Table 31. Explanatory Variables for Default and Prepayment Models

Variable Name	Description	Categorical Ranges
Options-Related Variables		
RS(t)	Relative spread between the note rate and the current average market rate. Entered as a (7x1)–vector of indicator variables for value categories. See text for explanation.	$RS \le -0.20$ $-0.20 < RS \le -0.10$ $-0.10 < RS \le 0.0$ $0.20 < RS \le 0.10$ $0.10 < RS \le 0.20$ $0.20 < RS \le 0.30$ $RS > 0.30$
PNEQ(t)	Probability of negative equity. Entered as an (8x1)–vector of indicator variables for probability of negative equity categories. See text for explanation.	$0.0 < PNEQ \le 0.05$ $0.05 < PNEQ \le 0.10$ $0.10 < PNEQ \le 0.15$ $0.15 < PNEQ \le 0.20$ $0.20 < PNEQ \le 0.25$ $0.25 < PNEQ \le 0.30$ $0.30 < PNEQ \le 0.35$ $PNEQ > 0.35$
Other Interest Rate Variables		
B(*t*)	Burnout factor. Defined as missed opportunity to refinance. This occurs if coupon on the mortgage was greater than 200 basis points above market rate during any two quarters over the past two years. Entered as an indicator variable for burnout effect.	No Chance to Refi Missed Chance to Refi
YS(*t*)	Yield curve slope. Entered as a (4x1)–vector of indicator variables for yield curve slope categories. Yield curve slope is defined as ratio of 10-year CMT to 1-year CMT.	$YS < 1.0$ $1.0 \le YS < 1.2$ $1.2 \le YS < 1.5$ $YS \ge 1.5$
Variables for Other Loan Characteristics		
AGE(t)	Mortgage age function. This variable is computed as a quadratic function of the number of quarters since origination. When combined with the constant term, this determines the baseline hazard function.	
LTV(*0*)	Original LTV. Entered as a (6x1)–vector of indicator variables for original LTV categories	$LTV \le 60$ $60 < LTV \le 70$ $70 < LTV \le 75$ $75 < LTV \le 80$ $80 < LTV \le 90$ $90 < LTV \le 100$

Table 31. Explanatory Variables for Default and Prepayment Models (Continued)

Variable Name	Description	Categorical Ranges
SEASON(*t*)	Season of the year. Entered as a (4x1)–vector of indicator variables for seasonal categories.	Winter Spring Summer Fall
OS	Occupancy status. Indicator variable for owner-occupancy status.	Investor Owner-Occupant
LOANSIZE	Relative loan size. Entered as a (6x1)–vector of indicator variables for original loan size relative to the state average loan size in the same year.	*LOAN SIZE* ≤ 0.40 0.40 < *LOAN SIZE* ≤ 0.60 0.60 < *LOAN SIZE* ≤ 0.75 0.75 < *LOAN SIZE* ≤ 1.00 1.00 < *LOAN SIZE* ≤ 1.25 *LOAN SIZE* > 1.50
Loan Product-Type Indicators		
BALLOON	Balloon Mortgages	Balloon / Non-Balloon
15-Year FRM	15-Year Fixed-Rate Mortgages	15 YR / Non-15 YR
20-Year FRM	20-Year Fixed-Rate Mortgages	20 YR / Non-20 YR
30-Year FRM	30-Year Fixed-Rate Mortgages	30 YR / Non-30 YR
GOVERNMENT	FHA/VA Mortgages	Government / Non-Government
SECONDS	Second Liens	Second liens/ first liens

Table 32. Comparison of Multinomial Logit Parameter Estimates for Quarterly Conditional Prepayment and Default Probabilities[1]

Explanatory Variables	30-Year FRM		ARM		Other Fixed-Rate Products	
	Prepay	Default	Prepay	Default	Prepay	Default
CONSTANT	−4.514 (0.000)	−6.985 (0.000)	−4.630 (0.000)	−5.218 (0.000)	−4.511 (0.000)	−7.045 (0.000)
AGE	0.072 (0.000)	0.118 (0.000)	0.061 (0.000)	0.057 (0.000)	0.078 (0.000)	0.139 (0.000)
*AGE * AGE*	−0.002 (0.000)	−0.002 (0.000)	−0.001 (0.000)	−0.002 (0.000)	−0.002 (0.000)	−0.002 (0.000)
LTV(0) *LTV* ≤ 60	0.169 (0.000)	-1.465 (0.000)	0.097 (0.000)	−1.424 (0.000)	0.117 (0.000)	−1.491 (0.000)
60 < *LTV* ≤ 70	0.069 (0.000)	−0.219 (0.000)	−0.008* (0.134)	−0.348 (0.000)	0.041 (0.000)	−0.219 (0.000)
70 < *LTV* ≤ 75	−0.024 (0.000)	0.426 (0.000)	−0.080 (0.000)	0.121 (0.000)	−0.027 (0.000)	0.374 (0.000)
75 < *LTV* ≤ 80	0.013 (0.000)	0.272 (0.000)	−0.071 (0.000)	0.191 (0.000)	−0.004* (0.106)	0.220 (0.000)
80 < *LTV* ≤ 90	−0.070 (0.000)	0.399 (0.000)	0.081 (0.000)	0.322 (0.000)	−0.049 (0.000)	0.412 (0.000)
90 < *LTV* ≤ 100	−0.157	0.587	−0.019	1.138	−0.078	0.704
PNEQ(t) *PNEQ* ≤ 0.05	0.234 (0.000)	−1.269 (0.000)	0.603 (0.000)	−1.206 (0.000)	0.328 (0.000)	−1.198 (0.000)
0.05 < *PNEQ* ≤ 0.10	0.199 (0.000)	−0.559 (0.000)	0.239 (0.000)	−0.413 (0.000)	0.174 (0.000)	−0.344 (0.000)
0.10 < *PNEQ* ≤ 0.15	0.196 (0.000)	−0.263 (0.000)	0.060 (0.000)	−0.292 (0.000)	0.132 (0.000)	−0.062* (0.055)
0.15 < *PNEQ* ≤ 0.20	0.169 (0.000)	−0.135 (0.000)	0.027 (0.037)	−0.043* (0.109)	0.074 (0.000)	−0.080 (0.040)
0.20 < *PNEQ* ≤ 0.25	0.015 (0.002)	0.254 (0.000)	−0.005* (0.736)	0.177 (0.000)	−0.042 (0.001)	0.164 (0.000)
0.25 < *PNEQ* ≤ 0.30	−0.207 (0.000)	0.563 (0.000)	−0.155 (0.000)	0.398 (0.000)	−0.125 (0.000)	0.404 (0.000)
0.30 < *PNEQ* ≤ 0.35	−0.249 (0.000)	0.647 (0.000)	−0.242 (0.000)	0.607 (0.000)	−0.169 (0.000)	0.421 (0.000)

Table 32. Comparison of Multinomial Logit Parameter Estimates for Quarterly Conditional Prepayment and Default Probabilities[1] (Continued)

