A Note on Behavioral Models for Managing Optionality in Banking Books

Antoine Frachot∗
Groupe de Recherche Opérationnelle, Crédit Lyonnais, France

October 19, 2001

Abstract

Banking books contain numerous implicit options such as prepayment options on mortgages, borrowing options, early withdrawal options etc. As these options may be exercised in response to market interest rate changes, they induce significant (non linear) interest rate risk. We propose here a class of behavioral models which are both tractable and sufficiently realistic.

1 Introduction

Banking books contain numerous implicit options such as prepayment options on mortgages, borrowing options, early withdrawal options etc. As these options may be exercised in response to market interest rate changes, they induce (non linear) interest rate risk that we shall call optionality hereafter. Unfortunately this feature is hardly manageable from an asset/liability management especially when it concerns retail products. The main difficulty comes from the heterogeneous behaviors of retail customers when exercising their options. It is then essential to build behavioral models relating exercise behavior with interest rate movements and to ensure that these models capture the observed heterogeneity.

Behavioral models are essential for ALM purposes. First these models are useful for risk managers to perform dynamic analysis of future cash flows and to estimate the likely path of future net interest income according to various financial scenarios including stress scenarios. Regulators are now enforcing financial institutions to be able to manage this task (see Basel Accord (2001) [1]). Secondly, sound behavioral models are critical for hedging interest rate risk. Thirdly, banks need to mark to market their banking books to comply with future accounting rules. Finally the set-up of a sound transfer pricing mechanism requires reliable behavioral models in order to spread economic-value-added commercial incentives across all business units.

As far as our knowledge two kinds of behavioral models have been proposed in past literature:

- structural models where customer behavior is modeled in a microeconomic way. In other words a specific utility function is assumed. The optimum response to market change is then derived. Unfortunately this class of models is generally subject to rather complicated equations leading to burdensome computations and black-box-type results. Though theoretically appealing, these models are often untractable from an ALM point of view.

∗Groupe de Recherche Opérationnelle, Immeuble Zeus, 4e étage, 90 quai de Bercy — 75613 Paris Cedex 12 — France; E-mail: antoine.frachot@creditlyonnais.fr
• statistical models where customer behavior is modeled in an econometric way. Taking the example of prepayment modeling, these models assume that prepayment rates depend on various factors (including of course market rates) through a simple, easy-to-estimate functional form. Customers’ propensity to prepay appears as (say) the product of various effects related to the age of mortgages, the refinancing incentive etc. These models are computationally easy to implement. As they can be plugged into a dynamical simulation framework they allow for sophisticated dynamical maturity gap analysis of the mortgage portfolio. Their main drawbacks concern their lack of economic foundations raising doubts about their ability to provide a comprehensive and reliable view of the “true” customer behavior.

This classification is comparable in nature with its counterpart in credit risk modeling. Indeed credit risk models fall into two categories: structural models such Merton-like models (or KMV in industry practices) and intensity-based models. The former corresponds to what we name also structural models while the latter refers here to statistical models.

This paper proposes some models designed to gather the advantages of the two approaches while simplifying away their main drawbacks. More specifically the models we propose are economically sound but sufficiently simple to be implemented for solving ALM issues. In short, we model optionality as the optimal exercise of an option with subjective, customer-specific strikes. Customers are therefore “rational” but their rationality differs from one another by their specific strikes whereas differences between subjective strikes capture the individual costs associated with customers’ decisions. We then assume a law of probability for the distribution of subjective strikes among the whole population of customers. Provided that sufficiently simple distributions are considered, a structural, option-based model is obtained but not at the expense of its tractability. Contrary to statistical models, the parameters of these “subjective strike” models have direct interpretation and so experts can gain intuition on their values without performing a heavy calibration process on historical data, which proves useful when historical data are missing. Finally, let us remark that these models appear more like a generalization of existing models, especially from prepayment models where it is now common practice to view the so-called burn-out effect as the result of heterogeneous sensitivities (here: heterogeneous strikes) to market rate changes.

