JOINT CENTER FOR HOUSING STUDIES
of Harvard University

PERFORMANCE OF LOW-INCOME AND MINORITY MORTGAGES

LIHO-01.10

Robert Van Order and Peter Zorn
September 2001

Low-Income Homeownership
Working Paper Series

Joint Center for Housing Studies

Harvard University

Performance of Low-Income and Minority Mortgages

Robert Van Order and Peter Zorn

LIHO-01.10

September 2001

We have received valuable help from Abdigani Hirad and Adama Kah and comments from Michael Bradley and Stuart Gabriel. What follows does not necessarily reflect the opinions of Freddie Mac.

This paper was prepared for the Joint Center for Housing Studies' *Symposium on Low-Income Homeownership as an Asset-Building Strategy* and an earlier version was presented at the symposium held November 14-15, 2000, at Harvard University. The Symposium was funded by the Ford Foundation, Freddie Mac, and the Research Institute for Housing America.

This paper, along with others prepared for the Symposium, will be published as a forthcoming book by the Brookings Institution and its Center for Urban and Metropolitan Policy.

ABSTRACT

This paper analyses the performance of low-income and minority loans (LIMLs) from a large sample of fixed-rate mortgages purchased by Freddie Mac in the 1990s. Our focus is on both default and prepayment behavior, especially prepayment. Prepayment is complicated because from a performance standpoint it matters whether a loan is prepaid after rates have fallen (taking advantage of the borrower's call option) or when rates are the same or have risen. Loans that prepay less rapidly when rates fall (i.e., when the option is "in the money") are more valuable to investors, but loans that prepay rapidly otherwise are less valuable.

We find that LIMLs generally prepay less rapidly than other loans when the option is in the money and prepay similarly when the option is out of the money, making them more valuable in terms of prepayment risk. We also find that they default more, offsetting some of the lower prepayment risk. In both cases the effect of race is larger than the effect of income. We analyze the value of both of these differences. While our results are inevitably imprecise, we find that the two effects (lower prepayment for LIMLs but greater default) are of similar value and if anything the lower prepayment risk has been of greater magnitude. Whether this will continue, given improvements in the ability to refinance, is unclear. We find that most, but not all, of the differences, especially in credit risk, can be explained by factors like loan-to-value ratio and credit history.

Table of Contents

I. Introduction

This paper analyzes the performance of low-income and minority mortgage loans relative to other mortgages for a large sample of fixed-rate mortgages originated in the 1990s and followed through 1999. Evaluating performance differences is complicated. For instance, it is not just a matter of credit risk. For fixed-rate mortgages, it is clearly the case that prepayment risk—the risk that comes from borrowers exercising their option to refinance when mortgage rates fall (which amounts to exercising a call option) —has a cost of at least the same order of magnitude as credit risk. Hence, low-income and minority loans may have had higher default rates than other loans, while also having different and more favorable prepayment characteristics.

We examine differences in prepayment and default behavior across groups and their effects on loan performance, which is measured by value to mortgage investors[1]. One way of tracking performance is to look at historic returns. We do not have market prices of the individual loans over time, so we cannot do this. We do, however, have data on the main things that effect performance, default, and prepayment. Both of these can be viewed as options that impose costs on investors, and we can ask questions about the differences in borrowers' propensity to exercise these options and use a generic pricing model to estimate "shadow prices" for the differences.

We first estimate models for prepayment, and we find that low-income and minority loans (LIMLs for short) have a lower propensity to exercise the prepayment option. We then analyze the extent to which this is the case: 1) when the option is "in the money" (i.e., is there a lower propensity to refinance when mortgage rates have fallen), and 2) when the option is not "in the money" (e.g., due to less mobility or a lack of access to other forms of raising money). From an investor's perspective it matters.

Clearly, if it is the case that LIML options are exercised less when they are in the money, then LIMLs are more valuable to investors because their "optionness" is less valuable to borrowers. However, if LIMLs prepay less rapidly when the option is either

[1] We mean "investor" in a rather broad sense. For instance, most mortgages now go into mortgage-backed securities, where the pool-issuer, e.g., Freddie Mac, takes the credit risk, but investors in the pools take the prepayment risk. By "investor" we mean of composite of all the stakeholders.

"close to the money" or "out of the money," then they are less valuable. If the option is close to the money, mortgage investors prefer borrowers who pay off quickly because they get a long-term rate for a short-term loan, and because quick prepayment means that the refinancing option will be outstanding for a shorter period of time; if the option is out of the money refinancing represents a windfall to lenders, so refinancing less rapidly lowers value to lenders.[2]

The issue of prepayment differences between LIML borrowers and others and their implications for pricing was raised in Chinloy and Megbolugbe (1994). Their argument was based on the notion that LIML borrowers are less mobile and less liquid than other borrowers, so they prepay less. That, in itself, does not get you to the proposition that LIML borrowers have less prepayment risk because lower mobility implies slower prepayment when the option is out of the money, which makes mortgages *less* valuable, and it has no particular implication for refinancing when the option is in the money. The liquidity problem will tend to offset this when the option is out of the money, because refinancing might be the only way for LIML borrowers to raise money, but it also suggests a propensity to exercise the option when it is in the money that is no less than anyone else's. Our contribution is the distinction between behavior when the option is in and out of the money and our use of a large data set to get an empirical handle on the problem. We do find that the basic proposition in Chinloy and Megbolugbe, that LIML borrowers have less prepayment risk, is correct.

We find in our data that, absent controls for loan characteristics, both low-income and minority loans are slower to prepay when the option is in the money than is the case for base case loans (more so for loans to Blacks and Hispanics than for low-income borrowers), but they are about the same in terms of other (close to or out of the money) prepayments. This suggests potentially important differences in value to investors not due to credit risk differences.

If we adjust for loan characteristics, particularly credit history, loan to value ratio (LTV) and loan amount, the results change, and low-income and minority loans are slow

[2] Rates on shorter-term mortgages are always lower than those on longer-term ones, in part because the prepayment option allows long-term borrowers to take advantage of any downward slope in the mortgage yield curve by converting their mortgages into shorter-term ones via prepayment. This is also reinforced by the general tendency of yield curves for noncallable debt to be upward sloping.

both in the money and out of the money by about the same multiple. This complicates performance analysis. In situations where the option is especially valuable (e.g., when rates fall, or in terms of *ex ante* valuation, when the yield curve is downward sloping, indicating market expectations of an interest rate decline, or rates are especially volatile), the LIMLs will be more valuable. However, they will be less valuable when rates are rising or, *ex ante*, in a sharply rising yield curve or low volatility environment.

We estimate similar models for default rates, and we also estimate determinants of differences in loss severity. Absent adjustment for loan characteristics, we find that low-income and minority loans default at significantly higher rates than base case loans. This, along with somewhat higher loss severity rates, implies higher default costs. When we adjust for loan characteristics, especially credit history and loan to value ratio, we find that borrower race/ethnicity has very little effect, but we do find explanatory power for borrower income and neighborhood minority composition and income.

