A Model of Prepayment for the French Residential Loan Market

Romuald Elie, Antoine Frachot, Pierre Georges, Pierre Neuvial*

—

Groupe de Recherche Opérationnelle, Crédit Lyonnais, France†

http://gro.creditlyonnais.fr

This version: June 19, 2002

Abstract

In this paper, we propose a prepayment model based on a set of residential loan terminations recorded on the French market. Prepayment is captured by loan-specific variables (loan age, loan type and loan balance) and market-driven variables. We scale down loan age as a percentage of the contractual maturity, so that our model is applicable to any loan, irrespective of its contractual maturity.

We first give a short description of our dataset and, secondly, derive a few historical statistics concerning past prepayments. We then discuss the models we have explored, with a special attention given to the shape of market rate dependence that is best suited to model prepayments. We finally provide calibration results and backtesting validations

*We gratefully thank Paul Demey and Gaël Riboulet for their comments and suggestions on prepayment modelling. All remaining errors are ours.

†*Address:* Crédit Lyonnais – GRO, Immeuble ZEUS, 4^{e} étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; *E-mail:* antoine.frachot@creditlyonnais.fr

1 Introduction

Prepayments have long been investigated in past literature and numerous papers have proposed models that relate prepayment behavior with different variables. Most of these papers focus on the North-American mortgage market while very few cope with European markets (see Charlier[3]) and as far as we know, the only model dealing with the French market is a default model (see Delfour et al.[5]). In the present paper we aim at providing a prepayment model based on a set of residential loan terminations recorded on the French market. The dataset we use is rather unique since it records a large number of loans issued by most prominent French banks.

Our data comes from Crédit Logement, which secures a large share of French banks portfolios. Crédit Logement works as follows: when a bank originates a loan, it may propose its customer to be "cautionné Crédit Logement" which means that Crédit Logement bears the default risk (in place of the lender) in exchange of a fee to be paid by the borrower. If the borrower eventually repays the entire loan without defaulting, he/she is refunded a certain percentage of his/her initial caution while the non-refundable part is calibrated to cover the loss experienced by Crédit Logement on its overall portfolio. This arrangement allows French banks to enhance the rating of their loan portfolios, but on the other hand, banks which arrange their loans that way also hold a stake in Crédit Logement's capital as shareholders. Put together, it means that banks spare a part of their Tier 1 capital which should normally cover their loan portfolios but at the expense of their stake in Crédit Logement. It is not however a zero-sum game: in the end, the total amount of economic capital that banks have to put against default risk is reduced since, by mutualizing banks portfolios, Crédit Logement enables to diversify away default risks which in turn diminishes the amount of unexpected losses. From a more pragmatic point of view, managing a mortgage generates higher operating expenses than securing a loan via Crédit Logement. All these reasons put together the Crédit Logement arrangement might allow to achieve more appealing returns on equity.

Our model is classical in the sense that prepayment is captured by usual variables such as loan age, coupon rate and loan type. Here the loan type refers to specificities of the French market which is segmented into different clusters of loans depending on whether the loan is partially subsidized by state aid or not. Beside these loan-specific variables, other market-driven variables are also included in the model in order to capture the refinancing incentive. As most of the loans issued on the French market are fixed-rate loans, banks have had a long experience of early terminations on their portfolios when interest rates tumble. Market-driven variables are, as usual, the spread between the coupon rate of the loan and market interest rates, i.e. the rate the borrower would be charged if he or she had the opportunity to refinance his/her loan. Finally, no borrower-specific variables are considered in our model, not because age, location, sex etc. do not convey any information about prepayment behavior - of course they do -, but essentially because our model is oriented toward asset-liability management issues. As a matter of fact, most Finance Divisions IT systems in French banks can hardly cope with customer-related variables, such as age, sex, marital status, location etc. For the very same reason, we have voluntarily ignored variables regarding the housing markets even though prepayments are certainly sensitive to housing price movements, but housing prices are not the type of risk (French) banks are accustomed to deal with or hedge against.

Regarding asset and liability management, we aim at improving the traditional (static) gap analysis. In static gap analysis prepayments are at best considered as deterministic functions of the age of the loan while the prepayment function has to be estimated "through the cycle", that is by ignoring the interest rate dependence. This approach is misleading in many respects. First it does not provide accurate forecasts of future prepayments and is therefore likely to miss any wave of prepayment flows resulting from a sharp and lasting fall of market interest rates. Secondly, it entirely ignores the prepayment option embedded in any fixed-rate loan, leading banks to overestimate the market value of their loan portfolios. As a result, banks tend to under-price the loans they originate, by setting to zero the premium of the options they have implicitly sold for free. Finally, as net interest margins can be affected by interest-related prepayments, it is all the more crucial to have a sound prepayment model since it allows to design a simulation scheme on future interest margins in response to various future interest paths. This in turn provides banks with better risk assessment and with a more reliable hedging strategy.

This paper is organized as follows: we first give a short description of our dataset and, secondly, derive a few historical statistics concerning past prepayments. We then discuss the models we have explored, with a special attention given to the shape of market rate dependence that is best suited to model prepayments. We finally provide calibration results and backtesting validations.

