HOW MUCH MORTGAGE POOL INFORMATION DO INVESTORS NEED?

Paul Bennett, Richard Peach, Stavros Peristiani*

January 29, 2001

Investors in pools of single-family mortgage loans may have only limited information about the individual loans within a pool. Would more information be useful? We use data on individual loans to estimate a model of sales, refinancings, and defaults. Then we construct hypothetical loan pools and examine their prepayment sensitivity to collateral and credit information not universally made available to investors. Simulations show that loan-level data can be extremely valuable in predicting pool durations. In particular, information on the distributions of homeowners' loan-to-value ratios -- and to a lesser extent on their credit scores – can be quite important in identifying fast- from slow-paying pools.

Send correspondence to: Paul Bennett, 33 Liberty Street, Main 3 East, Federal Reserve Bank of New York, New York, NY 10045. Tel: (212)-720-5647, Fax: (212)-720-1582.
Email: paul.bennett@ny.frb.org.

*The authors are economists at the Federal Reserve Bank of New York. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

1. INTRODUCTION

A primary consideration for the pricing of mortgage securities relative to treasury securities is prepayment risk. Even when mortgages are fully insured against credit loss, market participants are exposed to cash re-flows when homeowners sell a home or refinance a loan. Investors in pools of mortgages, in mortgage-backed securities (MBS), in mortgage servicing rights, or in collateralized mortgage obligations (CMO), rely heavily on formal prepayment models to value these option features of mortgages.

Industry prepayment models employ a variety of econometric techniques to estimate the influence of financial factors. The variable of interest is often a measure of prepayment speed for a well-defined cohort of MBSs. This dependent variable is usually regressed on the incentive to terminate a mortgage (interest rate spread), the average age of the pools, seasonality, and other relevant variables. Because of a variety of systematic influences, however, comparable MBSs often exhibit very different prepayment rates (Bykhovsky and Hayre 1992 and Abrahams 1997). Recent studies suggest that borrower-specific attributes, such as post-origination home equity or credit scores, may be at least as important in explaining mortgage termination rates as traditional variables such as relative interest rate measures. Due to data limitations, however, many models used by investors have not been able to control for these loan-specific effects on prepayment and duration. Asymmetric information increases the potential for investors with less data to incur losses, demand higher yields, or to avoid the market altogether. Recently, some issuers have begun reporting more detail on the make-up of pools, such as data on geographic distribution of the remaining balance of the pools.

This paper aims to shed light on the potential importance of loan-level information in

pricing MBSs more accurately. Having experienced several major prepayment cycles, market participants are more mindful of the limitations of traditional valuation approaches. While losses incurred by MBS investors in the early 1990s may have stemmed partly from their inability to anticipate interest rate changes, traditional models of prepayment also may have failed to capture factors influencing extension risk.

This article demonstrates that borrower characteristics can significantly alter the profile of prepayments in a pool. Loan-level information can therefore play an important role in more accurately pricing mortgage pools and passthrough securities, by implication tightening yield spreads and increasing liquidity in the secondary market. Overall, pool underwriters should be able to add value to investors and strengthen the market by making better pool-specific information available.

In the next section, we present a multinomial logit model of mortgage termination making use of borrower and loan-specific characteristics. Simulation experiments demonstrate how much borrower characteristics can reduce or extend the duration of passthrough securities, implying potentially large pricing differences among otherwise apparently similar securities.

2. A STATISTICAL MODEL OF MORTGAGE TERMINATION

The literature in housing finance examining home sales, refinancings, and defaults is voluminous.[1] Typically, home sales, refinancings, and defaults are lumped together because researchers cannot separately observe these prepayment outcomes. Refinancings and home sales account for the bulk of mortgage terminations, while mortgage defaults are rarer. Deng,

[1]The reader can refer to a number of useful articles such as Vandell [1993], Quercia and Stegman [1992], and Archer, Ling, and McGill [1995].

Quigley, and Van Order [2000] estimate from a Freddie-Mac database of prepayment and default histories that nearly 70 percent of borrowers sell their home or refinance after about 5 years. In contrast, the cumulative rate of default over the same horizon is typically less than 1 percent for loans with a loan-to-value ratio less than 90 percent. Despite the small likelihood of default, the option to prepay is not independent from the decision to default. As is shown by the literature on callable and defaultable bonds, the joint presence of two competing options (in our case, a put default option and a call prepayment option) alters the optimal exercise boundaries.