We have two main results:

1. LIMLs in our sample performed about the same as or better than other loans when prepayment risk is taken into account. That is, the (*ex ante*) shadow values of the default and prepayment differences from the base cases, taken from a generic pricing model, are, in absolute value, on the same order of magnitude. If anything the prepayment difference is more valuable, but the crudeness of our pricing calculations does not allow us to be sure.
2. Most of the performance differences between LIMLs and other loans can be explained by observable characteristics like downpayment and credit history. This is especially true for credit risk.

Whether the default or prepayment results will continue to hold in the wake of recent changes in mortgage markets is a different matter. We test for past stability by estimating separate models by exposure (calendar rather than origination) year. We find that during the 1990s it was the case that for prepayment the income effect fell, but the race effect changed little; and for default we find little change for race and some worsening of the income effect, but no big changes. In contrast, in the prepayment models with controls for mortgage and borrower characteristics, the coefficients of race/ethnicity were quite stable over time, but the income effect declined to virtually zero. Coefficients in default models with controls were generally not statistically significant in individual exposure years.

II. Models

It is by now well established that prepayment and default behavior can be viewed as exercising options. Prepayment amounts to exercising a call option, which gives the borrower the option to buy back the mortgage at a price equal to the mortgage balance, and default amounts to exercising a put option to sell the house to the lender at a price equal to the value of the mortgage. But these options are not perfectly or predictably exercised in the way that, say, corporate bond options are exercised. This is clearly true for default because exercising the put option involves significant costs to borrowers (e.g., worse credit history and diminished access to future credit). 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. Hence, a reason for exercising the call option is mobility. Furthermore, different borrowers have different access to other forms of credit. A borrower with limited access to other credit opportunities and/or lack of liquid assets might exercise an out of the money call option on a mortgage in order to refinance and take out equity in order to pay for something else.

Pricing options is a growth industry that has been extended to the mortgage business (see Hendershott and Van order (1987) and Kau et al., (1995), for surveys of option-type models as applied to mortgages). The methodology is in principle simple, but in practice can be very complicated. The basis of all pricing models is that the value of a mortgage is the (risk-adjusted) expected present value of its cash flows, taking account of the way borrowers exercise their options. The value of the prepayment option is the difference in value between a mortgage without a prepayment option and the value with it. This can be turned into an interest rate differential by comparing the coupon rate on a par-valued mortgage without a prepayment option to the (higher) coupon rate on a comparable mortgage with a prepayment option. This same methodology can also be used to answer questions about increases or decreases in mortgage rate due to differences in the extent to which different borrowers or borrower classes exercise the prepayment option. The methodology can be used in much the same way to price the level of credit risk and differences in credit risk across borrowers. In general it is easier to calculate up-front values (e.g., via expected present value calculations, using Monte Carlo techniques) than it is to calculate coupon rate differences (holding value constant), but there are simple rules of thumb that allow for simple conversions from one to the other (e.g., for a 30-ear fixed-rate mortgage, a one basis point (bp) increase in coupon rate generally leads to a four or six increase in up-front value, depending on the duration of the mortgage).

We estimate prepayment and default probabilities with proportional hazard models of the form:

(1) $h(t)=exp(Bx)$

where $h(t)$ is the probability over some small time interval 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 a vector of coefficients. The x's can take on a wide variety of forms. For instance, they can represent time trends or the age of the mortgage; e.g., some of the x's might be a series of dummy variables for the age of the loan or the quarter in which the loan was originated. An important property comes from the multiplicative nature of the model. For instance, if the x's are categorical variables, then there is an easy interpretation of the Bs as multipliers; $\exp(B_1)$ gives a multiplier for the effect of being in category 1, relative to some baseline.

In principle the two hazards, default and prepayment, should be modeled and estimated jointly (see Deng et al.). We estimate them separately, but we take account of jointness implicitly by using the same explanatory variables in the two equations. To the extent this presents problems it is likely to be in the estimates of default rather than prepayment, which is our main focus. This is because default is a very small number, typically 10 to 30 basis points per year, relative to prepayment, which fluctuates from around 10 to 40 percent per year. Hence, ignoring default in modeling prepayment is not likely to be quantitatively important.

A key variable is the extent to which the option is in the money, which in the case of prepayment can be measured by the difference (or ratio) between the rate on the mortgage and the current rate on a par mortgage.[3] The coefficient of this variable measures the propensity of the borrower to exercise the option as it goes into the money. Differences in this coefficient across groups will lead to different mortgage values across groups. We test for such differences by looking at the extent to which different groups have greater or smaller propensities to exercise their options for different categories of difference between mortgage rate and current coupon rate (i.e., different degrees of "in-the-moneyness").

We control for observable loan characteristics by treating them as categorical variables, which can be modeled as fixed effects. For each calendar quarter of originations we create fixed effects from "pseudo pools" by dividing originated loans into relatively homogeneous groupings based on observed characteristics—such as contract rate (50 bp buckets), LTV (four buckets), credit history measured by "FICO"[4] score (four buckets) and loan amount (three buckets). For each origination quarter this results in on the order of 200 pseudo pools. Each of these pseudo pools is then given a fixed-effect for each quarter it is "alive" (up to 27 quarters). With 12 origination quarters in the study, this amounts to a total of over 50,000 fixed effects in the model, which, because our data have well over one million loans and millions of loan-quarters, leaves plenty of degrees of freedom.

This structure is particularly good at accounting for the complex time-varying pool characteristics that plague traditional prepayment models. For instance, burnout (the notion that seasoned pools that have been exposed to one or more periods of low mortgage rates prepay at lower speeds than new pools) and seasoning effects are captured separately for each pseudo-pool by its quarter age fixed effect.

In the prepayment model we break the data down into quarters where the option is in the money, i.e., when the current coupon rate on a par mortgage is less than the average coupon of the mortgage pool, and out of the money, when the current coupon is higher than the average mortgage coupon of the pool, to varying degrees. We then

[3] In the case of default, it might be represented by difference between current house price and mortgage value; however, house prices are seldom observed over time, so that proxies like original loan-to-value ratio along with information about house price trends are typically used.

[4] This is a generic credit score developed by Fair Isaac Corporation, which has become widely used in mortgage credit-scoring models.

estimate versions of equation (1) with race and income and other explanatory variables for each of these samples. The coefficients of race and income tell us borrowers' propensity to exercise options when they are in the money and when they are out of the money. We have five in-the-moneyness groups, described below. We then use a generic pricing model, one which uses the expected present value methodology to price mortgages, but which was estimated with a different data set, to give back-of-the-envelope estimates of how much these differences in propensity to exercise options affect mortgage rates.

