



**SNS REAAL**

SNS REAAL: Group Risk Management

Supervised by Marek Jendrichovsky

**MORTGAGE PREPAYMENTS AND  
ECONOMIC CAPITAL ESTIMATION**

**Boris I. Ederov**

Vrije Universiteit Amsterdam: Faculty of Science

“Business Mathematics & Informatics”

Supervised by Fetsje Bijma

August 2009





## Executive Summary

This report describes a 6-month project at *SNS REAAL*, the major scope of which was to derive a mortgage prepayment modeling and forecasting methodology as well as a simple procedure for the estimation of economic capital for mortgage prepayments of a given mortgage pool.

The report begins with a study of the existing mortgage prepayment models in the literature. In general there are two types of models that are mostly used by financial institutions. The first group is the structural (option-based) models that consider each fixed-rate mortgage as a simple loan plus an option that the customer may prepay his mortgage notional amount (partially or fully) before the contractual mortgage expiration. A major drawback of such a model is the fact that it relies on the rationality of the customers. Option theory assumes that a customer will prepay his mortgage and get a new one every time when incentive for that exists (when he/she may get a new mortgage for a lower rate). Due to the fact that customers very often do not act rationally (they do not follow the mortgage rates regularly, act with delays, etc.), the rational prepayment modeling approaches provide lower accuracy than the other models. They are mainly used, because they are good in terms of functionality - the methodology works even in new economic environments and unseen before development of the mortgage markets. Another thing that makes them preferable is that their predictions are easy to interpret and explain. The second main type of prepayment models is the empirical (econometric) models. Those are used by most of the financial institutions due to their higher accuracy. This accuracy comes at the price of many explanatory variables used and for that reason the outputs of such models are hard to interpret and explain. The performance of the empirical models is doubtful in new economic environments, because the models are trained on past data. There are also mixed (hybrid) models that combine the above-mentioned types, but the literature about those is very scarce and they are usually even more complex than the empirical models.

Second, we proceed with a thorough analysis of the interest-only prepayments at *SNS REAAL* from the last 5 years. This allows us to find out what kinds of data are available in the database and gives us overview how it may be used for a development of a new prepayment model. We group the prepayments by the prepayment reason, we distinguish between refinancings (both external and internal), prepayment due to transfer of property and the remaining (we call those minimal) prepayments. The minimal prepayments are due to

destruction of the property, death or default of the owner. Each prepayment group we divide into sub-groups according to the type of penalty paid (full, partial, latency or none). The analysis helps us to conclude that refinancings are more, but movements are riskier for the bank (due to the fact that in the Netherlands no penalties are paid in case of transfer of property), that curtailments (partial prepayments) may be neglected by a future prepayment model and that the quantity of the available SNS data is insufficient for modeling (for that reason we use also data from *BLG Hypotheken*, which is part of the SNS Group).

Third, we give a complete (mathematical) overview of the models used by the *PRICING*, *ECAP* and *ALM* departments of *SNS REAAL* to define the requirements for a new prepayment model. This overview shows us that all of the departments use expected total conditional prepayment rate, but *ECAP* also uses separately only the expected movement CPR and *PRICING* uses only the expected minimal CPR. As the current proportional hazard model used in *SNS REAAL* forecasts only total CPR, the other rates are now extracted from it in an artificial (and thus not totally correct) way. As a conclusion we need to model refinancings, movement and minimal prepayments separately and then sum these up to get the total prepayment rate.

Fourth, we use a 3-factor Nelson-Siegel model to approximate historical yield curves. The three beta factors represent the level, slope and curvature of the curves. In this same Nelson-Siegel framework we use separate 1-lag autoregressive models to generate future scenarios for each of the three factors (by drawing values from the normal distribution) and thus we generate future scenarios for the complete yield curve.

