



The Performance of Low-Income and Minority Mortgages

Simon Firestone,* Robert Van Order** and Peter Zorn***

This article analyzes the performance of low-income and minority mortgages (LIMMs) from a large sample of fixed-rate conventional conforming mortgages. We find that low-income borrowers are less likely to prepay when it is optimal, whereas black and Hispanic borrowers prepay more slowly than other borrowers, regardless of the option's value. After controlling for equity, credit history and some other variables, LIMMs default slightly more frequently and have about the same loss severity as other loans. Our results suggest that, for most yield curve situations, differences in LIMM prepayment behavior have little effect on pricing.

Introduction

The purpose of this article is to analyze the performance of low-income and minority mortgage (LIMM) borrowers with respect to default and prepayment and to analyze the implications of these differences for mortgage pricing. Berkovec *et al.* (1994) and Van Order and Zorn (2000) examined the incidence of default by minority groups. Both found that default is higher, and thus performance is worse, among such borrowers. However, differences in prepayment speeds are an important factor in performance, and they must be included in any complete performance analysis. Default is a relatively rare event, occurring in approximately 0.6% of our sample; in contrast, over 92% of the mortgages in our sample prepaid during the period of observation. Slower exercise of prepayment options therefore could more than compensate for a greater frequency of default.

Our empirical analysis relies on loan-level data for over 1.2 million conventional conforming mortgages.¹ We find that, absent any controls for borrower or

*Housing Analysis and Research, Freddie Mac, McLean, VA 22102 or simon_firestone@freddiemac.com.

**University of Aberdeen and University of Michigan, Ann Arbor, MI 48109 or rvo@bus.umich.edu.

***Housing Analysis and Research, Freddie Mac, McLean, VA 22102 or peter_zorn@freddiemac.com.

¹ “Conventional” refers to loans that are not insured by the government. “Conforming” refers to loans that are eligible for purchase by Fannie Mae and Freddie Mac; this mainly has to do with loan size (a maximum that is indexed to house prices) and credit quality.

loan characteristics, LMM borrowers are more likely to default and less likely to prepay than the rest of our sample. These differences are robust to adding simple controls for credit-related and geographic variables. We exploit our large amount of loan-level data by analyzing the performance of “pseudopools” of loans matched by age, date of origination, credit quality, coupon, loan amount and loan-to-value (LTV) ratio. This approach allows us to control for complex path-dependent effects such as burnout, in which mortgages that survive repeated exposure to declines in interest rates prepay more slowly (see Hall 2000). It also allows us to control for seasoning, in which older mortgages are observed to prepay quickly relative to newer mortgages (see Richard and Roll 1989). We find that differences in prepayment speeds are large and robust to such controls.

In the next section we provide a brief review of the literature. In the third section we discuss our estimation strategy. The fourth section discusses the data. The fifth section discusses our default models, and the sixth section discusses our prepayment models and the implications of our results for pricing LIMMs. The seventh section concludes.

Literature Review

We extend the literature on the relationship between low-income and minority borrowers and the mortgage markets. It has long been recognized that low-income and minority households face multiple barriers to homeownership. Examples of such barriers include challenges to maintaining good credit and accumulating wealth, which Barakova *et al.* (2003) and others have identified as keys to attaining homeownership. Lacour-Little (1999) offers a good survey of the literature on discrimination in primary lending, which is another barrier often cited for these borrowers. Our results show that differences between LIMMs and other mortgages survive beyond the moment of tenure and mortgage choice and affect mortgage performance.

Archer, Ling and McGill (1996) analyze the prepayment behavior of low-income borrowers using American Housing Survey data and a logistic regression framework. They find that low-income households are not significantly different from others with regard to mortgage terminations. Deng and Gabriel (2006) analyze the behavior and pricing of LIMMs with Federal Housing Administration (FHA) loan data, controlling for both prepayment and default, and Goldberg and Harding (2003) analyze low- and moderate-income mortgage terminations with a sample of mortgages subsidized by a state housing finance authority. Both found a tendency for LMM-type loans to prepay more slowly than other loans, and Deng and Gabriel (2006) explored some pricing implications. We extend this analysis to a sample of 30-year fixed-rate conventional

conforming mortgages, which represents a much larger and more heterogeneous market.²

The issue of prepayment differences between LIMMs and other mortgages, and the implications for pricing, have been raised in Chinloy and Megbolugbe (1994). Their argument is based on the notion that low-income and minority borrowers are less mobile and less liquid than other borrowers, and as a result they prepay relatively slowly. However, that does not imply that LIMMs have a less valuable prepayment option. For example, lower mobility means that such borrowers have a longer time period over which they may exercise these options, which increases the options' value for the borrower. Banks, recognizing the borrower's higher option value, tend to raise rates on such mortgages. The key question is whether the options are exercised more ruthlessly, that is, when the prepayment option is in the money in the sense of the rate on the loan being greater than the current market rate. We find that LIMM borrowers generally refinance more slowly both when the prepayment option is in the money and when it is at or out of the money.³ This leads to somewhat ambiguous and small pricing implications.⁴

It is difficult to disentangle explanations for the slower prepayment speeds that we document. Peristiani *et al.* (1997) find that deterioration in credit and home equity impedes refinancing. Green and LaCour-Little (1999) show that a large fraction of households who fail to prepay when their option is in the money might be constrained by declining collateral values. Another explanation is a lower degree of mobility for such households, which would be consistent with

² We find the magnitude of the pricing differential between LIMMs and other mortgages is somewhat smaller than that in Deng and Gabriel's (2006) study. Deng and Gabriel (2006) analyzed portfolios of low credit risk and high credit risk mortgages, rather than comparing performance on the basis of borrower income or race. They also used FHA data where much of their sample was characterized by high LTV and debt-to-income ratios. Also, their results are for seasoned loans (see their Table 6). The newest pools they analyze are 3 years old. As seasoning declines, the differences in their results across groups become smaller and get closer to ours. They do not publish results for new pools, but interpolating their Table 6 suggests results that are within 15 basis points of ours.

³ Higher refinancing when the option is out of the money may be due to improvements in credit or a greater need to tap into home equity through cash-out refinancing. It could also be due to mismeasurement of the option's value.

⁴ This result is different from earlier drafts on this article in which in the raw data minority borrowers prepaid more slowly when the option was in the money but about the same otherwise. In revising this article we have recreated and extended the sample. We do find some difference in the direction of the results of the previous article for the pre-2000 period, though considerably smaller in magnitude. The results that we get for conditional estimates are very similar to our previous results.

the results in Goldberg and Harding (2003). Finally, low-income households may take out additional mortgages more frequently, and high LTV ratios (including all mortgage debt) may prevent refinancing, as discussed in Lacour-Little (2004).

Estimation Strategy

It is by now well established that prepayment and default behavior can be viewed as exercising options.⁵ Both prepayment and default are “American” options, meaning they can be exercised before maturity, but the options are costly to exercise and are not exercised in the way that, say, corporate bond options are exercised.⁶ This is clearly true for default because exercising the default option involves significant costs to borrowers (*e.g.*, worse credit history and diminished access to future credit, as well as moving costs). It is also true for prepayment; for instance, most mortgages are not assumable (the lender has the right to demand payment if the house is sold), so they are usually prepaid when the house is sold.