2 Data and market description

The "Crédit Logement" database contains more than 1 700 000 loans originated by French banks since January 1980 although the database used for calibration is limited to 1986-2000. It is worth noticing that most of recorded loans are fixed-rate loans, which reflects the traditional predominance of fixed-rate loans in the French market. For each loan, the following information is recorded: loan type, origination date, stated maturity, initial balance, coupon rate and the last payment date. As said before, loan type refers to the state subsidy which may exist at the benefit of the borrower. As a matter of fact, four significant types of loan exist:

- "prêt épargne logement" (PEL). This kind of loan is part of a more comprehensive product called "Plan/compte d'épargne logement". It is divided into two periods: a saving period of at least 4 years during which savings earn an interest rate lower than prevailing market rates and then a borrowing period during which customers are entitled to apply for a loan at a coupon rate fixed at the start of the saving period. This product is rather complicated and it is not worth entering contract details any further. The only thing to be kept in mind is that loans issued from "Épargne Logement" are fixed-rate loans with coupon rates generally below prevailing market rates on loans. As a result interest rate-related prepayments are expected to be lower than on other portfolios as refinancing incentive is often very poor.
- "prêt conventionné" (PC) and "prêt à taux zéro". These loans are partially or fully subsidized. In particular, "prêt à taux zéro" are interest-rate free loans. For obvious reasons, only low-income households are eligible to this type of loans.
- "prêt libre" (PBC, for "Prêt Bancaire Classique"). This last category contains all other loans which do not benefit from any subsidy. They are subsequently sold at prevailing market rates.

Clearly these different loan portfolios deserve specific treatment since each of them is expected to have different refinancing incentive as well as different prepayment behavior due to the characteristics of the population targeted by each class.

It is also worth mentioning that prepayment penalty can be legally demanded by lenders. This penalty amounts to either 3% of the remaining balance (as of at the time of prepayment) or a semester of interests, whichever is the lowest. In practice however many loans are penalty-free essentially because of commercial reasons related to a highly-competitive French market.

In many respects, the Crédit Logement database is much valuable thanks to its historical length and its wide scope of lenders. As far as we know, this dataset has not been explored yet. Unfortunately it is not perfect and suffers from a few pitfalls. First nothing ensures that borrowers belonging to the Crédit Logement portfolio can be considered as representative of the average applicant. As a matter of fact, not all loans are covered by a Crédit Logement guarantee while it is suspected that Crédit Logement is often dedicated to the best-rated applicants. Secondly, as Crédit Logement is mainly concerned with defaults, it does not record refinancing properly. Indeed, a customer being granted a lower coupon rate does not fundamentally change the default risk carried by Crédit Logement which in turn is left uninformed about it. As a consequence, it does not mean that refinancings are not observed in the Crédit Logement database. They actually are as soon as refinancing occurs along with early termination, and this is mainly the case when the borrower quits his/her bank, pays off his/her remaining balance and applies for a new cheaper loan in another bank. As a result refinancings *within* a bank are generally not recorded. Finally, we suspect the Crédit Logement dataset of missing a substantial part of partial prepayments.

Though not perfect, the Crédit Logement dataset is worth being investigated but the resulting model should be used with caution. In fact Crédit Logement should be seen as a convenient way to enrich internal database with external data, provided that all bias listed above are controlled in an effective way. Therefore a bank would not be careful if it tried to use directly our model without confronting our results with its own portfolio.

3 Preliminary statistical analyses

Following is a set of preliminary analyses aimed at investigating how prepayments are related with respect to loan characteristics and market conditions. Prepayment rates at time t are computed as the number of loans terminating at time t divided by the number of loans still alive just before time t. We focus on PEL, "Prêts bancaires classiques" (PBC) and "Prêts conventionnés" (PC) since they represent over 96% of the loans in our database. Table 1 exhibits simple descriptive statistics concerning these three types of loans.

Loan Type	Median duration (months)	Median initial balance (FF)	Average coupon rate (%)	Percentage
PEL	90	16000	5.79	56.4
PBC	144	38000	7.75	31.3
PC	144	28000	8.25	12.3
All	120	23 000	6.71	100

Table 1: *Descriptive statistics*

3.1 Loan type

Exhibit 1 shows the historical evolution of annual prepayment rates from 1998 to 2000 for three types of loans: PEL, PBC, PC. Prepayment rates globally tend to increase with time for the three types of loans. Not surprisingly PEL loans show lower prepayment rates than "Prêts bancaires classiques" and "Prêts conventionnés", which is in accordance with the fact that PEL loans generally have a lower coupon rate. Finally, "Prêts bancaires classiques" and "Prêts conventionnés" have similar prepayment rates.

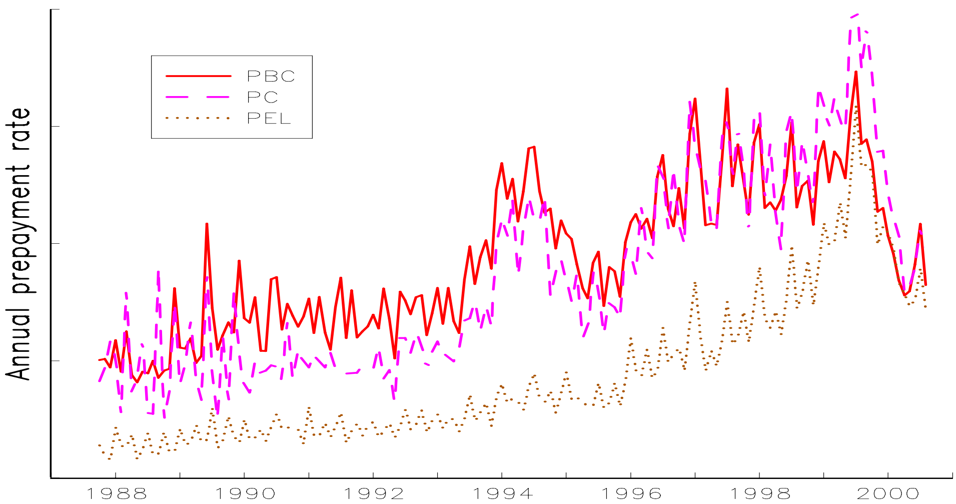

Figure 1: *Prepayment rate and loan type*

3.2 Loan age

Stated-maturities range from 2 to 25 years. Let us mention here that due to Crédit Logement information system, Crédit Logement does not properly record loan terminations that occur during the last year of the contractual life, which explains why prepayment rates fall down one year before contractual termination on figure 2.