We do something similar for default, although we do not have comparable data on whether the option is in or out of the money. In particular we estimate a hazard model like equation (1) for default to see if LIMLs default differently, and we also estimate determinants of loss severity rates. There is no available generic pricing model for credit risk, but we do have reasonable ideas about the likely expected present value of default costs of a baseline mortgage, to which we apply multipliers from our estimated models. This gives us estimates of price and implied coupon rate differences among groups.

III. Data and Methodology

We use two different data sets. One consists of all 30-year fixed-rate mortgages originated from 1993–1995 and purchased by Freddie Mac and for which key data are not missing. The data are used to model default; they contain about 1.4 million loans. The other data are the same except that they include loans originated from 1993–1997. This set contains about 2.8 million loans and is used to model prepayments. The reason the second set of data is not used for default modeling is that loans originated after 1996 have extremely low default rates; it is too soon for them to have defaulted.[5] However, loans originated in 1996 and 1997 have been exposed to one major rate decline, in 1998, and have a lot of prepayments.

In general our data are richer in prepayment experience, because the loans have all been exposed to at least one period of declining rates (there were sharp mortgage rate declines in 1993, 1995, and 1998). The default modeling suffers from excessively good times in the 1990s and relatively small levels of default. Happily for us modelers, the California economy did 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 end of 1999.

Equation (1) is defined for a model with continuous time. We group our observations into quarters. Prentice and Gloecker (1978) show that for data observed during discrete time intervals (1) becomes the complementary log-log model:

(2) $\log(-\log(1-h(t))=Bx$

where the observations of x happen over discrete intervals (in our case quarters), so each $h(t)$ is the hazard rate for a particular loan during a quarter.[6] This formulation has the advantage that the estimates are not affected by size of the interval (e.g., weeks vs. quarters). It is this equation that we estimate in order to obtain estimates of the Bs.

We estimate two sorts of models: simple and complicated. The simple models are standard hazard models that have race/ethnicity and income variables as the main x's, and baseline hazards that are fixed effects for loan age and, in the case of default, the state in which the house is located. These models give *average* or *unconditional* effects of these variables on default and prepayment. The complicated models add thousands of interactive fixed effects by creating pseudo pools of mortgages as described above; they give *marginal* or *conditional* effects of race and income on default and prepayment.

An analogy to typical panel data analysis is useful here. We can divide x into characteristics that vary within a group (pseudo pool) and those that vary only across groups. For the purposes of this study, we are interested in estimating within group variation in behavior—how LIML behavior varies conditional on observable characteristics (i.e., within a pseudo pool). To accomplish this we partition x as follows. Let

(3) $y = \log(-\log(1\text{-h(t))}=B\,x$

If we let y^{ij} be the value of y for the ith borrower in the jth pseudo pool, we can partition the right hand side into two parts, so that

[5] Given our estimation procedure, which we discuss below, we would throw out most of these observations anyway because they would introduce *log(0)* into the equations.
[6] See Agresti (1990) for a discussion of complimentary log-log models.

(4) $y^{ij} = B_1 x_1^{ij} + B_2 x_2^{j}$

where x_1^{ij} is a vector of the individual characteristics of the *i*th individual in the *j*th pseudo pool (the variables in which we are interested, borrower income and race/ethnicity) and x_2^{j} includes characteristics common to all borrowers in the *j*th pseudo pool (the fixed effect for the pseudo pool that the loan is in).

We are interested in estimates of B_1. Following the analogy with panel data analysis, this can be accomplished by including group (pseudo pool) level fixed effects to capture the effects of x_2^{j}.

Alternatively, this can be accomplished through the subtraction of group level means. Subtracting pseudo pool means from both sides, we can rewrite (4) as

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

where $\bar{y}^{j}$ is the fraction of loans in the loan's pseudo pool that prepaid (or defaulted) 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 pseudo pool.

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

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

which we can rewrite as

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

We estimate equation (7) using maximum likelihood.[7] Note that creating pseudo pools by fully interacting the control variables allows us to control for observable characteristics in a rather nonparametric way. We do not produce estimates of B_2. The creation of the pseudo pools allows us to control for their effects without estimating thousands of parameters. This is a very simple but also rather complete representation, which allows us to look at the effects of race, income, etc., within pools, holding effects at the pool level constant.

[7] Pseudo pools 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 $\bar{y}^{j}$. We use a SAS program for estimation of log-log models.

Borrower income is a categorical variable, indicating whether at origination the borrower is in one of four income groups relative to area median income. Our coefficients are estimated relative to the high income (greater than 120 percent of median), so that the coefficient of the lowest income group (less than 80 percent of median) represents a "multiplier," which tells us, for instance, the extent to which low-income borrowers are more or less likely to prepay when their option is in the money. Similarly, we measure race/ethnicity by four categorical variables: "Black," "Hispanic," "Other Minority," and "White." We suppress the White variable in our estimates so that the coefficients represent multipliers relative to White.

IV. Results

Figure 1 presents definitions and breakdowns of our major variables. Figure 2 presents simple cross tabs. Part A gives prepayment rates (percent that ever prepaid during the sample period) by race/ethnicity and income. It gives the basic story. Blacks and Hispanics prepay significantly slower than Whites and other minorities, and low-income borrowers prepay more slowly than high-income borrowers. On the default side, 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 instance, for Hispanics and Other Minorities defaults increase with income. These are, of course, crude statistics. For instance, we should at a minimum correct for the fact that these rates are averages over loans that were originated at different times and exposed to risks for different periods of time. We now turn to estimates of various forms of hazard model.

Figure 1
Prepayment Sensitivity

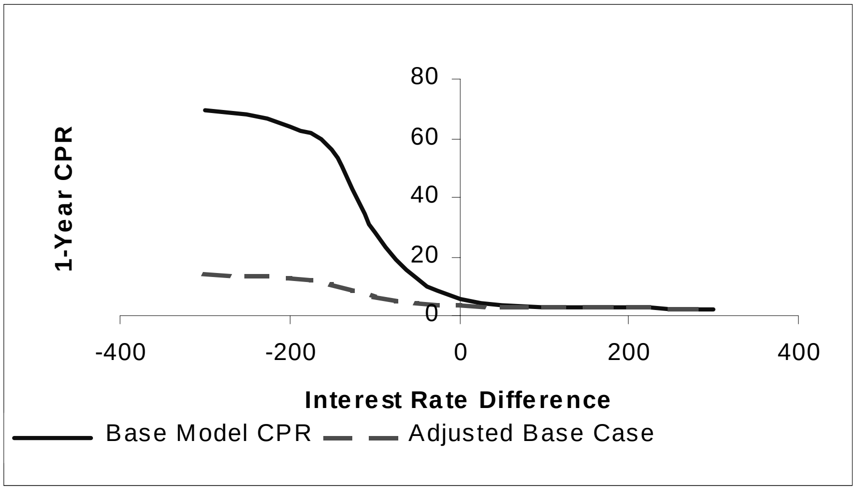

FIGURE 1: (continued) DATA

VARIABLE DESCRIPTIONS

Race/Ethnicity (0-1 dummy variables)