The rest of this paper translates these ideas in the case of three traditional issues regarding optionality in non-maturity deposits, prepayment options, and borrowing options.

2 Non-Maturity Deposit

Non-maturity deposits account for a substantial part of the liabilities of most banks. Typically, the holders of non-maturity deposits are free to increase or lower their balances at any time with possible ceilings on the allowed positive balance. Moreover the interest rate earned on these accounts is generally market-driven but with various speed of adjustment. In some cases like French checking accounts, the interest rate equals zero while some saving accounts (for example CODEVIs) have interest rates set at a national level and with low correlation with respect to market interest rates.

It has been empirically evidenced that aggregate balances move in response to changes in market interest rates. In practice, holders usually target a minimum positive balance on their accounts in order to meet their current liquidity or short-term savings needs. As the interest rate earned on their accounts is significantly below interest rate of other less liquid products, they are not however insensitive to changing market rates. Let us take the example of checking accounts with zero interest rate: when market rates move up, holders are likely to keep their checking accounts at their minimum balance and transfer their unnecessary liquidities or short term savings toward other more profitable assets. Conversely, when market interest rates are low, customers are more likely to accumulate savings in their checking accounts. In particular it implies that aggregate balances may respond asymmetrically to market rate changes. This is particularly the case when the interest rate earned on these deposits is repriced very rarely or fixed for ever (like for French checking accounts).
The sensitivity of aggregate balances vis-à-vis market interest rates is obviously a major source of interest risk in banks booking books and this risk has to be hedged adequately. Building a sound behavioral model of demand deposit is a fundamental requirement to tackle this issue.

2.1 A behavioral model

According to the guidelines described previously customers are defined by their subjective strikes, that is the market rate below which they keep their short term savings in their non-maturity accounts instead of redirecting them to more profitable but less liquid assets. Customers being different from one another, these strikes are not identical and subsequently are assumed to be drawn from a probability distribution whose parameters are to be calibrated on the population of holders. As in Selvaggio (1996) [7], we assume that holders modify their current balance \( b \) to target a level \( b^* \). When market rates are sufficiently low, the balance of the holder increases while “sufficiently” refers to the subjective strike. Mathematically, the non-maturity account balance at month \( t \) of a specific customer of strike \( k \) follows:

\[
b_t - b_{t-1} = \lambda \times [b^* - b_{t-1}] + \beta \times 1 (r_m(t) < k)
\]

where \( b^* \) is the targeted minimum balance, \( \lambda \) is the speed of adjustment toward this targeted balance, and \( \beta \) is the saving flow directed to the checking account when market rates, \( r_m(t) \), are too low (i.e. below customer’s strike). Consequently, \( 1 (r_t < K) \) equals 1 when \( r_t < K \) and 0 otherwise. Customers behave the following way: when market rates are above strike, the second term \( \beta \times 1 (r_m(t) < K) \) vanishes and the account balance converges progressively back to the targeted level \( b^* \). On the other hand account balance increases when market rates stay below strike.

Parameters \( b^*, \alpha, \beta, k \) are customer-specific and one should consider them as different from one customer to another. As this paper is focused on the heterogeneity caused by the different responses to market rate changes, we consider that the strike is the only customer-specific parameter. Subsequently parameters \( b^*, \alpha, \beta \) are set to their mean values. When aggregating equation (1) on the whole bank portfolio, we obtain the time-\( t \) balance \( b_t \) of an “average” customer’s account:

\[
b_t - b_{t-1} = \lambda \times [b^* - b_{t-1}] + \beta \times \int 1 (r_m(t) < k) f(k) \, dk
\]

where \( f(\cdot) \) is the (unknown) probability distribution of strikes among the whole population of bank customers. Examples of possible distributions will be given hereafter. Denoting \( F(\cdot) \) the cumulative distribution function, one can write a (non linear) econometric regression:

\[
b_t - b_{t-1} = \lambda \times [b^* - b_{t-1}] + \beta \times [1 - F(r_m(t))] \]

which can be calibrated using standard techniques.

