

## **An actuarial approach to pricing Mortgage Insurance considering simultaneously mortgage default and prepayment**

Jesús Alan Elizondo Flores  
Comisión Nacional Bancaria y de Valores  
[aelizondo@cnbv.gob.mx](mailto:aelizondo@cnbv.gob.mx)

Valeria Álvarez Navarro  
Sociedad Hipotecaria Federal  
[valvarez@shf.gob.mx](mailto:valvarez@shf.gob.mx)

Israel Sergio Valladares Cedillo  
Sociedad Hipotecaria Federal  
[ivalladares@shf.gob.mx](mailto:ivalladares@shf.gob.mx)

ICA 2010  
49 TRACK E - Financial Risks (AFIR)

### **Abstract**

Banks are exposed to potential losses originated from credit and prepayment risk embedded on mortgages. Both prepayment and credit risk are of opposing nature, not independent among them and generally lead to the termination of the mortgage contract.

Banks commonly take into consideration these risks for pricing, reserving and capitalization purposes. They may also acquire mortgage insurance which generates a way to alleviate and diversify credit risk on their portfolios by ceding part of the credit risk to a third party. However mortgage insurance does not provide coverage against prepayment risk and usually benefits the provider of protection as expected credit related losses diminish.

This work provides an actuarial framework to price and reserve mortgage insurance considering credit and prepayment risk as competing risks. By establishing the appropriate analogy, multiple decrement standard actuarial terminology is applied to mortgage insurance.

Appropriate formulas for pricing and reserving mortgage insurance are developed which consider simultaneously credit and prepayment risk. These formulas are subsequently compared to corresponding formulas that consider solely credit risk and hence the marginal effect of prepayment risk is derived on the final product price.

The paper further elaborates on the estimation of parameters needed to price mortgage insurance by using 2003 to 2008 default and prepayment risk experience from the Mexican mortgage sector.

Statistical relationships between a set of explanatory variables and default and prepayment experience are established through the use of separate standard logistic regressions. These relationships are further estimated simultaneously by using

multinomial logistic regression which is shown to refine credit and prepayment risk parameters while showing the competing nature of both risks.

Finally, mortgage insurance prices are calculated for a set of credit and prepayment parameters, mortgage and mortgage insurance product characteristics and differences are analytically explained by using the theoretical framework established in the document.

**Key Words:** default, prepayment, competing risks, mortgage insurance, pricing.

## 1. Introduction

Managing the risk of mortgage portfolios requires analytical tools that are able to consider a variety of factors that affect the cash flows of the loans. With adequate analytical tools, banks can establish appropriate reserves against credit losses, price their products to reflect the underlying risks and compare the risks and returns associated with portfolio segments.

Mortgage Insurance (MI) represents an important tool to manage the risk associated with the portfolio of mortgages and allows the bank to diversify the inherent credit risk. The price of the MI product is a key factor for banks to evaluate the cost and benefits of the product as the adoption of the insurance affects the risk and reward balance of the credit portfolio. Simultaneously, the price of the MI product represents the unique source of income for the mortgage insurance company.

The pricing of such contracts is a challenging task even when data is available. The task is generally more challenging in the case of emerging markets, where, in the case of borrower default, the process of repossession of loan collateral may last a variable number of years and where data on payment behavior is either unavailable or of poor quality.

It is a well documented fact that borrowers face two mutually exclusive alternatives when paying a mortgage, either to prepay the loan or to default on it. From a lender and insurer point of view, these are “competing risks” in the sense that they are mutually exclusive and the realization of one risk precludes the other.

For the MI provider, prepayment means cessation of cash flows in the form of any mortgage insurance premium and thereafter eliminates any possibility of default. Conversely, defaulting implies incurring in costs associated with the MI contract. Thus, one cannot calculate accurately the economic value of the cost of the insurance without considering default and prepayment.

There is a vast amount of literature related to mortgage pricing. The contingent claims model developed by Merton (1973) provides a motivation for borrower behavior using options theory. At the beginning most of the studies on default and prepayment used this approach to price mortgages; however they focused on default and prepayment as individual risks. For instance, Cunningham and Hendershott (1984) applied Black and Scholes (1973) option pricing model to price default risk considering default as a put option sold by the FHA and purchased by the homebuyer for the default protection of the lender. Schwartz and Torous (1989) and Quigley and Van Order (1990) provided empirical estimates of option-based prepayment models consistently with the contingent claims model. Kau and Kennan (1995) also provided a theoretical framework to price mortgages as derivative assets in a stochastic economic environment. They described prepayment and default as an American call option and a European compound put option respectively.

Kau, Keenan and Muller III, emphasized the importance of the jointness of prepayment and default options in the calculation of insurance prices for mortgages. The hypothesis presented is that a borrower that decides to default today is giving up his option to default in the future, but also is giving up the option to prepay the mortgage (Deng, Quigley, Van Order, 2000).

Deng, Quigley, and Van Order (1996) presented a joint model for default and prepayment in a proportional hazard framework using an option based approach. They also emphasized the importance of non economic trigger events in affecting default and prepayment behavior and

their relevance in the exercise of options to default or prepay on mortgages in the fully rational way predicted by finance theory.

Since the Cox Proportional Hazard (CPH) model was introduced by Green and Shoven (1986) to analyze mortgage termination by refinance, there has been a recurrent technique to model mortgage termination. Deng, Quigley and Van Order (2000) modeled the competing risks of mortgage termination empirically in a proportional hazard framework which allows correlated competing risks. However they concluded that unobserved borrower heterogeneity is important for predicting borrower behavior and CPH does not consider the heterogeneity among borrowers.

Recent research on mortgage borrower behavior has proposed several models for the competing risks of mortgage termination by prepayment or default. Deng, Quigley and Van Order (2000); Deng, Calhoun (2002) present default and prepayment as discrete choices where a Multinomial Logistic Model (MNL) would be an appropriate technique and an alternative to the Cox Proportional Hazard Model (CPH). The MNL avoids the proportional impact of the covariates in the exponential component implicit in CPH and allows direct competition among the choices: An increase in one termination probability must be offset by a decline in probability for one or more of the alternatives. However, the MNL cannot allow correlations among the termination risks, and one must assume independence from the alternatives.

In the present work, the theoretical and empirical analysis to estimate the competing risks of mortgage prepayment and default is based upon a discrete-time multinomial logistic model of mortgage termination. The empirical analysis is based upon a set of mortgage lending data from a housing government bank which concentrates information of most of the financial intermediaries that operated mortgage credit in the period 2003 - 2008. The dataset allows to study mortgage borrower prepayment and default behavior at loan level. The availability of loan-level data allowed to account for the dynamic of mortgage behavior based on payment history of individual mortgage loans. Logistic regression analysis was applied to establish the relationship between individual risks (default or prepayment) and explanatory variables. Secondly, the MNL simultaneously established the relationship between both risks and a set of explanatory variables. The resulting parameters illustrated the competing nature of risks.

Finally, once the conditional probabilities of default and prepayment were obtained, an actuarial approach for MI pricing is presented using single and multiple decrement (default and prepayment) tables. Pricing differences when considering only credit risk as compared to the simultaneous default and prepayment risks approach are explained and interpreted for a series of mortgage credit portfolio segments.

The remainder of this paper is organized as follows, section 2 presents the multinomial logistic model for default and prepayment risks. Section 3 presents Mexican experience and the estimated value of relevant parameters. Section 4 develops price and reserve formulas within an actuarial framework as applied to mortgage insurance. Section 5 discusses the differences between prices considering a single source of risks versus competing risks as applied to Mexican experience. Finally, section 6 presents conclusions.

## **2. Mortgage default and prepayment model**

Mortgage borrowers face different alternatives to pay their mortgage obligations at each payment date. Three mutually exclusive choices are considered in this document, i) failure

from the borrower to pay his loan obligations subsequently referred as *default*, ii) the completion of the loan obligations prior to the date originally set in the contract subsequently referred as *prepayment* and iii) the successful completion of the contract under the terms originally set. It is important to note that the three events lead to contract termination.

These choices can be modeled by using multinomial logistic regression in which a joint model for the probability of discrete choices can be estimated. Conditional probabilities of prepayment and default are also likely to depend on common factors so that there is some inherent simultaneity in the decisions of borrowers.

Assume there exist “p” covariates and a constant term. We denote the two proposed functions as:

$$\begin{aligned} g_1(x_1, \dots, x_p) &= \ln \left( \frac{P(Y = 1 / x_1, \dots, x_p)}{P(Y = 0 / x_1, \dots, x_p)} \right) \\ &= \beta_{10} + \beta_{11} \cdot x_1 + \dots + \beta_{1p} \cdot x_p \\ &= \sum_{j=0}^p \beta_{1j} \cdot x_{1j} \end{aligned}$$

And

$$\begin{aligned} g_2(x_1, \dots, x_p) &= \ln \left( \frac{P(Y = 2 / x_1, \dots, x_p)}{P(Y = 0 / x_1, \dots, x_p)} \right) \\ &= \beta_{20} + \beta_{21} \cdot x_1 + \dots + \beta_{2p} \cdot x_p \\ &= \sum_{j=0}^p \beta_{2j} \cdot x_{2j} \end{aligned}$$

It follows that the conditional probabilities of each outcome category given the covariate vector are:

$$\text{Default} \quad P(Y = 1 / x_1, \dots, x_p) = \frac{e^{g_1(x_1, \dots, x_p)}}{1 + e^{g_1(x_1, \dots, x_p)} + e^{g_2(x_1, \dots, x_p)}}$$

$$\text{Prepayment} \quad P(Y = 2 / x_1, \dots, x_p) = \frac{e^{g_2(x_1, \dots, x_p)}}{1 + e^{g_1(x_1, \dots, x_p)} + e^{g_2(x_1, \dots, x_p)}}$$

$$\text{Successful completion} \quad P(Y = 3 / x_1, \dots, x_p) = \frac{1}{1 + e^{g_1(x_1, \dots, x_p)} + e^{g_2(x_1, \dots, x_p)}}$$

Covariates “X” are described in annex 1 and were constructed considering the following categories:

1. Financial Characteristics of the loan	<ul style="list-style-type: none"> <li>▪ Term</li> <li>▪ Currency</li> <li>▪ Loan to Value at origination LTV</li> <li>▪ Subsidy</li> <li>▪ House Value</li> </ul>
2. Borrower information	<ul style="list-style-type: none"> <li>▪ Employment</li> <li>▪ Payment to Income (PTI)</li> <li>▪ Monthly Income</li> </ul>
3. Mortgage Behavior	<ul style="list-style-type: none"> <li>▪ Mortgage age</li> <li>▪ Loan Size</li> <li>▪ Current Loan to Value (CLTV)</li> </ul>
4. Financial environment	<ul style="list-style-type: none"> <li>▪ Mortgage Premium</li> <li>▪ Burnout</li> </ul>
5. Payment history	<ul style="list-style-type: none"> <li>▪ Delinquency</li> <li>▪ Outstanding Balance</li> </ul>

### 3. Estimation of mortgage default and prepayment MNL model with Mexican experience

The information used for the estimation of the model is organized in blocks conformed by 25 months of data divided in, i) a period of 12 months ( $t_{-12}, t_{-1}$ ) called the observation period, ii) a second period of 12 months ( $t_1, t_{12}$ ) called the performance period, and iii)  $T_0$  which is considered the time at observation.