Black

Hispanic

Other Minority (Othermin)

White (suppressed in estimates**)**

Borrower Income Percent of Area Median Income (0-1 dummy)

Inc1 0 to 80% of median

Inc2 81 to 100% of median

Inc3 101-120% of median

Inc4 121% or more of median (suppressed in estimates)

Percent Minority in census Tract (0-1 dummy)

Minority=Black+Hispanic+Other Minority

Min1 Minority share of population<10%

Min2 Minority share of population between 11 and 30%

Min3 Minority share of population between 31 and 50%

Min4 Minority share of population > 50% (suppressed in estimates)

Census Tract Income as Percent of Area Median (0-1 dummy)

Tractinc1 Tract income <80% of median

Tractinc2 Tract income between 81 and 100%

Tractinc3 Tract income between 101 and 120%

Tractinc4 Tract income>120% (suppressed in estimates)

In-the-Moneyness

Let $1-a$=*current coupon rate/coupon rate on mortgage.*

Then

$a<-0.035$=**Discount**

$-0.035<a<0.035$=**Current**

$0.035<a<0.100$=**Cusp**

$0.100<a<0.25$=**Premium**

$a>0.25$=S**uper Premium**

Other Variables

FICO: Credit Score from scoring mode of Fair Isaac inc.

LTV: loan to value ratio at origination

Coupon Rate on Mortgage

State (dummy variable)

Loan **Age**

Purpose of loan: Purchase or refinance (dummy variable)

Debt Ratio: Ratio of total borrower debt to borrower income

Distribution of Main Variables (Percent):
Prepayment Database (2.8 Million Loans)

Race	
Black	3
Hispanic	4
Other minority	6
White	88

Borrower Income (Percent of MSA Median)	
Inc1 (0-80%)	19
Inc2 (81-10%)	16
Inc3 (101-120%)	16
Inc4 (121%+)	50

Percent Minority in Census Tract	
Min1 (10% or less)	61
Min2 (11-30%)	28
Min3 (31-50%)	6
Min4 (50%+)	5

Census Tract Median Income (Percent of MSA Median)	
TractInc1 (0-80%)	7
TractInc2 (81-100%)	23

TractInc3 (101-120%)	31
TractInc4 (121+%)	38

Figure 2: Cumulative Prepayment and Default Rates
By Race and Income

A. Prepayments (Percent that Ever Prepaid)

Income	Black	Hispanic	Other Minorities	White	Total
0-80	29	31	40	42	41
81-100	31	33	43	46	44
101-120	32	34	45	47	46
120+	34	36	45	49	48
Total	32	34	44	47	46

B. Default (Percent that Ever Defaulted)

Income	Black	Hispanic	Other Minorities	White	Total
0-80	2.8	1.6	0.9	0.7	0.8
81-100	2.2	2.3	1.1	0.6	0.8
101-120	1.9	2.5	1.1	0.6	0.7
120+	1.4	2.3	1.2	0.5	0.6
Total	1.9	2.2	1.2	0.6	0.7

Prepayment Models

Average Effects

Figure 3 presents results for estimates of complementary log-log hazard models like equation (4). The x_1^{ij} variables are the race/ethnicity and borrower income variables; the x_2^{j} variable is the age of the mortgage, which is a series of dummy variables for number of quarters since origination. The age coefficients are not shown. Not surprisingly, it tells the same story as the cross tabs.

The right-hand column adds two location variables: the average income of households in the loan's census tract relative to the area median, and the minority (Black + Hispanic + Other minority) share of households in the census tract. Including these variables affects the race/ethnic coefficients, lowering them a bit. For instance the coefficient for Black increases from -0.48 to -0.34, and the coefficient for low minority concentration (Min1 [0 to 10]) is o.36, relative to high concentration (greater than 50 percent). 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. The results for income are that there is virtually no change in the explanatory power of individual income, but a small effect of neighborhood income on prepayment.

Figure 3: Basic Prepayment Results: Hazard Model Controlling for Age*

Dependent Variable: Prepayment Rate

Black	-0.48 (.01)	-0.34 (.01)
Hispanic	-0.40 (.01)	-0.24 (.01)
OtherMin	-.11 (.0004)	-0.02 (.004)
Inc1 (0-80)	-0.17 (.002)	-0.17 (.003)
Inc2 (81-100)	-0.09 (.003)	-0.10 (.003)
Inc3 (101-120)	-0.04 (.003)	-0.05 (.003)
Min1 (0-10)	-	0.36 (.005)
Min2 (11-30)	-	0.22 (.005)
Min3 (31-50)	-	0.11 (.006)
TractInc1 (0-80)	-	0.08 (.004)
TractInc2 (81-100)	-	0.02 (.002)
TractInc3 (101-120)	-	0.02 (.002)

*Standard errors in parentheses.

Because it matters whether lower prepayments rates happen when the prepayment option is in or out of the money, we divided our observations into loan quarters where the option was in or out of the money to varying degrees, as described above, and ran separate versions of the model in Figure 3 for each category. We have five categories. Going from most out of the money to most in the money, they are: 1) Discount, 2) Current, 3) Cusp, 4) Premium, and 5) Super Premium. Figure 1 gives definitions of the variables. In the interest of saving space and readers' time, we do not report results for "Cusp" or "Super Premium" and focus primarily on mortgages that were Discount (option out of the money) or Premium (option in the money).

Figure 4: Basic Prepayment Results: Hazard Model Controlling for Loan Age * Results by Exercise Category

	Discount	Current	Premium	
Black	0.01 (.01)	-0.28 (.01)	-1.57 (.02)	
Hispanic	-0.02 (.008)	-0.24 (.01)	-1.18 (.02)	
OtherMin	-0.13 (.006)	-0.09 (.008)	-0.22 (.01)	
Inc1	0.08 (.004)	-0.06 (.005)	-0.60 (.006)	
Inc2	0.004 (.004)	-0.06 (.006)	-0.21 (.006)	
Inc3	-(0.004) (.004)	-0.04 (.005)	-0.06 (.006)	

*Standard errors in parentheses.

Results are depicted in Figures 4 and 5. Figure 4 re-does the first column of Figure 3 and Figure 5 re-does the second column. The results are rather striking and similar in direction for both race and for income. In Figure 4 we see virtually no difference for Black and Hispanic prepayment rates (and a small decline for other minority) when the option is out of the money, a small difference when the option is close to the money, but a big decline when the option is in the money. For instance, the coefficient for "Black" implies that for premium loans Blacks are exp(-1.57) or 0.2 times as likely to prepay as Whites and exp(-0.28) or about 0.8 for current coupon loans. For the lowest income groups the results are similar but not as large. For instance, for the lowest income group prepayment speeds are exp(-0.60) or 0.5 times as likely to prepay as those with incomes more than 120 percent of median (about half the loans in the sample) with very little difference when the option is out of the money.