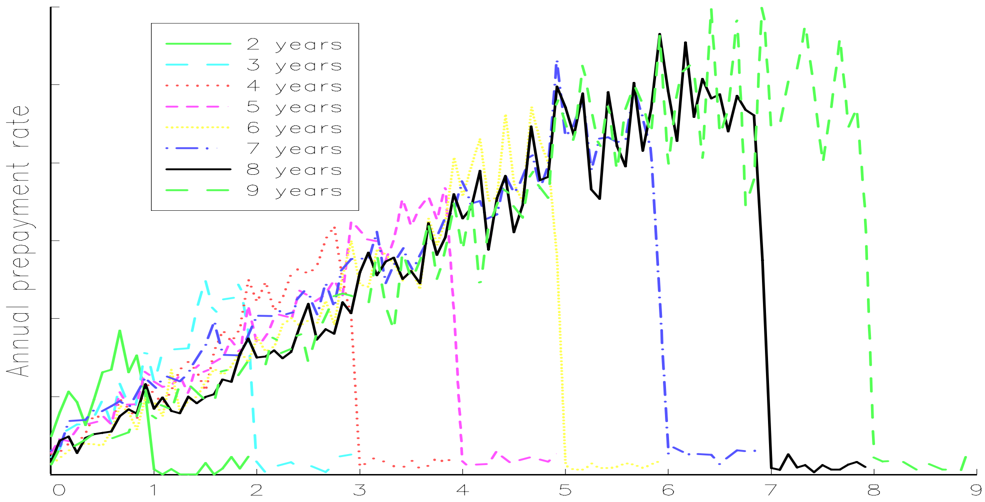

Figure 2: *Prepayment rate and loan age for different maturities*

As loans have different stated maturities, estimating a single curve relating prepayment with loan age is not that obvious. Therefore we scale loan age as a percentage of the contractual maturity *minus one year*. Taking the example of an 11-year loan, a termination occurring at month 24 will be reported as 0.2, that is 20 % of 10 (=11-1) years. So figure 3 shows prepayments with respect to loan age which thus ranges between 0 to 100 %. This is a very convenient way to obtain a single curve which can be used for any loan, at the expense of a slight loss of accuracy, since it would preferable to estimate as many curves as necessary.

Figure 3: *Prepayment rate and loan age— new time scale*

The shape of the curve we obtain is very similar to the Public Securities Association (PSA) curve which is the US industry standard (see Frachot and Gouriéroux[8]). Accordingly prepayment speed is quite small at the beginning of the loan, then increases almost linearly to a sort of long-term level and plateaus until termination (or termination minus one year) (see figure 3).

3.3 Loan balance

Empirical results confirm that loans with high balance show higher prepayment rates than loans with low balance. Figure 4 plots the empirical prepayment rates for each initial loan balance represented in our database. As recently-originated loans show higher prepayment rates than older ones, we compute prepayment rate for any loan balance as the proportion of prepaid loans among those whose contractual termination is less than the last month of record of our database, i.e. those whose contractual termination is before Sept 1^{st}, 2000.

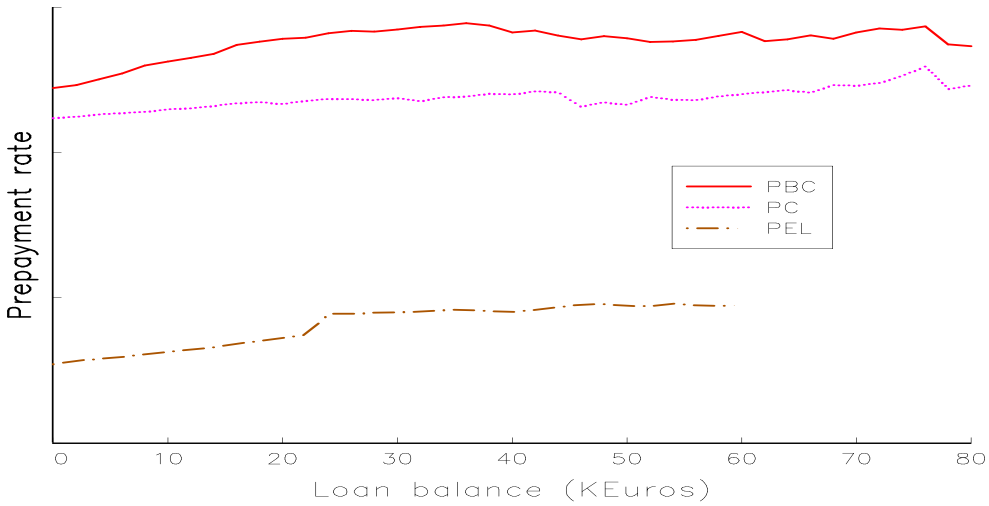

Figure 4: *Prepayment rate and loan balance*

Prepayment rates globally increase with initial loan balance for the three types of loan. Yet while loan balance influence PBC and PC in a similar way, PEL show significantly lower prepayment rates.

3.4 Seasonality

Exhibit 5 plots the average prepayment rate at each month for PBC, PEL and PC. PC and PEL seem to show a quarterly seasonality, while PBC rather show a slight semestrial seasonality. Yet our estimations did not evidence any statistically significant influence of seasonality on the prepayment rate.