The observation period facilitates the analysis of the characteristics of the loan and is used to build the explanatory variables that describe the future behavior of the loan. The performance period allows the observation of either default or prepayment from the borrower. The following definitions were considered for the observation of default and prepayment during the performance period.

**Default** is observed when the obligor is past due more than 90 days on any material credit obligation to the creditor during the performance period.

**Prepayment** is defined as the completion of the totality of the loan obligations prior to the date originally set in the contract during the performance period.

Loans that either defaulted or prepaid a  $t_0$  are not considered for the estimation of the model.

The database used for the estimation covers the period January 2004 to august 2007 and reflects the payment behavior of 131,667 mortgages.

Loan's distribution by origination date

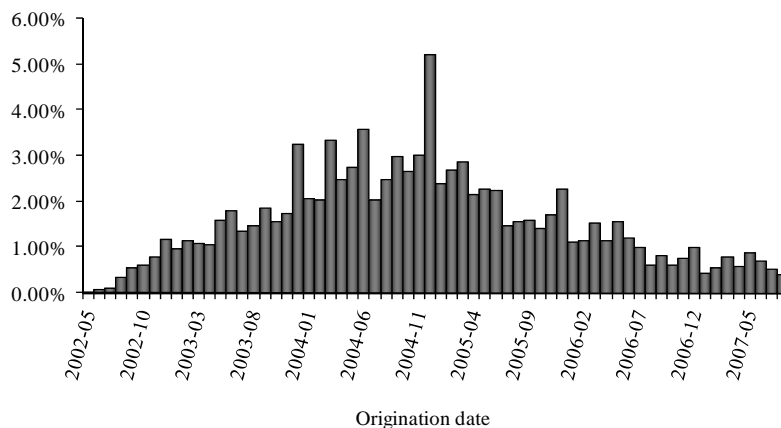


Figure 3.1: Mortgage's origination distribution in the total population database.

A random sample of 69,772 loans was taken for the estimation of the model. Figure 3.2 represent the time series for default and prepayment for the chosen sample and the population.

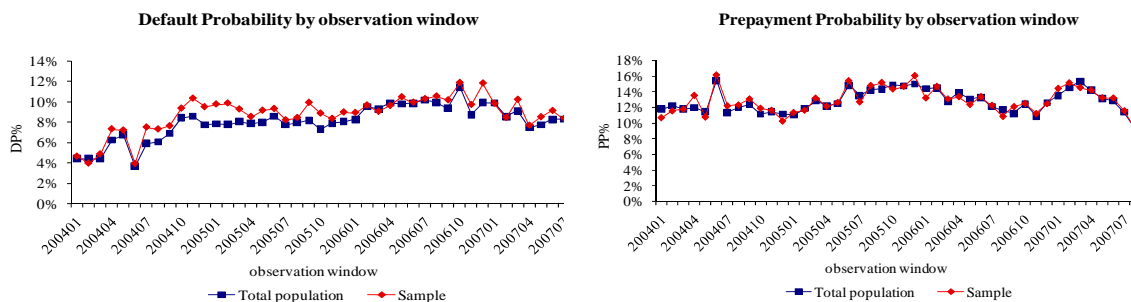
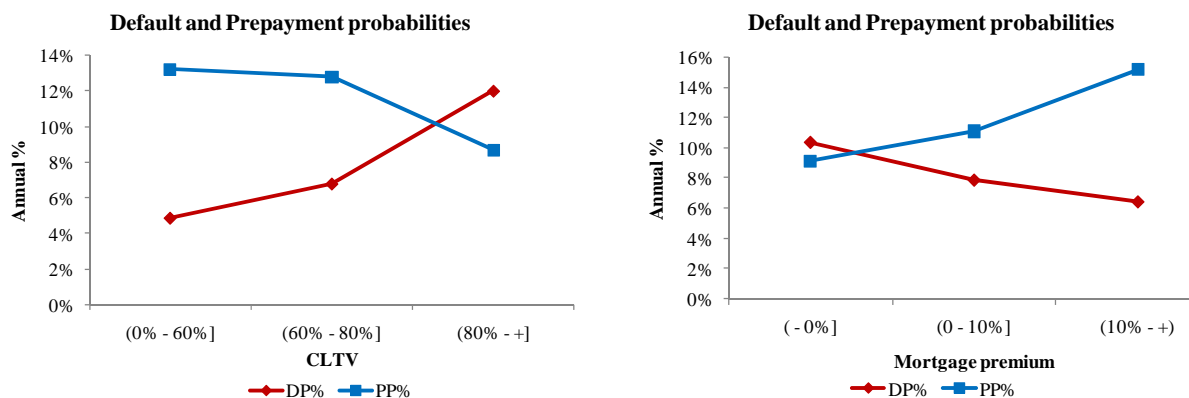


Figure 3.2: Default and prepayment probabilities by observation window total population vs sample.

For some variables the competing nature of both risks is made apparent through exploratory analysis. CLTV, Mortgage premium, num12\_0 (number of months that the loan has 0 past due payments in the last 12 months) and Burnout show a contrasting tendency towards default and prepayment.



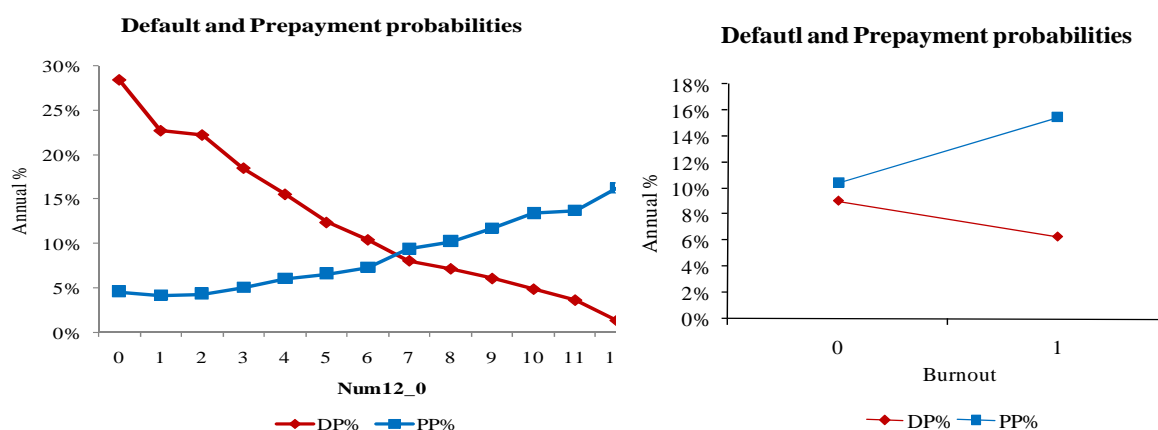


Figure 3.3: Default and prepayment observed rates.

The estimation of the model leads to the following selected significant variables:

- Current Loan to Value
- Mortgage Premium
- Number of months with 0 payments due in the last 12 months (num12\_0)

Analysis of Maximum Likelihood Estimates					
Parameter	Function	Estimate	Standard Error	Pr > ChiSq	
Intercept	Default	-1.0656	0.0361	<.0001	
	Prepayment	-3.0583	0.038	<.0001	
Mortgage Premium (- 0%]	(0 - 10%]	Default	-0.3054	0.0258	<.0001
	(0 - 10%]	Prepayment	0.1929	0.0233	<.0001
	(10% - +)	Default	-0.2552	0.0321	<.0001
	(10% - +)	Prepayment	0.4745	0.0242	<.0001
Current Loan to Value (0% - 60%]	(60% - 80%]	Default	0.2203	0.0376	<.0001
	(60% - 80%]	Prepayment	0.0361	0.025	0.149
	(80% - mayor]	Default	0.5718	0.0337	<.0001
	(80% - mayor]	Prepayment	-0.2086	0.0251	<.0001
Num12_0	Default	-0.2236	0.00296	<.0001	
	Prepayment	0.1087	0.00301	<.0001	

Table 1: Multinomial logistic regression estimates.