Figure 5: Basic Prepayment Results: hazard Model Controlling for Loan Age With Census Tract Variables*

	Discount	Current	Premium
Black	0.08 (.01)	-0.22 (.01)	-1.11 (0.2)
Hispanic	0.05 (.01)	-0.18 (.01)	-0.67 (.02)
Other Minorities	-0.08 (.006)	-0.05 (.01)	0.01 (.009)
Inc1	0.05 (.004)	-0.05 (.005)	-0.58 (.006)
Inc2	-0.01 (.004)	-0.05 (.006)	-0.21 (.006)
Inc3	-0.02 (.004)	-0.04 (.005)	-0.08 (.006)
Min1	0.22 (.008)	0.16 (.01)	1.11 (.02)
Min2	0.16 (.008)	0.11 (.01)	0.71 (.02)
Min3	0.07 (.009)	0.05 (.01)	0.42 (.02)
TractInc1	0.20 (.006)	0.04 (.008)	0.05 (.01)
TractInc2	0.11 (.004)	-0.03 (.005)	-0.03 (.005)
TractInc3	0.04 (.003)	-0.03 (.005)	0.05 (.005)

*Standard errors in parentheses.

Figure 5 re-does the estimates adding census tract variables to those in Figure 3. Results are very similar. They suggest, if anything, that differences are bigger, e.g., for a black borrower in a minority neighborhood.

Clearly this means that LIMLs were made more valuable by the difference in prepayment behavior. We do not have an easy way of converting our multipliers into value to investors, but we can use some back-of-the-envelope calculations to get orders of magnitude. First, we can look at some market rates. At the time of writing, current coupon (7.5 percent) Freddie Mac mortgage pools were trading at prices that corresponded to yields that were about 75bp greater than yields on Freddie Mac noncallable debt of comparable duration. This difference is not just due to the value of the call option on the mortgages because the debt is generally more liquid than pass through securities and on that account can be sold with a lower yield. A reasonable guess is that around 40 to 50bp represents the part of mortgage rate due to the prepayment option, a rough estimate of the maximum reduction in cost that could come from low prepayment response. The coefficient of 0.2 for Black mortgages would probably take away most of this cost difference.

To evaluate the effect on value a little more rigorously, we applied a pricing model from Salomon Brothers (see Hayre and Rajan [1995] for a description), which is a generic pricing model that is widely available, but proprietary. It uses Monte Carlo techniques combined with empirical prepayment models to compute the value of a mortgage as the expected present value of mortgage cash flows. A disadvantage of using the model is that because it is proprietary we do not know the details (coefficients) of the model, and our ability to tweak the model is limited. However, this is offset by the fact that it is widely used and we have the ability to change some of its parameters by multiples like the sort we estimate, so that we can compare changes in value due to changes in the propensity to exercise prepayment options. In particular, the Salomon model can be broken down into an option-exercising part and a part that takes account of other factors. To the extent we can identify these with our in-the-money and out-of -the money, coefficients we can use the model to predict pricing and mortgage rate differences given the multipliers we estimate. This is, of course, imprecise. First, our model does not have the same functional form as the Salomon model; second, it was estimated with an entirely different data set; and third, the multipliers we apply are quite low and imply prepayment functions that are probably outside the experience of the Salomon model.

We adjusted the Salomon prepayment model so that it was 0.2 times its baseline (the multiplier for premium mortgages) when the option is in the money, slightly lower when the option is close to money and the same when out of the money. We did this in two ways. First (call this scenario one), we multiplied their coefficient for "prepayment incentive," which is a sort of interest elasticity of prepayment speed, by 0.2 (The Salomon refinancing incentive variable is similar to ours; in particular it is in ratio rather than difference form). Because we do not know the functional form of the model, this may not be comparable to our multiplier, which multiplies the whole function by 0.2. Second (scenario two), we adjusted this coefficient until the Salomon prepayment function (which we can graph) was approximately 0.2 times the baseline function for options in the money. A version of the old and new prepayment functions in this case are depicted in Figure 6; the picture for scenario one is similar.

We analyze a current coupon 7.5 percent 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 a historically typical yield curve, which is upward sloping with 10-year Treasuries 125bp above three- month Treasuries. We ask 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 ask the model for the difference in mortgage coupon rate between the base case and the adjusted case, assuming both are priced at par. The later is the more interesting question; because in the Salomon model for scenario one a 1bp increase in coupon rate leads to a 5bp increase in value, the answer is approximately the price difference divided by five (six for scenario two). The answer for scenario one is a difference of about 25bp in mortgage coupon rate. We then chose a downward-sloping yield curve, like the one observed in September 2000, and a sharply upward sloping one. We got answers of about 35bp and 15bp respectively. For scenario two, which has a lower and flatter prepayment function, the results were close to 5bp greater.

For low-income borrowers, the multiplier is 0.5, and the effect is smaller. Repeating the above estimates gives a base case of about 15bp in mortgage rate, for scenario one, with a range of about 10 to 20bp.

Our estimated hazard models had a super premium category, which had a larger multiplier, 0.8, and a "cusp" (barely into the money) category with a multiplier 0.5, as well as "premium." Because these multipliers are both greater than the 0.2 used in the simulations, our procedure overestimates the price difference somewhat. We cannot tweak the Salomon model in a way that readily allows two or three multipliers, so we cannot incorporate these differences directly. However, we suspect that they are small. For instance the cusp variable applies only (at current, eight percent mortgage rates) to cases where rates are between 7.2 percent and 7.7 percent, i.e., where the option is not very far into the money; and the multiplier is quite low, 0.5, anyway (as we saw above, for the low income calculations, if the multiplier is five across the entire range the effect is still on the order of 15bp).

Super premium loans have a multiplier that is closer to one. We note that the super premium category is for loans for which the current mortgage rate is less than .75 times the rate on the mortgage. At eight percent rates this means the difference only kicks in when rates have fallen to six percent. In the Salomon model and in reality this has a low probability of happening, and even with a 0.2 multiplier many mortgages will have prepaid by the time rates get to six percent. We simulated the model several times for prepayment functions that became very steep after a two percent rate decline and found very small effects. A difference of 5bp would be high; hence we believe that our calculations are, in a back-of-the-envelope sense, not much affected by assuming a multiplier of 0.2 throughout, but they probably do err on the high side.