A borrower's multiple choice to sell a property, refinance, or default can be modeled using a logit model of competing risks. The multinomial logit model asserts that the conditional probability that borrower (i) will make choice (j) at time (t) assuming that no other event has occurred prior to that point of time is P_{tij}. The conditional probability depends on a number of exogenous factors. Using the log-odds ratio, we can simply express the multinomial logit as

$$\log\left(\frac{P_{tij}}{P_{ti0}}\right) = \beta_j \; x_{ti\bullet}$$

Here, the variable P_{ti0} represents the probability that the homeowner does not choose any type of prepayment, and $x_{ti\bullet}$ is the vector of explanatory variables.

Data

The data for this study were obtained through the Mortgage Research Group (MRG) of Jersey City, New Jersey. Aside from limiting the sample to complete observations, we further restricted it to a manageable size for computational purposes. We selected four clusters of

counties in different geographic regions of the country (New York/New Jersey, Central Florida, Chicago, and Los Angeles). To identify refinancings and defaults, we simply selected the most recent purchase transaction, going in some cases as back far as January 1984. The borrowers in many of these properties subsequently refinanced (or in some instances chose to default). The remaining properties had no further transactions recorded through the end of our sample period, December 1994. We identified home sales by isolating properties with two consecutive purchase transactions, which signify that the original purchaser of the property chose to sell the property to another individual. The resulting sample consists of 12,835 fixed-rate borrowers, of which 4,226 were refinanced, 148 experienced default, and 2,668 were eventually sold to a third party.

Variable Definitions

Our study uses the generalized multinomial logit framework to estimate a model of monthly mortgage termination. The variable of interest in logit analysis is the probability that the borrower will sell, choose to refinance, or default in a given month. An important market friction that affects all three choices is the underlying value of equity in the home. Clearly, the likelihood of default is greater when the value of the house is less than the market value of the mortgage. At the same time, however, home equity has an offsetting effect on the probability of selling or refinancing, since insufficient collateral makes refinancing or moving up to a costlier house difficult. To measure the effect of changes in home equity on the probability of refinancing, we include in the logit regression a proxy of post-origination loan-to-value ratio (LTV).[2]

[2]The numerator of the ratio is the amortized balance of the original first mortgage on the

The intrinsic value of the mortgage can be computed by comparing the noncallable value of the loan using the contract rate on the existing loan and the prevailing market rate. A simple measure of intrinsic value is given by the coupon rate spread (SPREAD). Although in some of our previous work on prepayment we have employed more complex ways to measure intrinsic value, in this paper we use the simple spread measure because it is more illustrative in the interest rate simulation experiments that will be presented in a later section. For all observations in our sample, the original rate is measured by the 10-year Treasury note rate for the month the loan was closed. Similarly, the prevailing market rate for the mortgage holder at month (t) is also the 10-year Treasury rate at month (t).

In addition to the intrinsic value of the mortgage loan, the decision to terminate a loan also depends on the time value of the embedded options. Although the values of the put or call options are not observable directly, we can proxy their key determinants (e.g., time-to-expiration and asset volatility). The volatility of the underlying assets plays a critical role because the mortgage options would be more valuable when volatility is expected to rise in the next period. For refinancings, an unbiased measure of volatility is provided by the implied volatility of the option price of a 10-year Treasury note futures contract that is traded on the Chicago Board of Trade. This implied volatility measure (denoted as σ_{BONDS}) is forward looking because it is based on underlying futures contracts. In defaults, the underlying asset is the value of the property. We used a simple econometric model of house returns to construct a forward-looking

---

property using standard amortization formulas for fixed rate mortgages and the interest rate assigned to that loan. The denominator is the original purchase price indexed using the Case-Shiller-Weiss repeat-sales home price index for the county in which the property is located.

measure of conditional variation in house returns (denoted as σ_{HOUSE}).[3] Finally, we control for the time-to-expiration of the mortgage options by the age of the loan (AGE).