Examples:

- uniform distribution, i.e. \( F(x) = \frac{x}{K_{\max}} \) for \( x \in [0, K_{\max}] \) and 1 for \( x \geq K_{\max} \) meaning that strikes are uniformly distributed in \([0; K_{\max}]\):

\[
b_t - b_{t-1} = \lambda \times [b^* - b_{t-1}] + \beta \times [K_{\max} - r_m(t)]^+
\]

with the convenient notation \([x]^+ = \max(x; 0)\) and with a slight change of notation (\( \beta \) switched into \( \beta K_{\max} \)). This equation is rather similar to those proposed in past literature except that the response to market interest rate changes is asymmetric. Of course it implies a non linear term instead of a linear one as proposed for example in Selvaggio (1996). Nonetheless it proves to generate only minor complications in the calibration process while providing a better fit of historical data.
• gaussian distribution, i.e $F(x) = \phi(x^m)$ where $\phi(.)$ is the cumulative distribution function of the standard gaussian law. The mean $m$ refers to the mean subjective market rate under which customer fuelled their account with their savings while $\sigma$ measures the degree of heterogeneity among the population.

**Remark 1:** We do not take the attrition rate into account as we only focus on the mean-balance of individual account. We implicitly assume that attrition is a separate, exogeneous process which is unrelated to market rates changes.

**Remark 2:** The previous model may be refined by increasing the number of factors which drive customer behaviors. It is for example well known that customers are likely to make arbitrage between short-term saving and liquidity needs, or between short-term and long-term savings. Capturing both effects would result in a two factor framework where customers would have two different subjective strikes for short-term market rates and for the slope of the yield curve.

### 2.2 Pricing non-maturity deposits

Once estimated, this behavioral model may serve for dynamical maturity gap analysis, for predicting future paths of aggregate balance in conjunction with market rate forecasts, for earnings-at-risk measurement and so on. Less obvious is the pricing issue. It is in fact quite simple within our model if we rely on a no-arbitrage argument like in Janosi, Jarrow and Zullo (1999) \[5\]. It is also worth noticing that our options have some analogy with binary (digital) options.

A no-arbitrage valuation can be performed for a specific customer (whose subjective strike is $k$) and then by integrating individual market values against the distribution of strikes. As such, the total balance is explicitly seen as a portfolio of individual embedded options. Equivalently, one can use the aggregate balance equation (2) and compute its market value. One advantage of our model is that both approaches (i.e. micro and macro) are consistent and provide the same result in a transparent way.

As an illustration, let us compute the individual market value that we shall express in continuous terms for notation convenience. Let us define the deposit account premium at time 0 for the period $[0, T]$ as:

$$V(k) = E \left[ \int_0^T ds \left( r_m(s) - c \right) \times b(s) \times \exp \left( \int_0^s r_m(u) du \right) \right]$$

where $E(.)$ denotes the risk-neutral expectation and $c$ the interest-rate earned on deposit. As a result, $V(k)$ is the market value of the discounted cash flows earned by the bank for a specific account. Using equation (1) in integrated form and denoting $DF(0, s)$ the price of zero-coupon bond maturing at $s$ (in particular, $DF(0, 0) = E \left( \exp - \int_0^0 r_m(u) du \right)$), we can re-arrange the previous equation to obtain (assuming $b(0) = b^*$):

$$V(k) = b^* \left[ 1 - DF(0, T) \right] + \beta \int_0^T ds E \left[ 1(r_m(s) < k) \exp \left( \int_0^s r_m(v) dv \right) \left( R^{(\lambda)}(s, T) - c \right) \int_s^T e^{-\lambda(u-s)} DF(s, u) du \right]$$

where $R^{(\lambda)}(s, T)$ is the time-s rate of a swap of maturity $T$ and whose nominal decays at a speed equal to $\lambda$. $R^{(\lambda)}(s, T)$ is thus defined as ($E_s(.)$ denotes the expectation operator conditionally to the information at time $s$):

$$E_s \left( \int_s^T r_m(u) - R^{(\lambda)}(s, T) \right) \times e^{-\lambda(u-s)} \times \exp - \int_s^u r_m(v) dv \right) = 0$$
The interpretation of equation (3) is the following: cash inflows per unit of time are equal to $\beta$ conditional to market rates being under the strike $k$; then the cash flows vanish progressively at a speed equal to $\lambda$. As a consequence, the bank margin is accrued by $(R^{(\lambda)}(s, T) - c) \times e^{-\lambda(u-s)}$ per unit of time.