We follow other papers, such as Deng, Quigley and Van Order (2000) and Van Order and Zorn (2000) in modeling option exercise in a proportional hazard framework, using variables that affect the probability that an option is in the money and other variables that capture the likelihood of trigger events that affect the probability of exercising the option. We estimate prepayment and default probabilities with models of the form:

$$h(t) = \exp(Bx(t)), \quad (1)$$

where $h(t)$ is the instantaneous probability of the borrower prepaying or defaulting conditional on having survived (neither prepaying nor defaulting) until time t , x is a vector of explanatory variables and B is a vector of coefficients. The multiplicative nature of the model means if, as is the case in most of our analysis, the x 's are categorical variables, then letting B_k be the coefficient of x_k , $\exp(B_k)$ gives a multiplier for the effect of variable x_k being in category 1, relative to a baseline case. We use the estimated multipliers to adjust existing pricing models or rules of thumb to estimate cost differences across groups for both options.

Because of the jointness of the options, default and prepayment should be modeled and estimated jointly (see Deng, Quigley and Van Order 2000 for

⁵ See Findley and Capozza (1977) and Dunn and McConnell (1981) for early discussions and Hendershott and Van Order (1987) and Kau and Keenan (1995) for additional analysis.

⁶ See Deng, Quigley and Van Order (2000) and Kau and Keenan (1995) for somewhat different versions.

a discussion). We control for the jointness of these options in two stages. We divide our sample of observations by the extent to which prepayment is rational, that is, the extent to which the prepayment option is in the money.⁷ In the first stage we estimate default hazards separately depending on the “moneyness” of the prepayment option and estimate loss severity rates conditional on default. In the second stage, we estimate our prepayment models, including the fitted probability of default from the first stage as a covariate.

We group our observations into quarters and assume that the hazard rates are constant within quarters. Let $h(T)$ be the continuous time hazard rate within quarter T and $H(T)$ be the probability of the hazard happening at some time during T . $S(T)$, the probability of surviving throughout the quarter conditional on having survived until the beginning of the quarter is

$$S(T) = 1 - H(T) = \exp(-h(T)\Delta T), \quad (2)$$

where ΔT is the length of the period. That is, the survival rate declines exponentially at a rate equal to the instantaneous hazard rate during the period. Letting $\Delta T = 1$ and taking logs twice we have the complementary log-log model:⁸

$$\log(-\log(1 - H(T))) = \log(h(T)) = Bx(T). \quad (3)$$

This formulation has the advantage that the estimates are not affected by size of the interval (*e.g.*, weeks vs. quarters). We use this equation to obtain estimates of the B 's.

We estimate both unconditional and conditional models. The unconditional models are hazard models that have race/ethnicity and income variables as the main x 's, with controls for time in the form of baseline hazards that are fixed effects for loan age. The conditional models add thousands of interactive fixed effects by creating the pseudopools of mortgages.

For each calendar quarter of mortgage exposure we create fixed effects using “pseudopools,” which are formed by dividing loans into relatively homogeneous groupings based on observed characteristics such as contract rate (50 basis-point buckets), LTV (four buckets), credit history measured by FICO⁹ score (four buckets) and loan amount (three buckets). For each origination quarter this results in on the order of 200 pseudopools. Each of these pseudopools is

⁷ See footnote 12 for a description of the method used.

⁸ See Agresti (1990) and Prentice and Gloeckler (1978) for discussions of complementary log-log models.

⁹ This is a generic credit score developed by Fair Isaac Corporation, which is widely used by lenders.

then given a fixed effect for each quarter it is alive (up to 51 quarters). With 20 origination quarters this amounts to a total of over 200,000 fixed effects, which, because our data have well over one million loans and millions of loan-quarters, still leaves a large number of degrees of freedom.

This is a very simple but also rather complete representation. It allows us to isolate the effects of race and income within pools, while holding effects at the pool level constant by allowing for extremely complicated interactions among the other explanatory variables. In particular, the procedure handles the burnout and other age-related heterogeneity problems discussed above.

An analogy to typical panel data analysis is useful in understanding the conditional models. We can divide x into characteristics that vary within a group and those that vary only across groups. For the purposes of this study we are interested in estimating within-group variation in behavior. To accomplish this we partition x as follows. Let

$$y \equiv \log(-\log(1 - H(T))) = Bx. \tag{4}$$

If we let y^{ij} be the value of y for the i th borrower in the j th pseudopool, we can partition the right-hand side into two parts, so that

$$y^{ij} = B_1 \mathbf{x}_1^{ij} + B_2 \mathbf{x}_2^j, \tag{5}$$

where \mathbf{x}_1^{ij} is a vector of the individual characteristics of the i th individual in the j th pseudopool, such as borrower income and race/ethnicity, and \mathbf{x}_2^j includes characteristics common to all borrowers in the j th pseudopool, including initial LTV ratio, date of origination, loan age and other characteristics.

We are interested in estimates of B_1 . Following the analogy with panel data analysis, this can be accomplished by including group-level fixed effects to capture the effects of x_2^j . Alternatively, this can be accomplished through the subtraction of group-level means. Subtracting pseudopool means from both sides, we can rewrite (5) as

$$y^{ij} - \bar{y}^j = (x_1^{ij} - \bar{x}_1^j) B_1 + (x_2^j - \bar{x}_2^j) B_2, \tag{6}$$

where \bar{y}^j is the average level of y in the loan's pseudopool in the quarter in question and \bar{x}_1^j and \bar{x}_2^j are the mean levels of x_1^{ij} and x_2^j in the j th pseudopool.

Because, by construction, $x_2^j = \bar{x}_2^j$, we have

$$y^{ij} - \bar{y}^j = (x_1^{ij} - \bar{x}_1^j) B_1, \tag{7}$$

which we can rewrite as

$$y^{ij} = (x_1^{ij} - \overline{x_x^j})B_1 + \overline{y^j}. \quad (8)$$

We estimate Equation (8) using maximum likelihood.¹⁰ We do not produce estimates of B_2 . The creation of the pseudopools allows us to control for their effects without estimating thousands of parameters.

We estimate the effect of differences in prepayment behavior on pricing by using YieldBook[®], a proprietary pricing model developed by Citigroup, and use rules of thumb to translate differences in default and severity into pricing. Finally, to understand the net effect of LIMM status on loan performance, we compare the pricing implications of differences in prepayment and default performance.

Data

Our data consist of all 30-year fixed-rate mortgages that originated during 1993–1997 and purchased by Freddie Mac for which key data are not missing or obviously inaccurate. The full data set contains about 2.7 million loans. There were sharp mortgage rate declines in 1993, 1995 and 1998, so our data are rich in prepayment experience. The default modeling suffers from excessively good times in the 1990s and relatively small levels of default. However, the California economy performed rather poorly in the early part of the period and provides us with some significant default data. The performance of all loans was followed through the third quarter of 2005.

Table 1 provides descriptive statistics for the major variables. We define income classes relative to the area median income, with low income as less than 80% of the area median. Of particular interest is the variable “in the moneyness,” which measures the extent to which the coupon rate on the loan is above or below the current market rate.¹¹ Table 2 presents simple cross-tabs. Part A gives prepayment rates (percent that was ever prepaid during the sample period) by borrower race/ethnicity and income. Blacks and Hispanics prepay at a slower rate than whites and other minorities, and low-income borrowers prepay more

¹⁰ Pseudopools with no prepayments or defaults are excluded from the analysis because there is no within-group variation to explain. Mathematically, this results in values of $\log(0)$ for $\overline{y^j}$. We use a program developed by the SAS Institute (Cary, NC) for estimation of log-log models.