The logit regression also controls for household transactions cost by including in the regression the average points and fees paid on conventional fixed rate loans closed in that month (POINTS). To analyze the impact of credit, we use the TRW credit report of the owner of the property. Since we only observe a snapshot of the individuals credit at one point in time, our analysis uses a worst-ever credit measure (CREDIT) that accumulates the history of the property owner across all credit lines.[4]

The role of transaction costs has been hotly debated in the mortgage default literature (for opposing views, see Kau, Keenan, and Kim 1991, and Quigley and Van Order 1992). Some researchers have argued that the evidence is consistent with a "ruthless" exercise of the default option whenever the value of the property is less than the value of the loan. Others contend that the decision to default is strongly influenced by a string of transaction costs (that is, reputation cost, moving cost), which inhibit the exercise of the default option.

The set of explanatory variables includes also the logarithm of the monthly mortgage payment as proxy of size (SIZE). Finally, the logit regressions account for regional variation in prepayments by including regional unemployment rate (UNEMPLOYMENT) and appropriate geographic dummies. Table 1 summarizes in more detail all the explanatory variables.

---

[3] To estimate the housing volatility, we use a generalized autoregressive conditional heteroskedasticity (GARCH) model. This approach allows us to model simultaneously the return as well as the variance of housing returns.

[4] Because in sales (and defaults) there is a change in the ownership of the property, we cannot say with certainty that the cumulative credit snapshot belongs to the original owner (seller) or the new occupant (buyer) of the house. As a result, the credit variable is not included in the sale and default regressions.

3. RESULTS OF MULTINOMIAL LOGIT MODEL

Table 2 presents maximum likelihood estimates of the multinomial logit model. Each column of the table represents the regression coefficients β_j, which measure the response of each explanatory variable for the three mortgage termination choices (represented by index j). The model chi-square statistics presented at the bottom of the table reject the null hypothesis that the parameter vector β_j is equal to zero.

The logit model for competing risks demonstrates that financial factors influence the three termination choices in different ways. The level of home equity in a property has a significant impact on the borrower's willingness to terminate an existing mortgage. As expected, the current LTV ratio coefficient is negative and very significant for refinancings. Homeowners with insufficient collateral may not be able to refinance their mortgage loan even if economic conditions are in their favor. The coefficients on the mortgage spread (SPREAD) for the refinancing is significant and positive, suggesting that the likelihood of refinancing is higher when the intrinsic value of the mortgage loan is bigger. The effect of points and fees on refinancing is negative and strongly significant, confirming that higher transaction costs raise the refinancing hurdle faced by homeowners. To a lesser extent, we find that a borrower's creditworthiness (CREDIT) has also an adverse effect on the probability of refinancing. More important, our analysis shows that expected interest volatility has negative and statistically significant effect on the likelihood of prepayment. This result confirms that the value of the call increases when the volatility of the noncallable asset is expected to rise. A borrower may be motivated to postpone refinancing because the benefit from refinancing may be greater in the next period.

Not surprisingly, the relationship between home sales and the explanatory variables is somewhat weaker as relocating homeowners are also affected by other unobservable idiosyncratic factors, such as job changes. Collateral value, however, continues to play an important role in the decision to sell a property. The negative coefficient on current LTV shows that the likelihood of a sale is greater for properties with a higher home equity value. The coefficient estimate of SPREAD has the right positive sign; however, the coefficient on SPREAD is smaller because home sellers may move for many reasons, diluting the pure interest rate effect, at least relative to the refinancing equation. Some sellers clearly would prefer to have a favorable interest rate environment, especially when they have to sell their property to upgrade into a large house. For these homeowners, the sale of the property is not very different than a refinancing.

The last column of Table 2 presents the logit equation for the likelihood of default. As expected, collateral value is the most important factor in the decision to default. The positive and significant coefficient on current LTV shows that borrowers with negative equity are more likely to exercise their "in-the-money" put option. The failure regression also reveals that the probability of default is negatively related to the expected volatility of the return on housing σ_{HOUSE}. This result implies that borrowers with valuable default options may choose to defer default seeking to gain a higher value for their put option from a further decline in house prices.