The final step would consist in integrating individual market values against the distribution of strikes. Except for very special cases, closed-form formulas cannot be attained and Monte-Carlo pricing is therefore necessary.

3 Prepayment option

We do not discuss here the various factors which drive customers to prepay their mortgages. See for example Hayre, Chaudhary and Young (2000) [4] for a comprehensive study of the causes of prepayment. Instead we focus on modeling the financing incentive using a “subjective strike” methodology in the same fashion as for non-maturity deposits. As a by-product we shall show that our modeling is naturally adapted to capture the so-called burn out effect.

3.1 A behavioral model

Following the same ideas as previously, the refinancing of a mortgage results from exercising an option under subjective strikes. In this case, the “subjective strike” model is closely related to other existing models (see for example Levin (2001) [6]). As usual we assume that mortgagors refinance as soon as the marked-to-market price of their mortgages exceed the remaining balance plus a subjective threshold $k$, i.e.:

$$MtM(t) > B(t) \times (1 + k)$$

where $k$ is the subjective strike expressed as a percentage of the current loan balance $B(t)$. $MtM(t)$ is the market value, that is the sum of remaining future cash flows discounted at the prevailing mortgage rate. By construction, a mortgage has not been refinanced at time $t$ if:

$$\max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right] < k$$

Therefore prepayment occurs at month $t$ if $\frac{MtM(t)}{B(t)} - 1 > k$ conditionally to the fact that prepayment has not occurred before, i.e. $\max_{s \leq t-1} \left[ \frac{MtM(s)}{B(s)} - 1 \right] < k$. Consequently, by aggregating on the whole population of mortgagors, we obtain the prepayment rate, exactly in the same way we computed an average balance of non-maturity deposit in the previous section. Expressed in terms of the cumulative distribution function $F(\cdot)$, the prepayment rate at time $t$ is equal to:

$$h(t) = \frac{F\left(\max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right]\right) - F\left(\max_{s \leq t-1} \left[ \frac{MtM(s)}{B(s)} - 1 \right]\right)}{1 - F\left(\max_{s \leq t-1} \left[ \frac{MtM(s)}{B(s)} - 1 \right]\right)}$$

This formulation is particularly well-suited to capture the so-called burn out effect. This effect is evidenced by the decline of pool sensitivities with respect to market interest rates over time. Indeed prepayment historical records show that after a sharp decline of market interest rates, prepayments may remain quite low if the pool of mortgazes have already experienced previous large exposure to refinancing opportunities. In other words, high-sensitive customers prepay first while the proportion

5
of low-sensitive customers increases in the remaining pool. Here high-sensitive (respectively low-sensitive) customers exactly refer to customers with low (resp. high) subjective strike. This is evidenced by the behavior of our prepayment function: if the refinancing incentive at time \( t \), i.e. \( \frac{\text{MtM}(t)}{B(t)} - 1 \), is below the past maximum incentive, i.e. \( \max_{s \leq t} \left[ \frac{\text{MtM}(s)}{B(s)} - 1 \right] \) then no prepayments occur at time \( t \), i.e. \( h(t) = 0 \). Time-\( t \) prepayments occur only if the time-\( t \) incentive exceeds the past maximum incentive.

In the same spirit as Hayre, Chaudhary and Young (2000) [4], one can construct a burn-out index as the average strike of a pool of mortgages over time. This index is derived in the appendix and it is shown that this index is obviously a non-decreasing process, that is the composition of the remaining pool is getting more and more biased towards less sensitive (i.e. high subjective strike) mortgagors.