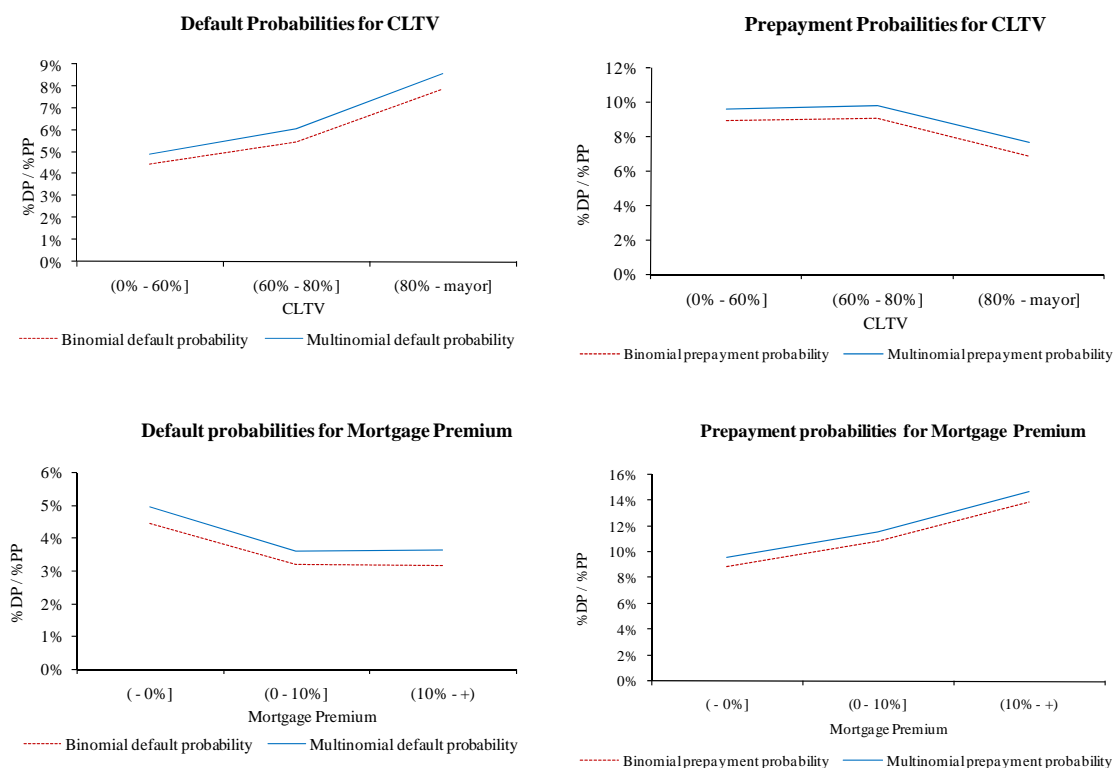
In order to assess the impact of the simultaneity of the estimation separate individual logit models for default and prepayment were run to compare the parameters involved in the estimation.

Variable	Reference Category	Categories	Function	$\beta$ joint model	$\beta$ binomial models
Intersection			Default	-1.0656	-1.1047
			Prepayment	-3.0583	-3.306
Mortgage Premium	(- 0%]	(0 - 10%]	Default	-0.3054	-0.3201
			Prepayment	0.1929	0.2137
		(10% - +)	Default	-0.2552	-0.2994
			Prepayment	0.4745	0.4935
Current Loan to Value	(0% - 60%]	(60% - 80%]	Default	0.2203	0.218
			Prepayment	0.0361	0.0289
		(80% - mayor]	Default	0.5718	0.5905
			Prepayment	-0.2086	-0.2394
Num12_0			Default	-0.2236	-0.2333
			Prepayment	0.1087	0.1282

Table 2: Comparison between multinomial vs. binomial logistic regression models.

As can be seen models provide very similar estimation values, however in some segments like mortgage premium and num12\_0 a larger value of default parameter is accompanied to a corresponding smaller estimate for prepayment parameter which suggests that whenever a segment of the population shows more tendency to default the joint model diminishes the probability to prepay reflecting the contrasting nature of events.

The following figures illustrate the comparison between probability estimates using a binomial logistic model versus a multinomial logistic model.



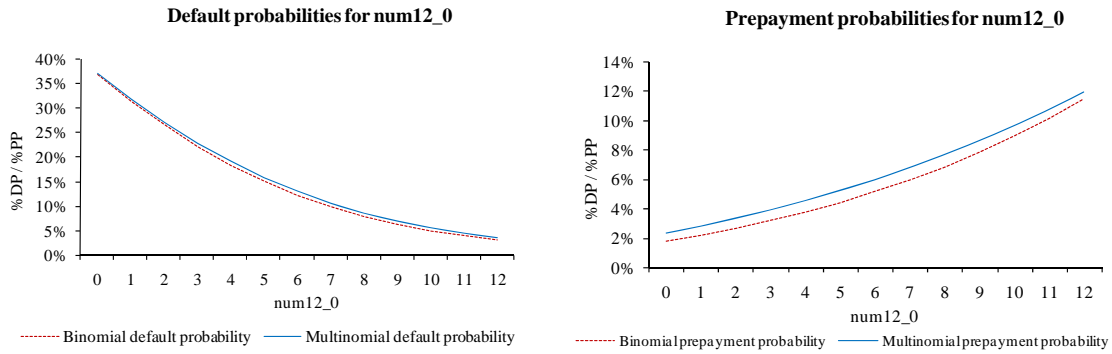


Figure 3.4: Comparison between binomial vs. multinomial default and prepayment probabilities

As can be observed, probability estimates do not differ significantly and MLN estimates appear to be higher than its logistic regression counterparts. It is interesting to note that in no case the sign or the interpretation of each covariate parameter is changed due to the estimation method used.

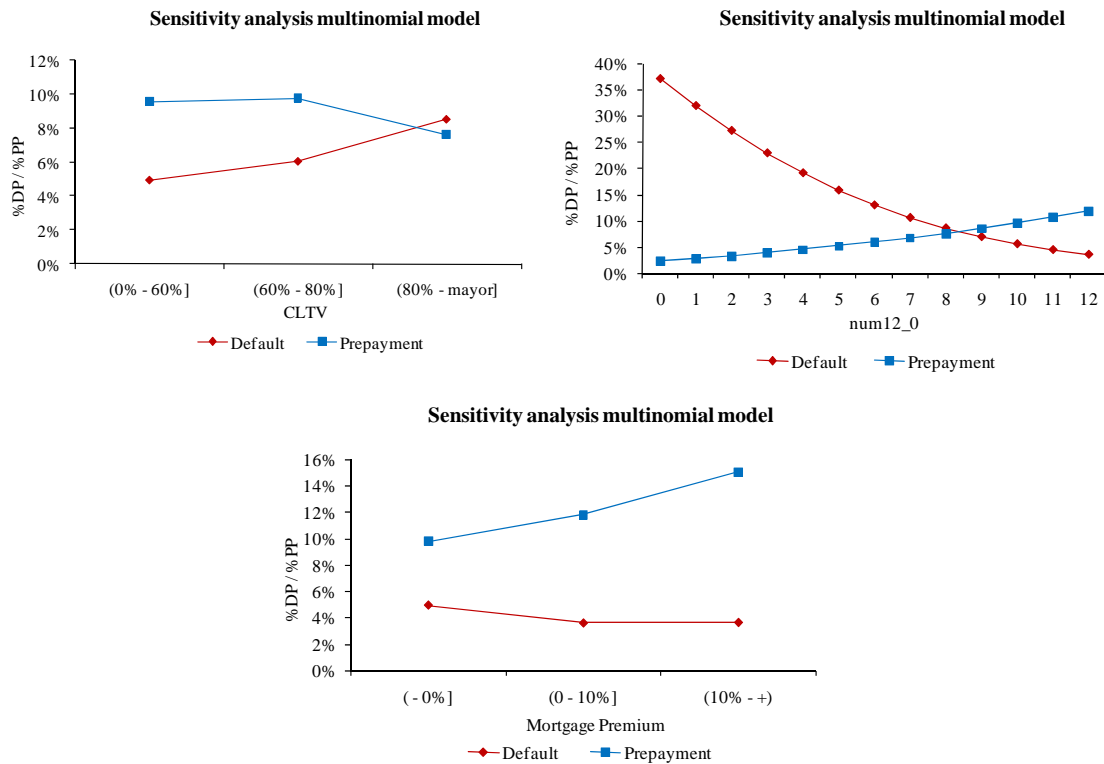


Figure 3.5: Sensitivity analysis for covariates in the model.

## 4. Actuarial framework for pricing and reserving mortgage insurance

### 4.1 Mortgage insurance and life insurance

Life insurance is a contract between the insurer and the *policy owner* whereby a *benefit* is paid to the designated *beneficiaries* if an *insured event* occurs. The policy owner agrees to pay a stipulated amount called a *premium*.

In mortgage insurance the *policy owner* is the financial institution that provides the mortgage loan, the *benefit* is an amount of money expressed as a percentage of the outstanding amount

of the loan at the moment of default and the *insured event* is the failure from the borrower to pay his loan obligations. The policy owner agrees to pay a *premium* that is paid either as a fixed amount at the beginning of the loan or periodically during the life time of the loan.

## 4.2 Terminology

The pricing and reserving framework for mortgage insurance benefits from the extensive literature developed for life insurance such as Bowers (1986)

To develop the corresponding formulas, the following definitions are necessary:

- $l_k$  number of loans that have not incurred in the insured event  $k$  periods after origination.
- $d_k$  number of loans that incur in the insured event when they reach  $k$  periods after origination. By definition:

$$d_k = l_k - l_{k+1}$$

- $p_k$  probability that a loan does not incur in the insured event from period  $k$  to  $k+1$  after origination.

$$p_k = \frac{l_{k+1}}{l_k}$$

- $q_k$  probability that a loan that has not incurred in the insured event in the first  $k$  periods after origination, incurs in the next.

$$q_k = \frac{d_k}{l_k} = \frac{l_k - l_{k+1}}{l_k} = 1 - \frac{l_{k+1}}{l_k}$$

- ${}_k p_0$  probability that a loan at origination does not incur in the insured event in the next  $k$  periods.
- ${}_k p_0 \cdot q_{x+k}$  probability that a loan at origination does not incur in the insured event in the next  $k$  periods and incurs one period after period  $k$ .

## 4.3 Contract termination

In life insurance the population of new born individuals incurs with certainty in the insured event (death) during a finite period ( $e$  years) which means that  ${}_e q_0 = 1$ .

In mortgage insurance where default triggers the payment of the insured amount, the population of loans not necessarily incurs in the insured event over the life time of the contract since only a proportion fails to pay his loan obligations.

However, a number of events may lead to the termination of the mortgage contract which in turn may represent costs to the policy owner. In what follows the definition of insured event will take the form of the events defined in section 2 (default, prepayment and successful completion of the contract).

All three events lead to the termination of the contract which completes the analogy with life insurance whereby  ${}_e q_0 = 1$

Each of the insured events leads to a different insured amount for the insurance company. Each cause of contract termination is considered a “decrement” which shows a different probability of occurrence.

