 125-144.
- Hayre, L. and R. Yong. "Anatomy of Prepayments." The Salomon Smith Barney Prepayment Model, In L. Hayre, ed., *Salomon Smith Barney Guide to Mortgage-Backed and Asset Backed Securities*, New York, NY: John Wiley & Company, 2001.
- Ho, T. and S.B. Lee. "Multifactor Interest Rate Arbitrage-free Model." Working Paper, 2005.
- Ho, T. and B. Mudavanhu. "Interest Rate Models' Implied Volatility Function Stochastic Movements." *Journal of Investment Management*, 2006.
- Levin, A., and A. Davidson. "Prepayment Risk and Option Adjusted Valuation of MBS." *The Journal of Portfolio Management*, Vol. 31, No. 4 (Summer 2005).
- Longstaff, F. "Optimal Recursive Refinancing and the Valuation of Mortgage-backed Securities." Working Paper, University of California, Los Angeles, 2002.
- Schwartz, E.S. and W.N. Torous. "Prepayment, Default, and the Valuation of Mortgage Pass-through Securities." *Journal of Business*, 65 (1992), pp. 221-239.
- . "Mortgage Prepayment and Default Decisions: A Poisson Regression Approach." *Journal of the American Real Estate and Urban Economics Association*, 21 (1993), pp. 431-449.

To order reprints of this article, please contact Dewey Palmieri at dpalmieri@iijournals.com or 212-224-3675

©Euromoney Institutional Investor PLC. This material must be used for the customer's internal business use only and a maximum of ten (10) hard copy print-outs may be made. No further copying or transmission of this material is allowed without the express permission of Euromoney Institutional Investor PLC.

The most recent two editions of this title are only ever available at <http://www.euromoneyplc.com>