

PN - ABM - 102  
ISN 78985

**Mortgage Prepayment  
in India**

by

**Raymond J. Struyk and Genevieve Kenney  
with  
Mark Friedman**

**The Urban Institute  
and  
The Housing Development Finance Corporation, Ltd.**

**Report Prepared for  
USAID/India**

**September 1988  
Paper No. 3641-07-3B**

## EXECUTIVE SUMMARY

Paying off mortgages completely before the full term of the loan--mortgage prepayments--has received great scrutiny in developed countries, because of the effect of prepayments on the profitability of holding mortgages or mortgage-backed instruments as investments. In developing countries, in contrast, very little attention has been given to prepayments, perhaps because prepayments are unexpected. But prepayments can be an important source of loanable funds and they can effect the attractiveness of selling mortgages or mortgage-back securities such as participation certificates to investors. Hence, knowing the patterns and causes of prepayments can improve the financial performance of a housing finance institution.

This paper presents an analysis of mortgage prepayments to the Housing Development Finance Corporation, Ltd. (HDFC), by far the largest private originator of home mortgages in India. We focus on loans originated between the beginning of operations (1978) and the end of 1985. Prepayment experience is monitored until July, 1987.

Before reviewing the results of the analysis, we can place the Indian experience in a general context by noting that the two most important factors driving prepayments in the United States--mortgage interest rate volatility and residential mobility--are very different in India. Residential mobility

rates are very low in India, and the general interest rate environment is very stable in the closely regulated capital markets.

There is a significant amount of mortgage prepayment in India. On average, about 4 to 5 percent of outstanding mortgages are being prepaid annually. This cumulates to a substantial share of originations quickly: for example, after 6.5 years, 23 percent of the loans originated in 1980 were prepaid. The general pattern, to the extent it can be discerned from the limited experience to date, is for the rate of prepayments to accelerate for the first three years after origination and then to level off at a more constant rate.

We estimated a model to predict the probability that a mortgage would prepay sometime over its life, given the characteristics of the borrower, property, and mortgage at the time of loan origination. The estimated model indicates that the longer the loan term or the higher the property value the less likely is prepayment. Likewise, borrowers older at the time of origination are somewhat less likely to prepay than their more youthful counterparts.

We also estimated a more complete model to analyze the probability that a mortgage would be prepaid in a particular year after origination. This model adds a variable for the rate of interest on several widely held assets: deposits made in company accounts, deposits made at commercial banks, or the rate of appreciation on gold. To deal with the combination of time

since origination and the fact that we do not have observations for the whole potential term of outstanding loans, we estimated separate models of the likelihood of prepayment in each year following origination for the first several years after origination. In general, the ability to predict and explain prepayment is much greater in the very early years of the mortgage term.

We found a strong, negative relation between the rate of return on other assets and the probability of prepayment. As interest rate on company deposits, for example, rises the likelihood of repayment falls, since the loan is then relatively cheap. This is consistent with our expectations. This result is also probably capturing other developments in the overall economy and therefore the absolute value of the estimated decrease in prepayments associated with rising interest rates should be treated with caution.

The results of this more complete model also show that the likelihood of prepayment falls steadily as the age of the borrower at loan origination increases, after controlling for other factors. For example, the effect of the borrower being 40 at the time of origination decreases the likelihood of prepayment by .28; for a borrower age 55, the effect is a decrease of .31.

Where within India loans are originated also has a distinct effect on the likelihood of prepayment: locations outside of Bombay are consistently less likely to prepay. This "Bombay

effect" seems to result from the generally greater financial sophistication of mortgagors in Bombay, as a result of which mortgagors more actively seek out and find cheaper, "friendly" financing to replace their HDFC mortgages.

Finally, we examined the source of funds which borrowers who prepaid told to HDFC. Sale of the unit was given as the source of funds for prepayment for a surprisingly large 25 percent of units on which mortgages were prepaid. The largest source (40 percent) was some form of savings. Other important sources were receiving a loan from the borrower's employer and payments from provident funds. Further analysis, detailed in the paper, found significant relationships between the likelihood of using particular sources of funds for prepayments and borrower and mortgage characteristics.

## CONTENTS

Introduction .....	1
Expectations .....	4
Analysis Plan .....	9
Data Employed .....	12
Patterns of Prepayment .....	12
Predicting Prepayment Over the Life of the Mortgage .....	14
Estimated Time Path of Loan Prepayment .....	17
Duration-Specific Probits of Prepayments .....	18
Reasons for Prepayment .....	25
Conclusions .....	28
References .....	31
 Annex A: Data Cleaning Procedures	

## Mortgage Prepayment in India

The prepayment of home mortgages prior to the end of the mortgage loan term has become a very closely studied phenomenon in many western countries. Mortgagees and secondary mortgage facilities have large financial stakes in being able to predict the likelihood of prepayment and to adjust points paid at closing and prepayment penalties so as to compensate them for this type of risk.[1] Indeed, a new type of mortgage-backed security--the Collateralized Mortgage Obligation--has been developed explicitly to deal with this risk (Roll, 1987).

In contrast, mortgage prepayment patterns in developing countries have received almost no attention. This neglect may stem from an assumption that prepayment is either a rare event among the lucky minority of households who succeed in obtaining formal mortgage financing in these countries or that it is not systematic in the sense of borrowers responding to financial incentives to prepay their loans.

For several reasons this lack of information may be a serious impediment to the development of viable financial intermediaries in the housing sector. First, studies of prepayment patterns in the U.S. attribute much of thrift institutions' financial problems of recent years to mispricing

-----  
1. See, for example, MacDonald (1986), Jacob et al. (1987), Dunn and McConnell (1981), and Curely and Guttentag (1977).

of prepayments. Hence, without such analysis in developing countries it may be impossible to define the risks involved with expected duration mismatching. Evidence on whether households respond to economic incentives with respect to prepayment would allow a much better understanding of the risks of intermediation.

Second, besides households responding to financial incentives, studies for the U.S. show that household characteristics can be very important in the prepayment decisions. If this is so in India which, like most developing countries, has a highly segmented credit system that rations credit, this type of analysis could allow lenders to discriminate in such a way that their rationed funds could be allocated to households who are more or less likely to prepay. Thus, this type of information could be important for institutional profitability and growth.

Other reasons for housing finance institutions to have a direct interest in prepayment can be mustered. For one, if these institutions could predict the pattern of prepayment with some accuracy, then they could more efficiently plan their lending programs. On the other hand, if they are constantly "surprised" by prepayment, then it is likely that they are originating fewer mortgages than they might otherwise. They may also be mispricing the mortgages that they do originate. Possible initiation of secondary mortgage markets is another

reason for housing finance firms to consider prepayments seriously. A key attribute of interest to potential investors in "pass through certificates," one of the simplest secondary instruments and therefore a strong candidate as an early offering, is the likely pattern of prepayments. The greater the uncertainty about or lack of information on the volume of prepayments, the less attractive such instruments will be for investors who do not want to be constantly reinvesting the unexpected proceeds from prepaid loans.

A broader reason for interest in prepayments concerns financial deepening. One of the objectives of developing a market-oriented housing finance system in India is the hope that the extension of more financial services to household will help encourage the continuation of the growth of financial assets as a share of savings (Chakravarty, 1985). Empirical analysis of whether and how households' mortgage prepayments respond to economic incentives from the financial market relative to other forms of wealth holding (e.g., gold) can help assess the impact that supplying mortgage credit could have on financial deepening. The question that can be answered is: Do the expected returns on non financial and financial assets affect household demands for financial instruments such as mortgages?