Let “ $k$ ” be the age of the loan and define  $T(k)$  and  $J(k)$  two random variables such that:

- $T(k)$  is the future life time of the loan aged  $k$  or alternatively the future life time while the loan does not incur in a contract termination cause and
- $J(k)$  the random variable describing the decrement cause for a loan aged  $k$ .

#### 4.4 Future life time of the loan

Let us define the density function of the random variable “future life time of the loan”  $T(k)$  as follows:

$$f_{T(k)} = {}_t p_k^{(\tau)} \cdot q_{k+t}^{(\tau)} / q_k^{(\tau)}$$

Where super index  $(\tau)$  denotes all decrement causes described above. The future life time of the loan is a function of the age of the loan and provides us with the probability that a loan aged  $k$  months terminates the contract in the interval defined by ages  $(k+t, k+t+1)$ .

The survival function of the variable “future life time of the loan” is:

$${}_t p_k^{(\tau)} = Pr(T_k^{(\tau)} > t) = 1 - Pr(T_k^{(\tau)} \leq t) = 1 - \sum_{i=1}^t q_k^{(\tau)} = 1 - {}_t q_k^{(\tau)}$$

#### 4.5 Joint distribution of future life time of the loan considering multiple decrements

The joint distribution of  $J(k)$  y  $T(k)$  is defined as follows:

$$f_{T,J}(t, j) = f_{T^{(\tau)}}(t) \cdot p_k(j, t) = {}_t p_k^{(\tau)} \cdot q_{k+t}^{(\tau)} \cdot p_k(j/t)$$

The term  ${}_t p_k^{(\tau)}$  denotes the probability that the contract remains active for  $t$  periods subject to all possible causes of termination at age  $k$ .

$q_{k+t}^{(\tau)} \cdot p_k(j/t)$  is interpreted as the probability that a contract aged “ $k$ ” is terminated due to cause “ $j$ ” at time “ $t$ ” given that the contract has not been terminated previously. This is denoted as:

$$q_{k+t}^{(j)} = q_{k+t}^{(\tau)} \cdot p_k(j/t)$$

Substituting:

$$f_{T,J}(t, j) = Pr(T(k) = t, J(k) = j) = {}_t p_k^{(\tau)} \cdot q_{k+t}^{(j)}$$

The joint distribution function of J and T is therefore defined as:

$$F_{T,j}(t, j) = Pr(T(k) \leq t, J(k) = j) = \sum_{i=1}^t p_k^{(\tau)} \cdot q_{k+1}^j =_t q_k^{(j)}$$

#### 4.6 Mortgage insurance pricing

The expected value of losses (EL) to the insurance company under mortgage insurance subject to multiple causes of termination of the contract is calculated as follow:

$$EL = \sum_{j=1}^{J \text{ term}} \sum_{k=1} B_k^j p_0^{(\tau)} q_k^j v^k$$

Where:

- $B_k^j$  Benefit paid at age k to the policy owner under termination of the contract j.
- $v^k$  discount factor calculated as  $v = \frac{1}{(1+i)^k}$ .
- $i$  funding interest rate
- $j$  number of causes of contract termination
- “term” denotes the legal term of the mortgage at origination

As mentioned before in this document J is proposed to be equal to 3 where  $j=1$  is the event of default,  $j=2$  is the event of prepayment of the loan and  $j=3$  is the successful completion of the contract. The benefit for the policy owner for each j is:

$B_k^1$  for first loss mortgage insurance is generally an amount equal to a percentage of the outstanding balance of the loan at the moment of default, therefore:

$$B_k^1 = c * OB_k$$

Where:

- $c$  is the percentage of the outstanding balance of the mortgage agreed to be paid in the event of default by the insurance company
- $OB_k$  Outstanding balance of the loan at age k.

For  $j=2$  (prepayment) and 3 (successful completion of the mortgage) the insurance company generally pays no benefit as mortgage insurance covers default and hence  $B_k^2 = 0$  and  $B_k^3 = 0$  for all k.

$EL$  represents the actuarial value of the upfront premium to be charged by the insurance company to the policy owner for the mortgage insurance contract.

The expected value of income for the insurance company under mortgage insurance subject to multiple causes of termination of the contract when the premium is expressed as a fixed percentage of the outstanding balance of the mortgage is calculated as follow:

$$EI = \sum_{k=1}^{term} ps \cdot OB_k \cdot p_0^{(\tau)} \cdot v^k$$

where:

- $ps$  represents the actuarial value of the premium expressed as a percentage of the outstanding balance of the loan.

It is important to note that  ${}_k p_0^{(\tau)} < {}_k p_0^{(j=1)}$  which implies that for the same premium  $ps$ ,  $EI < EI^{j=1}$  which implies that if other causes of termination different than default are not considered, the insurance company may overestimate  $EI$  which in turn would lead to subsequent losses for the mortgage insurance product.

$ps$  is estimated by deriving its value from the equation  $EL = EI$ .

#### 4.7 Mortgage insurance reserving

Reserves are defined as:

$$MIR_k^{(\tau)} = EL_k - EI_k$$

Where:

- $EL_k$  is the expected loss of the mortgage insurance contract at age  $k$ .
- $EI_k$  is the expected income of the mortgage insurance contract at age  $k$ .

Reserves may be estimated for each termination of the contract cause by estimating:

$$MIR_k^{(j)} = EL_k^{(j)} - EI_k^{(t)} = 0$$

Where:

$$EL_k^{(j)} = \sum_{i=1}^{term} B_i^j \cdot p_k^{(\tau)} \cdot q_i^j \cdot v^i$$

#### 4.8 Concept of a mortality table for default and prepayment

The values of  $l_k$  and  $d_k^j$  which represent the number of mortgages that has not terminated the contract for any cause and the number of contracts that terminates by cause  $j$  respectively can be observed from payment experience from a mortgage portfolio.

The probabilities of termination of the contract are therefore estimated as:

$$q_k^j = \frac{d_k^j}{l_k}$$

Similarly the subsequent survival probabilities are estimated accordingly:

$${}_k P_0^{(\tau)} \cdot q_k^j = \frac{d_k^j}{l_0}$$

Where no payment experience is available for the life time of the mortgage, probabilities of default and prepayment can be adjusted to mortgage market conventions such as de standard default assumption (SDA) and public securities association prepayment assumption (PSA).

## 4.9 Mortgage default and prepayment mortality conventions

### 4.9.1 Prepayment measures

#### Conditional Prepayment Rate (CPR)

The conditional constant prepayment rate (CPR) is a constant probability of prepayment of a mortgage portfolio. The CPR is established in terms of the percentage of the expected outstanding balance of the pool at the end of the year.

The CPR is an annualized rate; however the mortgage payments are received monthly so it can be expressed as a monthly rate known as “Single Monthly Mortality” (SMM). The SMM is the percentage of the loan’s outstanding balance at the beginning of month assumed to prepay in full during the month. It can be calculated as follows:

$$SMM = \frac{\text{Total payments, including prepayments} - \text{Scheduled interest payment} - \text{Scheduled principal payment}}{\text{Unpaid principal balance} - \text{Scheduled principal payment}}$$

The SMM and the CPR have the following relationship:

$$(1 - CPR) = (1 - SMM)^{12}$$

$$SMM = 1 - (1 - CPR)^{1/12}$$

One of the advantages of CPR is its easy calculation; however it is a constant rate that does not reflect accurately the behavior of mortgages in practice as it assigns a higher prepayment speed at the beginning of the mortgage.

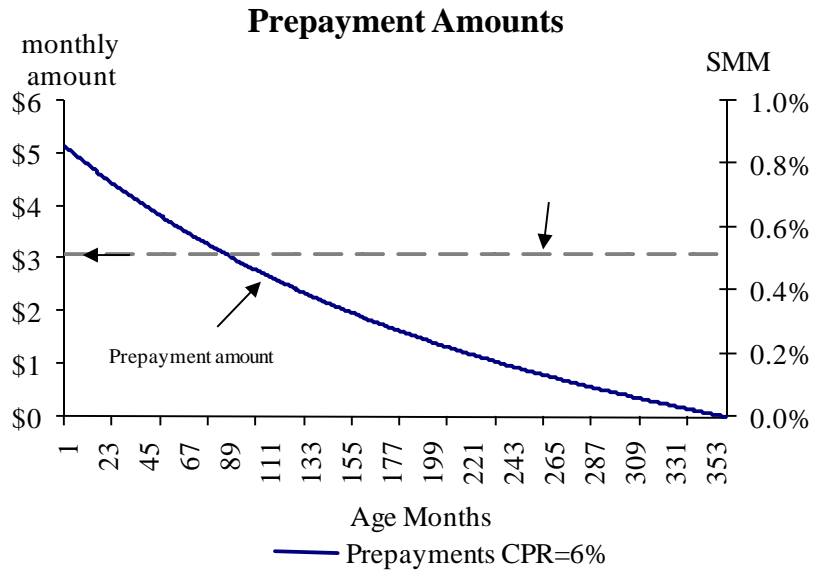


Figure 4.1: Prepayment behavior with a 6% CPR

**PSA standard**

The PSA standard is a benchmark developed by the Public Securities Association. It is built through a series of annualized prepayment rates changing upon the mortgage age in months.

The rates are CPRs. The PSA assumes that prepayment will slowly increase at the beginning of the amortization term then increase at a constant rate until month 30. After month 30, it will remain constant until maturity. The assumption is that prepayment rates are low for new originated mortgages and that it will increase as the mortgages become seasoned.

PSA can be described as follows (Fabozzi, Bhattacharya, Berliner, 2007):

- A CPR of 0.2% for the first month, increased by 0.2% per year per month for the next 29 months when it reaches 6% per year and 6% CPR for the remaining years.