Marginal Effects

The above looks at average experience. Our data allow us to control for loan to value ratio (LTV), credit history (measured by the Fair Isaac Corporations' credit scoring model, or FICO score) loan amount and other variables, as described above. Figures 6 and 7 report estimates of B_1 in equation (7). The new coefficients are, as before, for race/ethnicity and income by extent in and out of the money as in Figures 4 and 5, but now they are marginal effects, after controlling for loan characteristics listed in the figure, which define our pseudo pools.

Figure 6 corresponds to Figure 4. The controls have significant effects on the structure of the coefficients. In particular, for both Blacks and Hispanics the coefficients when the option is in the money fall in absolute value, and the coefficients when it is close to or out of the money become negative to the point where the three sets of coefficients are essentially the same. For low-income borrowers differences become quite small and almost the same across in-the-moneyness categories.[8] Figure 7 repeats Figure 5. The controls almost wipe out the income effects, but not the race/ethnicity effects.

[8] The FICO score is doing much of the work here. Apparently low FICO score borrowers prepay more when the option is out of the money and less when it is in the money. A hypothesis is that this reflects limited financing alternatives, so that low FICO score borrowers do not have access to other non-mortgage sources of funds and sometimes have to refinance a low-rate mortgage when they need money, but they are more likely to have trouble qualifying for a new loan when the option is in the money and they want to

Figure 6: Marginal Prepayment Results*
Hazard Model Controlling for Origination Quarter, Coupon, Age, FICO, LTV, and Loan Amount

	Discount	Current	Premium
Black	-0.58 (.02)	-0.56 (.02)	-0.62 (.009)
Hispanic	-0.53 (.02)	-0.47 (.02)	-0.59 (.008)
Othermin	-0.27 (.01)	-0.15 (.01)	-0.15 (.006)
Inc1	-0.13 (.007)	-0.07 (.007)	-0.07 (.004)
Inc2	-0.14 (.007)	-0.06 (.007)	0.02 (.004)
Inc3	-0.11 (.007)	-0.04 (.007)	0.04 (.004)

*Standard errors in parentheses.

These results complicate the pricing because we now have offsetting effects; for LIMLs, controlling for loan characteristics, prepayment is slower when the option is out of the money, which makes the loans less valuable and tends to offset the (now diminished) tendency to prepay less when the option is in the money. We redid the pricing exercise above, in this case multiplying the entire prepayment function by exp(-0.60) or about 0.6. For the current yield curve, which is the downward sloping (suggesting an expectation of falling rates) case, we found that the option part dominated and the difference in mortgage rates was about 10bp in scenario one. However, for the more normal yield curve (10-year Treasuries 125 bp above three-month Treasuries), where the option part is less important, we found the results approximately canceled out. For a sharply upward sloping yield curve the LIMLs are less valuable. For low-income borrowers the effects are in the same direction, but because the multipliers are quite small the effects are quite small.

refinance a high-rate mortgage. The correlation between FICO and race accounts for much of the change in coefficients. Loan balance is also a factor; low balance loans prepay in ways similar to low FICO loans.

Figure 7: Marginal Prepayment Results With Tract Variables*
Hazard Model Controlling for Origination Quarters, Coupon, Age, FICO, LTV, and Loan Amount

	Discount	Current	Premium
Black	-0.49 (.02)	-0.47 (.02)	-0.47 (.01)
Hispanic	-0.43 (0.02)	-0.36 (.02)	-0.41 (.01)
Other M	-0.21 (.01)	-0.08 (.01)	-0.06 (.01)
Inc^1	-0.12 (.007)	-0.06 (.007)	-0.05 (.004)
Inc^2	-0.13 (.007)	-0.06 (.007)	0.03 (.004)
Inc^3	-0.10 (.007)	-0.04 (.007)	0.04 (.004)
Min^1	0.24 (0.01)	0.27 (0.01)	0.39 (0.01)
Min^2	0.21 (0.01)	0.18 (.01)	0.21 (0.01)
Min^3	0.10 (.002)	0.10 (0.02)	0.13 (0.01)
$TractInc^1$	-0.03 (0.01)	0.03 (0.01)	-0.02 (.006)
$TractInc^2$	-0.10 (.006)	-0.05 (.007)	-0.03 (.004)
$TractInc^3$	-0.09 (.005)	-0.04 (.006)	0.02 (.003)

* standard errors in parenthesis

Stability Over Time

Mortgage markets have 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 LIML 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 above models for different exposure (i.e., calendar rather than origination) years. Results from some of the estimates are reported in Figures 8 and 9.

Figure 8: Average Prepayment Results Over Time
By Exposure Year

A. All Loans

	1993	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**
Black	-1.43 (0.2)	0.04 (0.03)	-0.13 (0.02)	-0.23 (0.01)	-0.27 (0.01)	-1.24 (0.02)	-0.45 (0.01)
Hispanic	-1.03 (0.1)	-0.06 (0.03)	-0.18 (0.02)	-0.24 (0.01)	-0.24 (0.01)	-0.97 (0.01)	-0.29 (0.01)
Inc^1	-1.57 (0.1)	0.05 (0.01)	-0.06 (0.01)	-0.03 (0.01)	-0.04 (0.01)	-0.44 (0.01)	-0.16 (0.01)
Inc^2	-0.78 (0.05)	-0.01 (0.01)	-0.06 (0.01)	-0.04 (0.01)	-0.04 (0.01)	-0.18 (0.01)	-0.08 (0.01)

B. Premium Loans

	1993	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**
Black	-1.49 (0.71)	-4.84 (11.31)	-2.39 (0.15)	-1.46 (0.07)	-0.96 (0.04)	-2.95 (0.05)	-0.72 (0.03)
Hispanic	-30.16 (2.42)	-4.61 (9.8)	-1.95 (0.12)	-1.51 (0.07)	-0.72 (0.03)	-1.92 (0.03)	-0.44 (0.02)
Inc^1	-1.56 (0.30)	-0.11 (0.8)	-1.65 (0.04)	-0.76 (0.02)	-0.37 (0.02)	-0.69 (0.01)	-0.33 (0.01)
Inc^2	-0.41 (0.18)	-0.70 (1.08)	-0.52 (0.03)	-0.23 (0.02)	-0.15 (0.02)	-0.22 (0.01)	-0.12 (0.1)

**Standard errors in parenthesis.*

Figure 8 presents results from the simple or *average* model, re-estimated for each exposure year; it presents only the coefficients for Black and Hispanic and the two lowest income groups. The results are comparable to those in Figures 3 and 4. Part A presents results for all loans, that is without separating the sample into in and out-of–the-moneyness (this is comparable to column one of Figure 3). Note that results for the Black and Hispanic coefficients have a similar pattern; the coefficients are especially big in 1998, about zero in 1994 and small, but negative, in the other years. The years with big effects, 1993 and 1998, were years of sharp interest rate declines and big refinancing waves; 1994 was a very small refinancing year because it was a year when interest rates went up. The other years were in between; 1995 saw rates drop, but many of the loans alive then were originated in 1993 and were out of the money (we need to control for age and in-the-moneyness). This is a pattern consistent with previous results: less exercise in years when the option is valuable and no difference when it is not.