Explanatory Variables	30-Year FRM		ARM		Other Fixed-Rate Products	
	Prepay	Default	Prepay	Default	Prepay	Default
0.35 > *PNEQ*	−0.357	0.762	−0.527	0.772	−0.372	0.695
RS(t) *RS* ≤ −0.20	−1.160 (0.000)		−1.473 (0.000)		−1.027 (0.000)	
−0.20 < *RS* ≤ −0.10	−0.822 (0.000)		−0.524 (0.000)		−0.810 (0.000)	
−0.10 < *RS* ≤ 0.0	−0.680 (0.000)		−0.328 (0.000)		−0.710 (0.000)	
0.0 < *RS* ≤ 0.10	−0.432 (0.000)		−0.162 (0.000)		−0.343 (0.000)	
0.10 < *RS* ≤ 0.20	0.633 (0.000)		0.414 (0.000)		0.628 (0.000)	
0.20 < *RS* ≤ 0.30	1.182 (0.000)		1.066 (0.000)		1.098 (0.000)	
0.30 > *RS*	1.279		1.007		1.164	
BURNOUT (*B(t)*) (No Chance to Refi)	0.106 (0.000)	−0.619 (0.000)	0.027 (0.000)	−0.468 (0.000)	0.087 (0.000)	−0.566 (0.000)
(Missed Chance to Refi)	−0.106	0.619	−0.027	0.468	−0.087	0.566
YS(t) *YS* < 1.0	−0.215 (0.000)		0.042 (0.000)		−0.214 (0.000)	
1.0 ≤ *YS* < 1.2	−0.228 (0.000)		−0.156 (0.000)		−0.211 (0.000)	
1.2 ≤ *YS* < 1.5	0.022 (0.000)		−0.101 (0.000)		−0.004* (0.197)	
1.5 ≤ *YS*	0.421		0.215		0.429	
SEASON(t) Winter	−0.154 (0.000)	−0.145 (0.000)	−0.151 (0.000)	−0.031 (0.020)	−0.158 (0.000)	−0.126 (0.000)
Spring	0.161 (0.000)	0.025 (0.000)	0.065 0.044 (0.000)	0.037 (0.004)	0.148 (0.000)	−0.010* (0.575)
Summer	−0.010 (0.000)	−0.052 (0.000)	0.009 (0.012)	0.010* (0.440)	−0.002* (0.421)	−0.050 (0.004)

Table 32. Comparison of Multinomial Logit Parameter Estimates for Quarterly Conditional Prepayment and Default Probabilities[1] (Continued)

Explanatory Variables	30-Year FRM		ARM		Other Fixed-Rate Products	
	Prepay	Default	Prepay	Default	Prepay	Default
Fall	0.003	0.172	0.077	−0.016	0.012	0.186
OCCUPANCY (*OS*) Investor	−0.140 (0.000)	0.244 (0.000)	−0.228 (0.000)	0.891 (0.000)	−0.142 (0.000)	0.269 (0.000)
Owner-Occupant	0.140	−0.244	0.228	−0.891	0.142	−0.269
LOANSIZE *LOANSIZE* ≤ 0.40	−0.531 (0.000)	−0.029* (0.084)	−0.399 (0.000)	−0.215 (0.008)	−0.506 (0.000)	−0.073* (0.082)
0.40 < *LOANSIZE* ≤ 0.60	−0.337 (0.000)	−0.043 (0.000)	−0.288 (0.000)	0.111 (0.000)	−0.321 (0.000)	−0.008* (0.779)
0.60 < *LOANSIZE* ≤ 0.75	−0.130 (0.000)	−0.039 (0.000)	−0.126 (0.000)	0.119 (0.000)	−0.131 (0.000)	−0.045* (0.092)
0.75 < *LOANSIZE* ≤ 1.00	0.051 (0.000)	−0.040 (0.000)	0.014 (0.005)	0.055 (0.004)	0.038 (0.000)	−0.040* (0.054)
1.00 < *LOANSIZE* ≤ 1.25	0.200 (0.000)	0.010* (0.174)	0.169 (0.000)	0.012* (0.528)	0.188 (0.000)	−0.009* (0.684)
1.25 < *LOANSIZE* ≤ 1.50	0.313 (0.000)	0.059 (0.000)	0.276 (0.000)	−0.036* (0.108)	0.300 (0.000)	0.089 (0.000)
1.50 < *LOANSIZE*	0.434	0.082	0.354	−0.046	0.432	0.086
PRODUCT TYPE Balloon					0.522 (0.000)	1.175 (0.000)
15-Year FRM					−0.046 (0.000)	−1.328 (0.000)
20-Year FRM					−0.059 (0.000)	−0.407 (0.000)
30-Year FRM					−0.042 (0.000)	−0.264 (0.000)
FHA/VA					−0.226 (0.000)	0.429 (0.000)
Second Liens					−0.149	0.395

[1] Note: All models were estimated by the method of maximum likelihood using the SAS CATMOD procedure. Empirical p-values are shown in parentheses. P-values are not shown for the imputed coeffieints (last category for each variable). An asterisk indicates that the coefficient is not statistically significant from zero at the five percent level for an asymptotic-normal hypothesis test. The coefficients burnout, occupancy status, product types, and the constants were modified for use in the regulation to reflect differently structure dummy variables.

9. References

Bartholomew, L., J. Berk, and R. Roll (1988). ‘‘Adjustable Rate Mortgages: Prepayment Behavior,’’ *Housing Finance Review*, 7:31–46.

Brennan, M. J., and E. S. Schwartz (1985). ‘‘Determinants Of GNMA Mortgage Prices,’’ *AREUEA Journal*, 13:209–228.

Buser, S. A., and P. H. Hendershott (1984). ‘‘Pricing Default Free

Mortgages,'' *Housing Finance Review,* 3:405–429.

Calhoun, C.A. (1996). ''OFHEO House Price Indexes: Technical Description,'' Washington, D.C., Office of Federal Housing Enterprise Oversight, April 1996.

Campbell, T.S. and J.K. Dietrich (1983). ''The Determinants of Default on Insured Conventional Residential Mortgage Loans,'' *Journal of Finance,* 38:1569–1581.

Case, K.E. and Shiller, R.J. (1987). ''Prices of Single Family Real Estate Prices,'' *New England Economic Review.* 45–56.

Case, K.E. and Shiller, R.J. (1989). ''The Efficiency of the Market for Single-Family Homes,'' *The American Economic Review,* 79, 125–137.

Costlett, S. (1981). ''Efficient Estimation of Discrete-Choice Models,'' pp. 51–111 in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications,* Cambridge, Massachusetts: MIT Press, 1981.

Cunningham, D., and C. Capone (1990). ''The Relative Termination Experience Of Adjustable To Fixed-Rate Mortgages,'' *The Journal of Finance,* 45(5):1687–1703.

Deng, Y. (1997). ''Mortgage Termination: An Empirical Hazard Model With Stochastic Term Structure,'' *The Journal of Real Estate Finance and Economics,* 14(3), forthcoming.

Deng, Y., J.M. Quigley and R. Van Order (1996). ''Mortgage Default And Low Downpayment Loans: The Costs Of Public Subsidy,'' *Journal of 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.

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,'' *AREUEA Journal,* 13(3):273–291.

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., D.C. Keenan, W. J. Muller, and J. F. Epperson (1985). ''Rational Pricing Of Adjustable Rate Mortgages,'' *AREUEA Journal,* 13(2):117–128.

Kau, J.B., D.C. Keenan, W. J. Muller, and J. F. Epperson (1990). ''The Valuation And Analysis Of Adjustable Rate Mortgages,'' *Management Science,* 36(12):1417–1431.

Lancaster, T., *The Econometric Analysis of Transition Data,* New York: Cambridge University Press, 1990.