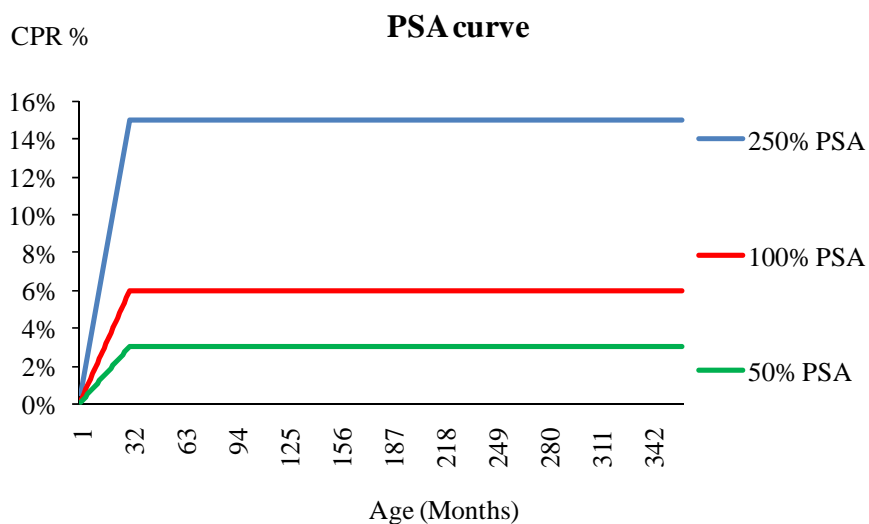


Figure 4.2: PSA curve

Other prepayment curves can be expressed through the PSA Standard. For example, a mortgage pool prepays in month 50 at 250%PSA meaning that it has an annual prepayment rate of  $15\% = 6\% * 2.5$ .

#### 4.9.2 Default measures

##### Conditional default rate (CDR)

CDR represents the monthly default rate expressed as annual percentage. It is an annualized value of the outstanding balance of the loans that defaulted in the current month as a percentage of the outstanding balance of the mortgage portfolio (Fabozzi, Bhattacharya, Berliner, 2007).

Mathematically it is calculated as follows:

$$MDR_t = \frac{\text{Default loan balance}_{month t}}{\text{Beginning Balance}_{month t} - \text{Scheduled Principal Payment}_{month t}}$$

To obtain de CDR, the MDR is annualized:

$$CDR = 1 - (1 - MDR_t)^{12}$$

The CDR is not a historic description of defaults behavior in the mortgage portfolio neither a predictive measure of the expected default rate, also it does not captures default variations over time. To assume a constant CDR may not be consistent with default behavior over the pool, however it is reference measure such as CPR.

##### Standard Default Assumption (SDA)

The SDA curve refers to a standard measure developed by “The Bond Market Association” it is a reference measure to default in a mortgage pool of 30 years fixed rate loans, upon the age of the loan. It is expressed in terms of the CDR.

The SDA assumes an initial CDR of 0.02% over the outstanding balance of the mortgage increasing in 0.02% monthly until month 30 where reaches a 0.6% CDR.

After month 30, the CDR remains constant till month 60 (fifth year), month in which it begins to decrease lineally in 0.095% each month until it reaches 0.03% in month 120. From age 120 the CDR remains constant until maturity (age 360).

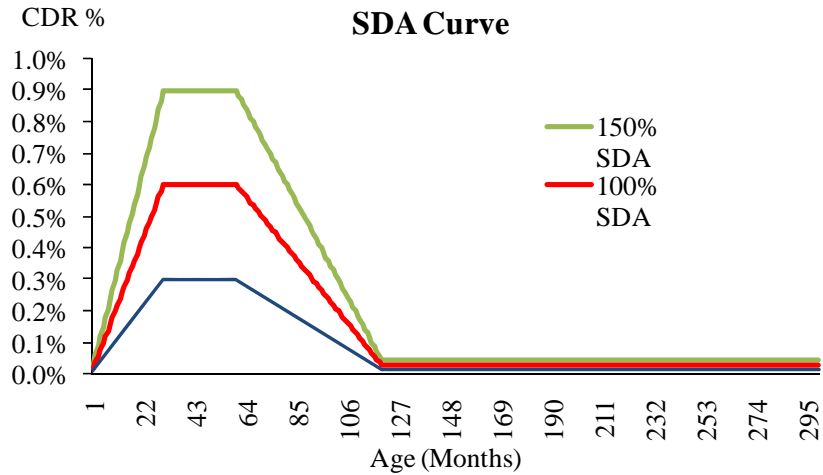


Figure 4.3: SDA curve

Both CPR and CDR can be translated to actuarial terminology assuming that loan amounts are the same as follows:

$$MDR_k = q_k^1 = \frac{d_k^1}{l_k}$$

$$SMM_k = q_k^2 = \frac{d_k^2}{l_k}$$

To illustrate the process consider the following mortality table that considers 100% SDA and 100% PSA standards.

#### 4.10 Illustration of a multiple decrement table

Age Months	$q_k^{(1)}$	$q_k^{(2)}$	$q_k^{(3)}$	$q_k^{(\tau)}$	$l_k^{(\tau)}$	$d_k^{(1)}$	$d_k^{(2)}$	$d_k^{(3)}$	$d_k^{(\tau)}$	$d_k^{(\tau)}/l_k^{(\tau)}$	${}_kP_0^{(\tau)}$	${}_kP_0^{(\tau)} \cdot q_k^{(1)}$	${}_kP_0^{(\tau)} \cdot q_k^{(2)}$	${}_kP_0^{(\tau)} \cdot q_k^{(3)}$	${}_kP_0^{(\tau)} \cdot q_k^{(\tau)}$
1	0.00%	0.00%	0.00%	0.00%	100,000	0	0	0	0	0.00%	100.00%	0.00%	0.00%	0.00%	0.00%
2	0.00%	0.00%	0.02%	0.02%	100,000	0	2	17	18	0.02%	100.00%	0.00%	0.00%	0.02%	0.02%
3	0.00%	0.00%	0.03%	0.04%	99,982	0	3	33	37	0.04%	99.98%	0.00%	0.00%	0.03%	0.04%
4	0.00%	0.01%	0.05%	0.06%	99,945	0	5	50	55	0.06%	99.94%	0.00%	0.00%	0.05%	0.06%
5	0.00%	0.01%	0.07%	0.07%	99,890	0	7	67	74	0.07%	99.89%	0.00%	0.01%	0.07%	0.07%
6	0.00%	0.01%	0.08%	0.09%	99,816	0	8	84	92	0.09%	99.82%	0.00%	0.01%	0.08%	0.09%
7	0.00%	0.01%	0.10%	0.11%	99,724	0	10	100	110	0.11%	99.72%	0.00%	0.01%	0.10%	0.11%
8	0.00%	0.01%	0.12%	0.13%	99,614	0	12	117	129	0.13%	99.61%	0.00%	0.01%	0.12%	0.13%
9	0.00%	0.01%	0.13%	0.15%	99,486	0	13	134	147	0.15%	99.49%	0.00%	0.01%	0.13%	0.15%
10	0.00%	0.02%	0.15%	0.17%	99,339	0	15	150	165	0.17%	99.34%	0.00%	0.01%	0.15%	0.17%
11	0.00%	0.02%	0.17%	0.18%	99,174	0	17	167	183	0.18%	99.17%	0.00%	0.02%	0.17%	0.18%
12	0.00%	0.02%	0.19%	0.20%	98,990	0	18	183	202	0.20%	98.99%	0.00%	0.02%	0.18%	0.20%
13	0.00%	0.02%	0.20%	0.22%	98,789	0	20	200	220	0.22%	98.79%	0.00%	0.02%	0.20%	0.22%
14	0.00%	0.02%	0.22%	0.24%	98,569	0	21	216	238	0.24%	98.57%	0.00%	0.02%	0.22%	0.24%
15	0.00%	0.02%	0.24%	0.26%	98,332	0	23	232	255	0.26%	98.33%	0.00%	0.02%	0.23%	0.26%
16	0.00%	0.03%	0.25%	0.28%	98,076	0	25	249	273	0.27%	98.08%	0.00%	0.02%	0.25%	0.27%
17	0.00%	0.03%	0.27%	0.30%	97,803	0	26	265	291	0.29%	97.80%	0.00%	0.03%	0.26%	0.29%
18	0.00%	0.03%	0.29%	0.32%	97,512	0	28	281	308	0.31%	97.51%	0.00%	0.03%	0.28%	0.31%
19	0.00%	0.03%	0.31%	0.34%	97,204	0	29	297	326	0.33%	97.20%	0.00%	0.03%	0.30%	0.33%
20	0.00%	0.03%	0.32%	0.35%	96,878	0	31	312	343	0.34%	96.88%	0.00%	0.03%	0.31%	0.34%
21	0.00%	0.03%	0.34%	0.37%	96,535	0	32	328	360	0.36%	96.54%	0.00%	0.03%	0.33%	0.36%
22	0.00%	0.04%	0.36%	0.39%	96,175	0	34	343	377	0.38%	96.17%	0.00%	0.03%	0.34%	0.38%
23	0.00%	0.04%	0.37%	0.41%	95,798	0	35	359	394	0.39%	95.80%	0.00%	0.04%	0.36%	0.39%
24	0.00%	0.04%	0.39%	0.43%	95,404	0	37	374	410	0.41%	95.40%	0.00%	0.04%	0.37%	0.41%
25	0.00%	0.04%	0.41%	0.45%	94,994	0	38	389	427	0.43%	94.99%	0.00%	0.04%	0.39%	0.43%
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
356	0.00%	0.00%	0.51%	0.52%	16,634	0	0	86	86	0.09%	16.63%	0.00%	0.00%	0.09%	0.09%
357	0.00%	0.00%	0.51%	0.52%	16,548	0	0	85	86	0.09%	16.55%	0.00%	0.00%	0.09%	0.09%
358	0.00%	0.00%	0.51%	0.52%	16,463	0	0	85	85	0.09%	16.46%	0.00%	0.00%	0.08%	0.09%
359	0.00%	0.00%	0.51%	0.52%	16,378	0	0	84	85	0.08%	16.38%	0.00%	0.00%	0.08%	0.08%
360	99.48%	0.00%	0.51%	100.00%	16,293	16,209	0	84	16,293	16.29%	16.29%	16.21%	0.00%	0.08%	16.29%

Table 4.1: Multiple decrement table.