This paper presents the results of analyzing mortgage prepayment in India. The data used come from the Housing Development Finance Corporation (HDFC), far and away the

country's largest private originator of home mortgages. HDFC began operations in 1977/1978 with about 250 individual loans. For its 1986/1987 fiscal year it originated about 25,500 individual mortgages, for a cumulative total of 103,000 such mortgages. Over its decade of operations, HDFC has financed about 53,000 additional units under loans to corporations. HDFC has achieved impressive expansion and has an excellent record of loan collections and management (Buckley et al., 1985; Boleat, 1985). The data analyzed are for individual home mortgages originated from the start of operations in 1978 through the end of 1985; prepayment experience on these mortgages until July 1987 is analyzed.

As documented below, prepayment is a significant factor at HDFC. So more precise information on prepayment patterns is of immediate interest to its management--both for its normal lending operations and in the context of its fledgling secondary market (pass through certificate) program.

### **Expectations**

Since the prepayment experience of the United States is perhaps the best documented and analyzed, we can use it to point out important characteristics of India's situation. In the United States the primary reasons for prepayments are declines in mortgage interest rates, which can make refinancing attractive, and residential mobility (both local and long

distance), which in turn depends on the state of the economy and demographic patterns[2]. In contrast, neither of these factors is likely to be very important in India. Residential mobility rates in India are extremely low--apparently under 2 percent per year. Rates for homeowners are significantly below the overall average. A recent study of housing in Ahmedabad, for example, found that 65 percent of all households had never moved from the dwelling occupied at the time the household was formed.[3] As to the interest rate environment, stability is the hallmark of India's highly managed economy. This stability is reflected in the data in Table 1 on movements in interest rates on various savings alternatives over the 1979/80 to 1985/86 period.

Thus, in general we expect the prepayment decision for HDFC borrowers to be much more associated with their attributes and those of the mortgage loans than is the case in the United States. A number of Indians involved in housing finance have stated that Indians feel much more strongly about extinguishing debt on their homes than other forms of debt. Moreover, these experts are doubtful that there are any systematic patterns in mortgage prepayment.

Still, several borrower and mortgage characteristics appear to have the potential to be significantly related to prepayment.

-----  
2. See Roberts (1987) for a general discussion of the causes of prepayment in the United States.

3. Mehta and Mehta (1987), tables 4.9 and 4.11; also see Schenk (1986).

Table 1  
Trends in Selected Interest Rates in India,  
1979/80 to 1985/86

	1978/79	1980/81	1981/82	1982/83	1983/84	1984/85	1985/86
Commercial Banks 1-year deposits	7.0	7.5	8.0	8.0	8.0	8.0	8.5
U.T.I. dividend rate <sup>c</sup>	10.0	11.5	12.5	13.5	14.0	14.25	15.25
Company rates <sup>a</sup> 1-year deposits	9.0/13.0	9.0/13.5	9.0/13.5	9.0/15.0	10.0/15.0	9.0/15.0	10.0/15.0
Gov't of India bonds annual yield for 5-15 year bonds <sup>b</sup>	4.70/5.74	5.8/6.8	5.8/7.0	6.2/7.8	6.7/9.0	6.5/9.0	6.5/9.5

a. By "well established companies" for deposits accepted from the public.

b. Annual redemption yield calculated on assumption that average price relates to the middle of the period.

c. Unit Trust of India, a national mutual fund.

Source: Report on Currency and Finance, 1985/86, (Bombay: Reserve Bank of India, 1987), Statement 50.

Among borrower attributes, age is likely to be important. Older middle-aged households have the possibility of obtaining lump sum settlements from provident fund contributions and whole life insurance policies that could be used to pay off their loans. Also, self-employed individuals may have greater possibilities than salaried workers for high income years to generate the funds for prepayment. On the other hand, as documented by Kozel (1987) for the Ivory Coast, the self-employed have higher savings, some of which must be kept in accessible form as a hedge against uncertainty. If assembling the down payment for home purchase caused them to draw down their reserves to an insufficient level, then "extra income" will go to rebuilding reserves rather than to loan repayment. In addition, since profit taken out of the business would be subject to the personal income tax, the borrower would be discouraged from taking a large amount of funds out of the business in a single year, since this would place him in a higher tax bracket.

After controlling for other factors, income relative to mortgage payments may be important, with those with lower housing expense burdens better able to accumulate funds for prepayment. The ethnicity of the borrowing household might be significant. Given the cultural diversity of India, it might well be that one ethnic group has a higher propensity to prepay than others, e.g. the Tamils compared to Bengalis. HDFC's

national system of 13 branches (1985) insures a wide range of locations; we approximate ethnicity by analyzing prepayment with the location of the branch office in which the loan was originated. Lastly, while less clear, the sex of the borrower may influence investment behavior, with men or women systematically more likely to diversify their investments away from housing.

In terms of the characteristics of the mortgage itself, we should begin by noting that with some quite minor exceptions, HDFC originates self amortizing fixed interest rate mortgages.[4] The average loan to value ratio is about 0.4, and the average term is about 13 years, with loans rarely being made for more than 15 years. Interest rates vary directly with the size of the loan from 12.5 percent to 14.5 percent. There are prepayment penalties, computed in months of interest payments on the loan. The highest penalty of 12 months interest is applicable to borrowers who are selling their homes. Lower penalties apply to stationary borrowers; for example, there is a single month's penalty if the payment source is receipt of a

-----  
4. There are two alternative instruments of note. One is used when a middle-aged father and son jointly purchase a home. The loan is structured so that payments are higher in early years and then shift to a lower level at the time of the father's expected retirement. This is called the "FLIP" mortgage. The second alternative instrument allows a middle-aged borrower to legally assign the anticipated proceeds of a whole life insurance policy or provident fund contributions to HDFC. The loan principal is reduced by the discounted present value of the assignment. Payments then are on the reduced principal for the term of the loan. This is termed a "balloon" mortgage by HDFC.

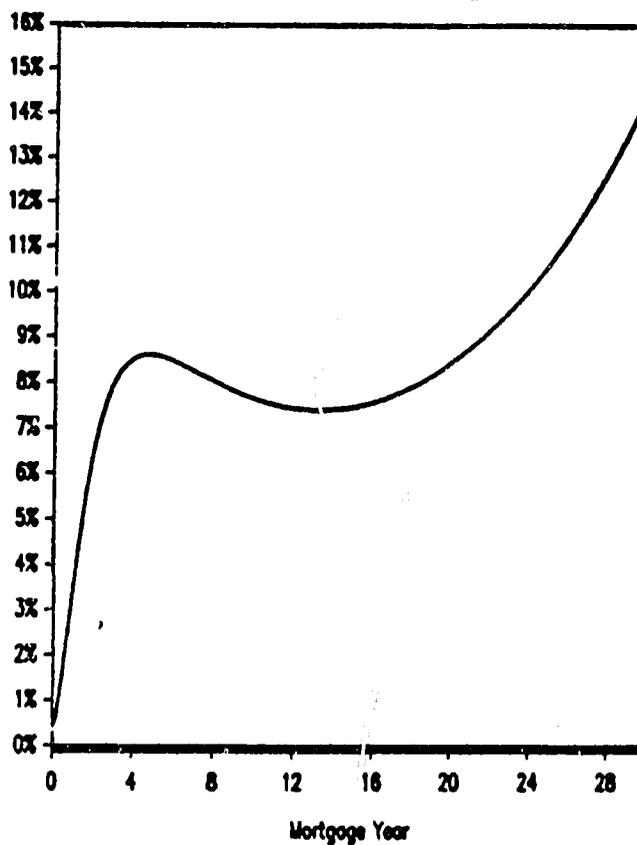
lump sum payment from a provident fund.[5]

One expects the propensity to prepay to vary directly with the mortgage interest rate, i.e., the lower the rate the more attractive it is to invest "extra funds" in other assets, and for the likelihood of prepayment to decline if the rates of return on other investments increased. One also expects prepayment to be more likely the lower the ratio of loan amount to income. The lower this ratio, the easier it is for the household to repay if it chooses to do so. Lastly, loan term may be a factor. If the loan term is quite short, then the borrower may be less anxious to eliminate this debt than if the term is long. Similarly, on a mortgage with a longer term, as the number of years to final payment becomes small, the borrower may feel less compelled about prepaying.