Manski, C.F. and D. McFadden (1981), ''Alternative Estimators and Sample Designs for Discrete Choice Analysis,'' pp. 2–50 in C.F. Manski and D. McFadden (eds.), *Structural Analysis of Discrete Data with Econometric Applications,* MIT Press, 1981.

McFadden, D. (1976), ''Quantal Choice Analysis: A Survey,'' *Annals of Economic and Social Measurement,* 5:363–390, 1976.

Price Waterhouse. *An Actuarial Review of the Federal Housing Administration's Mutual Mortgage Insurance Fund,* Washington, DC: Price Waterhouse, 1990.

Richard, S.F. and R. Roll (1989). ''Prepayments on Fixed Rate Mortgage Backed Securities,'' *Journal of Portfolio Management,* 15(3):73–82.

Schwartz, E.S. And W.N. Torous (1989). ''Prepayment And The Valuation Of Mortgage-Backed Securities,'' *The Journal of Finance,* 44(2):375–392.

Vandell, K.D. ''Handing Over the Keys: A Perspective on Mortgage Default Research,'' *AREUEA Journal,* 21(3):211–246.

Zorn, P., and M. Lea (1986). ''Adjustable Rate Mortgage, Fluctuations In The Economic Environment And Lender Portfolio Change,'' *AREUEA Journal,* 14:432–447.

C. Single Family Loss Severity

1. Introduction

This supplementary material provides information on the estimation and application of statistical models for the single family loss severity component of the proposed risk-based capital stress test and regulation. With one exception, all cost and revenue elements of loss severity are calculated as averages of historical Enterprise experience with foreclosed mortgages. The one exception is that a statistical regression model was developed to project the sale proceeds on foreclosed (real estate owned, or REO) properties. This regression model uses the same property valuation process that was used to create a probability of negative equity variable in the default/prepayment analysis. However, in projecting REO sales proceeds, the process is used to create a variable that measures the average equity of performing loans that have the same characteristics (other than equity) as defaulting loans. The regression then describes the relationship between average equity of performing loans and average (negative) equity of defaulting loans. One minus the projected negative equity on defaulting loans gives the projected REO sale proceeds. This regression analysis allows stress test loss severity rates to reflect economic conditions and provides an opportunity to reasonably relate loss severities on current Enterprise portfolios to the benchmark experience.

With the exception of government insured loans, OFHEO's loss severity analysis does not make explicit distinctions by loan product type. Differences by loan products are captured in the basic loan terms—coupon rate, LTV, and amortization term-that factor into loss severity equations.

The Enterprises rely upon various counterparties to provide credit enhancements that offset gross severity rates. An explanation of how credit enhancements are modeled in the stress test can be found in the appendix to the regulation.

The remainder of this supplementary material is organized as follows: section 2 provides the conceptual framework for single family loss severity analysis; section 3 describes the data used in the analysis; section 4 discusses the statistical analysis; section 5 examines adjustments made to the severity equations to reasonably relate the results to the historical benchmark experience identified in the first NPR; and section 6 explains how the results of the statistical analysis are applied in the stress test.

2. Conceptual Framework

In determining the approach to use in modeling loss severity rates, OFHEO reviewed four research studies. None of these attempted to analyze the various components of loss severity, but rather used simple regressions of some measure of a gross severity rate on original loan-to-value and loan age. These studies provide little guidance, as they do not provide frequency distributions of observed severity rates, nor do they provide averages y loan types.[233]

OFHEO chose to analyze defaulted loan severity rates in three parts: loss of loan principal, transaction costs, and

[233] These studies are: Clauretie (1990), Lekkas, Quigley, and Van Order (1993), Crawford and Rosenblatt (1995), and Berkovec, *et al.* (1997). The Berkovec, *et al.* study is not focused on loss severities, but rather analyzes them as part of a broader study of potential lending discrimination. These four studies are reviewed by Capone and Deng (1998), who themselves are interested in variations in loss severity rates across defaulted loans that can be explained by the tenets of option pricing theory. *See also* Kau and Keenan (1997) for the one example of severity analysis in a theoretical mortgage pricing model.

funding cost. This decomposition was used for three reasons. First, the loss of unpaid principal loan balance (UPB) is a function of the loss of property value before and during the default period, which can be statistically modeled as a function of economic conditions. The second reason for a decomposition analysis is to accommodate the timing of various cash flows during the period between initial default (month of first missed payment) and final property disposition. In the stress test, all default losses are accounted for in the month of default. The loss severity rate accounts for the timing of income and expenses after the default month. The timing of post-default cash flows is captured using present value discounting techniques. This method also captures funding costs of the nonearning assets-first the mortgage, and then the REO. Finally, the stress test calibrates the severity component related to loss of principal balance to the economic conditions of the BLE, as will be discussed in section 5. The stress test also uses BLE data for the elapsed time between default and foreclosure completion, and between foreclosure completion and property disposition.

Loss severity is most frequently expressed as a rate rather than a dollar amount. The most accurate representation of the magnitude of losses is to express loss severity as a percentage of the UPB at the time of default. Therefore, OFHEO has chosen to calculate all costs and revenues associated with loss severity as a percentage of the UPB. This will result in the computation of loss severity rates rather than dollar amounts, but they become dollar amounts when the stress test multiplies both default and loss severity rates against loan balances.

3. Data

Loan level data on Enterprise single family REO properties were used to analyze the components of single family loss severity rates. The data contain all defaulted mortgages on single family (1–4 unit) properties that were both originated and had a last-paid-installment date between January 1980 and December 1995. After removing incomplete records, over 116,500 valid records remained in the analysis database. These records consist of loan terms, event dates (default, foreclosure, disposition), and various expense and revenue fields.

A second analysis database was created consisting of only those loans in the historical REO analysis database that met benchmark criteria. Those criteria singled out conventional, 30-year fixed-rate loans on single family properties (single unit, owner-occupied, detached properties) that originated in 1983 and 1984 in the States of Arkansas, Louisiana, Mississippi, and Oklahoma, and defaulted within ten years of origination. This benchmark database (789 loans) was used to create an adjustment factor that provides consistency between the loss severity rates projected in the stress test and the benchmark loss rates. This process is discussed in section 5, Consistency with the Benchmark Loss Experience, below.

Other data used in the analysis of loss severity rates includes historical Census division level HPI indices and their associated volatility parameters, which come from the *OFHEO HPI Report,* 1996:3.

4. Statistical Analysis

The primary statistical analysis performed for single family loss severity rates measured the impact of market conditions on REO sale proceeds. This is the one dynamic element of loss severity in stress test application. It relies upon original LTV, loan amortization, and Census division level house price growth. OFHEO performed a statistical regression analysis to model negative equity for defaulted loans as a function of the average equity of similar, but performing, loans. All other statistical analyses involved calculating average historical experience by loss severity element. The two elements with values computed as historical averages are foreclosure expenses and a combination of REO expenses, revenues (other than disposition proceeds), and property selling expenses. In addition, average times to foreclosure and time in REO were computed for use in calculating the net present value of revenues and expenses in the month of default.