¹¹ Note that we assume that borrowers can obtain mortgages at the average contract rate. To the extent that borrower credit or other borrower or property characteristics make the contract rate available for refinancing a loan deviate from the average rate, we measure the in the moneyness with error.

Table 1 ■ Descriptive statistics.

Panel A		Borrower Income by % of Area Median Income			
Borrower Race	All	Inc1 (0–80%)	Inc2 (81–100%)	Inc3 (101–120%)	Inc4 (>120%)
Black	3.20	4.51	3.20	2.87	2.64
Hispanic	2.93	4.44	3.39	2.85	2.04
Other minority	4.12	4.17	4.53	4.34	3.85
White	89.75	86.88	88.88	89.94	91.47
Total	100.00	100.00	100.00	100.00	100.00

Panel B		Borrower Race			
Borrower income (percent of area median)	All	Black	Hispanic	Other	White
Inc1 (0–80%)	22.73	32.13	34.50	23.10	22.01
Inc2 (81–100%)	16.05	16.12	18.60	17.69	15.90
Inc3 (101–120%)	15.39	13.84	15.00	16.28	15.43
Inc4 (>120%)	45.83	37.90	31.87	42.93	46.66
Total	100.00	100.00	100.00	100.00	100.00

Table 2 ■ Cumulative prepayment and default rates.

Borrower Income	Borrower Race				
	Black	Hispanic	Other	White	Total
Panel A: Cumulative Default Rate					
Inc1 (0–80%)	2.80	1.35	0.75	0.83	0.94
Inc2 (81–100%)	2.06	1.60	0.81	0.62	0.71
Inc3 (101–120%)	2.01	1.63	0.93	0.52	0.62
Inc4 (> 120)	1.44	1.40	0.84	0.37	0.44
Total	2.05	1.45	0.83	0.54	0.63
Panel B: Cumulative Prepayment Rate					
Inc1 (0–80%)	72.87	84.45	90.00	88.94	88.05
Inc2 (81–100%)	79.81	88.13	92.82	92.04	91.56
Inc3 (101–120%)	81.04	88.66	92.86	93.11	92.63
Inc4 (> 120)	84.04	89.45	93.53	94.37	93.97
Total	79.36	87.36	92.50	92.62	92.03

Note: Percent of total mortgages of each type that defaulted or prepaid during the observation period.

slowly than high-income borrowers. Blacks and Hispanics have higher default rates than whites. Low-income borrowers tend to default more, but the differences are not very large, and the relationship does not hold for all groups. For example, defaults by Hispanics and other minorities increase with income. These are, of course, crude statistics without basic controls. To address this concern, we now turn to estimates of various forms of hazard models.

Default Models

Panel A of Table 3 presents a basic hazard model of default. We approximate the extent to which the default option is in the money with an estimate of current loan-to-value (CTLV), and we control for the value of the prepayment option by separately estimating the model according to the degree to which prepayment is in the money.¹² We estimate CLTV by using Freddie Mac's Conventional Mortgage Home Price Index (HPI) to control for home price appreciation at the Metropolitan Statistical Area (MSA) level.¹³ Our other explanatory variables

¹² Let $a = 1 - \text{current coupon rate}/\text{coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{Cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$.

¹³ Where a match with MSA was not possible (for less than 1% of the data), state-level HPI values were used.

Table 3 ■ Default model.

Panel A: Model Coefficients															
By Prepayment "In the Moneyness"															
Variable	Discount			Current			Cusp			Premium			Super-Premium		
	Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE	
CTLV	0.0170	0.0028	0.0142	0.0026	0.0119	0.0021	0.0053	0.0014	0.0072	0.0026	0.0014	0.0053	0.0014	-0.0072	0.0026
FICO < 620	2.6419	0.1160	2.7488	0.1016	2.6615	0.0877	2.6839	0.0622	1.8821	0.1299	0.0622	2.6839	0.0622	1.8821	0.1299
FICO 620-679	2.0130	0.0960	1.8634	0.0915	2.0612	0.0776	2.0562	0.0580	1.5167	0.1194	0.0580	2.0562	0.0580	1.5167	0.1194
FICO 680-720	1.0773	0.1078	1.1336	0.1001	1.1023	0.0889	1.1991	0.0659	0.9128	0.1369	0.0659	1.1991	0.0659	0.9128	0.1369
Debt 0-30	-0.5236	0.0861	-0.5876	0.0770	-0.5181	0.0628	-0.4726	0.0429	-0.3439	0.0932	0.0429	-0.4726	0.0429	-0.3439	0.0932
Debt 31-36	-0.0669	0.0865	-0.2395	0.0787	-0.2415	0.0643	-0.2230	0.0441	0.0127	0.0941	0.0441	-0.2230	0.0441	0.0127	0.0941
Loanamt < 76 k	0.3708	0.0909	0.4314	0.0834	0.4229	0.0685	0.3498	0.0489	0.2828	0.1248	0.0489	0.3498	0.0489	0.2828	0.1248
Loanamt 76-125 k	0.1387	0.0860	0.1689	0.0805	0.1090	0.0673	0.0740	0.0498	0.1209	0.1309	0.0498	0.0740	0.0498	0.1209	0.1309
Purp = purchase	-0.4954	0.0775	-0.5147	0.0694	-0.5886	0.0565	-0.5012	0.0376	-0.3983	0.0802	0.0376	-0.5012	0.0376	-0.3983	0.0802
Number of observations	6,109,543		4,454,309		5,342,525		6,793,213		816,575		6,793,213		816,575		
Log likelihood	-7,539.64		-8,670.10		-12,614.25		-25,543.90		-5,160.14		-25,543.90		-5,160.14		

Panel B: Multipliers															
Variable	Discount			Current			Cusp			Premium			Super-Premium		
	Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE	
FICO < 620	14.04		15.62		14.32		14.64		6.57		14.64		6.57		
FICO 620-679	7.49		6.45		7.86		7.82		4.56		7.82		4.56		
FICO 680-720	2.94		3.11		3.01		3.32		2.49		3.32		2.49		
Debt 0-30	0.59		0.56		0.60		0.62		0.71		0.62		0.71		
Debt 31-36	0.94		0.79		0.79		0.80		1.01		0.80		1.01		
Loanamt < 76 k	1.45		1.54		1.53		1.60		1.33		1.60		1.33		
Loanamt 76-125k	1.15		1.18		1.12		1.08		1.13		1.08		1.13		
Purp = purchase	0.61		0.60		0.56		0.61		0.67		0.61		0.67		

Note: Panel A shows hazard models for the probability of default, estimated separately by the value of the prepayment option. Let $a = 1 - \text{current coupon rate} / \text{coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Panel B shows base of the natural logarithm raised to the value of the relevant coefficient. CLTV = current loan-to-value.

are credit history, as measured by the FICO score at origination, the ratio of borrower debt payments to income, loan amount, loan purpose and loan age.

The model works very much as expected. FICO score and CLTV have strong effects on default. Coefficients do not vary much by in the moneyness. The exception is across CTLV ratio, where the effect decreases as in the moneyness increases. Borrowers are less likely to default on low-rate loans; the low rate means that the market value of the mortgage is below par, so that mark-to-market equity is higher than is given by our calculation of CLTV.