Not mentioned thus far but clearly a factor is the time since origination. The more time passes, the greater the household's opportunity to amass the funds for prepayment--and the smaller the balance to be paid. Figure 1 illustrates the effects of "pure aging" of mortgages on prepayment of FHA insured mortgages (Roberts, 1987, fig. 1). Prepayments are sharply higher in the first and the last years of the mortgage term. The high early prepayments are associated with

-----  
5. These are the penalties in effect for the analysis period. There is now active consideration to revise them to a less costly schedule.

**FIGURE 1**  
**The Pure Aging Effect:**  
**Annualized Percentage Prepay Rates**



Note: Prepayment rate defined as prepayments as a percent of outstanding balance yearly.

Source: Roberts (1987), figure 1.

8A

mobility and households "trading up" to better homes.

### **Analysis plan**

After reviewing the overall pattern of prepayments at HDFC, we report below on three distinct analyses, two predicting the likelihood of prepayment and one analyzing the sources of funds given by prepaying borrowers for doing so. In the first likelihood model the dependent variable equals one if the mortgage was prepaid at any time during the observation period and is zero otherwise. The independent variables, listed in Table 2, are of the type discussed above on borrower and mortgage attributes.

The estimated model will be useful for predicting, at the time of origination, which mortgages will be prepaid. Note that it excludes the rate of return on other investments, since it presumably is this rate in the future--when the borrower has the possibility for prepayment--that is relevant.

The second model of prepayment likelihood goes a step further by predicting the likelihood of prepayment in any given year. This model includes the rate of return on alternative, widely available investments--defined here to be the interest rate on deposits paid by companies, the interest rate on one year time deposits at commercial banks (see Table 1), and the appreciation in the price of gold, another widely held investment. As described more fully with the results, we also

**Table 2**  
**Definitions of Independent Variables**

<u>name</u>	<u>description</u>
TERM	loan term in months
LOANRATE	interest rate on loan
PAYINC	ratio of montly loan payment to monthly household income at time of origination
VIP	var=1, if loan made to shareholder
VALUE	value of property on which loan taken
LOANVAL	ratio of loan to property value
AGE	age of borrower; age squared also used.
SEX	sex of borrower; 1 = female
SELFEMP	var=1, if borrower is self employed
LOCATION	set of dummy variables for state (branch office) in which loan was originated; defined relative to Bombay (head-quarters).
TIME	number of months between origination and prepayment for prepaid loans only in cross section model; in cross section-time series model, months since origination.
COMPRATE	interest rate paid by well established companies for deposits from the public
YEARRATE	interest rate paid by commerical banks on one year time deposits.
GOLDRATE	annual percentage increase in price of gold in the Bombay market.

stratified the observations into estimation files by the number of years since loan origination to help deal with data censoring problems, i.e., the problem arising because we do not have information on whether a mortgage will prepay at any point over its full term, since our observations stop prior to the complete term for any mortgage under analysis.

The third component of the analysis explores the sources of funds used for prepayment by studying the sources stated by those prepaying. In particular, at the time a borrower notifies HDFC that it will repay its loan, HDFC inquires about the source of funds, which have been categorized into seven general types: (1) savings, (2) sale of the unit, (3) a loan received from the employer at a favorable interest rate, (4) "friendly," low interest loans from family or friends, (5) payments from a provident fund, (6) lump sum payment from a life insurance policy, and (7) other diverse sources. For this analysis we have collapsed these into three groups on the basis of whether the source of the prepayment was (a) sale of the unit; (b) savings, payments from a provident fund, life insurance, or attaining a mortgage loan from an employer (both of which might have been foreseen but with less than total certainty); or (c) other sources which appear to be less anticipated (a "friendly" loan). We examine the incidence of these reasons for prepayment and then estimate some simple regression models to explore the factors associated with prepayment.

Thus, this analysis is of the likelihood that the borrower has prepaid using funds from one of these three broad sources. Two models are estimated: one in which the dependent variable is 1 if prepayment is associated with home sale and is zero otherwise, and one in which the dependent variable is 1 if prepayments were due to receipt of provident fund payments, or other reasonably predictable sources. We do not estimate a third model, as we view the receipt of these sources as essentially random.

We employ ordinary least squares regression estimation using the same set of borrower and loan attributes as the independent variables to explore the relationship between these factors and the source of funds. In this case, however, these variables have somewhat different rationales than they do in the likelihood-of-prepayment models. For example, we expect older borrowers to prepay because they have received lump sum payments from provident funds or life insurance policies. Those with the greatest mortgage burdens (relative to income) may be more likely to prepay using receipts from the sale of their homes, in order to relieve general financial pressures on the family.

#### **Data employed**

Two different bodies of data are used. The computation of time profiles of prepayment use data on a random sample of approximately 5,100 mortgage originations and a random sample of

about 2,000 prepayments. The samples are of all mortgages originated from the start of HDFC's operations and the end of calendar year 1985. The prepayment experience of these mortgages is followed until July 1, 1987. After various data cleaning operations, described in Annex A, the final samples are 4,552 and 1,919, respectively.