3.5 Market interest rate and cost of refinancing

The way refinancing incentive is modelled will be discussed in the next section. Here we just plot the relationship between the prepayment speed and the spread between coupon rate and prevailing market rates. However it is quite a difficult task to define a single index which could capture market rate movements in the loan market. We would rather use the 10-year swap interest rate, assuming that coupon rates for newly originated loans are set to the swap rate plus an approximately fixed commercial margin and a constant default spread. Under this assumption, the spread between coupon rate and current swap rate can be considered as a proxy for refinancing incentive.

4 Prepayment modelling

Prepayment models have two distinct parts, one related to refinancing incentive and the other one which gathers all other causes such as house sale, change of marital status, and so on. Because the two parts are really different in nature, they should deserve specific treatment. The first one should be modelled by choosing an appropriate variable capturing refinancing incentive. The second part is modelled as a function of loan type, initial balance, loan age, considering that it may

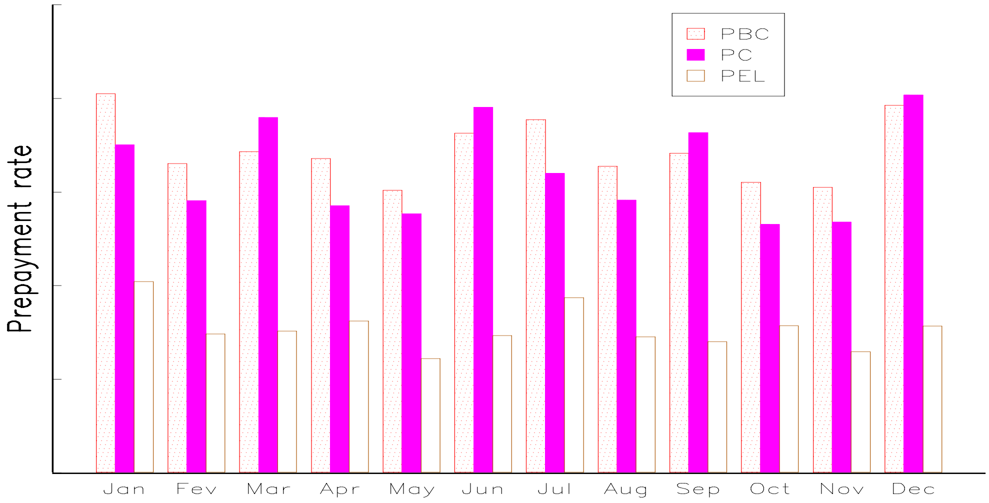

Figure 5: *Seasonality of prepayment rates*

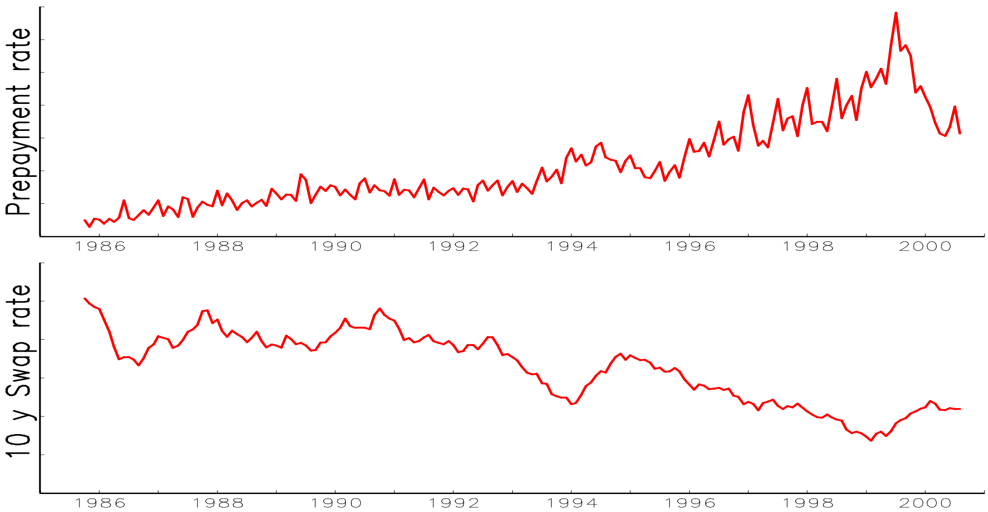

Figure 6: *Prepayment rate and interest rates*

suffice to capture all other prepayment causes unrelated to market interest rates and refinancing purposes.

Furthermore the shape of the speed of prepayment is a very classical one. The prepayment rate at time t, $h(t)$, is expressed as the product of all relevant variables:

$$h(t) = h_0(\text{loan age}) \times g(\text{initial balance, loan type}) \times \text{refinancing incentive}$$

where $h_0(.)$ is the baseline hazard function capturing how the prepayment incentive changes when the loan is ageing. The following subsections describe how the components entering the prepayment rate function have been specified.

4.1 Refinancing incentive

Option theory should be a relevant source for modelling: interest rate-related refinancing should be viewed as the exercise of an interest-rate option by the borrower at a time he/she may find it rational to do so (see Calhoun and Deng[2]). It is traditionally objected that such an approach is useless for practical applications because borrowers are far from being rational, meaning that exercise behavior is unpredictable from one borrower to another. Tentative solutions have been considered by means of a borrower-specific threshold with the following mechanism: the borrower does not refinance as long as market rates remain above his/her specific threshold while he/she does as soon as this threshold is reached. Assuming that individual thresholds are distributed according to a Gaussian distribution (or any other distribution) among the population of borrowers, one can derive the shape of the refinancing incentive for any pool of loans (see Demey, Frachot and Riboulet[6], and Frachot[7]).

However such specifications as well as those proposed in past literature are left for further research. Instead we consider a more classical way of modelling (see Bennett et al.[1]. The shape of the prepayment rate with respect to market interest rate is specified as the exponential of a piecewise linear function of the spread between the loan coupon rate and the prevailing market rate (cf. figure 1). This shape is sufficiently flexible to match almost any kind of dependence with respect to interest rates. In particular, it is well-suited for capturing asymmetrical effects which are empirically encountered as prepayment rate sensitivities behave differently whether interest rates move down or go up.