When averages were computed for loss elements, a two-step procedure was used. First, the average experience of each firm was calculated using UPB as a weighting factor. This weighted average provides a good measure of portfolio-wide performance, although the analysis is based on individual loans. The second step was to give equal weight to the experience of each firm by taking a simple average of the experience of the two Enterprises. This procedure is also consistent with the procedure used to find the benchmark loss severity rate reported in NPR1.[234]

The averages of the foreclosure and the REO expense/revenue elements are based on the entire national, historical sample of Enterprise experience. Benchmark experience was not used by itself because it was evident from an analysis of the data that there were significant numbers of records with missing expense components. The magnitudes of these expense items should not vary between the benchmark region and other areas of the country for two reasons. First, the benchmark region has a variety of foreclosure laws, by State, so that the average foreclosure expense rate for the benchmark region is similar to averages from other regions of the country, and to the average for the nation as a whole. Second, OFHEO computed these loss components as percentages of the outstanding loan balance, rather than as actual dollar amounts. Thus, the fact that the benchmark region may have had lower property values than the national average, and therefore lower dollar losses per loan, will not be material. Average loss rate components from other regions of the country should be comparable to what would be found in the benchmark loan data, if those records were complete.

OFHEO does, however, base time frames on benchmark experience. Because the benchmark region does have a variety of foreclosure laws, these time frames are actually very close to those of the entire national experience of the Enterprises.

a. Predicting REO Sale Proceeds

The REO sale proceeds, as a percentage of the defaulting UPB, measures the impact of erosion of property value over time, both prior to and after default. To begin the analysis of REO sale proceeds, OFHEO computed negative property equity, the difference between the defaulting UPB and the gross property sale proceeds, as a percentage of the UPB.[235] This amount was regressed against average equity for similar, but non-defaulting loans. The resulting regression coefficient provides the relationship between average equity of performing loans and average (negative) equity of defaulting loans. The nuance here is that average equity of performing loans is first transformed into a standardized normal distance, or what is commonly called a z-score, before being used in the regression. This is a widely used statistical technique for

[234] *See* 61 FR 29592, 29597, June 11, 1996. Procedures here differ from those of the first NPR by calculating loss severity as a percentage of the outstanding loan balance at time of default, rather than a percentage of the original loan balance.

[235] The one expense that OFHEO does net from sale proceeds here is property repairs undertaken by the Enterprises during the REO period. Because these expenses reflect part of the loss of property value that occurred prior to foreclosure completion, it is appropriate that they be included in the estimation of the loss of UPB due to property value deterioration.

creating a standard unit of measure for comparisons across many different variables and/or value levels.

To measure average (performing loan) equity, the property value underlying each defaulting mortgage was adjusted using the change in the (Census division) OFHEO HPI from origination to the last-paid-installment date, and using loan amortization schedules.[236] This adjustment provides average expected equity for each loan, if it were performing. But these loans are not performing, and rather than having average house price growth, they will generally have lower-than-average house price growth. In fact, defaulting loans come from the lower tail of the equity distribution, so the statistical analysis must capture just how far into the tail defaulting loan properties will be, on average. OFHEO analyzed several measures of the house price distribution to find which gave the best prediction of the difference between average performing loan equity and average non-performing loan equity. The best predictor was the z-score, identifying the distance between the expected (performing loan) house price and the (actual defaulting) loan balance. The z-score transforms the actual difference between (expected) house price and (actual) loan balance into the number of standard deviations there are between the two values, where the standard deviation is of house prices in the Census division. The z-score tells how far below the average property value growth in the Census division must the growth of any individual property value be, before all borrower equity is eliminated. The difference of actual growth of defaulting loans from average growth for performing loans will be larger than this, on average, because the z-score distance gives the minimal difference needed to eliminate borrower equity. The z-score equation is:

$$z = \frac{ln(HPI_{d,\,q,\,t}) - ln(B)}{\sigma_{d,\,t}} \qquad (Eq.\ 13)$$

where:

z = standardized distance of the loan balance from the average house price at the time of default

$HPI_{d,q,t}$ = House Price Index value for properties in Census division d, whose loans originated in quarter q and defaulted at age *t* (in quarters). This is created by dividing the HPI value for the calendar quarter of the last-paid-installment date by the HPI value in the calendar quarter of loan origination.

B = the ratio of outstanding loan balance at default to the original house price. This captures the equity generated from both the original downpayment and loan amortization over time.

$\sigma_{d,t}$ = standard deviation of HPI growth rates for properties in Census division *d*, after *t* quarters. This is the square root of ($\alpha t + \beta t^2$), where α and β are the two volatility parameters for each HPI series (published in the *OFHEO HPI Report*).

In their continuous rate forms, the cumulative growth rate factors are found by taking the logarithm of the HPI, as is done here. The log of *HPI* gives average price appreciation, and the difference between that and the log of the loan balance, *B*, gives the expected loan equity due to price appreciation, downpayment, and amortization.[237]

These standardized distances, or z-scores, are the key values used to compute the expected negative property equity (as a percent of the outstanding loan balance) when a foreclosed property is sold. Larger z-scores reflect some combination of large downpayments, loan amortization, and high levels of (average) house price growth since loan origination. In these circumstances, loans that do default should have relatively good rates of property sale proceeds as a percent of the mortgage UPB (small rates of negative equity). In other environments, where z-scores are small, there are low rates of appreciation in the market, and/or low downpayments and a lack of significant amortization. The small z-score indicates that there is a wide range of property values in the market area that are below the loan balance. Therefore, REO sale proceeds will be low and the negative property equity will be high.

The statistical equation used to predict negative property equity (*L*) was estimated using ordinary least squares (OLS) regression of actual rates of UPB loss on the z-scores computed for each loan. The regression dataset was limited to historical REO observations where ($-0.50 \le z_t \le 4.0$), because sample sizes outside this range were very thin.[238] Log-transformed values of negative property equity (ln(*L*) + 1)) were used in the regression to account for a change in the relationship between negative equity and z-scores as those values change. The estimated regression equation is:

[236] The last-paid-installment (LPI) month is the month directly prior to the month of default, when the first payment is missed. Loan amortization ends at LPI, and because the HPI index is updated quarterly rather than monthly, the choice of LPI month or default month for loan seasoning is immaterial.

[237] Taking the logarithm of *B* transforms owner-invested equity (downpayment plus amortization) into an implied HPI growth rate factor. It is the cumulative (negative) growth of HPI necessary to eliminate all positive equity in the property. By transforming *B* into its continuous rate counterpart in this fashion, the z-score variable can measure the amount by which the growth of property value on loan properties must be less than the average growth rate of performing loans before default is a real possibility (the point of zero equity). The regression then measures the relationship between actual below-normal growth on REO properties and the minimumly required below-normal house price growth needed to trigger default.

[238] In stress test application, outliers are given predicted equity loss values measured at the boundary points of the z-score range employed in the regression.

$$ln(L + 1) = 0.241325 - 0.076959 \cdot z \qquad (Eq.\ 14)$$

where:

t-statistic for z coefficient = -102

R^2 = 0.09

One-half the regression variance (0.029104) is added to the regression equation to provide the median-to-mean adjustment factor for log-normal models.[239] The result is:

$$\ln(L + 1) = 0.27043 - 0.076959 \cdot z \qquad (Eq.\ 15)$$

so that:

$$L = \exp(0.27043 - 0.076959 \cdot z) - 1 \qquad (Eq.\ 16)$$

The low R-squared value for the regression indicates a wide variance of actual loss rates around the average, predicted rates. OFHEO has analyzed this variance and believes that using the simple regression equation that captures average loss rates at each z-score value is more appropriate for the stress test than is a more complex model that would capture deviations around that average loss rate. Average rates provide an appropriate simplification because loss severity rates will be applied to groups of loans.