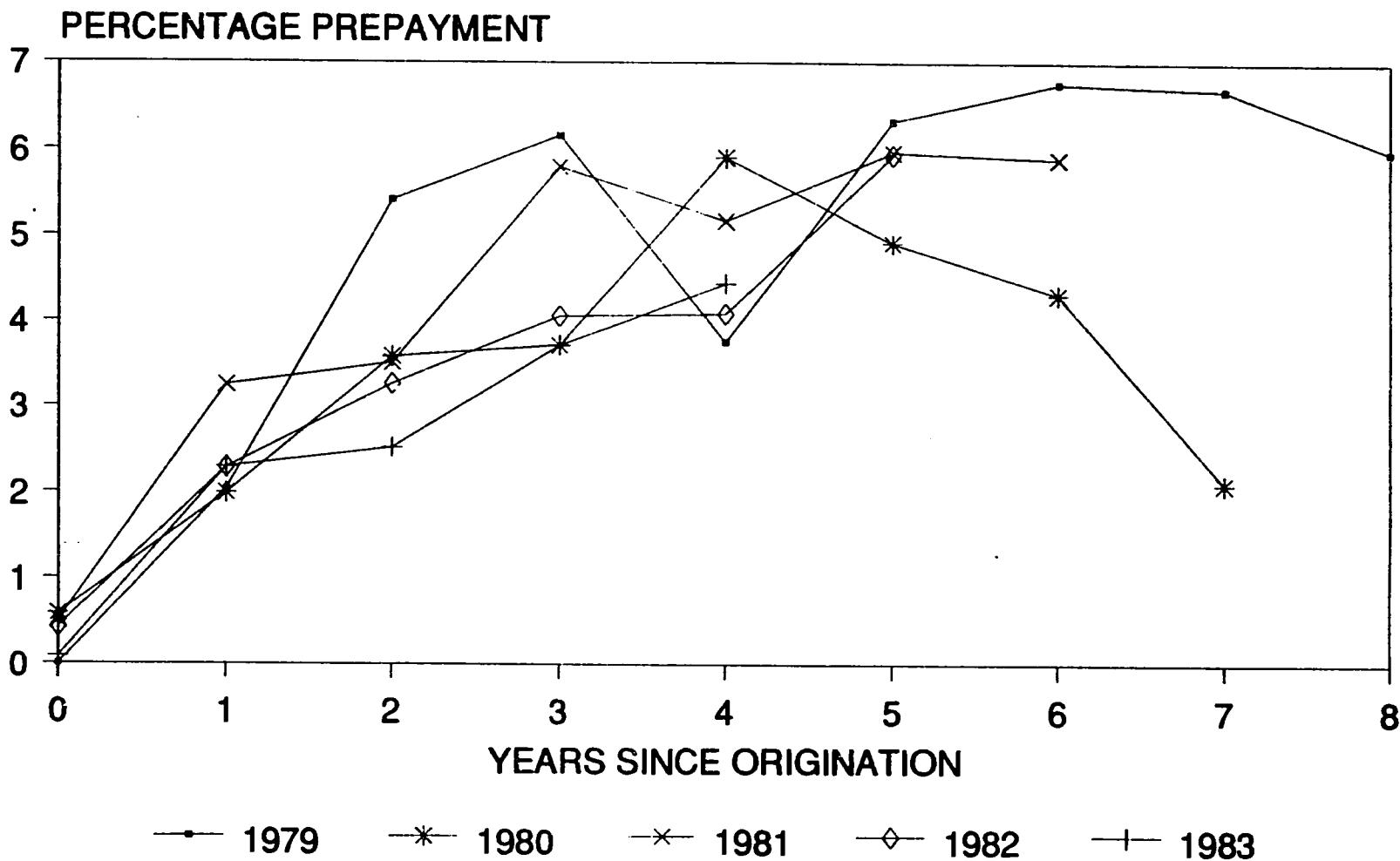
Data on the borrower and the mortgage at the time of origination were obtained from the sanctions and disbursement file. The date of prepayment and reason for prepayment were added to the record for prepaid mortgages. Different sampling weights were used in drawing the two samples, and the tabulations presented below weight observations by the inverse of the applicable sampling rate, after correction for dropped observations.

### Patterns of prepayment

The pattern of prepayment is shown in Figure 2 and in Table 3. In these, prepayments are calculated as a percent of the remaining mortgages outstanding each year. The general pattern--which is clearest in the last column of Table 3 where unweighted averages across the years for which we have data are given--is for there to be negligible prepayments in the year the loan is originated, followed by an increasing rate of prepayments over years 1 to 3, with the rate of prepayment then leveling off.

Figure 2

# PERCENTAGE OF PREPAYMENTS BY YEAR OF ORIGINATION



**Table 3**  
**Percentage of Outstanding Mortgages Prepaying**  
**Each Year, by Years of Origination <sup>a</sup>**

Years Since Origination	Origination Year					Unweighted Average
	1979	1980	1981	1982	1983	
0	—	.58	.50	.43	.08	.32
1	2.04	1.99	3.25	2.28	2.29	2.37
2	5.41	3.58	3.50	3.26	2.51	3.65
3	6.16	3.71	5.80	4.05	3.71	4.68
4	3.75	5.91	5.17	4.08	4.44	4.67
5	6.33	4.91	5.97	5.94		5.78
6	6.76	4.31	5.88			5.65
7	6.69	2.10				4.39
8	5.98					5.98
Number of Loans Originated <sup>b</sup>	892	3,129	5,063	8,483	12,880	
Cumulative percent Prepaid	33.8	23.3	24.2	15.9	10.3	

a. Last year figures are annualized rates based on first six months.

b. Based on weighted observations in the final analysis file.

c. Actual experience to July 1, 1987.

Source: Urban Institute tabulations of HDFC data.

One distinct variation of this pattern is visible: the prepayment rates of loans originated in 1979 are consistently higher than those for later years. It may be that the first customers for loans from HDFC were in fact quite different from those of later years.

The contrast between the prepayment patterns for HDFC loans and those for U.S. mortgage loans insured by FHA (Figure 1) is dramatic. While prepayment rates increase in the first three years following origination in both cases, the rates accelerate to substantially higher rates in the U.S. before leveling off. Prepayment rates in the U.S.--driven by high mobility rates and shifting interest rates--are over 8 percent by the fourth year, compared with an average rate of about half that for the HDFC mortgages. The U.S. pattern is for a gradual but clear decline thereafter, while the parallel pattern for HDFC is not yet clear.

Lastly, note the cumulative percentage of loans prepaid for each year of origination (last row on Table 3). Even after allowing for the loans originated in 1979 prepaying at an unusually high rate, it appears that a quarter of all loans originated will pay off in the first eight years after origination. Hence, early prepayments can be expected to be a major source of funds for additional lending by HDFC and probably other housing finance institutions in India.

## Predicting prepayment over the life of the mortgage

The first model for whether or not the loan was prepaid over the sample period is estimated using the probit technique, and the results are presented in Table 4.[6] These results point to the importance of the age of the borrower in determining the likelihood of prepayment. As the age of the borrower increases, the likelihood of prepayment decreases, but at a decreasing rate to age 55, after which the likelihood increases slightly. Neither the employment status nor sex of the borrower had a significant impact on prepayment. Borrowers who are shareholders tend to repay at higher rates than non shareholders, other things equal. The binary variables included for the city in which the loan was originated indicate that those issued in Bombay (the excluded location in the specification) have a greater likelihood of prepayment than loans originated elsewhere.[7] All of the city variables are highly statistically significant.

Among the variables that characterize the terms of the loan--the property cost, the term of the loan, and the rate of interest--have coefficients that are statistically different from

-----  
6. The dependent variable takes on the value of one if the loan was prepaid over the observed period, and zero if the loan was not prepaid. The ordinary least squares linear probability model (OLS) is inappropriate when the dependent variable is binary because the prediction probabilities generated by the linear model probability model can be greater than one or less than zero and the disturbance term is no longer homoskedastic which introduces bias to the estimated standard errors.

7. Because of the very small number of observations for loans originated in Jaipur and Lucknow (5 in total), these two locations were dropped from the analysis.

**Table 4**  
**Results for Model of Prepayment Ever Occurring<sup>a</sup>**  
**(dependent variable = 1, if mortgage is prepaid**  
**over the observation period)**

<u>Independent Variables</u>	<u>Coefficient</u>
Constant	8.00*
Loan term	-.002*
Rate of interest	-.072*
Payment/monthly income	-.232
Borrower is shareholder, yes=1	.194*
Property cost (lakhs)	-.001*
Loan to value ratio	-.212
Age of borrower	-.086*
Age squared	.001*
Borrower is female, yes=1	.108
Borrower is self-employed, yes=1	-.036
Location of originating bank <sup>b</sup>	
Ahmedabad	-.691*
Bangalore	-.540*
Calcutta	-1.18*
Cochin	-.756*
Hyderabad	-.769*
Madras	-.687*
New Delhi	-.609*
Parel	-.446*
Pune	-.702*
Pearson Goodness of Fit, p	.423
Chi square	6455
d.f.	6441

**Notes:**

- \* Significant at the .05 level or higher
- + Significant between the .05 and .10 levels
- a. Model estimated using probit analysis
- b. Relative to Bombay (headquarters)

zero. Neither the ratio of monthly payments to income nor the loan to value ratio seem to have a significant impact on the likelihood of prepayment. Longer terms and higher property costs appear to be linked to lower probabilities of prepayment, other things equal, while higher mortgage interest rates appear to be associated with lower probabilities of prepayment. (Most of the results are discussed somewhat further later in the context of the more refined model.)