Finally for  $j=1$  which denotes the termination of the contract due to default, the probability of occurrence from age 0 to age  $k$  is:

$${}_k q_0^{(1)} = {}_k P_0^{(\tau)} \cdot q_k^{(1)} = \frac{d_k^{(1)}}{l_0^{(\tau)}}$$

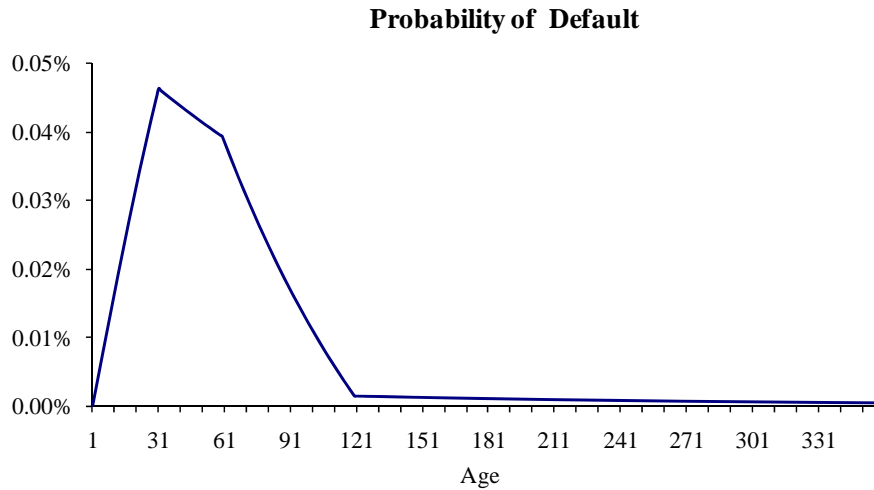


Figure 4.4: Probability of default

For  $j=2$  which denotes the termination of the contract due to prepayment the probability of occurrence from age 0 to age  $k$  is:

$${}_k q_0^{(2)} = {}_k P_0^{(\tau)} \cdot q_k^{(2)} = \frac{d_k^{(2)}}{l_0^{(\tau)}}$$

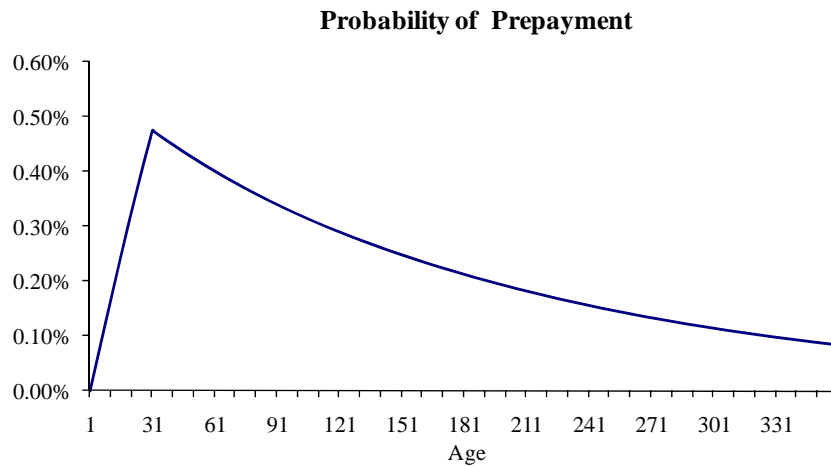


Figure 4.5: Probability of prepayment

The probability that the contract terminates for any cause is represented as follows:

$${}_k q_0^{(\tau)} = {}_k P_0^{(\tau)} \cdot q_k^{(\tau)} = \frac{d_0^{(\tau)}}{l_0^{(\tau)}}$$

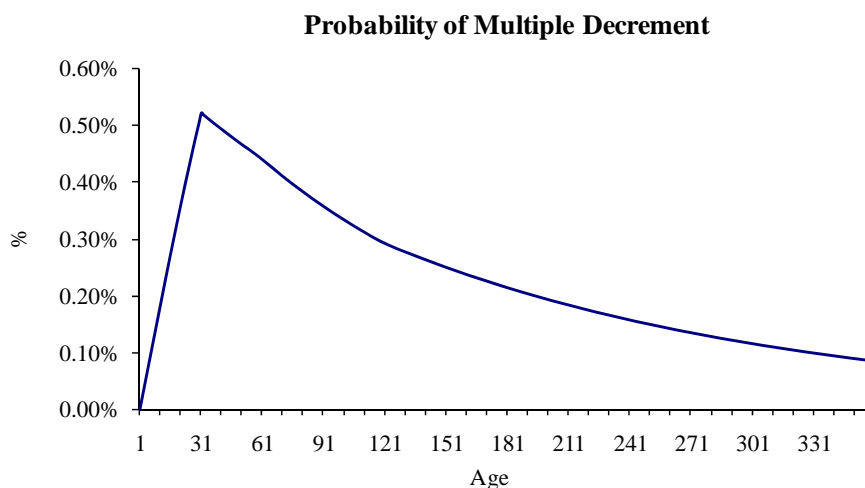


Figure 4.6: Probability of multiple decrement

## 5. Price sensitivity

### 5.1 Probability estimation with Mexican experience

Default and prepayment curves were obtained as a multiple of the PSA and SDA original curves described in section 3. As default and prepayment probabilities using the multinomial logistic model are given in terms of annual probabilities no transformations were needed to homologate them to CDR and CPR standards. The annual CDR used for 100% SDA is 0.6% and the annual CPR for 100% PSA is 6%.

Finally, the default and prepayment SDA and PSA curves for probabilities were introduced to the pricing model described in section 4 to obtain the upfront and level premiums for mortgage insurance. The premiums were differentiated according to different values of the explanatory variables considered by the MNL model.

The following tables show the default and prepayment probabilities for a selected covariate and SDA and PSA factors associated to each of the buckets in which the variable was categorized.

- Current loan to value

<b>Current Loan to Value (CLTV)</b>	<b>DP%</b>	<b>PP%</b>	<b>PSA Factor</b>	<b>SDA Factor</b>
<b>(0% - 60%]</b>	4.87%	13.20%	8.11	2.20
<b>(60% - 80%]</b>	6.78%	12.81%	11.30	2.13
<b>(80% - +]</b>	11.99%	8.67%	19.99	1.44
<b>Total</b>	8.74%	10.99%	14.57	1.83

Table 5.1: Default and Prepayment probabilities and factors for CLTV.

The loan to value ratio serves as an indicator of the current equity position of the borrower in the property. Higher values of CLTV increase the default probability and lower the prepayment probability.

- Mortgage Premium

<b>Mortgage Premium</b>	<b>DP%</b>	<b>PP%</b>	<b>SDA Factor</b>	<b>PSA factor</b>
<b>( - 0%]</b>	10.32%	9.14%	17.20	1.52
<b>(0 - 10%]</b>	7.84%	11.10%	13.07	1.85
<b>(10% - +)</b>	6.42%	15.16%	10.69	2.53
<b>Total</b>	8.74%	10.99%	14.57	1.83

Table 5.2: Default and Prepayment probabilities an factors for Mortgage Premium.

Mortgage premium is the difference of the present value of the future stream of mortgage payments discounted at the current market rate and the present value of the mortgage evaluated at the mortgage fixed rate. For lower values of mortgage premium the default probability will increase and the prepayment probability will decrease. When the market rate is lower than the mortgage rate the difference between rates will be positive and the borrower has an incentive to prepay because he will be able to find a cheaper loan in the market. If the market rate is higher the difference between present values will be negative indicating that the borrower is less likely to prepay, his current loan has better conditions than any loan in the market.

- Number of months the loan has 0 payments due in the last 12 months.

<b>Number of times with 0 payments due in the last 12 months</b>	<b>DP%</b>	<b>PP%</b>	<b>SDA Factor</b>	<b>PSA factor</b>
<b>0</b>	28.33%	4.60%	47.21	0.77
<b>1</b>	22.64%	4.20%	37.74	0.70
<b>2</b>	22.16%	4.36%	36.93	0.73
<b>3</b>	18.45%	5.03%	30.75	0.84
<b>4</b>	15.52%	6.08%	25.86	1.01
<b>5</b>	12.37%	6.61%	20.61	1.10
<b>6</b>	10.43%	7.30%	17.38	1.22
<b>7</b>	8.08%	9.38%	13.47	1.56
<b>8</b>	7.19%	10.23%	11.98	1.70
<b>9</b>	6.10%	11.68%	10.16	1.95
<b>10</b>	4.91%	13.39%	8.18	2.23
<b>11</b>	3.68%	13.70%	6.14	2.28
<b>12</b>	1.40%	16.17%	2.33	2.69
<b>Total</b>	8.74%	10.99%	14.57	1.83

Table 5.3: Default and Prepayment probabilities factors for number of times with 0 payments due in the last 12 months.

The number of months the loan reported 0 payments due in the last 12 months is a behavior variable constructed with the payment history of the loan in the last year. As the loan remains current (0 payments due) the probability of default lowers and prepayment probabilities are higher.

## 5.2 Price estimation with Mexican experience

Considering the results shown in tables 5.1 to 5.3, SDA and PSA curves are adjusted by a factor of 15 and 1.8 respectively. The following mortgage is chosen to exemplify the pricing of the MI product:

<b>Loan Amount</b>	<b>MXN \$10,000</b>
<b>Loan term</b>	<b>360 months</b>

<b>Interest Rate</b>	<b>10%</b>
<b>Mortgage Insurance</b>	<b>30%</b>
<b>Monthly Payment</b>	<b>MXN\$ 87.76</b>
<b>Discount Rate</b>	<b>5.00%</b>

Using the methodology described in Section 4, we obtain two different prices for the MI product i) annual premium applicable to the outstanding balance of the loan, ii) upfront premium.

<b>Pricing</b>	
<b>Annual</b>	<b>Up Front</b>
<b>1.61%</b>	<b>7.98%</b>

The present value of the expected income and losses over the lifetime of the loan can be seen in the figure below:



Figure 5.1: Expected Loss vs. Expected Income

The difference between the two curves is the cash flow that generates the reserve for the product. Over time this difference can be positive or negative, but it adds up to zero at origination.