Multipliers relative to a baseline are presented in Panel B. With the exception of the super-premium category, a loan in the lowest FICO class is about five times as likely to default as one in the medium (680–720) class. Smaller original loan sizes are also associated with a higher frequency of default. Multipliers are fairly consistent for all categories of the prepayment option, with the exception of a decline in the impact of FICO for super-premium mortgages.

In Table 4, we show the results of adding minority status and income to the default model. Recall that absent controls, LMM borrowers had much higher default rates, on the order of three times as high for Hispanic and four times as high for black. In this model, with controls for borrower and loan characteristics, the effect of LMM status is smaller. For instance, the multiplier for “black” (relative to the base case, “white”) is approximately 1.6, with the exception of the super-premium category. A similar reduction is evident for Hispanics. Most of the minority status effect, therefore, is eliminated with controls for other characteristics, and none of the coefficients on the other variables are affected much by adding minority and income.

To analyze default cost we need to model loss severity rates as well. Table 5 presents ordinary least squares results from regressing log of loss severity¹⁴ divided by mortgage balance on the basic minority status and income variables. It suggests small differences by minority but bigger differences by income. Table 6 controls for LTV, FICO, *etc.* and adds census tract variables. It suggests that the major explanatory factors are census tract median income and minority percentage.¹⁵

¹⁴ The loss used to compute severity is based on internal Freddie Mac calculations, which include collateral deficiency as well as lost interest, transaction costs, legal expenses and selling expenses.

¹⁵ It is intriguing that the control variables effectively eliminate race/ethnicity effects for the default severity model but do not do so for the default incidence model. Likely this reflects the fact that severity is critically dependent on state laws and local house price appreciation, both of which are reasonably well captured by our estimation specification.

Table 4 ■ Default model with borrower race and income.

Panel A: Model Coefficients											
By Prepayment "In the Moneyness"											
Variable	Discount		Current		Cusp		Premium		Super-Premium		
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	
CTLV	0.0170	0.0029	0.0141	0.0026	0.0117	0.0021	0.0048	0.0014	-0.0067	0.0027	
FICO < 620	2.6125	0.1191	2.7078	0.1038	2.6011	0.0899	2.6140	0.0632	1.8705	0.1317	
FICO 620-679	2.0160	0.0981	1.8367	0.0933	2.0470	0.0791	2.0153	0.0585	1.5292	0.1202	
FICO 680-720	1.0857	0.1099	1.1317	0.1015	1.0931	0.0905	1.1747	0.0664	0.9150	0.1376	
Debt 0-30	-0.4419	0.0909	-0.4644	0.0809	-0.4232	0.0660	-0.4014	0.0448	-0.2900	0.0965	
Debt 31-36	-0.0239	0.0880	-0.2198	0.0804	-0.1994	0.0654	-0.2033	0.0448	0.0428	0.0952	
Loanamt < 76 k	0.2730	0.1076	0.2736	0.0981	0.2719	0.0806	0.2039	0.0568	0.1284	0.1396	
Loanamt 76-125 k	0.1213	0.0926	0.0740	0.0866	0.0304	0.0720	-0.0046	0.052	0.0204	0.1367	
Purp = purchase	-0.4972	0.0787	-0.5036	0.0704	-0.5840	0.0573	-0.4961	0.0381	-0.3882	0.0811	
Black	0.4373	0.1359	0.5198	0.1156	0.5742	0.0892	0.4454	0.0570	-0.1163	0.1242	
Hispanic	0.5180	0.1482	0.5888	0.1262	0.6849	0.0978	0.3375	0.0707	-0.2171	0.1689	
Other minority	0.7039	0.1331	0.5823	0.1284	0.4462	0.1126	0.3851	0.0802	-0.2379	0.2292	
Inc1 (0-80%)	0.3389	0.1023	0.4298	0.0958	0.3176	0.0767	0.2839	0.0517	0.2724	0.1088	
Inc2 (81-100%)	0.0663	0.1099	0.3948	0.0973	0.2395	0.0797	0.2193	0.0552	0.1965	0.1234	
Inc3 (101-120%)	0.1065	0.1076	0.3945	0.0968	0.1250	0.0829	0.1192	0.0581	0.2243	0.1280	
Number of observations	6,109,543		4,298,857		5,429,207		6,922,004		834,126		
Log likelihood	-7,297.97		-8,403.15		-12,258.92		-24,956.04		-5,068.82		
Panel B: Multipliers											
Variable	Discount		Current		Cusp		Premium		Super-Premium		
Black	1.55		1.68		1.78		1.56		0.89		
Hispanic	1.68		1.80		1.98		1.40		0.80		
Other minority	2.02		1.79		1.56		1.47		0.79		
Inc1 (0-80%)	1.40		1.54		1.37		1.33		1.31		
Inc2 (81-100%)	1.07		1.48		1.27		1.25		1.22		
Inc3 (101-120%)	1.11		1.48		1.13		1.13		1.25		

Panel A shows hazard models for the probability of default, estimated separately by the value of the prepayment option. Let $a = 1$ - current coupon rate/coupon rate on mortgage. We define prepayment option value as follows: $a < -0.035$ = discount, $-0.035 < a < 0.035$ = current, $0.035 < a < 0.100$ = cusp, $0.100 < a < 0.25$ = premium, $a > 0.25$ = super-premium. Panel B shows base of the natural logarithm raised to the value of the relevant coefficient. CLTV = current loan-to-value.

Table 5 ■ Loss severity by borrower race/income.

Variable	Coefficient	SE
Black	0.0201	0.0053
Hispanic	0.0127	0.0052
Other	-0.0199	0.0071
Inc1	0.0873	0.0038
Inc2	0.0426	0.0043
Inc3	0.0208	0.0044
Number of observations		22,038
R^2		0.0279

Table 6 ■ Loss severity by race/income/tract variables.

Variable	Coefficient	SE
Black	-0.0024	0.0054
Hispanic	0.0018	0.0053
Other	-0.0236	0.0070
Inc1	0.0218	0.0042
Inc2	0.0026	0.0043
Inc3	-0.0032	0.0042
Tract%Min1 (<10%)	-0.0298	0.0068
Tract%Min 2 (11–30%)	-0.0357	0.0068
Tract%Min 3 (31–50%)	-0.0442	0.0080
Tract Med Inc1 (0–80%)	0.0584	0.0057
Tract Med Inc2 (80–100%)	0.0012	0.0044
Tract MedInc3 (101–120%)	-0.0240	0.0045
Number of observations		22,038
R^2		0.1763

Note: Includes controls for LTV, Orig Amt, # Units, FICO, State.
 ordinary least squares regressions: Loss severity is defined as (unpaid principal recovery + administrative and transaction costs)/original principal balance.

Pricing

We do not have a well-developed pricing model for credit risk. However, from Freddie Mac history we can approximate a baseline level of default to which we can apply our estimated multipliers. In this sample the median loan has an LTV just under 80%. Freddie Mac history suggests that loans like these have about a 1–2% chance of ever defaulting; this was higher in the early 1990s during the recession and was smaller later, during the housing boom. Average loss severity rates on these mortgages have been about 30%. This suggests average losses of about 0.3–0.6% of loan balance, which, discounted to the present, implies

an expected present value of about 0.25–0.5% of loan balance. This decline in value can be converted into an equivalent required increase in coupon rate on the mortgage.