The prediction success of the model is not great. Of the 1,919 prepayments in the sample, only 364, or 19 percent, were correctly predicted to be prepayments; and of the 4,552 non prepayments, 94 percent were correctly predicted not be prepayments for an overall success rate of 71 percent. One caveat is that this probit was not run with weighted data which could affect prediction outcomes. The large share of prepayers which are misclassified reveals just how difficult it is to model prepayment.[8]

These results must be regarded as tentative because of two sources of model misspecification. First, this model assumes that the impact of the explanatory variables does not depend on the term of the loan, i.e., loans with different terms are analyzed together, without explicitly permitting interactions between the term of the loan and other variables. Second, the

-----  
8. We also did OLS estimates of this model with one half of the sample and predicted the classification of the other half. The predictive power of this model was about the same as that reported in the text.

model ignores the fact that all loans that were not prepaid are censored at a duration that equals the difference between the survey and loan origination dates.

To examine the role that term plays in affecting prepayment, we estimated seven separate models of prepayment for loans of 180, 144, 120 months, and other terms. We used the ordinary least squares technique (OLS) because the one-to-one comparison of the probit and OLS results of prepayment ever occurring (the model in Table 4) generated almost identical estimated coefficients and significance levels. We found no evidence that the coefficients of the explanatory variables were different in the equations for loans of different terms. The F statistic had a value that was less than 1.0 and was not significant at even a 10 percent level. Therefore, in what follows, we will continue to include loan term as an explanatory variable, but we will not estimate separate equations for loans of different terms.

The second problem associated with the model presented in Table 4 is that we do not have complete histories for many of the loans in the sample. For these loans, we know that prepayment has not occurred within the time between the loan origination date and the survey date, but we do not know anything about prepayment beyond that point. This problem is called "right censorship" and it does not affect all loans in the same way: it is a greater problem for more recent loans.

To the extent that some of the borrower or loan attributes have changed systematically over time, spurious relationships can emerge from ignoring the fact that more recent loans have had much less opportunity to be prepaid than loans originated earlier. Specifically, because HDFC's schedule of interest rates has increased over time, this could lead to an estimated negative relationship between interest rate on the loan and likelihood of prepayment even when there is no such underlying relationship. Given that the loans with the lower interest rates have had greater opportunity to prepay, whereas the more recent loans with higher interest rates would have a smaller opportunity, a negative relationship could arise due to the greater censorship of the more recent observations. Any systematic change over time in an attribute variable could produce inconsistent estimates related to this censorship in the model presented in Table 4.

An additional reason for moving beyond the model in Table 4 is that no time-varying independent variables can be included in this specification. It was argued earlier that prepayments on loans should rise when the prevailing interest rate falls, other things equal, because of the heightened incentives to refinance the loan.

#### **Estimated Time Path of Loan Prepayment**

Table 5 gives the estimated "survival" proportions of all

**Table 5**  
**Cumulative Survival Proportions by Mortgage Term**

<u>Interval</u> (years)	<u>Loan Term</u>					
	<u>All Terms</u>		<u>15 Years</u>		<u>10 Years</u>	
	proportion surviving	std. dev.	proportion surviving	std. dev.	proportion surviving	std. dev.
0 to 1	.988	.001	.988	.002	.988	.004
1 to 2	.959	.002	.963	.003	.955	.007
2 to 3	.927	.003	.936	.004	.916	.010
3 to 4	.890	.005	.903	.005	.871	.013
4 to 5	.848	.006	.862	.008	.830	.018
	<u>8 Years</u>		<u>7 Years</u>		<u>5 Years</u>	
0 to 1	.983	.009	.984	.011	.985	.010
1 to 2	.941	.017	.962	.017	.909	.023
2 to 3	.893	.024	.915	.026	.845	.031
3 to 4	.868	.028	.854	.038	.759	.042
4 to 5	.825	.038	.806	.050	.644	.064

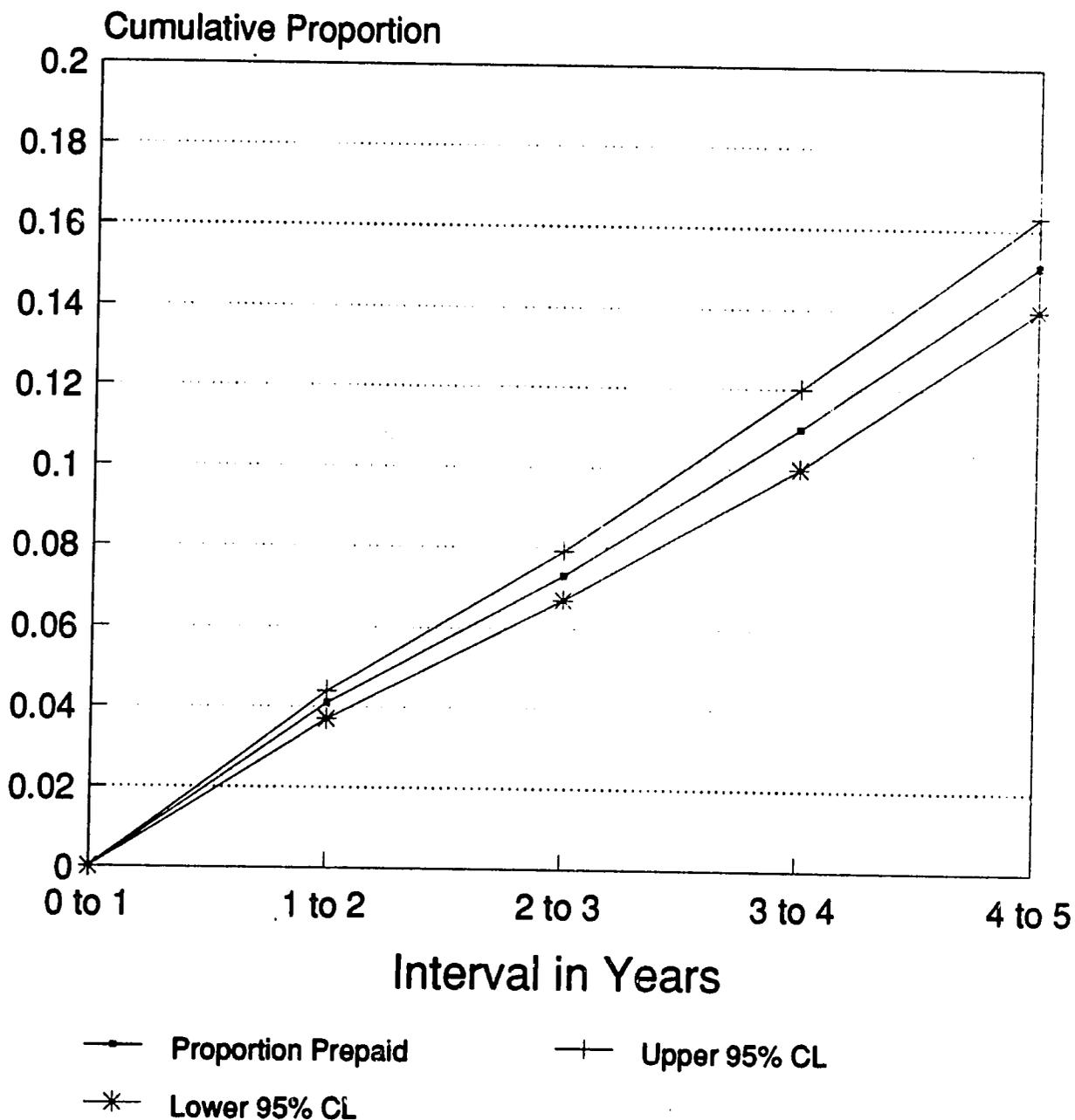
loans in the sample and for loans of six different terms separately. The sample weights have been used to calculate these life table survival proportions. There are six pairs of columns of results. Within each pair, the first column gives the proportion of loans that are estimated to remain outstanding at the end of the interval (i.e., that have not been prepaid), and the second gives the standard error of the estimated proportion. While there seems to be a smooth prepayment path across durations, the prepayment rates are higher for shorter terms, most markedly for loans of five year terms.