The following illustrates the sensitivity of the annual and up front premiums keeping the same probability of default but changing the probability of prepayment.

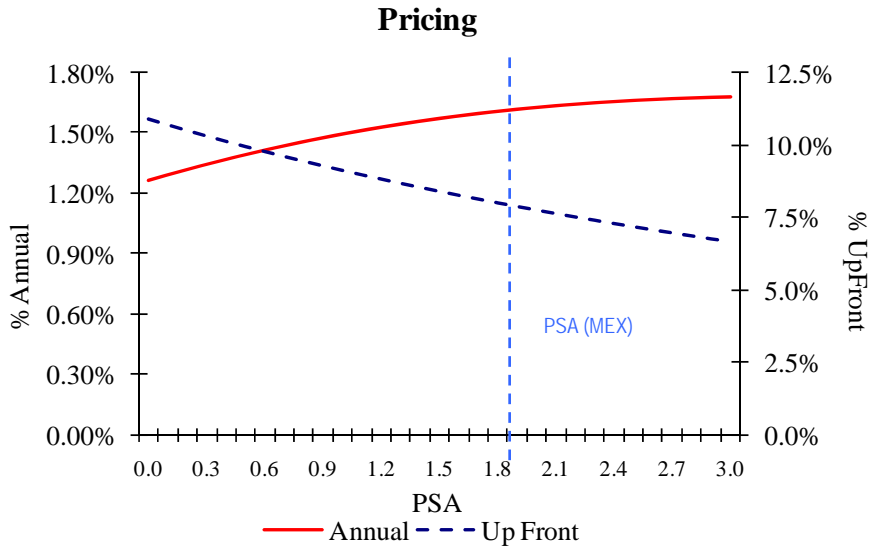


Figure 5.2: Pricing Sensitivity.

A higher value of PSA affects the cash flows of the product and generates a lower value of the upfront premium. This effect is explained by the fact that a lesser number of loans default which in turn generates lesser expenses to the insurance company.

In contrast, the same effect works differently for the annual premium as it rises when prepayment rises. This in turn is explained by the fact that a higher rate of prepayment generates a lesser amount of premium charged during the life time of the portfolios of loans and hence requires a larger premium to compensate.

The following figures show this effect for the present value of expenses and income of the MI product as a function of PSA.

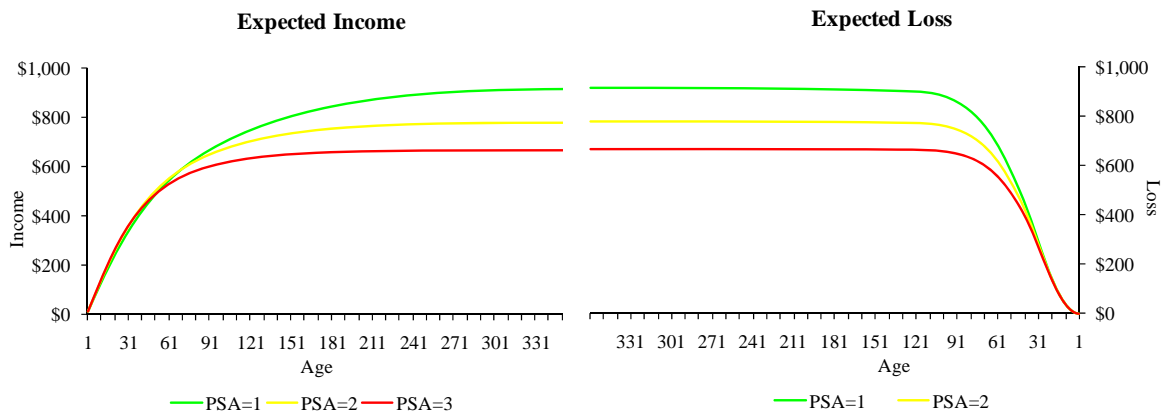


Figure 5.3: Expected Income vs. Expected Loss

It is interesting to note that the prices that result from using the original curves of SDA and PSA are:

Pricing	
Annual	Up Front
0.09%	0.74%

The following figures show how the price obtained with Mexican experience moves when only one assumption is changed.

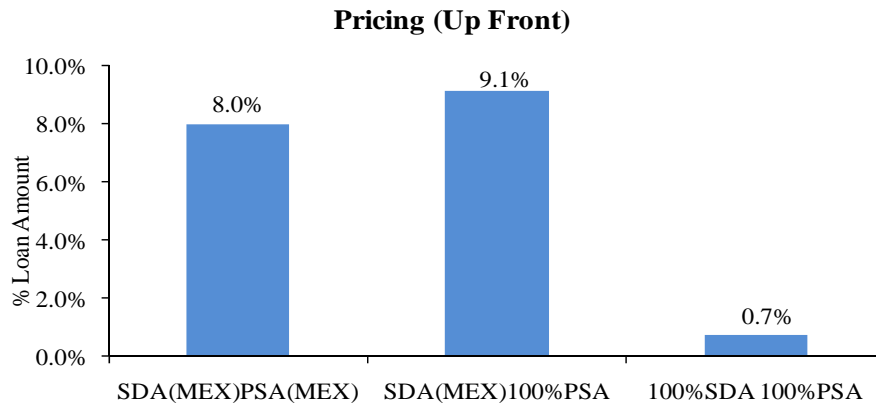


Figure 5.4:Pricing Up Front

The price for the original PSA and SDA curves is significantly lower than the price estimated with Mexican experience PSA and SDA curves. However, when the prices is estimated by moving one parameter at the time we observe that for the same SDA estimated with Mexican experience (SDA(MEX)) but 100%PSA (Mexican experience PSA(MEX) implies 180% PSA) the price is higher implying that a lesser frequency of prepayment results in a larger number of default and hence a higher up front premium. However when SDA is updated to 100% factor the impact is non ambiguous and risk and up front premium lower significantly.

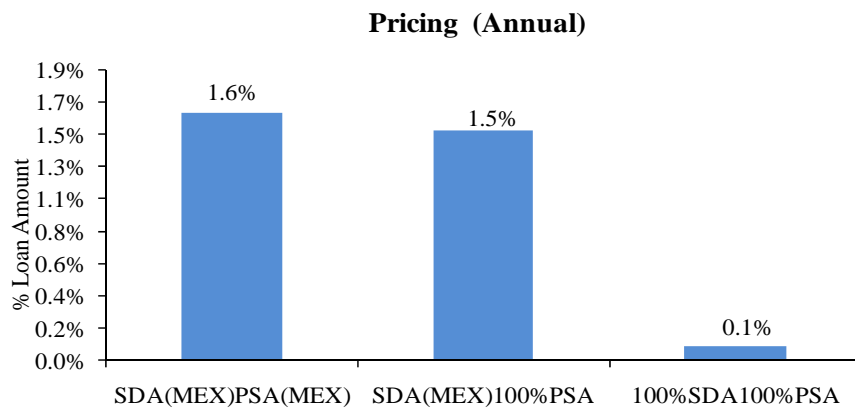


Figure 5.5: Pricing Annual.

As for the annual premium, when PSA is updated, the insurance company expects to receive a larger amount of income as a larger number of loans remain current. This allows for a lesser premium. Similarly, when SDA is updated the premium lowers significantly by taking into account the lesser risk incurred by the insurance company.

To illustrate the price sensitivity under different default assumptions Figure 5.6 graphs the resulting price taking several scenarios of probability of default and maintaining fixed prepayment (PSA 100%). The effect in this case follows the expected direction for both type of premiums.

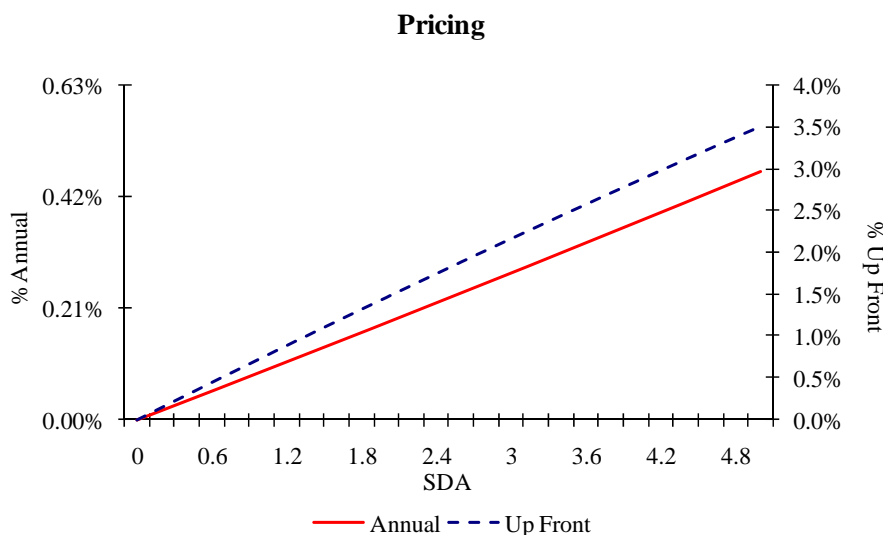


Figure 5.6:Pricing

Based on estimates of default and prepayment probabilities, the final proposed prices are:

Mortgage Premium ( - 0%)					Pricing	
Current Loan to Value	DP%	PP%	SDA Factor	PSA Factor	Annual	Up Front
(0% - 60%]	4.92%	9.58%	8.21	1.60	0.84%	4.94%
(60% - 80%]	6.04%	9.78%	10.07	1.63	1.04%	5.90%
(80% - mayor]	8.55%	7.62%	14.25	1.27	1.46%	8.38%

Table 5.4: Mortgage premium

As can be seen, the annual price may differ more than 0.6% for different values of CLTV. These prices provide the insurance company with some flexibility to simultaneously recognize the risk inherent in the portfolio of mortgages and remaining a profitable business.

## 6. Conclusions

The empirical analysis of this document considered Mexican loan level information to estimate a multinomial logit model to simultaneously explain default and prepayment. The explanatory variables reflect the design of the mortgage (CLTV), the current cost of the mortgage as compared to external financial conditions typically associated to prepayment (mortgage premium) and payment history of the individual which tends to describe its willingness to pay.