The boundary values of L are computed at the boundary points of z used in the regression sample, 4.0 and −0.5. When $z = 4.0$, $L = -0.04$. This suggests that, on average, REO sales prices are 4 percent higher than the mortgage UPB in areas with significant house price appreciation and/or for loans that have substantial amortization. That is, the average default (and there will be relatively few) will actually have a small amount of positive equity, though generally not enough to pay the costs of selling the property. At the other extreme, where $z = -0.5$, the predicted value of $L = 0.36$. This is a situation where average property values on performing loans are 36 percent below their associated mortgage balances. This extreme was reached in several areas of the country at various times during the study period. Such a loss of loan principal can cause the total loss severity to exceed 60 percent of UPB.

b. Foreclosure Expenses

Foreclosure expenses vary principally by property State and by the rate of bankruptcy filings among defaulted borrowers.[240] The average expense rate in the historical observation period is five percent of UPB. Unlike other loss components, this component is based solely on Fannie Mae experience because Freddie Mac did not break out foreclosure expenses from REO expenses in its data systems.

c. REO Holding and Disposition Expenses

Property (REO) holding costs include such items as property maintenance, utilities, property taxes, and hazard insurance. OFHEO calculated the average total REO holding expenses, plus selling costs (principally, realtor fees), less miscellaneous revenues to produce a final REO expense loss severity factor of 13.7 percent.[241]

d. Time Frames

There are two time frames of interest: time from default to foreclosure completion, and time from foreclosure completion to property disposition. A mean expected value for each of the time periods of interest was calculated from BLE data. The mean benchmark foreclosure time (period from default to foreclosure) was 13 months. The mean benchmark REO/property sale time was seven months. These time frames are used in the stress test to discount the various default-related cash flows to the month of default.

5. Consistency With the Benchmark Loss Experience

The equation for negative equity of defaulted loans (equation 14) was estimated on all historical REO experience of the Enterprises. Using this broad range of data assured that the equation would be appropriate for loans entering the stress test with a wide range of loan amortization and cumulative HPI experience. The equation used in the stress test includes an adjustment that calibrates the results to the BLE.

The procedure for calibrating equation 16 to the benchmark experience parallels the procedure used by OFHEO to calibrate the single family default equations to the BLE. A database of defaulted loans meeting benchmark criteria was input into the negative equity equation to compute the projected negative equity, by loan. The z-score variable values were computed by assuming that all loans originated in the first quarter of 1984, using the West South Central HPI series, for purposes of assigning house price appreciation rates. These predicted rates of negative equity were then averaged by Enterprise, using UPB as a weighting factor. Finally, a simple average of these Enterprise averages was computed to arrive at a mean expected value for the benchmark REO database.

[239] The logarithmic equation used in the regression implies a lognormal distribution of potential negative equity values around predicted values. The point estimates from the regression, therefore, produce median rather than mean value estimates of loss of principal balance. The adjustment to arrive at the mean is the additive constant (0.029104), one-half the variance of the regression residuals.

[240] To process foreclosures when defaulting borrowers file for Bankruptcy Court protection requires further legal expenses to gain release from the bankruptcy "stay" on debt collection actions.

[241] As noted earlier, the Freddie Mac foreclosure expense rate is imputed from the Fannie Mae experience (five percent). Therefore, the REO holding costs used to create the average rate shown here use total expense for Freddie Mac less imputed foreclosure expense for Fannie Mae.

This final mean rate of negative equity on defaulted loans was then compared with the actual, historical mean rate across the two firms' benchmark experience. The average projected rate of negative equity using equation 16 and this averaging method was 21.30 percent. The actual historical experience average was 31.64 percent. The difference, 10.34 percent, reflects the nature of the benchmark experience: that defaulting benchmark loans tended to have larger losses, on average, than did loans from other regions of the country that experienced the same housing market conditions. The adjusted negative equity equation is:

$$L = (\exp(0.27043 - 0.076959 \cdot z) - 1) + 0.1034 \qquad \textit{(Eq. 17)}$$

Proceeds from REO sale are then computed as one minus the projected negative property equity for the defaulting loans in each loan group.

6. Application to the Stress Test

Stress test application of loss severities begins with the results of the statistical analysis of severity components discussed here, but then adds components for loss of loan principal, servicer claim payments, mortgage insurance, and seller/servicer recourse. OFHEO's approach is to account for all default related cash flows at one of three points in time: 120 days delinquency, foreclosure, and property disposition. The stress test then calculates the effective loss severity rate as a net present value of all cash flows, in the month of loan default. The month of default is one month after the last paid installment (LPI) date, the month of the first missed payment.

There is a difference in the treatment of sold and retained loans when computing stress test loss severity rates. For retained loans, defaulting UPB is not a cash outlay and, therefore, is not discounted. For sold loans, however, the defaulting UPB represents the current expense of repurchasing a defaulted loan from a security pool. It is, therefore, a cash-flow element that should be discounted.[242] This expense is normally incurred in the fourth month of default. Sold loans in default also involve four months of interest passthroughs to the investors while the loans remain in the security pools. The interest passthroughs are not immediate expenses of the Enterprises because they are initially matched by passthroughs made by the seller/servicers to the Enterprises. However, all post-default interest payments received by the Enterprises are reimbursed to servicers in the post-foreclosure claim filing. Therefore, all interest passthroughs between seller/servicers and Enterprises are ignored. Only the passthrough by the Enterprise to security holders is counted as an expense in the stress test, and it is included with the seller/servicer claim payment at time of foreclosure.

The stress test provides that, at the time of foreclosure, the Enterprises make servicers whole for expenses incurred on the loan and property, including foreclosure costs, and receive proceeds from any available mortgage insurance. When mortgage insurance is present, mortgage insurance payments will generally be larger than the servicer claim payment and provide net inflows of funds to the Enterprises at foreclosure.

Also, any available seller/servicer recourse is applied to reduce the final loss severity rate. There are some smaller sources of credit enhancements that further reduce Enterprise losses, and these are added once dollar losses are computed in the cash flow component of the stress test.[243]

7. References

Berkovec, James A., Glenn B. Canner, Stuart A. Gabriel, and Timothy Hannan. 1998. Discrimination, Competition, and Loan Performance in FHA Mortgage Lending, *Review of Economics and Statistics*, forthcoming.

Capone, Charles A. and Yongheng Deng. 1998. ''Loss Severities and Optimal Put Exercise: An Examination of Negative Equity in Mortgage Foreclosure,'' unpublished manuscript. OFHEO: Washington, DC, January 1988.

Clauretie, Terrence. 1990. ''A Note on Mortgage Risk: Default vs. Loss Rates,'' *AREUEA Journal* 18 (2), 202–206.

Crawford, Gordon and Rosenblatt, Eric. 1995. ''Efficient Mortgage Default Option Exercise: Evidence from Loss Severity,'' *Journal of Real Estate Research* 19 (5), 543–555.

Kau, James B. and Donald C. Keenan. 1993. ''Transaction Costs, Suboptimal Termination, and Default Probabilities for Mortgages,'' *AREUEA Journal* 21(3), 247–63.

Kau, James B. and Donald C. Keenan. 1997. *Patterns of Rational Default,* unpublished working paper, University of Georgia.

Lekkas, Vassilis, John M. Quigley and Robert Van Order. 1993. ''Loan Loss Severity and Optimal Mortgage Default,'' *AREUEA Journal* 21 (4, Winter), 353–372.