4. SIMULATION RESULTS

With the growth of the secondary mortgage market, market participants have devoted considerable resources to developing formal statistical models of prepayment. A prepayment model is essential to the mortgage valuation process because it enables investors to determine

cash reflows (see Fabozzi 1992 and Hayre, Chaudhary, and Young 2000). To further illustrate the importance of borrower/loan characteristics in pricing MBSs, this section presents a number of simple simulation examples. In a way, our approach is very similar to the option adjusted spread (OAS) methodology, which is used extensively in mortgage valuation (Hayre 1990).

The simulation approach can be divided in three broad steps. First, we use the multinomial logit model to construct an artificial pool of mortgages. Using the sampling distribution of the logit coefficient estimators, we create pools that mimic the underlying sample of homeowners. We allow for randomly heterogeneous prepayment propensities, including the responses to exogenous determinants of prepayments. In other words, the "average" prepayment behavior of borrowers, defined by the empirical logit equations (Table 2), is randomly shocked to generate heterogeneity across borrowers. For simplicity, we assert that each pool contains 1,000 mortgages of equal original balance having an average coupon rate of 9 percent.[5]

The second critical step in the simulation involves the specification of the interest rate diffusion process. For our simulations, we use long-term treasury rate scenarios generated on actual historical yields. The drift in the estimated auto-regressive model for long-term rates reflects the declining interest rate environment characteristic of the sample period. We also apply this approach to generate values for points and fees that are again consistent with historical trends.

In the third step of our simulation, we combine the interest rate and transaction costs realizations with the behavioral logit equations to determine pool-level prepayments. We then compute the weighted average life (WAL) and a Macaulay modified duration measure for the

[5]We also assume a pool-servicing fee of 0.5 percent, meaning that the pass-through rate

simulated pools. The artificial pool of 1,000 mortgages is evaluated at 500 different interest rate and transaction paths that start at month 1 and terminate at month 360. At each monthly node, we determine the probability of refinancing, sale, and default for each mortgage conditional on the value of the interest rate spread, and points and fees. The decision to terminate is evaluated at five levels of initial LTV (50, 60, 70, 80 and 90 percent) and three values of credit quality (CREDIT=1 good credit, CREDIT=90-days delinquent, and CREDIT=400 default). The original levels of LTV evolve according to the two house appreciation scenarios presented in the two panels of Table 3. All other explanatory variables are evaluated at their means, except the contribution of the variable AGE, which is determined by the value of the month. The mortgage is assumed to terminate if one of the three competing termination choices has a probability greater than 0.5. Note that by definition the sum of the probabilities for all four outcomes (refinancing, sale, default, no action) must be one. A probability cutoff of 0.5 therefore ensures a unique outcome at each monthly node. Once the mortgage is terminated, it is excluded from the pool and the cash flows are adjusted.

Table 3 summarizes the WAL, modified Macaulay duration, and cumulative prepayment of the pool for different levels of LTV and CREDIT. Not surprisingly, the simulations show that pools with high quality borrowers exhibit greater prepayment speeds. However, the extent of the difference in duration is quite striking. Looking at the top panel of the table, we find that the WAL of these artificial pools extends from 5.42 years for the highest quality pools to 18.94 years for pools composed of borrowers with an initial loan-to-value ratio of 90 percent and poor credit. In a similar fashion, the modified Macaulay duration rises from 3.59 years to 8.59 years.

---

to investors is 8.5 percent.

The bottom panel of Table 3 summarizes simulation experiments that assume an environment of declining housing prices. A drop in housing values leads to gradual deterioration of collateral value, resulting in a further duration extension of the passthrough securities. In addition to the two duration measures, the table provides the average cumulative prepayment for each level of LTV and CREDIT. Looking again at the rising house prices example (top panel), we find that pools composed of high quality credit borrowers and an initial LTV of around 70 percent exhibit a cumulative prepayment of 58 percent. Overall, the cumulative rate of prepayment varies from 79 percent for scenarios with the highest quality pools (initial LTV=50, good credit) to 11 percent for simulations based on collateral- and credit-constrained households (initial LTV=90, poor credit).