Our model also considers the time lapse between a shift in market rates and the reactions of the borrowers to this shift. Figure 6 of our descriptive analysis suggests this time lapse typically ranges from 4 to 7 months. Therefore, in order to model prepayment at time t, we compare the loan coupon rate to the average market rate from month $t-4$ to month $t-7$. In short, if $\delta(t)$ denotes the spread between the loan coupon rate and the 10-year swap rate at time t, the refinancing incentive is captured by:

$$\overline{\delta}(t) = \frac{1}{4}\{\delta(t-4) + \delta(t-5) + \delta(t-6) + \delta(t-7)\}$$

Finally, the refinancing incentive is modelled as the exponential of the piecewise linear function $l(\overline{\delta})$:

$$h(t) \propto \exp[l(\overline{\delta}(t))]$$

where $l(x) = \exp[\alpha_1(x-u_1)^+ + \alpha_2(x-u_2)^+]$ [1]

Let us also mention that, theoretically speaking, the difference between the Mark-to-Market and the remaining balance should be a better measure of refinancing incentive. In practice however this variable does not perform better than the interest-rate spread together with the initial balance.

[1] x^+ denotes $x \times 1_{x>0}$.

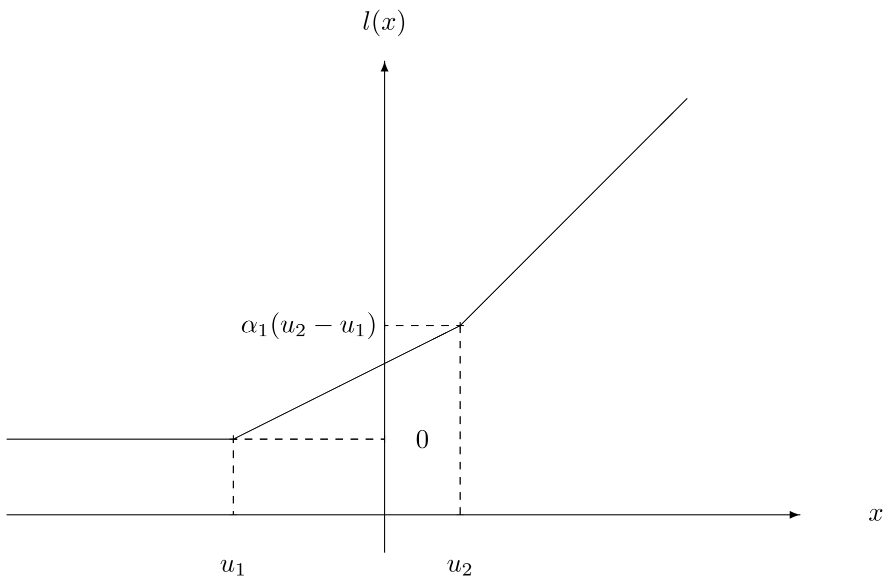

Figure 7: Refinancing incentive

4.2 Loan age dependence

Our aim is to build a model applicable to any loan irrespective of its contractual maturity. To achieve this, one must construct a one-size-fits-all master scale obtained by scaling down loan age as a percentage of the contractual maturity. It prevents from calibrating one model for every class of maturity. In practice, if loan age is d, the effect of age on prepayment rate is taken as a function of $\frac{d}{M-1}$ where M is the initial contractual maturity. With this time-scale, a 5-year maturity loan behaves the same way with respect to age (all other variables being equal) after two years as a 15-year maturity loan after 6 years. Our statistical analyses tend to let us being confident about this assumption, in comparison with the benefits it provides in terms of simplicity of our resulting model.

Using this time scale, one can plot prepayment rates with respect to the ratio $d/(M-1)$. Exhibit 3 shows that loan age dependence may be approximated using a piecewise linear function in exactly the same way as in the industry standard model PSA used in the US market. Formally speaking, three parameters suffice to define the shape of loan age dependence: the prepayment rate right after origination π_0, the time ℓ it takes for the prepayment rate to climb up to its plateau, and the level π_∞ of the plateau itself. In short:

$$h_0(d) = \begin{cases} \pi_0 + \frac{\frac{d}{M-1}}{\ell}(\pi_\infty - \pi_0) & \text{for } \frac{d}{M-1} \leq \ell \\ \pi_\infty & \text{for } \frac{d}{M-1} > \ell \end{cases}$$

In our dataset, ℓ typically ranges from 30% to 40 %.

4.3 Type of loan

Two different models are estimated, one on the PEL portfolio and the other one on both standard loan portfolio (PBC) and subsidized loans (PC), as the last two portfolios do not show significantly different behaviors. However, we add a dummy variable to the model on the standard or subsidized loans portfolio in order to capture the slight differences of prepayment rates between standard and subsidized loans:

$$h(t) \propto \exp(\beta_1 1_{\text{Subsidized loan}})$$

We expect $\beta_1 > 0$ in accordance with the fact that PBC show slightly lower prepayment rates than subsidized loans.

4.4 Initial loan balance

The easiest way to model loan balance dependence is to use a linear function of the initial balance. We ceil the initial balance to 75000 euros (in order to minimize the influence of extreme values).

$$h(t) \propto \exp(\beta_2 \min(\text{Initial Balance},75000))$$

We expect $\beta_2 > 0$ as prepayment rates increase with loan balance (cf. section 3.3). Let us mention here that we have also considered a piecewise linear function of the initial balance in order to capture non-linearities, but it increases the numbers of parameters to estimate, while it provides only poor enhancement in model accuracy.