D. Multifamily Default/Prepayment

1. Introduction and Conceptual Framework

This section describes how OFHEO developed its model of multifamily default and prepayment rates for use in the risk-based capital stress test. The same theory that underlies the single family default/prepayment models, financial options theory, also underlies OFHEO's modeling of mortgage performance for multifamily loans. However, the single family approach is modified to account for the importance of property cash flows in the default decisions of investors. This theoretical framework treats mortgage terminations as a function of their financial value to the borrower. Both the single family and multifamily default/prepayment models also use a multinomial logistic specification to estimate the impact of explanatory variables on default and prepayment rates. Beyond these similarities in general approach, however, there are significant differences in the specifics of model construction and estimation.

Many of these differences reflect special features of multifamily mortgages. For these loans, the borrowers are all investors, and that affects the determinants of credit risk. Two key financial ratios are used in commercial mortgage underwriting: the DCR and the LTV. DCR is a property's net operating income (NOI) divided by the mortgage payment.[244] DCR indicates how much cash there is available for loan repayment after operating expenses are paid. LTV is the ratio of the UPB to the value of the property; it measures

[242] Such loans become part of the Enterprise retained portfolios once they are bought out of the security pools.

[243] These lesser sources of credit enhancements are items where the amount of recourse available to the Enterprises is not a function of per loan losses, but rather it is available in total dollar amounts for pools of loans.

[244] NOI is a measure of the differernce between full potential rent at market prices and operating expenses (including vacancy losses).

borrower equity.[245] Lenders concentrate on these two ratios at loan underwriting, and all major credit rating agencies start their analysis of the credit support levels needed to receive various rating grades with the DCR and LTV values of the loan collateral.

Multifamily mortgage modeling should also recognize the special features that differentiate commercial loans from single family residential loans. Commercial loans have prepayment restrictions, usually in the form of yield maintenance clauses, that severely reduce the value of refinancing during the early years of a mortgage. Commercial loans are also dominated not by fully amortizing 30-year loans, but by balloon mortgages with maturities of up to 15 years. These two product distinctions—yield maintenance and balloon terms—create different borrower incentives and different mortgage performance patterns for multifamily mortgages.

Previous research on multifamily mortgage performance has generally made simplifying assumptions to avoid having to deal with all of these issues in one model. First, research has tended to ignore DCR and only concentrate on LTV. Even then, without readily available property value indexes, researchers have not updated LTV over time to capture local market conditions.[246] Some studies have captured property cash flows, but they omitted LTV and had no mechanism for updating property cash flows for projection purposes.[247] One study that recognized the need for both DCR and LTV for predicting default rates, defined them to be perfectly correlated so that only one financial variable needed to be included in the model.[248] Another shortcoming of past research has been that default and prepayment have not been analyzed together.[249] Either defaults are assumed not to matter because of agency guarantees, or else prepayments are ignored because of yield maintenance terms. Most studies model defaults without prepayments, but prepayment studies are starting to appear, with three in 1997 and one in 1998.[250] In both default and prepayment studies, little work has been done to understand the dynamics of yield maintenance and balloon terms.[251] But even with all of these limitations in current research, the greatest concern is that researchers most often resort to pooling multifamily mortgages with loans on other commercial property types in order to have sufficient sample sizes.[252]

The broad conceptual framework chosen by OFHEO corresponds to the dominant paradigm in mortgage research, financial options theory. Studies that apply financial options theory to commercial mortgage performance have generally emphasized the role of borrower equity (LTV) in default rate estimation, but have not seriously modeled the role of cash flows (DCR).[253] However, because both DCR and LTV are critical credit risk dimensions, an appropriate multifamily mortgage performance model should also treat cash flows and equity as essential elements.[254]

For the default option to be in the money, the property must have both negative equity (LTV>1) and negative cash flow (DCR<1). The two sources of income for an investment property owner are rental (current) income and capital gains. Rental income can be thought of as dividend payouts from the property. Capital gains result when the property is sold. The owner holds the property until the expected annual rate of return from both dividends and capital gains becomes less than the return that could be earned by selling the property and investing the proceeds into another investment. However, if the rental market declines, and property equity becomes negative, then default becomes a viable option. This option will not be exercised as long as the dividend payout is positive. If property owners/borrowers were to default in the presence of positive cash flows, they would give up valuable cash flow streams. Therefore, default is only optimal if both equity and cash flow are negative. This implies that the dual condition, LTV> 1 and DCR<1, is required for default to occur.[255]

Prepayment options are in some ways simpler and in others more complex than default options. The simplicity arises because the financial value of prepaying a mortgage is directly measured by the mortgage premium value, the difference between the present value of future mortgage payments discounted at the current note rate, and present value of those same payments discounted at the current market rate. When interest rates fall, there is negative value to holding onto the existing mortgage, measured by a negative mortgage premium value. However, measuring the premium value itself is complex because of yield maintenance and balloon terms. When a

[245] Commercial loan underwriting also includes examinations of borrower credit, servicing capability, site and engineering reviews, and cost certifications for new construction. Market condition reports are part of the appraisal process used to estimate LTV at loan origination.

[246] Vandell (1992) and Vandell, et al. (1993) develop models of commercial mortgage default that update LTV over time using a national property-value index, along with the property-value diffusion process introduced by Foster and Van Order (1984) for single family mortgages.

[247] *See* ICF (1991) and Pedone (1991). These studies adapt the work of Edward Altman (1981, 1983) to predict corporate bankruptcy to model multifamily defaults. Capone (1991) discusses the application of bankruptcy models to multifamily mortgages, and provides a review of this literature. A related line of literature discusses the relationship between lender and borrower in the default/bankruptcy process. Kahn (1991) and Mahue (1991) study the impact of foreclosure laws on the balance of borrower and lender bargaining strength at these crucial junctures. Riddiough and Wyatt (1994a, 1994b) explore the power of lender signals of intent to pursue debt collections on distressed-loan foreclosure.

[248] Abraham (1993b).

[249] The first known attempt outside of OFHEO to model default and prepayment rates simultaneously was by Boyer, Follain, Ondrich, and Piccirillo (1997), who studied FHA insured mortgages.

[250] Abraham and Theobald (1997), Elmer and Haidorfer (1997), Follain, *et al.* (1997), and Capone and Goldberg (1998).

[251] In a theoretical pricing model, Kau, *et al.* (1990) do attempt to show how prepayment restrictions impact both default and prepayment options with balloon mortgages.

[252] The lack of historical data has often been cited as a major obstacle to research on multifamily and commercial loan credit risk (DiPasquale & Cummings, 1992; Standard & Poors, 1993; and Vandell, *et al.,* 1993). Studies that combine multifamily with other commercial mortgage types include Vandell (1992), Vandell, *et al.* (1993), Barnes and Gilberto (1994). Studies that use only multifamily data tend to model FHA-insured loans (Goldberg, 1994; ICF, 1991; Follain, *et al.,* 1997). Exceptions to this include Abraham (1993a, 1993b), who used multifamily loan data from Freddie Mac to study defaults, and Abraham and Theobald (1997), who use Freddie Mac data to model multifamily prepayment rates. Elmer and Haidorfer (1997) use Resolution Trust Corporation data to study multifamily prepayment rates. Researchers at OFHEO have published a default study based on Enterprise data (Goldberg and Capone, 1998).

[253] Even theoretical ''pricing'' models that simulate default rates on a pool of newly originated mortgages make simple assumptions that cash flow to the property owner is a fixed percentage of property value (Titman and Torous, 1989; Kau, Keenan, Epperson, and Muller, 1987 and 1990). They also treat cash flow as something negative (detracts from potential future property value) rather than something positive to the investor/owner/borrower.