Figure 9: Marginal Prepayment Results Over Time by Exposure Year*

A. Premium Loans

	1993	**1994****	**1995**	**1996**	**1997**	**1998**	**1999**
Black	-0.47 (0.37)	- -	-0.77 (0.05)	-0.72 (0.04)	-0.73 (0.03)	-0.63 (0.01)	-0.51 (0.02)
Hispanic	-1.17 (0.38)	- -	-0.76 (0.04)	-0.75 (0.03)	-0.62 (0.03)	-0.63 (0.01)	-0.43 (0.01)
Inc^1	-0.18 (0.17)	- -	-0.17 (0.02)	-0.12 (0.02)	-0.05 (0.01)	-0.07 (0.01)	-0.04 (0.01)
Inc^2	0.07 (0.44)	- -	0.00 (0.02)	0.02 (0.01)	0.03 (0.01)	0.02 (0.01)	0.03 (0.01)

B. Discount Loans

	1993	**1994**	**1995**	**1996**	**1997**	**1998**	**1999**
Black	0.32 (0.71)	-0.50 (0.07)	-0.60 (0.05)	-0.64 (0.04)	-0.59 (0.04)	-0.63 (0.07)	-0.51 (0.03)
Hispanic	0.51 (0.59)	-0.55 (0.06)	-0.62 (0.05)	-0.59 (0.04)	-0.59 (0.03)	-0.51 (0.07)	-0.40 (0.03)
Inc^1	-0.48 (0.34)	-0.30 (0.02)	-0.24 (0.02)	-0.17 (0.01)	-0.12 (0.01)	-0.09 (0.037)	-0.01 (0.01)
Inc^2	-0.42 (0.34)	-0.21 (0.02)	-0.22 (0.02)	-0.17 (0.01)	-0.14 (0.01)	-0.11 (0.03)	-0.05 (0.01)

*Standard errors in parenthesis.
** Not enough premium loans to estimate the model.

Part B gives results for the premium loans. This corresponds to column 3 of Figure 4. The central result is that while coefficients do change from year to year there is little apparent pattern. Aside from 1994 when there were few premium loans and the results are insignificant, the year with the biggest coefficient is 1998.[9] The only hint of a trend is one point; the lowest year is 1999. Hence, the lower rate of option exercise by Blacks and Hispanics when the option is in the money appears not to be changing much.

[9] The 1998 result is consistent with the notion that the differences are biggest in heavy refinancing years (the second biggest multiplier was in 1995, which was also a year when mortgage rates dropped and refinancing increased), which suggests that the effect is not proportional and that by under-weighting the heavy refinancing years, we are underestimating the shadow price difference. Note, however, that this pattern does not hold up when we add our controls in Figure 10.

The effect of income does appear to be declining, however. For the lowest income group the multiplier was about 0.7 in 1999, vs. about 0.5 for the sample as a whole.

Figure 9 depicts the results with control variables added, as in Figure 6. Part A presents results for premium loans. The coefficients for Black and Hispanic are quite stable across exposure years, but with a small decline in 1999. The coefficients for low-income borrowers go to close to zero. We have estimated exposure year variants of all the models in Figures 6 and 7, which are not shown here. We found that minority census tract concentration has the same sort of effect as before, but no surprising trends were found.

Our basic results then are that there is some tendency for prepayment behavior of low-income borrowers to converge to that of high-income borrowers; there is little or no tendency for that to be the case for Black and Hispanic borrowers; and the equations with the controls added are rather stable over time and are similar to those reported for the sample as a whole.

Figure 10: Basic Average Default Results:
Hazard Model controlling for State and Age*

Black	1.01 (.04)	0.77 (.04)
Hispanic	0.77 (.03)	0.53 (.04)
Other minorities	0.34 (.03)	0.24 (.03)
Inc1	0.25 (.03)	0.06 (.03)
Inc2	0.20 (.03)	0.08 (.03)
Inc3	0.16 (.03)	0.08 (.03)
Min1	-	-0.26 (.04)
Min2	-	-0.18 (.04)
Min3	-	0.02 (.04)
TractInc1	-	0.77 (.03)
TractInc2	-	0.56 (.03)
TractInc3	-	0.33 (.03)

*Standard errors in parentheses

.

Default Models

Average Effects. To analyze performance with respect to default we go through similar exercises. Figure 10 presents a model similar to that in Figure 3 except that it explains defaults, controlling for age of loan and state. Again, not surprisingly, it tells the same story as the simple cross tabs in Figure 2. In the left-hand column, the model implies a multiplier for Black of exp(1.01), or about three, and for low-income borrowers of exp(.25), or about 1.3. The right-hand column adds census tract variables. Note that the right-hand column implies a rather large effect for neighborhood income rather than individual borrower income. This result is similar to that in Van Order and Zorn (2000).

Figure 11: Marginal Default Results*
Hazard Model Controlling for Origination Quarter, State, Age, FICO, LTV, Debt Ratio, Loan Amount, and Purpose

Black	0.06 (.04)	-0.08 (.04)
Hispanic	0.13 (.04)	-0.02 (.04)
OtherMin	0.15 (.04)	0.08 (.04)
Inc1	0.32 (.03)	0.24 (.03)
Inc2	0.16 .03	0.11 (.03)
Inc3	0.08 (.03)	0.05 (.03)
Min1	-	-0.25 (.04)
Min2	-	-0.23 (.04)
Min3	-	-0.07 (.04)
TractInc1	-	0.40 (.04)
TractInc2	-	0.26 (.03)
TractInc3	-	0.13 (.03)

*Standard errors in parentheses.

Figure 12: Loss Severity
OLS Estimates for Log of Loss/Loan Balance*

Black	0.14 (3.51)
Hispanic	0.06 (1.69)
OtherMin	-0.03 (-0.90)
Inc1	0.38 (13.31)
Inc2	0.20 (6.42)
Inc3	0.11 (3.56)

**t* ratios in parentheses.

Figure 13: Marginal Loss Severity
OLS Estimates of log of severity rate, controlling for LTV, Origination Amount, Units, FICO, and State

Black	0.04 (0.98)	-0.03 (0.7)
Hispanic	0.02 (0.65)	-0.04 (-1.01)
OtherMin	-0.06 (-1.63)	-.09 (-2.52)
Inc1	0.05 (1.49)	0.00 (0.08)
Inc2	-0.03 (-0.87)	-0.06 (-1.72)
Inc3	-0.02 (-0.49)	-0.04 (-1.12)
Min1	-	-0.08 (-1.83)
Min2	-	-0.08 (-2.07)
Min3	-	-0.1 (-0.29)
$TractInc^1$	-	0.32 (7.92)
$TractInc^2$	-	0.13 (4.16)
$TractInc^3$	-	0.08 (2.64)

**t* ratios in parentheses.

To analyze default cost we need to model loss severity rates as well. We do this in Figures 12 and 13. These figures present OLS results from regressing log of loss severity (including all sorts of transaction and opportunity costs) divided by mortgage balance on the same variables as in Figure 10. It suggests small differences by race but bigger differences by income. Figure 13 controls for LTV, FICO, etc., and adds census tract variables. It suggests that the major explanatory factor is census tract income.