While loans of lower terms appear to be prepaid at faster rates in absolute terms, loans of higher terms appear to be prepaid at higher rates relative to the term of the loan. For example, one-fifth the way into a five year loan, an estimated 1.5 percent of the loans are prepaid, while one-fifth of the way into a 15 year loan, an estimated 14 percent of the loans are prepaid.

Figure 3 gives the 95 percent confidence interval around the estimated cumulative proportion prepaid for all loans included in the sample. Because the overwhelming majority of loans have 15 year terms, this graph resembles very closely the time path of prepayment within the first five years of these loans. While the life table shows the estimated prepayment schedule for loans in the sample, the estimates presented in the next section provide insight into how borrower and loan

Figure 3

# Cumulative Proportion Prepaid Loans of All Terms



95 % Confidence Intervals

attributes and the prevailing company deposit rate affect the likelihood of prepayment at different durations.

### **Duration-Specific Probits of Prepayment**

We next examine the probability of prepayment within one year of a loan being issued, and the probability of prepayment between the first and second years after the loan has been issued (conditional on the fact that the loan was not prepaid in the first year), etc., and how that prepayment probability relates to loan and borrower attributes. The prepayment schedule was presented in Table 3 by examining the prepayment path for loans issued in different years. The prepayment path after loan issue should be independent of the date of issue, holding other factors equal such as borrower and loan attributes and macroeconomic conditions.

We have estimated a series of probit models for prepayment probabilities in successive years after loan issue. We have pooled the loans across different terms and included loan term as a shift variable. We have estimated a probit for prepayment with the first year of the issue date, which includes all the loans for which a complete year of data are available. We then estimated four successive probits, the first being for prepayment with the period from the first to second year following issuance, up to the last being for prepayment within the fourth to fifth years after issuance. While we had

durations of up to nine years, the sample sizes were not large enough to do separate analyses for these higher durations.[9]

Only one time-varying explanatory variable was included for the five duration-specific probits: the company deposit rate. All loans issued in a certain year have the same observed company deposit rate at each duration which is different from the company deposit rate that prevails at that duration for loans issued in another year. Because no other time-varying explanatory variables are included, the coefficient estimated on the company deposit rate must be interpreted with caution.[10]

The estimated coefficients for the five probit models are given in Table 6. Even with a very high number of iterations, no convergence was obtained for the probit models for the first two durations which is likely due to the very small percentage of cases that were prepaid in these periods. Therefore, we present OLS results for these two durations. The OLS linear probability results for other durations are very similar to the probit results in terms of significance levels and magnitude of the

-----  
9. We have been conservative in setting up the criteria for including loans in the duration-specific analysis by only including loans for which we observed the whole length of the duration. This reduced the sample sizes, especially for the higher durations. We also estimated the models with the loans that were not observed over the whole interval. Our more conservative approach led to fewer significant coefficients, especially at higher durations; but the pattern of findings did not differ.

10. We also estimated models in which the appreciate rate for gold or the interest rate on commercial bank deposits were included instead of the company deposit rate. Results were similar to those for the company rate.

**Table 6**  
**Results of Duration-Specific Models of Loan Prepayment**

<u>Independent Variables</u>	<u>Year of Prepayment</u>				
	<u>0 to 1<sup>a</sup></u>	<u>1 to 2<sup>a</sup></u>	<u>2 to 3<sup>b</sup></u>	<u>3 to 4<sup>b</sup></u>	<u>4 to 5<sup>b</sup></u>
Constant	2.74*	2.96*	18.2*	17.2*	5.51
Company deposit note	-.213*	-.227*	-.965*	-.802*	.067
Loan term	-.001*	-.001*	-.004*	-.004*	-.004*
Rate of interest	.024*	.020*	.061*	.006	.019
Payment/monthly income	-.042	-.014	-.083	-.532	.164
Borrower is shareholder, l=yes	-.008	-.006	.097	.0001	-.050
Property cost (lakhs)	-.019*	-.017*	-.111*	-.081+	-.106
Loan to value ratio	-.091*	-.061*	-.318+	-.026	-.002
Age of borrower	-.011*	-.011*	-.010*	-.096*	-.055
Age squared	.0001*	.0001*	.001*	.001*	.001
Borrower is female, l=yes	.016	.017	.089	.065	.153
Borrower is self-employed l=yes	-.017	-.013	-.106	-.114	-.089
Location of originating bank:					
Ahmedabad	-.090*	-.079*	-.572*	-.450*	-.409+
Bangalore	-.099*	-.075*	-.417*	-.331*	-.268+
Calcutta	-.160*	-.146*	-1.29*	-1.34+	-1.21*
Cochin	-.094*	-.094*	-.647*	-.622*	-.365
Hyderabad	-.107*	-.088*	-.755*	-.619*	-.590+
Madras	-.110*	-.099*	-.774*	-.702*	-.918*
New Delhi	-.079*	-.062*	-.473*	-.532*	-.557*
Parel	-.064*	-.060*	-.351*	-.106	.200+
Pune	-.094*	-.081*	-.638*	-.680*	-.775*
R <sup>2</sup>	.070	.061			
F	24.1	19.0			
d.f.	6423	5858	3964	2337	1203
Pearson Goodness of Fit Chi square			.51	.45	.43
			3961	2345	1211

**Notes:**

- a. Estimated with ordinary least squares.
- b. Estimated with probit technique.
- c. Relative to Bombay (headquarters).
- \* Significant at .05 level or higher.
- + Significant .10 to .05 level.

estimated coefficients.

The probit coefficients do not have the same interpretation as in the OLS linear probability model; in fact, the effect of a change in an explanatory variable on the probability of prepayment within the specified duration depends on the probability of prepayment itself in the probit specification. In general, the impact is greatest in absolute terms at a prepayment probability of .5 and decreases as one moves toward probabilities of zero and one. For prepayment probabilities in the range of .3 to .7, one can multiply the estimated probit coefficient by .4 to approximate the coefficient in the linear probability model which has the interpretation of the change in the probability due to a one unit change in the explanatory variable. For prepayment probabilities outside that range, the probit coefficients should be scaled down further. The proportion of cases prepaying rise with duration for these five duration-specific probits, implying that the estimated coefficients in the low durations would be scaled back further than for higher durations when converted to a linear probability approximate.