4.5 Burnout

Burnout effect has been extensively discussed in past literature. In short, burnout effect is evidenced by the empirical fact that pools of loans which have experienced large exposure to refinancing opportunities tend to have lower prepayment rates, other things being equal. This can be explained by assuming that the borrowers who are the most sensitive to refinancing opportunities prepay rapidly as soon as it is profitable. After a period of low refinancing rates, only the least sensitive borrowers still remain in the pool and prepayment rates decay. This phenomenon is well documented in the US market (see Chaudhary, Hayre and Young[4] or Jegadeesh and Ju[9]).

Going beyond intuition is not that easy since sensitivity to interest rates is not a character which is easy to observe for one specific individual. As a result, it is quite difficult to quantify the number of "sensitive" borrowers who leave as time goes by. However as burnout can be related to past experience of interest rate movements, one should observe that the time-t prepayment depends on market interest rates *before* t. More specifically, if past refinancing opportunities have been much appealing then loans which are still alive at present are probably owned by unsensitive people. Once we have acknowledged that time-t prepayment rate should be taken as a decreasing function of the most appealing past refinancing opportunity, one has to provide a proxy for this variable. It casts no doubt that the minimum of past interest rates could play as a good proxy. Unfortunately, we have not evidenced any statistically significant effect of this proxy on the prepayment rate. As a result our model relates prepayment rates to contemporaneous interest rate and is free of any path dependency. Let us note however that most models sold with ALM software are alike.

5 Results

Our model has been estimated with a traditional maximum likelihood procedure on 75 % of our database, while the remaining 25 % are left for back-testing and validation. The log-likelihood is not that obvious in the sense that truncation and censorship must be properly tackled (see Appendix A and B). In this section we give the results of our calibration on the PEL portfolio, and on the PBC and PC portfolio.

5.1 Prêts Épargne Logement (PEL)

We plot the observed prepayment rates and the predicted prepayments rates based on our model (see figure 8). In the declining interest rates environment of the late eighties and the nineties, observed and predicted prepayment rates has increased sharply. As the trend shifts back to higher interest rates at the very end of the nineties, prepayment rates tumble down from summer 1999 to summer 2000.
Figure 8 shows that prepayment rates estimated by our model closely track the observed pattern. The model reacts quite well to the shift in the evolution of interest rates in the late nineties: observed and estimated rates both fall in August 1999, i.e. 6 months after the shift in the interest rates, which leads us to the conclusion that the time of reaction of the borrowers to a shift in the interest rates is well calibrated in our model.
However our model overestimates real prepayment rates during summer 1994. In fact borrowers seemingly did not react to the sharp fall of interest rates in the end of 1993 (see figure 6), whereas they did in 1999: we explain this phenomenon by the increase in borrowers sensitivity to the evolution of interest rates. Our model better fits present day reactions of the borrowers.

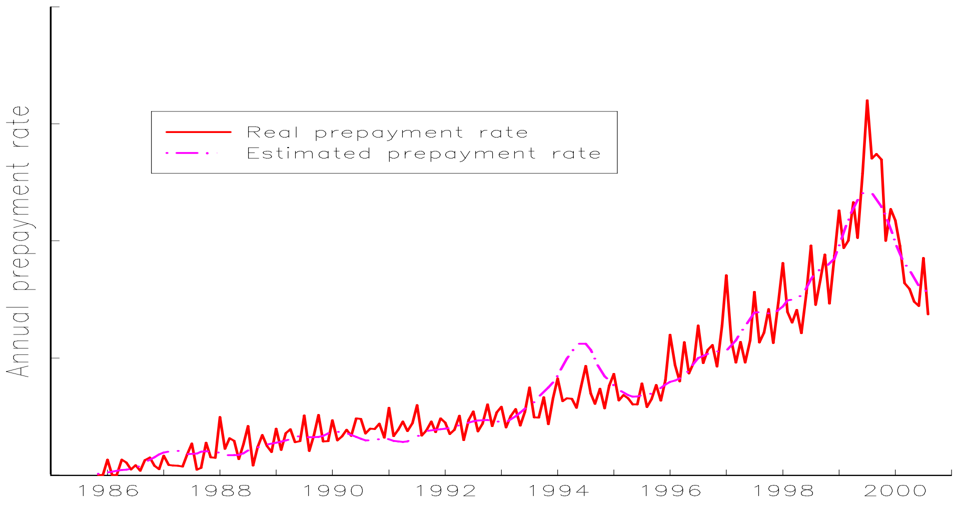

Figure 8: *Goodness of fit of the model — PEL*

5.2 Standard or subsidized loans (PBC and PC)

Just as for PEL, we plot the observed prepayment rates and the predicted prepayments rates based on our model (cf. figure 9). As coupon rate of PBC and PC are generally higher than for PEL, we expect higher prepayment rates than for PEL as well.