5. CONCLUSION

The large disparity in duration among these different passthrough simulation examples raises some important implications for market participants seeking to hedge a portfolio of mortgage securities. Investors in MBSs and other mortgage-based derivative products aim to alleviate interest rate risk by offsetting their long MBS positions with short positions in treasury securities. For example, as interest rates rise, a typical hedging strategy would be to counterbalance the increasing duration of the MBS assets by extending the duration of the short position in treasuries. Our simulation analysis, however, shows that borrower characteristics can significantly alter the extension risk of a portfolio of passthrough securities. Although the path of interest rates was exactly the same for the different simulation scenarios, lower quality mortgage pools prepaid at a much slower pace, resulting in significantly greater extension risk. Having very limited information on the underlying characteristics of mortgage securities,

investors are less likely to fully price these unobserved prepayment factors, resulting in a greater duration mismatch between their MBSs and treasury holdings.

REFERENCES

Abrahams, Steven W. "The New View in Mortgage Prepayments: Insight from Analysis at the Loan-by-Loan Level." *Journal of Fixed Income*, June 1997, 1-21.

Archer, Wayne, David Ling, and Gary McGill. "The Effect of Income and Collateral Constraints on Residential Mortgage Terminations." *Regional Science and Urban Economics*, pp. 235-261.

Bykhovsky, Michael, and Lakhbir S. Hayre. "Fact and Fantasy about Collateral Speeds." *Journal of Portfolio Management*, Summer 1992, pp. 63-66.

Deng, Y., J. Quigley and R. Van Order. "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options." *Econometrica*, March 2000, pp.275-307.

Fabozzi, Frank J. *Handbook of Mortgage-Backed Securities*. Chicago, Illinois: Probus Publishing Company, 1992.

Hayre, Lakhbir S. "Understanding Option-Adjusted Spreads and Their Use." *Journal of Portfolio Management*, Summer 1990, pp. 68-69.

Hayre, Lakhbir S., Sharad Chaudhary, and Robert A. Young. "Anatomy of Prepayments." *Journal of Fixed Income*, June 2000, 19-49.

Kau, James B., Donald C. Keenan, and Taewon Kim. "Transaction Costs, Suboptimal Termination and Default Probabilities." *American Real Estate and Urban Economics Association Journal*, Fall 1993, pp. 247-63.

Monsen, Gordon. “The New Thinking on Prepayments,” Mortgage Banking, October 1992, pp. 51-56.

Quercia, Roberto G. and Michael A. Stegman. “Residential Mortgage Default: A Review of the Literature.” *Journal of Housing Research*, 1992, pp. 341-70.

Quigley, John M. and Robert Van-Order. “Efficiency in the Mortgage Market: The Borrower's Perspective.” *American Real Estate and Urban Economics Association Journal*, Fall 1990, pp. 237-52

Peristiani, Stavros, Paul Bennett, Gordon Monsen, Richard Peach, and Jonathan Raiff. “Effects of Household Creditworthiness on Mortgage Refinancings.” *Journal of Fixed Income,* December 1997, pp. 7-22.

Vandell, Kerry D. “Handing over the Keys: A Perspective on Mortgage Default Research.” *American Real Estate and Urban Economics Association Journal*, Fall 1993, pp. 211-46.

TABLE 1. Variable Definitions for Explanatory Variables

Variables	Description	Mean			
		Refinancing	Home Sale	Default	No Choice
SPREAD	Spread between coupon rate and prevailing market rate (percent)	1.88	1.17	2.52	1.37
LTV	Current loan-to-value (percent).	54.93	56.45	81.57	70.50
RATE VOLATILITY, σ_{BONDS}	Implied volatility on options on the 10-year treasury note futures (basis points).	6.63	7.04		7.02
HOUSE VOLATILITY, σ_{HOUSE}	One-month ahead forecast of the volatility of housing returns (basis points).			6.13	17.64
POINTS	Initial fees and point changes on conventional home mortgages. National average for all major lenders (percent).	1.71	1.83	1.68	1.85
CREDIT	Worst delinquency ever (1 = good credit, 30, 60, 90, 120, 150, 180, 400=default).	64.79			113.74
SIZE	Logarithm of sale price of the house (in thousands of dollars).	148	115	169	111
UNEMPLOYMENT	County unemployment rate (percent)			9.46	7.28
Number of monthly observations		4,099	2,668	148	404,446

NOTES: The last column of the table represents only homeowners that have not terminated their mortgage.

TABLE 2. Multinomial Logit Model for Mortgage Termination
(Numbers in parentheses represent Wald chi-square statistics).