Exploratory analysis on these variables showed the contrasting behavior of prepayment and default in absolute terms and for different values of the covariates. The MNL model confirmed this behavior as parameters for these variables resulted significant and showed the expected sign.

Two single logistic regressions were estimated to compare the behavior of the MLN model and differences were observed in parameter values where the MNL model tended to associate larger default parameters with smaller prepayment parameters which is interpreted to be a tendency to take into account the competing nature of discrete choices involved in the model.

The importance of considering prepayment for mortgage insurance pricing is described as prices not only shift importantly according to the prepayment assumption taken into consideration but for the case of the annual premium move in opposed direction to the

upfront premium which implies hidden second round effects that the insurance company needs to take into consideration in its pricing policy.

Valuing the risks of default and prepayment together is highly recommended at two instances, first the pricing formulas which should consider default and prepayment but second, and most important, at the parameter estimation level where it was shown that a simultaneous estimation of default and prepayment provides different parameters and probabilities estimation which are considered to contain more information than their single risk parameter estimation counterparts.

This document applies one of the multiple approaches proposed in the extensive literature on the topic of simultaneous estimation of prepayment and default risk and adds to the growing field of pricing mortgage insurance by proposing a decrement table actuarial approach using mortgage market PSA and SDA conventions. It also represents a first reference for mortgage insurance pricing with Mexican experience.

## References

- Blood, R. (1997). *Mortgage Default Insurance Study: Final Report*. Unpublished Report for FOVI.
- Blood, R. (2001). *Mortgage Default Insurance: Credit Enhancement for Homeownership*. Housing Finance International.
- Bowers, Gerber, Hickman, Jones y Nesbitt (1986). *Actuarial Mathematics*. Society of Actuaries. EE.UU.
- Calhoun, C. and Deng, Y. (2002). *A Dynamic Analysis of Fixed- and Adjustable-Rate Mortgage Terminations*, Journal of Real Estate Finance and Economics, 4 (1-2), 9-33.
- Chen, K. *The Role of Mortgage Insurance in Risk Management*. International Journal of Real Estate Finance, Volume 1, No. 2.
- Cunningham, D.F., and P.H. Hendershott (1984): *Pricing FHA Mortgage Default Insurance*, Housing Finance Review, 13, 373-392.
- Deng, Y, Quigley, J.M. and R. Van Order, (2000), *Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options*, Econometrica 68(2): 275-307.
- Deng, Y. and S. Gabriel, (2002). *Modeling the performance of FHA-insured loans: borrower heterogeneity and the exercise of mortgage default and prepayment options*. Office of Policy Development and Research, U.S. Department of Housing and Urban Development. May, 2002.
- Deng, Y., J. M. Quigley, and R. Van Order (1996): *Mortgage Default and Low Down-payment Loans: The Cost of Public Subsidy*. Regional Science and Urban Economics, 26(3-4), 263.285.
- Fabozzi, F., A.K. Bhattacharya, and W.S.Berliner, (2007) *Mortgage-Backed Securities: Products, Structuring and Analytical Techniques*. The Frank J. Fabozzi Series. New York: Wiley
- Foster, C. and Van Order R. (1984). *An Option Based Model of Mortgage Default*. Housing Finance Review 3 (4): 351-72.
- Foster, C., and R. Van Order (1984): *An Option-Based Model of Mortgage Default*. Housing Finance Review, 3(4), 351.372.
- Gangani, S (1998). *MBS Structuring: Concepts and Techniques*. The Securitization Conduit. Vol. I, No. 3. 1998: 26-37. Deloitte & Touche LLP.
- Green, J., and J. B. Shoven (1986): *The Effect of Interest Rates on Mortgage Prepayments* .Journal of Money, Credit and Banking, 18, 41.50.

- Hosmer, D.W. and S. Lemeshow, (2000). *Applied Logistic Regression* (2nd edition). New York: Wiley.
- Kau, J.B. and D.C. Keenan (1995): *An overview of the Option-Theoretic Pricing of Mortgages*, Journal of Housing Research, 6(2), 217-244.
- Kau, J.B., D.C. Keenan and W.J. Muller III (1993): *An option-Based Pricing Model of Private Mortgage Insurance*. The Journal of Risk and Insurance, 60(2), 288-299.
- Lundstedt, K. (2005). *The competing risks of mortgage prepayment and default*. Andrew Davidson & Co., Inc. 2005.
- Mercer Management Consulting. (2000). *Credit Risk and Mortgage Default Insurance: Issues in Developing Mortgage Markets*. The World Bank: Mortgage Finance.
- Merton, R. C. (1973): *Theory of Rational Option Pricing*. Bell Journal of Economics and Management Science, 4, 141.183.
- Pickering, Natalie. (2000). *The SOFOLES: Niche Lending or New Leaders in the Mexican Mortgage Market?*. Joint Center for Housing Studies. Harvard University.
- Pierre, David. (2004). *The Role of Mortgage Insurance in Developing a Mortgage Market*. CMHC & SCHL Canada. Russia, Dubna.
- Pierre, David. (2006). *Mortgage Insurance International Lessons Housing Finance in Emerging Markets*. CMHC & SCHL Canada.
- Polverini, Sacha. (2002). *GE Mortgage Insurance: Benefits of a Mortgage Insurance System in Transition Economies*. Third Workshop on Housing Finance in Transition Economies.
- Quigley, J. M., and R. Van Order (1990): *Efficiency in the Mortgage Market: The Borrower's Perspective*. Journal of the American Real Estate and Urban Economics Association, 18(3), 237.252.
- Schwartz, E. S., and W. N. Torous (1989): *Prepayment and the Valuation of Mortgage-Backed Securities*. The Journal of Finance, 44(2), 375.392.
- Smith, L.D. and Lawrence, E.C. (1996): *A Comprehensive Model for Managing Credit Risk on Home Mortgage Portfolios*. Decision Sciences; Spring 1996; 27,2, 291-317.

## Annex 1

### *Financial characteristics of the loan:*

- **Term:** Original loan term in months
- **Currency:** Mexican pesos or inflation indexed currency (UDIS)
- **Loan to Value at origination (LTV):** The ratio between original loan amount and the house value at origination (t=0).

$$LTV = \frac{\text{Original Loan's Amount (0)}}{\text{House Value (0)}}$$

- **Subsidy:** describes the existence of financial support provided by the government.
- **House value:** classification of the house according to its value:
 

<i>Type I:</i>	Up to USD 20,000
<i>Type II:</i>	From USD 20,000 to USD 30,000.
<i>Type III:</i>	More than USD 30,000

### *Borrower information*

- **Employment:** main source of income of the borrower.
  - **Formal employee:** affiliated to social security benefits
  - **Independent professionals:** borrowers who may work by their own in a personal business or provide professional services.
  - **Informal workers:** borrowers who work in the informal economy.
- **Payment to Income (PTI):** ratio between the monthly payment and the borrower and co-borrower (if it is the case) income at origination.

$$PTI = \frac{\text{Mortgage monthly payment}}{\text{Borrower's Income}}$$

- **Monthly Income:** monthly income expressed as number of times the minimum wage according to the current minimum wage at origination date.
 

0 to 10:	Monthly minimum wages	0 up to usd 1,200 approximately.
10 to 20:	Monthly minimum wages	usd 1,200 up to usd 2,500 approximately
20 or more:	Monthly minimum wages	More than usd 2,500

### ***Mortgage behavior***

- ***Mortgage Age:*** Number of months from origination to age of analysis.

Default tends to increase in the first years following the origination, find its maximum somewhere during early years and then decline thereafter until it remains constant.

- ***Loan Size:*** ratio between the original loan amount at origination and mean loan amount in the same geographic location originated in the same fiscal year (Calhoun and Deng, 2002).

$$Loan\ Size = \frac{Original\ Loan\ Amount}{Mean\ Loan\ Amount_{State\ j}}$$

- ***Current Loan to Value (CLTV):*** ratio between the current outstanding balance of the mortgage and the house value at origination.

$$CLTV = \frac{Outstanding\ Balance(t)}{House\ Value(0)}$$

### ***Financial environment***

- ***Mortgage Premium  $MP(t)$***

The mortgage premium at time (t) is a function of the difference between the present value of the cash flows that the borrower will pay (payments) discounted at the market rate  $r^M(t)$  and the present value of the future payments discounted at the mortgage rate  $r^H(t)$  both for the remaining term of the loan (T-t). Mathematically the mortgage premium is calculated as:

$$MP(t) = \frac{\sum_{\tau=1}^{T-t} Payment(t) \left( \frac{1}{1+r^M(t)} \right)^\tau - \sum_{\tau=1}^{T-t} Payment(t) \left( \frac{1}{1+r^H(t)} \right)^\tau}{\sum_{\tau=1}^{T-t} Payment(t) \left( \frac{1}{1+r^M(t)} \right)^\tau}$$

Where  $T$  is the original mortgage term in months and  $Payment(t)$  is the fixed monthly payment of the mortgage.

For simplicity mortgage premium can be calculated as the difference between the mortgage rate and the market rate at time (t) (Deng, Quigley and Van Order, 2000). The expression for mortgage premium can be calculated as follows:

$$MP(t) = \frac{r^H(t) - r^M(t)}{r^M(t)}$$

- **Burnout:** For this work, the variable will take the value 1 if in the last 24 months the mortgage interest rate exceeded in at least 200 basis points the market interest rates in at least 2 occasions.

### ***Payment history***

- **Delinquency:** This category includes a number of variables constructed using information of the loan's payment history in the 12 months prior to observation. The variables use the delinquency history of the loan that can be helpful to establish or analyze behavior patterns.

The following illustrates the construction of some variables using the loan's payment history in the last 12 months prior to observation:

- ***Number of months with 0 payments due in the last 12 months (num12\_0):*** counts the times that the borrower was current in the loan's payments.
- ***Number of months with delinquency > 0 in the last 6 months (mayor\_0\_6m):*** using a shorter time series (6 months) it counts the times the loan had delinquency greater than 0 or was delayed in the payment for more than 30 days.
- ***Maximum delinquency observed in the last 12 months (max12).***