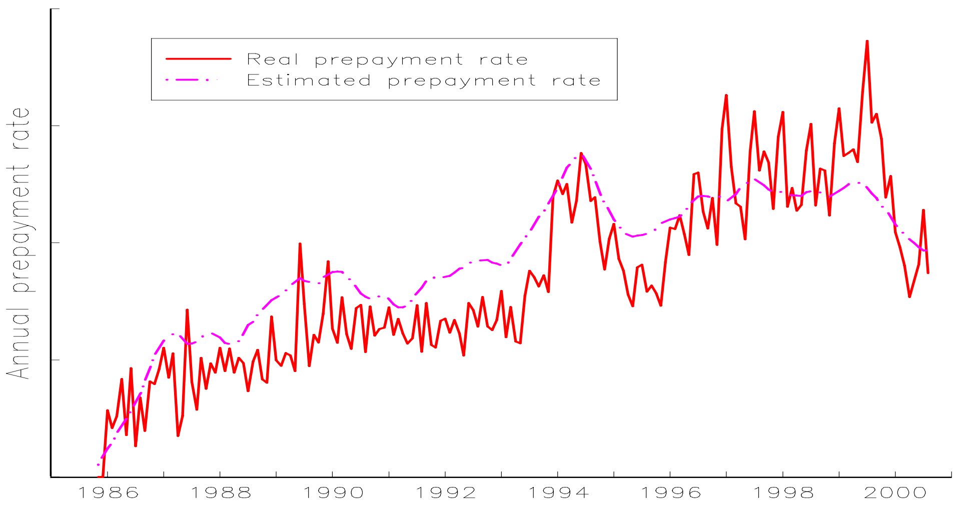

Figure 9: *Goodness of fit of the model — PBC and PC*

Prepayment rates estimated by our model correctly follows the shifts in observed prepayment rates. Yet our model slightly overestimates the prepayment rates until 1995 and fits it better afterwards. Here again this evolution can be explained by the fact that borrowers are becoming more and more reactive to the evolution of the interest rates: had borrowers been as rational as they have been since the late nineties, observed prepayment rates would have been higher.

6 Back-testing

In this section we test the robustness of our model by quantifying the influence of the forecasting time scope and that of the knowledge of the market rates on the goodness of fit of the model.

6.1 Influence of forecasting time scope

Figure 10 plots the observed prepayment rates and the prepayment rates predicted by the model one, two, and three years in advance.

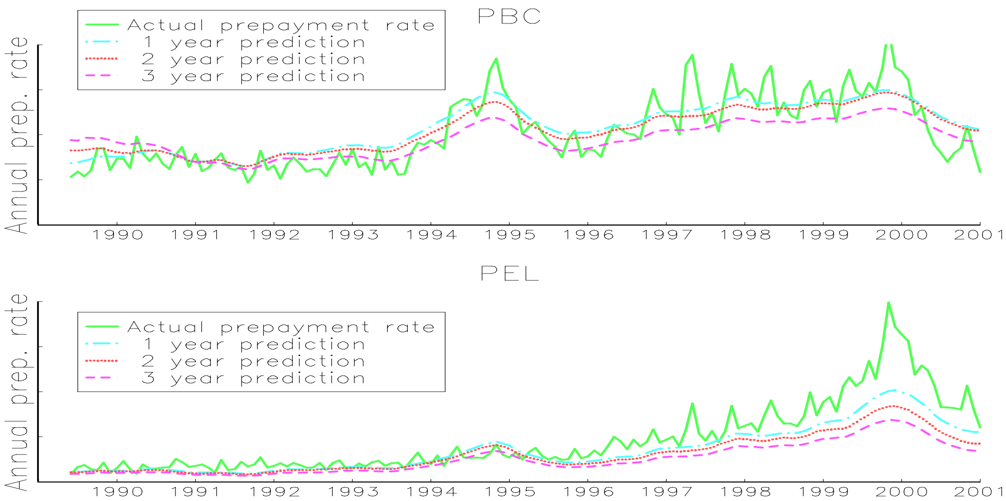

Figure 10: *Influence of forecasting time scope*

6.2 Influence of market rates

Choosing a one-year forecasting horizon, we test the sensitivity of the model to interest rates by comparing the observed prepayment rates to the prepayment rates predicted by the model with actual market rates and with actual market rates ± 20 or 40 bp (see Figure 11).

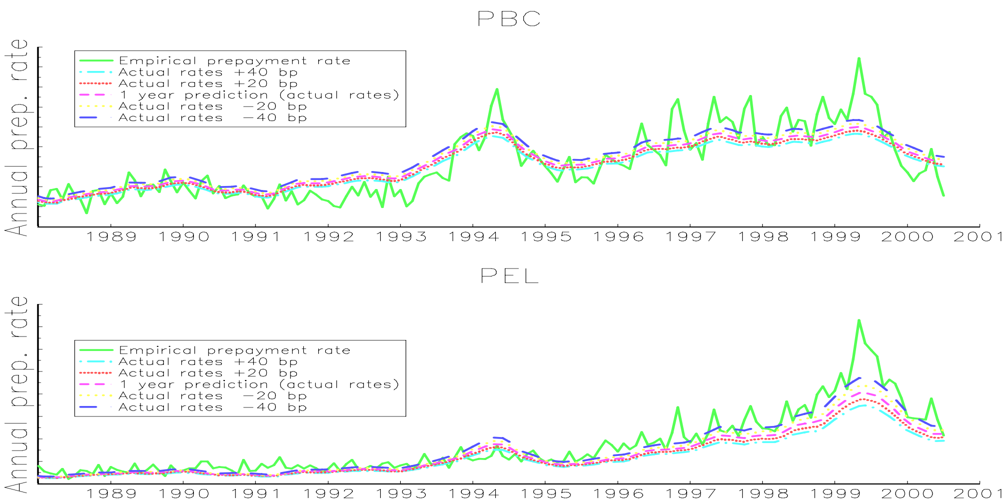

Figure 11: *Influence of market rates*

7 Concluding Remarks

When calibrated on the Crédit Logement database, our prepayment models perform quite well and capture most changes of trend. The remaining issue concerns the quality of the Crédit Logement database itself as it has serious flaws. First it cannot be considered as totally representative of the French residential loan market in terms of credit risk but we are quite confident that the related bias is of second order. More disturbing is the way Crédit Logement receives information on partial prepayments. Though it is certainly a serious deficiency, it can however be corrected in our calibration process provided we dispose of sufficient internal data. Indeed we are currently investigating how to scale both external and internal databases to one another, so as the partial prepayment issue be properly corrected.