In practice prepayments come as a result of other reasons than refinancing. One can think of seasoning, geographical location, seasonality etc. Capturing all other effects can be done easily through a competing-risk framework where a mortgage is still “alive” at time \( t \) only if it has neither refinanced before nor prepaid for other reasons. Assuming independence between refinancing and “non-refinancing”-based prepayment, the survival rate factors into two terms and the total prepayment rate can then be written as (see appendix):

\[
h(t) = 1 - (1 - h_0(t)) \times \frac{1 - F \left( \max_{s \leq t} \left[ \frac{\text{MtM}(s)}{B(s)} - 1 \right] \right)}{1 - F \left( \max_{s \leq t-1} \left[ \frac{\text{MtM}(s)}{B(s)} - 1 \right] \right)}
\]

where \( h_0(t) \) is the base prepayment rate designed to capture all “non-refinancing” based reasons. This form should be compared to usual models where prepayment is modelled as a multiplicative function of a base rate and an incentive index for refinancing. Here modeling is very similar except that the multiplicative decomposition relates to the probability of \textit{not} prepaying.

It is not necessary to detail the prepayment rates corresponding to each possible cumulative functions \( F(.) \). The same examples as for checking accounts modeling still hold. Hence \( F(.) \) can be chosen as the cumulative density function of the uniform law, the gaussian or log-normal laws etc. Each example provides a set of parameters to be calibrated on actual customer data. Once again, one of the main advantages of this approach relates to the fact that this set of parameters has a direct, meaningful economic interpretation such as the mean or maximum strike of refinancing, the dispersion of subjective strikes and so on. This implies that, in absence of a reliable or available database, one can however obtain a sound prepayment model provided that the parameters are set according to practical intuition. As an example, let us imagine that the mean threshold is assumed (by a panel of experts) to be (say) 10 %, that is the average customer prepays when the marked-to-market of its mortgage goes up above 10 % of the remaining balance. Let us further assume that the fraction of “irrational” customers among the whole population amounts to (say) 5 %, that is 5 % of mortgagors have a negative strike. It is then easy to check that one can fit a gaussian cumulative distribution function satisfying these two requirements.

### 3.2 Option pricing

One can price the market value of a mortgage whose termination follows the previous law with a specific subjective strike. Let us consider a fixed-rate mortgage with constant cash-flows. Scheduled cash flows are received by the lender until early termination, that is as long as the market value \( \text{MtM} \) remains below the remaining balance plus the subjective strike. Hence it is very similar to a barrier option. More specifically denoting \( T \) the (random) termination time, the market value of the mortgage with its embedded option can be written as:

\[
V(k) = E \left[ \sum_{s \geq 1} m.1 (T > s) \cdot e^{-\int_0^T r_u \, du} + B(T) \cdot e^{-\int_0^T r_u \, du} \right]
\]
where \( m \) is the (here constant) monthly payment and where, according to our previous framework, \( T \) can be defined as:

\[
T > t \iff \max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right] < k
\]

with \( k \) the subjective strike and:

\[
MtM(t) = \sum_{s > t} m.DF(t, s)
\]

In practice \( T \) is obviously capped by the contractual term to maturity but this cap can be straightforwardly taken into account. \( E(.) \) denotes the risk-neutral expectation.

The total value of all embedded options is obtained by summing up all market values against the distribution of subjective strikes. See Demey, Frachot and Riboulet (2000) [3] for the implementation of pricing formulas.

### 4 Borrowing option

In the following section, we focus on a less widespread implicit option although it is heavily present in french banking books. In the french retail market, this option is embedded in a saving contract. After a saving effort of (say) 4 years, the owner is entitled to apply for a mortgage at an interest rate known at the origination of the contract. Mortgage specifications (such as the nominal) depend on the saving effort which is measured by the total amount of interest earned during the saving period. This product is named Plan d’Epargne Logement in the french market.

This product is rather complex because it combines many optional features generally encountered separately. We do not tackle the probably impossible task of modeling all the implicit options embedded in the product (see for example Baud, Demey, Jacomy, Riboulet and Roncalli (2000) [2]). Instead we restrict ourselves to the borrowing option which is likely to be one of the most valuable options.