We do not have a generic pricing model to use to assess the extra cost of the higher default and severity rates; there is not much of a market for trading credit risk. But from recent Freddie Mac history we can approximate a base line level of default to which we can apply our estimated multipliers. In this sample the median loan has an LTV just under 80 percent. History suggests that these loans have about a two percent chance of ever defaulting; this was higher in the early 1990s during the recession and is smaller currently during the housing boom. Average loss severity rates on these have been about 30 percent. This suggests average losses of about 0.6 percent, which discounted to the present implies an expected present value of about 0.5 percent of loan balance. Using a “divide by five” rule of thumb, this implies an annual charge of about 10bp. A reasonable range around this is probably five to15bp (currently at the lower end). For loans to Blacks, the overall multiplier (including severity rates) is around 3.2, suggesting a range of cost of 16 to 48 bp and a difference from the baseline of roughly 11 to 33 bp. For low-income borrowers the multiplier, including severity rate differences, is about 1.7, which suggests a mean of 17bp and a range of default costs of 11 to 26 bp with a differential of six to 11 bp.[10]

An implication of this is, that while LIMLs do indeed have higher default costs, the lower costs from exercising the prepayment option have at least offset these for our loan sample.

Marginal Effects

[10] A factor not included is capital costs. To the extent that riskier loans require more capital, this can increase costs. Note that capital cost might increase cost differences for both credit risk and prepayment risk.

As with prepayment, we want to control for other determinants of default. We do not have the ability to measure in-the-moneyness at the loan level as we did with the prepayment model, so we must create proxies. We created pseudo pools to control for LTV, FICO, age, the ratio of borrower debt payments to borrower income, loan amount, loan purpose (refinance or purchase), and state in which the property is located. Again, we have pseudo pools for every combination of these, and we report the new coefficients for race/ethnicity, etc., (comparable to Figure 6) in Figure 11. The outstanding characteristic of Figure 11 is how little is left for the race variables to explain after controlling for the other variables. In the first column the coefficients for Black and Hispanic drop precipitously from those in Figure 10. In the second column, the sign for Black turns negative, although there is a significant effect for minority composition of the neighborhood. It is, however, also the case that the controls do not lower the income effect.

Redoing the pricing reveals almost no effect for race and a smaller (because of the new severity rate coefficients) effect for income. Of some interest is the effect of neighborhood (census tract) variables. Our data set does not allow us to say much about why we might see these effects. Two possibilities are that they capture information about property values (e.g., volatility might differ across neighborhoods) or they are a proxy for some better measure of (permanent) income.

Stability Over Time

As in the prepayment section, we re-ran our models by exposure year to see if default behavior had changed over time. This yields less information. First, the sample is smaller because in the default sample we look only at loans originated from 1993–1995, and, second, the exposure years 1993 and 1994 have very few defaults (the loans are too new to default, especially in a growing economy). Because of the thinness of the results we present a broad summary of the results, rather than figures. Our basic results for 1995 through 1999 are as follows:

1. Coefficients for Black and Hispanic in the simple, average model (corresponding to Figure 10) were somewhat below average in 1995 and 1996 and were stable at about their sample wide levels from 1997–1999. Hence there was little change; if anything there was a slight increase.

2. The effect of borrower income (i.e., the tendency of low-income borrowers to default more) increases over time, and the coefficient for the lowest income group was 0.34 in 1999 (vs 0.25 for the period as a whole; see Figure 10)
3. Tract income had a stable positive effect of about the same size as in Figure 10.
4. Minority tract composition had a slightly increasing effect over time.
5. When we added controls as in Figure 9, Black and Hispanic coefficients fell drastically as before, but we could not discern a pattern across time because the coefficients in individual years were not significant.
6. After controls, the income effect showed some sign of declining. Minority tract composition effects increase slightly over time.

Overall there is little reason to believe there have been significant changes in default behavior during the period.

V. Comments

Our major results are:

1. Our data imply that LIMLs generally performed the same as or better than other loans. This is because they were significantly less likely to refinance when mortgage rates dropped. The crudeness of our pricing calculations does not allow us to be confident about the extent to which the prepayment effect is larger.
2. Much of this can be explained by loan characteristics like credit history and LTV.

A question is whether this can be projected into the future. A major change in mortgage markets over the past decade has been the increased quickness with which borrowers refinance and the increased ability of riskier borrowers to get loans. Given the change that has taken place it seems unlikely that, to the extent the low prepayment multiplier is due to lack of sophistication or market opportunity, it will continue to be so low. On the other hand, much of the gap in prepayment behavior can be explained by variables like FICO. If those variables are “fundamental”, then we might project long-run prepayment differences to come from projections that eliminate the race and income effects in Figures

6 and 7, but this might be mitigated by increased ability of low FICO borrowers to get loans causing the FICO effects to decline over time.

When we re-estimated by exposure year, we found evidence that income effects on prepayment have fallen over time and, after adding controls, have about vanished, but we found very little reason to believe there have been changes in prepayment coefficients for Black and Hispanic borrowers. Changes in default behavior over time are more difficult to be sure of, but our estimates by exposure year suggest little change over time for Black or Hispanic borrowers with ambiguous (and small) results with and without the controls for the income effect.

References

Agresti, A. *Categorical Data Analysis*, New York: Wiley, 1990.

Chinloy, P., and I. Megbolugbe, "Hedonic Mortgages," *Journal of Housing Research* 5, no. 1 (1994): 1–21, Fannie Mae.

Deng Y.H., J. Quigley, and R Van Order. “Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options,” *Econometrica* (2000).

Hayre, L., and A Rajan. “Anatomy of Prepayments: The Salomon Brothers Prepayment Model,” Salomon Brothers, 1995.

Hendershott, P., and R. Van Order, “Pricing Mortgages: An Interpretation of Models and Results,” *Journal of Financial Services Research* 1 (1987).

Kau, James, and Donald Keenan. “An Overview of Option-theoretic Pricing of Mortgages,” *The Journal of Housing 6,* no. 2 (1995): 217–244.

Prentice, R., and L.A. Gloeckner. “Regression Analysis of Grouped Survival Data with Application to Breast Cancer Data,” *Biometrics* 34 (1978): 57–67.

Van Order, R., and P. Zorn.” Income, Location and Default: Implications for Community Lending,” *Real Estate Economics* (2000).