Let us finally recall that our models are dedicated to asset-liability management and thus use only loan-specific variables and interest-rate related variables. In particular, neither borrower-specific variables nor real-estate price index are included in the models, as we left this issue for further research. Nonetheless, we strongly believe in the major breakthrough implied by our model in terms of balance sheet management, and we are currently developing a wide set of applications designed to forecasting prepayments and Net Interest Income, hedging embedded prepayment options, pricing residential loans, building a sound fund transfer pricing etc.

Appendix

A Measure of prepayment

In our model, loan duration is denoted by the discrete random variable T, and $f(t)$ represents its probability density function. The survival function $S(t)$ is defined as the probability that a particular loan survives at least until time t, or formally $S(t) = P(T \geq t)$. The hazard function is defined as

$$h(t) = \frac{f(t)}{S(t)}$$

$h(t)$ represents the probability that a given loan will be prepaid between t and $t+1$, conditional on its being still alive at t. Denoting Z_1 the time dependent covariates and Z_2 the time independent covariates, we model the hazard function $h(t)$ as:

$$h(t, \theta) = h_0(t, \theta_0) \times \exp(\theta_1 Z_1(t)) \times \exp(\theta_2 Z_2)$$

where h_0 is the baseline hazard function that captures loan age dependence.

B The likelihood function

The contribution l_i of loan i to the likelihood function depends on whether loan i was still alive at the last period recorded by the database (i.e. loan i is truncated) and on whether loan i prepaid. Therefore, denoting M_i the *contractual* duration of loan i, and t_i its *observed* duration,

$$l_i(\theta) = \begin{cases} f_i(t_i, \theta) & \text{if loan } i \text{ prepaid} \\ S_i(M_i, \theta) & \text{if loan } i \text{ expires without prepaying} \\ S_i(t_i, \theta) & \text{if loan } i \text{ is truncated} \end{cases}$$

Thus, remarking that $t_i = M_i$ if loan i expires without prepaying, $l_i(\theta)$ can be rewritten :

$$l_i(\theta) = f_i(t_i, \theta)^{d_i} \times S(t_i, \theta)^{1-d_i},$$

where $d_i = 1$ if loan i prepaid, and 0 otherwise.
Since $f = hS$, we have $l_i(\theta) = h_i(t_i, \theta)^{d_i} \times S(t_i, \theta)$ and the contribution L_i of loan i to the log-likelihood function is

$$L_i = \log(l_i) = d_i \log(h(t_i, \theta)) + \log(S(t_i, \theta))$$

Besides, the survival function $S(t)$ and the hazard function $h(t)$ have the relationship $S(t) = \prod_{u=0}^{t-1}(1 - h(u))$. As h represents monthly prepayments rates, it is small enough for us to assume $\log(1 - h(t, \theta)) \# - h(t, \theta)$. Thus

$$\begin{aligned} \log(S(t_i, \theta)) &= \sum_{u=0}^{t_i-1} \log(1 - h_i(u, \theta)) \\ &\# \ -\sum_{u=0}^{t_i-1} h_i(u, \theta) \end{aligned}$$

and L_i can therefore be rewritten in function of the history of the covariates, given the specification of the hazard $h(t)$:

$$\begin{aligned} L_i &= d_i \log(h(t_i, \theta)) - \sum_{u=0}^{t_i-1} h_i(u, \theta) \\ &= d_i(\log(h_0(t_i, \theta_0)) + \theta_1 Z_1(t_i) + \theta_2 Z_2(t_i)) - \sum_{u=0}^{t_i-1} h_0(u, \theta_0) \exp(\theta_1 Z_1(u) + \theta_2 Z_2(u)) \end{aligned}$$

Finally, denoting N the total number of loans in our database, the log-likelihood of the model is

$$L = \sum_{i=1}^{N} d_i(\log(h_0(t_i, \theta_0)) + \theta_1 Z_1(t_i) + \theta_2 Z_2(t_i)) - \sum_{u=0}^{t_i-1} h_0(u, \theta_0) \exp(\theta_1 Z_1(u) + \theta_2 Z_2(u))$$

References

[1] P. Bennett, G. Monsen, R. Peach, S. Peristiani, et J. Raiff. *Credit, Equity, and Mortgage Refinancings. FRBNY Economic Policy Review*, pages 83–99, july 1997.

[2] C. A. Calhoun et Y. Deng. *A Dynamic Analysis of Fixed- and Adjustable-Rate Mortgage Terminations. Journal of Real Estate Finance and Economics*, pages 9–33, 2002.

[3] E. Charlier. *Quantifying the Refinancing Incentive and Losses from Prepayments. The Journal of Fixed Income*, pages 55–64, june 2001.

[4] S. Chaudhary, L. S. Hayre, et R. A. Young. *Anatomy of Prepayments. The Journal of Fixed Income*, pages 19–49, june 2000.

[5] O. Delfour, H. M. Heberlein, P. Richard, et E. Rcordeau. *French Residential Mortgage Default Model. Fitch IBCA, Duff & Phelps*, june 2001.

[6] P. Demey, A. Frachot, et G. Riboulet. *Évaluation de l'Option de Remboursement Anticipé. Banques & Marchés*, 2000.

[7] A. Frachot. *A Note on Behavioral Models for Managing Optionality in Banking Books*[2]. 2001.

[8] A. Frachot et C. Gouriéroux. *Titrisation et remboursements anticipés*. Economica, 1995.

[9] N. Jegadeesh et X. Ju. *A Non-Parametric Model and Valuation of Mortgage-Backed Securities. The Journal of Fixed Income*, pages 19–49, june 2000.

[2] http://gro.creditlyonnais.fr/content/rd/rd_gb/alm.htm#paper9