[254] Abraham (1993b), Goldberg (1994), and Quercia (1995) have all questioned the sufficiency of net equity as a default trigger.

[255] The wealth-maximizing borrower should default if the property expects to have negative equity and negative cash flow from this point on. If there are negative cash flows, delaying default would lower wealth. If negative equity and negative cash flow were expected to be only temporary conditions, default would not be optimal. In principle one should incorporate expectations regarding rental markets and interest rates, simulate wealth over time, and have the borrower default only if it maximizes wealth over some long-run investment horizon. This was viewed as an overly complex, expensive, and therefore unfeasible approach. Theory notwithstanding, researchers typically construct the default option value variable using just current year information. This is also the approach taken by OFHEO. For relevant theoretical studies, *See* Kau *et al.* (1987, 1990), Brennan and Schwartz (1985), Dyl and Long (1969), Joy (1976), and Robichek and VanHorne (1967).

fixed-rate loan is under yield maintenance, it may refinance, but it will not accrue any value from the transaction until the yield maintenance period expires.[256] With balloon loans, there is the added uncertainty surrounding the contractual requirement to find new funding at loan maturity. Risk averse borrowers, therefore, may desire to refinance in the pre-balloon period even if the call option is not in the money.

An additional consideration for modeling prepayment speeds is that investors desire to leverage their investments to maximize return on equity. Interest rate spreads do not, therefore, provide the only incentive for refinancing a mortgage. To maximize leverage requires maximizing LTV ratios, within bounds set by lenders. Over time, investors will engage in cash-out refinancings in order to rebalance the ratio of debt to equity in the property. This second prepayment incentive can be captured by the LTV of the mortgage.

In modeling multifamily mortgage default rates, OFHEO distinguishes among the various programs of the Enterprises. Conventional multifamily loan purchases by the Enterprises began in 1983, and include ''cash'' and ''negotiated'' programs. Under the cash programs, the Enterprises purchased newly originated individual loans underwritten according to their own guidelines. Historically, most of these loans were retained in the portfolios of the Enterprises. Some ''cash'' loans were swapped for MBS, and this type of transaction is becoming more common. In a negotiated transaction, an Enterprise swaps pools of seasoned (i.e., aged and performing) loans for securities. These loans need not meet the underwriting guidelines of cash programs, and they are priced according to the risk of the loans in the pool. In negotiated transactions, unlike cash purchases, an Enterprise often requires credit enhancement from the seller/servicer to cover expected credit losses.

The initial cash programs exposed the Enterprises to significant credit risk in the late 1980s and into the 1990s. This exposure was due to generous appraisal practices used in the 1980s and to other significant weaknesses in those programs that do not exist today. Fannie Mae changed its cash program in 1988. Freddie Mac continued to build a portfolio of less-than-investment-grade mortgages through 1990. The poor performance of this portfolio led to a three-year moratorium on Freddie Mac's new purchases of multifamily loans, and a complete overhaul of the multifamily operations of the Enterprise.

Prepayment rates were modeled by loan characteristics product type rather than program type. This breakdown captures the differences in financial incentives to prepay that exist when yield maintenance penalties are or are not in effect, and the impact on defaults of balloon mortgage maturity. Balloon maturity is a significant multifamily modeling issue for the stress test because, in an up-rate interest rate environment, balloon loan borrowers are often required to pay off the existing mortgage and refinance, at much higher interest rates than property financials are currently supporting. In order to refinance at the balloon point in the up-rate scenario, property income must be higher than the minimum necessary to qualify for a new loan under the original interest rates. Therefore, it is important to model both the expected default and payoff rates of loans at balloon maturity for the stress test.

Section 2 of this supplementary material on multifamily default/prepayment provides a review of the historical data used to estimate the statistical models, and section 3 reviews the statistical procedures employed. Section 4 completes the description of the statistical model with explanations of the development of the explanatory variables. Section 5 presents and reviews the results of statistical estimations, and section 6 concludes with a discussion of how the estimated statistical equations are applied in the stress test.

2. Historical Data

a. Enterprise Loan Records

OFHEO used the combined historical experience of the Enterprises, 1983–1995, to estimate the statistical model of default and prepayment rates. This experience provided a large and rich data base that encompasses three different programs: the initial cash purchase programs that had high default rates; negotiated purchase (or transactions) programs where securities were swapped for pools of seasoned and performing mortgages; and new cash purchase programs that corrected flaws in the original programs and have experienced low default rates.

The historical data includes 35,759 conventional multifamily loans.[257] After eliminating missing or erroneous records, the sample includes observations on 21,994 loans: 12,845 from Freddie Mac and 9,149 from Fannie Mae. Of these, 61 percent are cash purchases and 39 percent are negotiated purchases. The final cash purchase sample is more complete than the negotiated purchase sample because, in negotiated programs, the Enterprises have relied more on buying seasoned portfolios with (limited) credit risk recourse to the seller/servicer, rather than on gathering enough property financial characteristics to re-underwrite the loans.[258]

The database was expanded by creating annual observations from loan acquisition to the termination year, or to 1995 if no termination occurred. The loan-year file includes 89,577 loan-year observations for cash purchases, and 59,415 observations for negotiated purchases. Cash purchases appear in the database with origination years from 1983 to 1995. The negotiated loans, however, have origination years as early as 1970 because they were often highly seasoned at time of acquisition. Annual observations are used, rather than monthly or quarterly observations, because of the relatively small number of multifamily termination events. If quarterly or monthly event histories were used, there would be significant numbers of time periods in which there were no terminations.

To avoid any possible statistical bias resulting from not having records of loan terminations prior to 1983, negotiated purchase loans enter the database starting in the acquisition year, rather than the origination year. But they enter at their proper age and are not treated as new originations at the time of acquisition. The same issue of potential ''left censoring'' bias also appears for certain cash purchase programs, where the Enterprises did not begin to maintain systematic records of loan terminations until 1991. For such programs, the loans do not enter the statistical estimation sample until 1991.[259]

[256] ARM loans have minimal penalties, and they have prepaid much more often in the early years after loan origination.

[257] Fannie Mae has maintained a portfolio of FHA-insured multifamily mortgages over time. OFHEO chose not to model performance of these loans, but rather to assign default and prepayment rates according to conventional loans with similar features. Because FHA pays for nearly 100 percent of default losses, the stress test imposes no credit losses on FHA-insured mortgages on the stress test.

[258] Ninety percent of cash purchases are retained in the final sample, while only 41 percent of negotiated purchases had enough loan characteristics data to be kept in the sample. For the 41 percent of negotiated purchase loans in the sample, DCR values at time of acquisition were estimated by OFHEO by first estimating net operating income (NOI) as NOI=value at origination divided by an estimate of the average CAP rate multiplier for the year, divided by the mortgage payment amount.