The results indicate that the explanatory variables have the strongest effects (in terms of magnitude and significance levels) in the earliest years after the loan is issued. While all coefficients but those on sex of borrower, employment status, shareholder status, and ratio of payment to income are

significant at the 5 percent level or greater in the equations for prepayment within the first two years, only five coefficients--loan term, and four locational variables--are significant at this level in the equations for prepayment within the fourth and fifth years. In general, one's ability to predict prepayment in a particular year declines with the seasoning of the mortgage.

For durations of one, two, and three years, the results indicate that as age of the borrower increases, the probability of prepayment within the specified period decreases, but at a decreasing rate, echoing the results in Table 4. Neither age variable was significant for durations for four or five. For the early years, a 10 year increase in the borrower's age (from 40 to 50) has a moderate impact on prepayment: it decreases the probability of prepayment from by .02 during the year 1 to 2 period after origination.

The impact of the city of origination on prepayment of different durations, when significant, is consistent across durations and confirms what emerged from the model estimated for prepayment ever occurring: the prepayment probability in Bombay seems consistently higher, other things equal, than in the other cities. This "Bombay effect" may well be due to borrowers there being more financially astute and therefore more aggressive in seeking out cheaper loans from "friendly sources" to replace their HDFC mortgages.

The higher the property cost and the higher the ratio of the loan to property value, the lower are the probabilities of prepayment at least for durations of one, two, and three. Both results reflect the greater difficulty in repaying a larger loan from savings or in borrowing the relatively larger amount for prepayment. For the two higher durations, neither coefficient is significant. In the model of prepayment ever occurring, only the property cost variable was significantly different from zero.

There are two other variables that characterize the terms under which the loan was issued, the loan term and the rate of interest. The coefficient estimated on the loan term has the same negative sign as in the model of prepayment ever occurring, while the coefficient on the rate of interest, when it is significant, is positive, while it was negative in the earlier model. Loan term seems to be consistently linked to lower probabilities of prepayment, even after controlling for duration. This result is very robust with respect to model specification. The quantitative impact, however, is modest: a five year increase in term decreases prepayment probabilities in the early years by about .005. The mortgage rate of interest, in contrast, has a positive effect on prepayment for durations of one, two, and three and appears to have no impact thereafter. This result is in keeping with our expectations: borrowers with higher interest rate mortgages are anxious to pay them off. The

negative relationship estimated between the mortgage rate of interest and the likelihood of prepayment in the model presented in Table 4 could be due to the systematic censorship of more recent loans from the sample.

Finally, the results indicate that in the years when the company deposit rate is higher (a change in the rate structure), other things equal, the likelihood of prepayment falls for loans of duration one, two, three, and four. The impact is often very large. For the first two durations a decrease of 100 basis points raises probabilities of prepayment by .2. The general result jibes with our expectation that refinancing incentives increases as prevailing interest rates on other assets falls. However, since the interest rate is likely to move along with other macroeconomic variables, the coefficients estimated on the market interest rate are likely embodying the effects of changes in other variables as well.[11]

While it seems appropriate to include a variable for the market interest rate of widely held assets in the analysis of the timing of prepayment, future modeling and estimation should include other time-varying macroeconomic and locational variables as well. Our interpretation of the present result is

-----  
11. Additionally, when we exclude the company deposit rate from these models, the signs and significance patterns resemble those from the probit of prepayment ever occurring. The estimated coefficient on the mortgage rate of interest becomes negative and significant. The coefficient on loan value loses its significance while the coefficient on the borrower being a shareholder becomes significant.

that macroeconomic conditions as embodied in interest rates on company deposits have a powerful effect on the likelihood of prepayment.

### Reasons for Prepayment

We now turn to an examination of the sources of funds for repayment stated by those borrowers who have prepaid their mortgages. The following tabulation gives the percentage distribution of the number of units whose prepayment was financed from each of these sources (where only the primary source is indicated, if more than one was reported).

source of funds -----	percentage of total prepaying -----
sale of unit	25.0
savings	39.5
loan received from employer	13.1
"friendly loan"	5.3
payments from provident funds	14.4
lump sum from insurance policies	.5
other	2.0

In light of the very low residential mobility rates in India, a finding that one-fourth of all those prepaying are doing so by selling their home is rather surprising. This seems to lend some credence to the idea that some HDFC borrowers buy a house in a developing suburban area and sell it a few years later with substantial capital gains, and buy another unit possibly again

with HDFC financing.[12] In a sense, then, HDFC is encouraging mobility and "trading up" to better housing.

The other somewhat surprising entry in this tabulation is the large share (13 percent) of those prepayers obtaining loans from their employers. It appears that at least some of these borrowers are using loans from HDFC to move up the time at which they can acquire a home compared to waiting until they might obtain a low rate loan from their employer. Still, by far the largest share of prepayments--40 percent--are financed by household savings in one form or another.

As discussed in the analysis plan, we have estimated two regression models for the likelihood that a particular source of financing would be used to prepay a mortgage. The sources of finance of particular interest are (a) the sale of the unit and (b) savings combined with payments from provident funds and insurance companies, and employer loans. OLS linear probability models have been used in this exploratory analysis.

The results for the models estimated for the likelihood that the mortgage will have prepaid from either home sale or from receipt of funds provident funds, other savings or employer loans are given in Table 8. The observations are restricted to mortgages that have been prepaid. The coefficients in this table are interpreted similarly to the earlier ones, as giving the effect at the margin on the probability that the loan will

-----  
12. This idea was suggested by Dr. M. Metha.

**Table 7**  
**Results for Regression of Likelihood of**  
**Home Sale and of Savings/Provident Funds Being the**  
**Source of Funds for Prepayment**

<u>Independent Variable</u>	<u>Home Sale</u>	<u>Savings &amp; Provident Funds</u>
Constant	.537*	.399*
Loan term (months)	.0007*	.0009*
Rate of interest	.016*	-.029
Mortgage payment/income	-.183	.182
Shareholder (1=yes)	.008	.001
Cost of home (lakhs)	-.037*	.026
Loan to value ratio	-.092	-.049
Age of borrower	-.022*	.031*
Age squared	.0002	-.0003*
Borrower is woman, 1=yes	-.079*	.082*
Borrower is self-employed, 1=yes	.130*	-.104*
Location of HDFC branch <sup>a</sup>		
Ahmedabad	.161*	-.142*
Bangalore	-.117*	.076*
Calcutta	-.253*	.244*
Cochin	-.197*	.195*
Hyderabad	.104	.030
Lucknow	-.327	-.001
Madras	-.221*	.265*
New Delhi	-.274*	.313*
Parel	-.207*	.111*
Pune	-.157*	.173*
R <sup>2</sup>	.110	.100
R <sup>2</sup> (Adj.)	.101	.090
F	11.79	10.50
Sign. of F	.0001	.0001

Notes:

\* Statistically significant at the .05 level or higher.

a. Headquarters (Bombay) is omitted location.

be prepaid from a particular source, after controlling for other factors.

The results for the two models are almost completely divergent. Without exception, the signs of those coefficients which are statistically different from zero (as indicated by an asterisk in the table) are the opposite from each other in the two models. Moreover, the coefficients are generally of a similar order of magnitude in absolute value. The strength and consistency of the differences between the two model is somewhat surprising given that there is a third source of loan repayments --"friendly loans" and a small miscellaneous category--which accounts for about 7 percent of all prepayments.