In the next section we make use of YieldBook[©] for pricing mortgages with prepayment risk. It has the property, which is also approximately true in the market for mortgage-backed securities, that a 1-basis-point increase in coupon rate will lead to a 5-basis-point increase in mortgage value and vice versa. Hence, we can divide the value decline by five to get an estimate of the required increase in coupon rate to compensate for the higher default cost. This implies an average annual charge of 5–10 basis points.

For loans to blacks and Hispanics the overall unconditional multiplier estimate (including severity rates) is around three, suggesting a range of cost of 15–30 basis points and a difference from the baseline of 10–20 basis points. The multipliers from the model that incorporates FICO and loan characteristics are much smaller for both groups, on the order of 1.5, consistent with a difference from the baseline of 2.5–5 basis points. For low-income borrowers the unconditional multiplier, including severity rate differences, is about two, which suggests a range of default costs of 10–20 basis points with a differential of 5–10 basis points.¹⁶ Inclusion of other borrower and loan characteristics reduces both the multiplier and effect on price by one-quarter, consistent with a differential of 3–8 basis points.

Prepayment Models

The Unconditional Model

Panel B of Table 2 presents cross-tabs for prepayment by income and minority status. It suggests that minority status is negatively associated with prepayment, with a similar effect associated with income. Table 7 presents results for estimates of complementary log-log models, controlling only for loan age. Results in the first column tell the same story as in Table 2: Black, Hispanic and low-income borrowers tend to prepay at a slower rate.

The right-hand column adds two neighborhood characteristics, the median income of households in the loan's census tract relative to the area median as well as the minority (black, Hispanic and other minority) share of households in the census tract. Including these variables affects the race/ethnicity coefficients,

¹⁶ A factor not included is capital costs. Riskier loans require more capital, which will increase costs.

Table 7 ■ Basic prepayment results.

Variable	Without Census Tract Characteristics		With Census Tract Characteristics	
	Coefficient	SE	Coefficient	SE
Black	-0.5301	0.0059	-0.4510	0.0062
Hispanic	-0.2539	0.0058	-0.1814	0.0060
Other	-0.0120	0.0048	0.0343	0.0049
Inc1 (0–80%)	-0.2681	0.0025	-0.2445	0.0025
Inc2 (81–100%)	-0.1179	0.0028	-0.1047	0.0028
Inc3 (101–120%)	-0.0617	0.0028	-0.0539	0.0028
Tract%Min1 (01–10)			0.1491	0.0051
Tract%Min2 (11–30)			0.1067	0.0052
Tract%Min3 (31–50)			0.0867	0.0060
Tract Med Inc1 (0–80%)			-0.1001	0.0039
Tract Med Inc2 (81–100%)			-0.0787	0.0026
Tract Med Inc3 (101–120%)			-0.0319	0.0023
Number of observations	23,268,202		23,268,202	
Log likelihood	-4,200,761.21		-4,199,191.26	

lowering them a bit. For instance, the coefficient for black with current mortgages increases from -0.53 to -0.45 , and the coefficient for low minority concentration (Min1 (0–10)) is 0.15 , relative to high concentration (greater than 50%). Hence, the result that minorities tend to prepay less is partly explained by the racial composition of the neighborhood as well as the race of the borrower. Adding neighborhood characteristics creates virtually no change in the coefficient for individual income, but there is a small effect of neighborhood income on prepayment.

Table 8 presents estimates of a hazard model for the probability of prepayment as a function of how far into the money the option is. In-the-moneyness categories are as described in Table 1: discount, current, cusp, premium and super-premium, in which the latter is the furthest into the money and the first the furthest out of the money.¹⁷ There are also controls for loan age and origination year, which are not shown. Premium loans prepay about three times as often as current loans. Current loans have historically had annual prepayment speeds on the order of 10%; so a premium loan will tend to have a speed in excess of 30%.

The effect of prepayment speeds on pricing depends on when prepayment speeds differ. If a group prepays relatively slowly when the option is

¹⁷ Note that the multiplier for super-premium is set to 1, and the prepayment speed of all other in-the-moneyness categories are measured relative to it.

Table 8 ■ Simple prepayment model.

Variable	Coefficient	SE	Multiplier
Discount	-1.4255	0.0058	0.24
Current	-1.0742	0.0052	0.34
Cusp	-0.5947	0.0045	0.55
Premium	0.0138	0.0037	1.01
Number of observations		23,749,189	
Log likelihood		-4,188,724.60	

Note: Hazard models for prepayment speeds. Controls for loan origination year and age are not shown. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Multipliers are the base of the natural logarithm raised to relevant coefficient.

in the money, the investment value of the mortgage increases, whereas slower prepayment when the option is out of the money decreases its value to investors. Given the importance of the value of the option, we perform the rest of the estimation separately by the option value category.

Table 9 repeats the analysis of the first column of Table 7, with the estimation performed separately by option value category. The results differ for race and income. In Table 9 we see virtually the same difference for black prepayment rates (and a small decline for Hispanic and other minority) whether the option is in or out of the money. The coefficient for “black” implies that for premium loans blacks are approximately $\exp(-0.4805)$ or about 0.6 times as likely to prepay as whites, regardless of option value. For the lowest income group, results depend on the option value. There is virtually no difference in prepayment speeds when the option is out of the money, whereas they are about 0.7 times as likely to prepay as those with incomes more than 120% of median (about half the loans in the sample) when the option is in the money. Table 10 shows the results of adding census tract variables. The relationship between prepayment speeds and minority or low-income status is slightly smaller, but otherwise the results are similar.

We next control for LTV, credit history (FICO), loan amount (given our controls for LTV, this is equivalent to controlling for property value) and other variables by forming pseudopools as described above. Tables 11 and 12 report estimates of B_1 in Equation (7). As before, we show coefficients for race, ethnicity and income groups separately for different categories of option value. However in this table the coefficients measure effects that are conditional on the loan

Table 9 ■ Basic results by exercise category.

Panel A: Model Coefficients											
By Prepayment "In the Moneyness"											
Variable	Discount		Current		Cusp		Premium		Super-Premium		
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	
Black	-0.5467	0.0260	-0.5148	0.0225	-0.5159	0.0158	-0.4805	0.0081	-0.4978	0.0139	
Hispanic	-0.4211	0.0252	-0.4144	0.0223	-0.3008	0.0149	-0.2645	0.0078	-0.2033	0.0149	
Other	-0.3553	0.0194	-0.2152	0.0167	0.0181	0.0107	0.0450	0.0064	0.0164	0.0157	
Inc1 (0-80%)	-0.0196	0.0091	-0.0663	0.0084	-0.2722	0.0060	-0.3540	0.0034	-0.2495	0.0079	
Inc2 (81-100%)	-0.0560	0.0101	-0.0866	0.0094	-0.1514	0.0065	-0.1433	0.0037	-0.0698	0.0093	
Inc3 (101-120%)	-0.0728	0.0102	-0.0609	0.0094	-0.0792	0.0064	-0.0682	0.0038	-0.0297	0.0099	
Pr. (default)	246.78	10.79	202.84	7.31	44.86	5.32	-214.11	2.83	-249.22	4.57	
Number of observations	5,909,264		4,298,857		5,231,741		6,656,882		801,435		
Log likelihood	-430,403.35		-456,723.19		-835,197.26		-1,983,261.96		-291,851.26		
Panel B: Multipliers											
Variable	Discount	Current	Cusp	Premium	Super-Premium						
Black	0.58	0.60	0.60	0.62	0.61						
Hispanic	0.66	0.66	0.74	0.77	0.82						
Other minority	0.70	0.81	1.02	1.05	1.02						
Inc1 (0-80%)	0.98	0.94	0.76	0.70	0.78						
Inc2 (81-100%)	0.95	0.92	0.86	0.87	0.93						
Inc3 (101-120%)	0.93	0.94	0.92	0.93	0.97						

Note: Panel A shows hazard models for prepayment speeds estimated separately by prepayment option. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Probability of default is a fitted value from the default hazard model. Panel B is the base of the natural logarithm raised to the relevant coefficient. Controls for loan origination year and age are not shown.