[259] The left-censoring bias would result if the statistical model used complete loan-history records

For cash loans, the default outcome of record is a foreclosure or foreclosure alternative that still provides for the property to be liquidated.[260] For most Fannie Mae negotiated purchase loans, however, the default event of record is a 90-day delinquency. This is because, for Fannie Mae negotiated transactions, the loan is repurchased by the seller/servicer if it becomes 90-days delinquent. The seller/servicer then bills Fannie Mae for resolution costs, and these are deducted from a limited recourse pool originally established with funds from the seller/servicer at time of acquisition. OFHEO recognizes that 90-day delinquencies cannot be treated as full default events, and makes adjustments in the statistical model.

b. Rents and Vacancies

OFHEO uses a unique approach to property valuation that uses local market indexes of rent growth rates and vacancy rates to update net operating income, and through that, update DCR and LTV over time. Rent growth rates came from the residential rent component of the CPI for each of the four Census regions, and for the 29 MSAs covered by Bureau of Labor Statistics (BLS) surveys. Most MSA level CPI series produced by BLS start in 1970, but some do not begin until the 1980s. The regional CPI series are available beginning in 1978, so percent changes for these can only be computed starting in 1979. To capture rent growth rates for each year, partial MSA series were completed with regional series starting in 1979 and national series before that. The regional series themselves were also filled in for the pre-1979 period with percent changes in the national CPI residential rent series.

Vacancy rates were obtained from the Bureau of the Census H–111 series. These are available for the same MSAs as is the CPI residential rent series (back to 1970), and for Census regions, and, beginning in 1986, for the 50 States plus the District of Columbia.[261] As with rent growth rates, the most disaggregated index available was used for each loan, in each calendar year.

c. Tax RatesOFHEO required tax rate data for calculating the present value of depreciation writeoffs (see discussion of the explanatory variable, *DW*, below). In order to compute weighted average tax rates, OFHEO used Internal Revenue Service (IRS) data on the income distribution of taxpayers with net capital gains. For 1983–90, data on adjusted gross income for taxpayers with net capital gains were obtained from the IRS publication, *Individual Income Tax Returns* (annuals). For 1991–95, data were obtained from IRS, *Statistics of Income Bulletin* (quarterly). These income-class weights were used to compute weighted average tax rates for both capital gains and ordinary income.

The marginal tax rate on ordinary income used here is for Married Filing Jointly taxpayers (Schedule Y–1). Five percent was added to the Federal tax rate for State income taxes. Schedule Y–1's for 1983–95 were obtained from Internal Revenue Service, *Package X* (annual publications 1983–95). Data on capital gains tax rates were obtained from IRS's *Package X*, for 1983–95. No adjustment was made for State taxes on capital gains.

Data on depreciation schedules is for newly constructed residential rental property, from the IRS publication, *Depreciation 1992,* Publication 534. This publication includes accelerated schedules for years 1983–92. Accelerated depreciation was assumed in years in which it was an option. Because there were no changes in the tax code affecting depreciation after 1992, the schedule for 1992 was used for 1993–95.

3. Statistical Estimation

The statistical estimation involves binomial logistic regressions of subsets of the data. There are two separate regressions for default rates and five separate regressions for prepayment rates. This breakdown accommodates programmatic differences between cash and negotiated purchases in the default equations, and the changing nature of prepayment incentives across various products and loan terms. The results are matched together so that the end result is trinomial logistic probability equations that provide the same result as if defaults and prepayments were estimated simultaneously for each loan program and product.[262]

The logistic model is founded on assumptions that the utility of each borrower payment choice—make payment, prepay, or default—is a function of its contribution to wealth and that, each observation period, borrowers make the choice that maximizes wealth. The regressions compute weights (coefficients) that estimate the influence of each explanatory variable on the net wealth effect of one choice over another. These models estimate the log-odds of choosing a mortgage termination over continuing to make loan payments as a function of the explanatory variables. In particular,

$$ln\left(\frac{\text{probability of default}}{\text{probability of continuing payments}}\right) = X\beta \qquad \textit{(Eq. 18)}$$

and

$$ln\left(\frac{\text{probability of prepayment}}{\text{probability of continuing payments}}\right) = Y\Gamma \qquad \textit{(Eq. 19)}$$

for all loans, when some groups of loans only enter the sample if they survive to a certain point (e.g., time of acquisition by the Enterprise). If the sample were not censored at the acquisition point, the model could severely underestimate the rates of loan termination in the early years of a mortgage.

[260] Foreclosure alternatives include third party sales where a ''third party'' purchases the property at the foreclosure auction; short sales, where the Enterprise finds a buyer for the property prior to completion of foreclosure; and note sales, where the mortgage itself is sold to another investor.

[261] Census also added more MSAs starting in 1986. These were not used in OFHEO's statistical analysis.

[262] This is the three-choice logit model, though the more generic model is known as the multinomial logit, or MNL.

where:

ln = natural logarithm

X = matrix of explanatory variables (columns) by loan record (rows)

β = (column) vector of coefficients (weights) to be estimated

Y = matrix of explanatory variables (columns) by loan record (rows)

Γ = (column) vector of coefficients (weights) to be estimated

And the resulting equations for calculating probabilities are transformations of these equations:

$$\text{Probability (default } X, Y) = \frac{e^{X\beta}}{1 + e^{X\beta} + e^{Y\Gamma}} \qquad (Eq.\ 20)$$

and

$$\text{Probability (prepayment } X, Y) = \frac{e^{Y\Gamma}}{1 + e^{X\beta} + e^{Y\Gamma}} \qquad (Eq.\ 21)$$

If *X* and *Y* are matrices of all event-history records, then the resulting probabilities will be (column) vectors of estimated probabilities for each of these records, for each observed time period. Because of the relatively small number of loan defaults in the data, OFHEO used annual observations to estimate the equations. Economic variables are averages for each calendar year, and the logistic equations estimate probabilities of default and prepayment for all loans surviving to the beginning of the next year.

The probabilities of default and prepayment are interdependent, and normally the equations would be estimated using simultaneous equations methods. However, because there are two default equations and five prepayment equations, doing so would be quite complex. Following Begg and Gray, OFHEO estimated the system using single equation methods in which separate binomial log-odds equations are estimates for default and prepayment.[263]

4. Explanatory Variables

The multifamily mortgage performance model has separate sets of explanatory variables for default and prepayment analysis. They are described separately here.

a. Default Equations

OFHEO estimated two separate logit default equations, one for cash purchases and one for negotiated purchases. This decomposition serves three purposes. First, significant numbers of negotiated purchase loans did not enter the Enterprise portfolios until after the Tax Reform Act of 1986. That statute greatly changed the value of depreciation allowances to new purchasers of investment real estate. OFHEO desired to model the effects of tax law changes on default rates, but could only do this with the cash purchase loans, where there are significant numbers of observations both before and after tax reform. The second reason for separating cash from negotiated purchase loans is that negotiated loans did not undergo the same change of quality as did cash purchases. It is easier to separate the effects of movements by the Enterprises from original to new cash-purchase programs if these are isolated from the negotiated purchases for default analysis. A third reason for separating the two programs into two separate default equations is that the majority of negotiated purchase loans have seller/servicer repurchase provisions, which required use of 90-day delinquency as the default event of record. OFHEO decided that capturing the difference between 90-day delinquencies and full defaults was best achieved through an estimation that involved only negotiated purchases.

Table 33 provides a list of the explanatory variables used in each default equation. Each variable listed in the Table will be described and developed more fully below.

[263] *See* Begg and Gray (1984). To do this, one must be sure to censor competing termination events from the regression samples. That is, for default rate log-odds estimation, all prepayment observations must be censored in the period of the prepayment (and vice versa). This censoring assures that the estimation is of the log-odds of defaulting (or prepaying) versus remaining current on the mortgage. The underlying principle of logistic regression analysis that allows for this approach to modeling the competing risks of default and prepayment is called the independence of irrelevant alternatives. This principle means that logistic analysis assumes that the log-odds of default versus remaining current are not influenced by the log-odds of prepaying versus remaining current.