In any event, in reviewing the results we concentrate on the results for the model in which the dependent variable has the value of 1 if selling the home was the source of funds used to pay off the mortgage (see the first column in Table 8). The principal findings are:

-- The longer the loan term the more likely prepayment from these sources. The probability increases by .04 when the term increases from 10 to 15 years.

-- The higher the cost of the home, the lower the probability of prepayment from home sale: a one lakh increase in cost lowers the probability by .04.

-- Older borrowers are distinctly less likely to prepay from sale of the unit and more likely to prepay from provident funds. In interpreting the effect of age, one must consider both age terms in the model, which together capture a non linear relationship between age and prepayment source. (The age squared term is significant at the .10 level.) A borrower age 40 at origination is .56 less likely to be prepaying from a home sale compared to a

very young borrower, while a borrower age 55 at origination has a .61 lower probability.

-- Women borrowers are about .08 less likely to prepay from home sales than men.

-- Self employed borrowers (who would not have access to employer loans or provident fund payments) have a .13 higher likelihood of prepaying from home sale, but a lower likelihood to prepay from savings or provident fund payments. This suggests that the self employed may be using their home equity for businesses purposes in a pinch and putting other forms of savings into their business rather than into their home.

-- Where the loan is originated definitely makes a difference in source of prepayment funds. For home sales, with the exception of loans originated in Ahmedabad, loans originated outside of the Bombay headquarters are generally significantly less likely to use home sale as a source of funds.

In short, borrower attributes of age, type of employment, and ethnicity (as indicated by location of loan origination) make a distinct difference in the source of funds used to prepay loans, as do the term of the loan, rate of interest and the cost of the unit purchased.

## Conclusions

This analysis clearly demonstrates that mortgage prepayments are a potent factor in housing lending at HDFC and presumably at other housing finance institutions in India. For loans originated by HDFC in 1979, some 34 percent had been prepaid by mid-1987. Obviously, prepayments of this magnitude have important impacts on effective interest rates, on the match between the terms (or "durations") of mortgages and liabilities raised to finance the loans, and on the ability of the

institution to make loans in the future.

The pattern of prepayments was found to be significantly related to several borrower and mortgage attributes, but such relationships were established less frequently than we had anticipated. Prominent among the factors significantly related to the likelihood that a mortgage would be prepaid are: the length of the loan term, the time since the loan was originated, the age of the borrower, and the location within India where the loan was originated (a proxy for ethnicity and possibly market effects)--prepayments of loans originated in Bombay were exceptionally high. At the same time, the results show a strong negative relationship between the likelihood of prepayment and changes in the interest paid on widely held assets--a relationship which probably captures the effects of broader macroeconomic changes on prepayment as well as the pure interest rate effect. All of the relationships just noted are strongest for the first couple of years following loan origination. In later years only the results for the term of the loan and the location of origination continue to be significant.

Finally, those repaying are predominantly relying on some type of savings or the receipt of payments from provident or insurance funds. Surprisingly, in light of the low residential mobility rates in the country, about 25 percent of those repaying are doing so with the proceeds from selling the home.

As far as we know, this is the first analysis done of

mortgage prepayments in a developing country. As noted, the patterns uncovered have immediate implications for efficient management of housing finance institutions. Replicating this analysis for other countries to determine if the experience of HDFC is repeated elsewhere could pay large dividends in terms of the kind of advice which could be given to housing finance institutions in many countries.

## References

- Boleat, M. (1985). National Housing Finance Systems. London: Croom Helm.
- Buckley, R., J. Khadduri, and R. Struyk (1985). The Housing Development Finance Corporation of India: Evaluation of the Housing Guaranty Loan. Washington, D.C.: The Urban Institute, Report to the Office of Housing and Urban Programs, USAID.
- Chakravarty, S. (1985). Report of the Committee to Review the Working of the Monetary System. Bombay: Reserve Bank of India.
- Curley, A. J. and J. M. Guttentag (1977). "Value and Yield Risk on Outstanding Insured Residential Mortgages," Journal of Finance, vol. 32, no.2, pp. 403-16.
- Dunn, K.B. and J.J. McConnell (1981). "A Comparison of Alternative Models for Price GNMA Mortgage-Backed Securities," Journal of Finance, vol. 36, no.2, pp. 471-84.
- Jacob, D.P., G. Lord, and J. Tilley (1987). "Price, Duration and Convexity of Mortgage-Backed Securities," in F. J. Fabozzi (ed.), Mortgage-Backed Securities. Chicago: Probus Publishing.
- Kozel, V. (1987). Income, Consumption and Savings: Behavior in the Cote d' Ivorie. Cambridge: MIT, Department of Urban Studies and Planning, PhD dissertation.
- MacDonald, G. (1986). "A Household/Lender Based Analysis of Prepayment Penalties." Charlottesville, Virginia: Department of Economics, University of Virginia, processed.
- Mehta, M. and D. Mehta (1987). Metropolitan Housing Market: Housing Supplies, Demand, and Residential Behavior in Ahmedabad. Ahmedabad: School of Planning, Centre for Environmental Planning and Technology.
- Roberts, R. B. (1987). "The Consequences of the Pure Aging Effect on the Yields of Mortgage-Backed Securities," in F. J. Fabozzi (ed.), Mortgage-Backed Securities. Chicago: Probus Publishing.
- Roll, R. (1987). "Collateralized Mortgage Obligations: Characteristics, History, Analysis," in F. J. Fabozzi (ed.), Mortgage-Backed Securities. Chicago: Probus Publishing.
- Schenk, H. (1986). "Residential Immobility in Urban India," Geographical Review, vol. 76, no.2, pp. 184-94.

**Annex A**  
**DATA CLEANING PROCEDURES**

While the data coded and keypunched from the HDFC files generally arrived in good condition, we performed a number of checks to insure that extreme values did not enter the data set and thereby cause the analysis to produce invalid results.

An early perusal of the data showed that values for incomes, property values and loan amounts had not been consistently coded, in the sense that some values were clearly in lacs and others in whole numbers. The values were recoded to a consistent basis.

Following these observations were dropped if they met any of the tests listed below:

- a. improper codes for branch designation, sex of household head, or age of household head were clearly wrong, e.g., household head of age 5
- b. loan term recorded as less than 12 months
- c. income under Rs.20 per month
- d. monthly payments in excess of 50 percent of income
- e. property cost or loan amount of zero or very small values
- f. date of loan before 1978
- g. interest rate of less than 10 or greater than 15 percent
- h. loan origination date after prepayment date for prepaid loans
- i. ratios of loan value to property cost of less than 10 or greater than 90 percent
- j. loan amounts greater than 60 times monthly income
- k. the employment code was miss coded.

Three observations were also dropped which contained mostly missing data.

A summary of the reasons for dropping observations is presented in Table A.1 on the next page. Note that listed is the first reason an observation was rejected; had we continued to analyze the observation it might have had multiple problems.

**Table A.1**  
**Number of Observations Dropped by Reason**

reason	number dropped
age too low	2
sex out of range	3
branch out of range	8
date of loan	5
employment	14
income too small	6
loan amount of zero	1
loan term too short	8
(loan/property value) high	71
(loan/property value) low	40
loan too high for income	328
monthly payment	42
(monthly payment/income) high	116
prepayment date	9
property cost	1
rate of interest	8
most data missing	3

Memorandum item:

	outstanding		prepaid	
	no.	%	no.	%
original observations	5115	70	2024	30
revised	4552	72	1919	28