Table 10 ■ Basic results by exercise category with tract variables.

Variable	By Prepayment "In the Moneyness"											
	Discount		Current		Cusp		Premium		Super-Premium			
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE		
Black	-0.4803	0.0270	-0.4496	0.0235	-0.4572	0.0164	-0.3875	0.0085	-0.4091	0.0151		
Hispanic	-0.3678	0.0260	-0.3599	0.0230	-0.2465	0.0154	-0.1708	0.0081	-0.1153	0.0158		
Other	-0.3056	0.0201	-0.1665	0.0173	0.0568	0.0112	0.1027	0.0066	0.0661	0.0162		
Inc1 (0-80%)	-0.0192	0.0093	-0.0649	0.0086	-0.2582	0.0061	-0.3165	0.0035	-0.2116	0.0080		
Inc2 (81-100%)	-0.0534	0.0102	-0.0846	0.0095	-0.1390	0.0065	-0.1229	0.0038	-0.0499	0.0094		
Inc3 (101-120%)	-0.0704	0.0103	-0.0593	0.0095	-0.0712	0.0065	-0.0562	0.0038	-0.0182	0.0099		
Tract%Min1 (01-10)	0.2070	0.0203	0.1923	0.0181	0.1423	0.0124	0.1679	0.0069	0.1191	0.0143		
Tract%Min2 (11-30)	0.2077	0.0203	0.1793	0.0182	0.1098	0.0125	0.1010	0.0070	0.0921	0.0143		
Tract%Min3 (31-50)	0.1446	0.0234	0.1265	0.0210	0.0877	0.0145	0.0792	0.0082	0.0632	0.0169		
Tract Med Inc1 (0-80%)	0.0193	0.0149	0.0740	0.0134	-0.0509	0.0094	-0.2041	0.0053	-0.2381	0.0116		
Tract Med Inc2 (81-100%)	-0.0087	0.0096	-0.0168	0.0088	-0.0795	0.0061	-0.1426	0.0035	-0.1460	0.0086		
Tract Med Inc3 (101-120%)	-0.0476	0.0084	-0.0235	0.0078	-0.0516	0.0054	-0.0514	0.0031	-0.0718	0.0082		
Pr. (default)	-248.09	10.83	204.35	7.33	53.07	5.32	-198.10	2.83	-240.88	4.564		
Number of observations	5,909,264		4,298,897		5,231,741		6,656,822		801,435			
Log likelihood	-430,314.87		-456,647.54		-834,988.43		-1,980,981.81		-291,450.93			

Note: Hazard models for prepayment option speeds were estimated separately by prepayment option. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Probability of default is a fitted value from the default hazard model. Controls for loan origination year and age are not shown.

Table 11 ■ Pseudopools, effect of race and income variables.

Panel A: Model Coefficients												
By Prepayment "In the Moneyness"												
Variable	Discount			Current			Cusp			Premium		
	Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE	
Black	-0.6458	0.0978		-0.5484	0.0977		-0.6918	0.0945		-0.7786	0.1317	
Hispanic	-0.4577	0.0779		-0.6010	0.0854		-0.7302	0.0808		-0.3738	0.0949	
Other	-0.4278	0.0536		-0.4577	0.0540		0.0189	0.0396		-0.0755	0.0600	
Inc1 (0-80%)	-0.1980	0.0281		-0.1432	0.0327		-0.2788	0.0305		-0.1986	0.0413	
Inc2 (81-100%)	-0.1551	0.0262		-0.1926	0.0312		-0.1433	0.0267		-0.0498	0.0351	
Inc3 (101-120%)	-0.1170	0.0259		-0.1257	0.0299		-0.0263	0.0241		0.0252	0.0326	
Pr. (default)	1986.82	402.90		2411.606	271.05		3,677.97	320.04		1457.39	372.80	
Number of observations	1,122,739			544,372			389,784			85,387		
Log likelihood	-69,737.78			-52,460.04			-57,612.57			-24,227.51		
Panel B: Multipliers												
Variable	Discount			Current			Cusp			Premium		
	Coefficient	SE		Coefficient	SE		Coefficient	SE		Coefficient	SE	
Black	0.52			0.58			0.50			0.46		
Hispanic	0.63			0.55			0.48			0.69		
Other minority	0.65			0.63			1.02			0.93		
Inc1 (0-80%)	0.82			0.87			0.76			0.82		
Inc2 (81-100%)	0.86			0.82			0.87			0.95		
Inc3 (101-120%)	0.89			0.88			0.97			1.03		

Note: Hazard models for prepayment option speeds are estimated separately by prepayment option. Pseudopools are formed by grouping loans according to LTV, borrower credit, contract rate, date of origination and loan amount. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Probability of default is a fitted value from the default hazard model. There were not enough observations to conduct analysis of super-premium mortgages.

Table 12 ■ Pseudopools, effect of race and income variables with tract variables.

Variable	By Prepayment "In the Moneyness"							
	Discount		Current		Cusp		Premium	
	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Black	-0.5302	0.0996	-0.4387	0.0994	-0.6202	0.0956	-0.6545	0.1330
Hispanic	-0.3576	0.0794	-0.5018	0.0870	-0.6664	0.0817	-0.2361	0.0967
Other	-0.3377	0.0552	-0.3632	0.0558	0.0779	0.0411	0.0235	0.0617
Inc1 (0-80%)	-0.1861	0.0284	-0.1343	0.0329	-0.2608	0.0307	-0.1689	0.0415
Inc2 (81-100%)	-0.1460	0.0263	-0.1868	0.0313	-0.1305	0.0268	-0.0367	0.0352
Inc3 (101-120%)	-0.0113	0.0260	-0.1231	0.0299	0.0186	0.0241	0.0293	0.0327
Tract%Min1 (01-10)	0.3370	0.0568	0.3342	0.0591	0.2441	0.0526	0.3007	0.0710
Tract%Min2 (11-30)	0.3390	0.0568	0.2965	0.0590	0.1369	0.0528	0.1356	0.0714
Tract%Min3 (31-50)	0.2194	0.0645	0.1919	0.0673	0.1456	0.0600	0.0500	0.0834
Tract Med Inc1 (0-80%)	0.0621	0.0403	0.1084	0.0447	0.0301	0.0388	-0.0554	0.0559
Tract Med Inc2 (81-100%)	-0.0401	0.0248	-0.0051	0.0274	-0.0743	0.0239	-0.1348	0.0329
Tract Med Inc3 (101-120%)	-0.0712	0.0209	-0.0205	0.0226	-0.0371	0.0197	0.0077	0.0264
Pr. (default)	2,035.22	403.65	2,449.15	271.28	3,843.7	321.29	1,662.96	373.87
Number of observations	1,275,635		544,372		389,784		85,387	
Log likelihood	-76,236		-52,439.11		-57,581.71		-24,177.65	