#### 4.1 A behavioral model

When the 4 year constrain is satisfied, the owner of the contract is allowed to apply for a mortgage at predefined conditions. Let us note that this right can be exercised at any time, adding an american-style feature. The owner exercises his or her option if the (predefined) rate is below the current market rates for similar mortgages. We then naturally define a subjective strike in the following way:

\[
\text{exercise} \iff r_m > r_c + k
\]

where \( r_m \) is the market rate, \( r_c \) the predefined mortgage rate attached to the contract and \( k \) a subjective strike.

Defining the conversion rate \( h(t) \) as the proportion of eligible applicants who exercise their borrowing options at time \( t \), \( h(t) \) equals\(^1\):

\[
h(t) = F \left( r_m(t) - r_c(t) \right)
\]

In the french retail market, application for PEL mortgages (i.e. mortages related to the contract) can not be rejected by banks provided as the eligibility conditions are satisfied (essentially the 4 year requirement). In particular, rejection for bad rating is forbidden. As a result, PEL mortgages might be under certain circumstances the only source of financing, leading to apparently “irrational” exercise.

Taking a uniform distribution, \( h(t) \) can be written as:

\[
h(t) = 1 - \beta \times \left[ K_{\text{max}} - (r_m(t) - r_c(t)) \right]^+\]

which is, in practice, sufficient to match the historical record of conversion rates. \( \beta \) is thus the minimum spread (between market and contractual rates) demanded by customers.

\(^1\)We limit ourselves to the case where the customer closes its contract in order to immediately apply for a mortgage. In practice, he or she can either close the contract and invest the money elsewhere, or keep closing until market conditions improve.
4.2 Option Pricing

The exercise of the borrowing option costs to his/her bank the difference between the marked-to-market and the nominal amount of the mortgage. Considering the behavioral model described above, the payoff can be expressed as:

\[
[\text{MtM} - B] \times 1 \left( \frac{\text{MtM}}{B} - 1 > \tilde{k} \right)
\]

As we assume that the closing of the contract does not depend on interest rates, the price of this option is thus:

\[
V(\tilde{k}) = E \left[ \sum_u q_u \times [\text{MtM}_u - B] \times 1 \left( \frac{\text{MtM}_u}{B} - 1 > \tilde{k} \right) \right]
\]

where \( q_u \) is the (deterministic) proportion of customers who close their contracts and apply for a mortgage. Integrating this value against the distribution of subjective strikes gives the total market value of the embedded options.
References


Appendix A : A Burnout Index

Let us consider a pool of identical mortgages. Mortgagors still present in a pool at time $t$ represent a fraction $1 - F\left(\max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right] \right)$ of the initial population. As a result, the average subjective strike at time $t$ is equal to:

$$I(t) = \frac{\int_{k > k_{\max}(t)} k dF(k)}{1 - F(k_{\max}(t))}$$

with:

$$k_{\max}(t) = \max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right]$$

As $\max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right]$ is necessarily non-decreasing over time, so is the burn-out index $I(t)$.

Appendix B : Competing Risk Prepayment Rate

Let us define $T_r$ (respectively $T_{nr}$) the (random) time before prepayment due to refinancing reasons (resp. non refinancing related reasons). The actual time before prepayment is then defined as:

$$T > t \iff T_r > t \text{ and } T_{nr} > t$$

$T_r$ is defined such as $T_r > t$ if and only if $\frac{MtM(s)}{B(s)} - 1 > k$ and $T_{nr}$ is defined through a given deterministic hasard rate $h_0(t)$. Assuming independence between the two durations (conditionally to market rates):

$$\Pr(T > t) = \Pr(T_r > t) \times \Pr(T_{nr} > t)$$

which implies:

$$1 - h(t) = \left[ 1 - h_0(t) \right] \times \frac{1 - F\left(\max_{s \leq t} \left[ \frac{MtM(s)}{B(s)} - 1 \right] \right)}{1 - F\left(\max_{s \leq t-1} \left[ \frac{MtM(s)}{B(s)} - 1 \right] \right)}$$