Variable	Refinancing	Home Sale	Default
INTERCEPT	2.985*** (76.58)	4.439*** (66.41)	-15.31*** (35.12)
CREDIT	-0.0009*** (46.82)		
SPREAD	0.291*** (110.84)	0.125*** (76.91)	0.392*** (12.43)
LTV	-0.025*** (1331.5)	-0.027*** (427.54)	0.040*** (52.47)
VOLATILITY, σ_{BONDS}	-0.122*** (29.49)	0.171*** (109.28)	
VOLATILITY, σ_{HOUSE}			-0.127*** (52.82)
UNEMPLOYMENT			0.431*** (33.76)
POINTS	-4.243*** (698.6)	-4.027*** (2024.9)	0.919 (2.17)
SIZE	0.145*** (29.95)	-0.187*** (23.74)	-0.134 (0.80)
AGE	0.053*** (57.68)	0.048*** (284.66)	0.007 (0.06)
AGE^2	-0.0008*** (22.82)	-0.0005*** (205.74)	0.0009 (2.01)
AGE^3	1.61×10^{-6} (1.96)	1.26×10^{-6} *** (143.91)	-7.2×10^{-6} ** (4.29)
χ^2 test $H_0: \beta_j = 0$	5,560.1***	4,196.66***	531.80***
Termination obs	4,099	2,668	148
No choice obs	458,663	457,585	399,557

NOTES: The Logit regressions also include geographic and quarterly dummy regressors. The symbols (***), (**), and (*) indicate statistical significance at the 1-, 5-, and 10-percent level, respectively. Table 1 describes in more detail the explanatory variables. The sample size for the group of borrowers with no choice also includes monthly observations for each mortgage termination choice (refinancing, sale, and default) up to the month before the homeowner terminates the mortgage.

TABLE 3. Duration and cumulative prepayment of simulated pools for different levels of credit and collateral values.

		CREDIT		
		GOOD	90-DAYS LATE	DEFAULT
		A. Rising Home Prices		
Original LTV				
50	DURATION	3.59	3.84	4.85
	WAL	5.42	6.03	8.60
	PREPAYMENT	78.28	75.41	62.83
60	DURATION	4.34	4.64	5.84
	WAL	7.25	8.05	11.25
	PREPAYMENT	69.45	65.59	49.68
70	DURATION	5.25	5.61	6.89
	WAL	9.64	10.58	14.14
	PREPAYMENT	57.73	52.83	35.07
80	DURATION	6.28	6.65	7.85
	WAL	12.42	13.47	16.85
	PREPAYMENT	43.75	38.55	21.49
90	DURATION	7.32	7.65	8.59
	WAL	15.31	16.27	18.94
	PREPAYMENT	29.15	24.37	10.79
		B. Declining Home Prices		
Original LTV				
50	DURATION	4.10	4.38	5.47
	WAL	6.55	7.24	10.15
	PREPAYMENT	72.41	69.09	54.83
60	DURATION	4.99	5.35	6.56
	WAL	8.95	9.82	13.23
	PREPAYMENT	61.47	56.68	39.99
70	DURATION	6.10	6.45	7.66
	WAL	11.88	12.86	16.24
	PREPAYMENT	46.47	41.62	24.35
80	DURATION	7.20	7.53	8.49
	WAL	14.98	15.94	18.65
	PREPAYMENT	31.11	26.37	12.43
90	DURATION	8.14	8.39	9.01
	WAL	17.68	18.40	20.08
	PREPAYMENT	17.77	14.12	5.15

NOTES: Pools are assumed to have a maturity of 30 years. DURATION = Modified Macaulay duration (in years); WAL = Weighted average life of pool (in years); and PREPAYMENT = Cumulative prepayment of the pool (percent). The value DEFAULT in credit indicates that a borrower has defaulted at least once on any credit line, including secondary lines of credit such as retail credit cards. Rising house price simulations assume that home values increase by 6 percent in the first two years, 2 percent in the third and fourth year, and 3 percent for remaining life of the pool. Declining house price simulations assume that home values decline by 6 percent in the first two years, 2 percent in the third and fourth year, and remain unchanged for the remaining life of the pool. Simulations assume that, on average, interests rates would decline from 9 percent to 6 percent after 5 years.