Note: Hazard models for prepayment option speeds were estimated separately by prepayment option. Pseudopools are formed by grouping loans according to LTV, borrower credit, contract rate, date of origination and loan amount. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define prepayment option value as follows: $a < -0.035 = \text{discount}$, $-0.035 < a < 0.035 = \text{current}$, $0.035 < a < 0.100 = \text{cusp}$, $0.100 < a < 0.25 = \text{premium}$, $a > 0.25 = \text{super-premium}$. Probability of default is a fitted value from the default hazard model. There were not enough observations to conduct analysis of super-premium mortgages.

characteristics that define the pseudopools. Table 11 corresponds to Table 9. The magnitude is slightly larger for the minority controls. The control for low income is similar regardless of the value of the prepayment option. Table 12 repeats Table 10. Here, the additional controls have relatively little effect.¹⁸

Stability Over Time

Mortgage markets changed rapidly in the 1990s, particularly with respect to prepayments. It has become increasingly easy to refinance, and it may well be the case that it has become increasingly easy for LMM borrowers to get loans. Hence, it may be that the coefficients estimated above have changed over time. Because of the large size of our data set we can test this by re-estimating the models for different exposure (*i.e.*, calendar rather than origination) years.

The first panel of Table 13 presents results from the unconditional model, re-estimated for each exposure year. It presents only the coefficients for black and Hispanic and the lowest income groups. There does appear to have been a decline in the magnitude of the effects, particularly in the last 2 years of the sample. The magnitude of all coefficients dips during 1999 and 2000, years of relatively slow refinance activity. Panel B of Table 13 depicts the conditional results. Adding pseudopools does not substantially affect the coefficients on black and Hispanic borrowers, but coefficients on low-income borrowers become smaller in magnitude during the later period of the time series.

Pricing

Our estimates suggest that there are significant differences in prepayment speeds across groups. In addition, the exercise of prepayment options appears to be somewhat less ruthless among low-income groups. During recent refinance booms, prepayment speeds have often been in the range of 40% annual rates, so small relative differences in prepayment speeds can have large effects on the value to investors.

To evaluate the effect on value we applied YieldBook[©] (see Hayre and Rajan 1995 for a description of the prepayment model incorporated in this software). This model is widely available, but proprietary. It uses Monte Carlo techniques combined with empirical prepayment models to compute the value

¹⁸ FICO score has an important effect on prepayment. Low-FICO borrowers have a multiplier less than 1 when the option is in the money, but a multiplier greater than 1 when it is out of the money. This makes low-FICO borrowers relatively desirable prepayment-wise. This is consistent with results in Deng and Gabriel (2006), who find a relatively small impact of this on pricing for new loans, but a big effect if the loans are seasoned.

Table 13 ■ Stability over time.

Year	All Mortgages				Premium Mortgages			
	Black	Hispanic	Inc1 (0–80%)	Inc2 (81–100%)	Black	Hispanic	Inc1 (0–80%)	Inc2 (81–100%)
Panel A: Unconditional Model*								
1995	-0.5335	-0.5097	-0.2922	-0.1646	-0.7783	-0.6872	-0.6686	-0.3156
1996	-0.4755	-0.4586	-0.1253	-0.0786	0.7019	-0.6380	-0.3867	0.1689
1997	-0.4431	-0.3319	-0.1135	-0.0988	-0.7265	-0.4599	-0.3696	-0.1858
1998	-0.5242	-0.3660	-0.2904	-0.1134	-0.6783	-0.4859	-0.4138	-0.1620
1999	-0.2921	-0.1335	-0.1357	-0.0595	-0.4219	-0.2181	-0.2915	-0.1130
2000	-0.2900	-0.1614	0.0792	-0.0042	-0.4233	-0.1801	-0.0867	-0.0642
2001	-0.4940	-0.1452	-0.1506	-0.0653	-0.6243	-0.2372	-0.2607	-0.1139
2002	-0.5269	-0.1849	-0.3139	-0.1363	-0.6331	-0.2863	-0.4139	-0.1820
2003	-0.5487	-0.2220	-0.3573	-0.1531	-0.5057	-0.2360	-0.3805	-0.1676
2004	-0.2326	-0.0175	-0.2086	-0.1025	-0.2355	-0.0060	-0.1895	-0.1048
2005	-0.0652	0.1223	-0.1483	-0.0877	-0.0466	0.0304	-0.1317	-0.1123
Panel B: Model with Pseudopools†								
1994	-0.7204	-0.7194	-0.3530	-0.2633				
1995	-0.7762	-0.6826	-0.2339	-0.1167	-0.7550	-0.7374	-0.2244	-0.0469
1996	-0.8306	-0.7552	-0.1942	-0.1085	-0.6760	-0.7983	-0.1034	0.0114
1997	-0.7816	-0.5884	-0.1236	-0.1060	-0.7638	-0.5736	-0.1147	-0.1009
1998	-0.6733	-0.5730	-0.1010	-0.0125	-0.5796	-0.5205	-0.0711	0.0160
1999	-0.5377	-0.3700	-0.0823	-0.0424	-0.4324	-0.2986	-0.0306	0.0268
2000	-0.5739	-0.3979	-0.0373	-0.0820	-0.4694	-0.3495	-0.0050	-0.0165
2001	-0.6626	-0.3151	-0.0566	-0.0309	-0.6109	-0.2798	-0.0421	-0.0101
2002	-0.5050	-0.2373	-0.0783	-0.0103	-0.5214	-0.2661	-0.0920	-0.0180
2003	-0.4220	-0.2420	-0.0775	0.0013	-0.4316	-0.2740	-0.0900	-0.0099
2004	-0.2683	-0.0568	-0.0387	-0.0291	-0.3278	-0.1148	-0.0302	-0.0323
2005	-0.1282	0.1079	0.0000	-0.0765	-0.2467	-0.0400	-0.0422	-0.1048

Note: *Coefficients from hazard models were estimated separately for each year of the sample. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define premium as $0.100 < a < 0.25$.
†Coefficients from hazard models estimated separately for each year. Let $a = 1 - \text{current coupon rate/coupon rate on mortgage}$. We define premium as $0.100 < a < 0.25$. Coefficients in 1994 for premium mortgages omitted due to insufficient number of observations. Pseudopools are formed by grouping loans according to loan-to-value ratio, borrower credit, contract rate, date of origination and loan amount.

of a mortgage as the expected present value of mortgage cash flows. A disadvantage of using a proprietary model is that we do not know the details of the model, and our ability to modify the model is limited. However, the Yield-Book[©] model has been widely used, and we have the ability to change some of its parameters by multiples of the sort we estimate. This allows us to compare changes in value due to changes in the propensity to exercise prepayment options.

In particular, the model can be broken down into an option-exercising part and a part that takes account of other factors. To the extent that we can identify these components with our in-the-money and out-of-the money coefficients,

Table 14 ■ Effect of prepayment speeds on required yield spread.

In Basis Points	Yield Curve Slope		
	Flat	Up	Down
Borrower Type			
Black	2	-2	-15
Hispanic	1	0	-6
<80% Area median income	1	-4	-9

Note: The table shows the premium (positive numbers) or discount (negative numbers) in basis points of coupon adjustment required for an LMM of the appropriate type to sell at par. Values were calculated by using the information about prepayment speeds from our hazard models, adjusting the prepayment model in YieldBook^(c) to reflect these differences in prepayment speed.

we can use the model to predict pricing and mortgage rate differences given the multipliers we estimate. This is, of course, imprecise as our prepayment model does not have the same functional form as the Citigroup model and it was estimated with an entirely different data set.

Black and Hispanic borrowers have slower prepayment speeds regardless of whether the option is in or out of the money. Consequently, we adjusted both the “turnover multiplier” and the “refinancing multiplier.” The turnover multiplier corresponds to prepayment in states of the world where refinancing motives other than rational prepayment, such as home sales, are the primary reason for prepayments, whereas the refinancing multiplier relates to states where it is financially beneficial to prepay. The effect of differences in the frequency of this kind of prepayment on prices is ambiguous, as prepayment when the option is in the money decreases the price that the investor is willing to pay, whereas prepayment when the option is out of the money should increase the price. A model is required to determine which effect dominates.

We analyze a current coupon 30-year fixed-rate mortgage. The model requires inputting the yield curve and a measure of interest rate volatility. We do not adjust the model’s volatility numbers, but we do explore prices for different yield curves. Our base case uses the current (Q1, 2007) flat yield curve. We asked the model to give us the difference between a base case price for the mortgage and the one adjusted for the new prepayment model. We then asked the model for the difference in mortgage coupon rate between the base case and the adjusted case assuming both are priced at par. We then chose scenarios with a sudden change to a downward-sloping and the more typical upward-sloping yield curves and considered the effects on yield spreads.

Our results are in Table 14. For the most common scenario, an upward-sloping term structure, the required yield on mortgages held by black borrowers should be discounted by 2 basis points, whereas that for low-income borrowers should be discounted by 4 basis points. The effect on yields for mortgages with Hispanic borrowers is less than a basis point. For a downward-sloping yield curve, in which the anticipated financial benefits for the borrower are substantial, the effects are larger: a 6-basis-point discount for Hispanic borrowers, a 6 for low-income borrowers and 15 for black borrowers. In the current regime of a relatively flat yield curve, the net result is a required addition to yield of 1 basis point for low-income and Hispanic borrowers, and 2 basis points for black borrowers. Results are not much affected by whether or not we use conditional or unconditional results.

Conclusions

Our main result is that there is a tendency for the prepayment behavior of low-income borrowers to be less ruthless than that of high-income borrowers; Black and Hispanic borrowers prepay more slowly than whites regardless of the value of their prepayment option. This result is relatively stable over time, although it shows signs of decline in the last 2 years of our data (2004 and 2005). The pricing implications of L IMM prepayment performance is relatively small for upward-sloping and flat yield curves, but for downward-sloping yield curves it is almost large enough to offset the effect of higher expected default.

Although our data set is both larger and more representative than other studies that have exclusively analyzed subsidized loans, it is still limited to the conventional conforming 30-year fixed-rate mortgages. As selection across mortgage types is surely not a random decision, L IMM borrower performance may differ for jumbo mortgages and other kinds of products, such as adjustable rate mortgages. Many mortgage products outside the scope of this article are designed to enhance affordability, making their relationship to historically underserved populations of particular interest. Within the range of conventional conforming loans, a very large part of the market, it appears to be the case that differences in prepayment behavior by L IMM borrowers are significant, but not very important in terms of cost and pricing.

We have received valuable help from Abdigani H irad and Adama Kah and comments from Michael Bradley, Darryl Getter, Stuart Gabriel and Nancy Wallace and seminar participants at Berkeley, the Philadelphia Fed and the University of Pennsylvania. The views expressed in this article are our own and do not necessarily reflect the views of Freddie Mac or its Board of Directors.

References

- Agresti, A. 1990. *Categorical Data Analysis*. New York: Wiley.
- Archer, W., D.C. Ling and G.A. McGill. 1996. The Effect of Income and Collateral Constraints on Residential Mortgage Terminations. *Regional Science and Urban Economics* 26: 235–261.
- Barakova, I., R. Bostic, P. Calem and S. Wachter. 2003. Does Credit Quality Matter for Homeownership? *Journal of Housing Economics* 12: 318–336.
- Berkovec, J.A., G.B. Canner, S.A. Gabriel and T.H. Hannan. 1994. Race, Redlining, and Residential Mortgage Loan Performance. *The Journal of Real Estate Finance and Economics* 9: 263–294.
- Chinloy, P. and I. Megbolugbe. 1994. Hedonic Mortgages. *Journal of Housing Research* 5: 1–21.
- Deng, Y.H. and S. Gabriel. 2006. Risk-Based Pricing and Enhancement of Mortgage Credit Availability among Underserved and Higher Risk Populations. *Journal of Money, Credit, and Banking* 38: 1431–1460.
- Deng, Y.H., J. Quigley and R. Van Order. 2000. Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options. *Econometrica* 68: 275–307.
- Dunn, K.B. and J.J. McConnell. 1981. Valuation of GNMA Mortgage-Backed Securities. *The Journal of Finance* 36: 599–616.
- Findley, M.C. and D. Capozza. 1977. The Variable Rate Mortgage: An Option Theory Approach. *Journal of Money, Credit and Banking* 9: 356–364.
- Goldberg, M. and J. Harding. 2003. Investment Characteristics of Low and Moderate Income Mortgage Loans. *Journal of Housing Economics* 12: 151–180.
- Green, R.K. and M. Lacour-Little. 1999. The Truth about Ostriches: Who Doesn't Prepay and Why They Don't. *Journal of Housing Economics* 8: 233–248.
- Hall, A. 2000. Controlling for Burnout in Estimating Mortgage Prepayment Models. *Journal of Housing Economics* 19: 215–232.
- Hayre, L. and A. Rajan. 1995. *Anatomy of Prepayments: The Salomon Brothers Prepayment Model*. New York: Salomon Brothers.
- Hendershott, P. and R. Van Order. 1987. Pricing Mortgages: An Interpretation of Models and Results. *Journal of Financial Services Research* 1: 19–55.
- Kau, J. and D. Keenan. 1995. An Overview of Option-theoretic Pricing of Mortgages. *The Journal of Housing Research* 6: 217–244.
- Lacour-Little, M. 1999. Discrimination in Mortgage Lending: A Critical Review of the Literature. *Journal of Real Estate Literature* 7: 15–49.
- . 2004. Equity Dilution: An Alternate Perspective on Mortgage Pricing. *Real Estate Economics* 32: 359–384.
- Peristiani, S., P. Bennett, G. Monsen, R. Peach and J. Raiff. 1997. Credit, Equity, and Mortgage Refinancings. *FRBNY Economic Policy Review* 3: 83–99.
- Prentice, R. and L.A. Gloeckler. 1978. Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data. *Biometrics* 34: 57–67.
- Richard, S.F. and R. Roll. 1989. Prepayments on Fixed-Rate Mortgage-Backed Securities. *Journal of Portfolio Management* 15: 73–82.
- Van Order, R. and P. Zorn. 2000. Income, Location and Default: Implications for Community Lending. *Real Estate Economics* 28: 385–404.