Fifth, we present an innovative and yet simple yield-curve-based model for forecasting total mortgage CPR. The actual relation of total prepayments and the yield curves is made through refinancings, as minimal and movement prepayments do not depend on the interest rates. We adopt a linear regression for future scenario generation of the refinancing rates, by means of the future scenarios for the beta Nelson-Siegel parameters that we generated earlier. To forecast the minimal and movement prepayments we use autoregressive models (with 3 and 2 lags respectively) and generate scenarios based on the error terms. At the end we combine the scenarios for the 3 separate CPRs into scenarios for the total prepayment rate.

Sixth, for each of the generated future total prepayment scenarios, we estimate the discounted cash-flows of a given mortgage pool. The pool under consideration consists of mortgages of the same type, with similar customer rates and remaining fixed interest periods.

We then use a number of zero-coupon bonds and swaptions to replicate the discounted cash-flows of the pool. For the purpose we use a linear regression to estimate how many of the above mentioned financial instruments we need to include in our replicating portfolio, such that the discounted cash-flows of the mortgage pool match as good as possible those of the replicating portfolio in all generated scenarios. In the example we give we achieve accuracy of 85.7%. This allows us to estimate economic capital for the pool under consideration, based on the replication.

Last but not least, we present an improved prepayment model. The improvement is done in a very simple and straightforward way. The only thing we need to change in the already presented model is the input. Instead on historical interest rates we apply the Nelson-Siegel fitting procedure on the historical mortgage rates (rates for origination of new *SNS REAAL* mortgages in the past). Thus the modeling methodology is the same with the small difference that the beta parameters used before to model refinancings are now different. We also do not provide scenario generation for this model and assume constant mortgage curves for the future. Worst-case scenarios may still be manually plugged in the model. The improvement increases the refinancings model accuracy from 29% to 77%. The modeling of minimal and movement prepayments in the improved model remains the same as in the yield-curve-based one.

## Acknowledgements

I would like to thank Mr. Eelco Scheer for giving me the opportunity to join *SNS REAAL* and for proposing this interesting research topic that allowed me to apply my academic knowledge to satisfy the industry needs. Many thanks go also to my company supervisor Mr. Marek Jendrichovsky and my university supervisor Ms. Fetsje Bijma for their guidance, their permanent availability and helpful suggestions and comments. Thanks to Mr. Frans Boshuizen, Ms. Lydia Snelder and all the others from *SNS REAAL Group Risk Management* for the precious discussions, the help and the support throughout my internship. Another thanks goes to Mr. Toon Jansen from *BLG Hypotheken* for the assistance and the provided historical data, so much needed for the development of the prepayment model. Last, but not least, I thank my family and my beloved one for their persistent support and encouragement.

# Table of Contents

- EXECUTIVE SUMMARY..... 3**
- ACKNOWLEDGEMENTS..... 6**
- TABLE OF CONTENTS ..... 7**
- 1 INTRODUCTION..... 9**
- 2 PREPAYMENT MODELS IN THE LITERATURE ..... 10**
  - 2.1 STRUCTURAL (OPTION-BASED) APPROACHES ..... 10
  - 2.2 EMPIRICAL (ECONOMETRIC) APPROACHES..... 12
- 3 ANALYSIS OF THE INTEREST-ONLY MORTGAGE PORTFOLIO OF SNS REAAL ..... 15**
  - 3.1 THE DATA ..... 15
  - 3.2 GENERAL RESULTS ..... 15
  - 3.3 REFINANCINGS (CATEGORIES 1 & 3) ..... 16
  - 3.4 PROPERTY TRANSFERS (CATEGORY 2) ..... 16
  - 3.5 GENERAL PREPAYMENTS (CATEGORY 4)..... 16
  - 3.6 PARTIAL PREPAYMENTS ..... 16
  - 3.7 THE 3% CAP..... 16
  - 3.8 GENERAL CONCLUSIONS ..... 16
- 4 TYPES OF PREPAYMENT CURVES AND THEIR USAGE BY SNS REAAL ..... 17**
  - 4.1 INTRODUCTION ..... 17
  - 4.2 COMMON FEATURES ..... 17
  - 4.3 ALM DEPARTMENT ..... 17
  - 4.4 ECAP DEPARTMENT ..... 17
  - 4.5 PRICING DEPARTMENT ..... 17
  - 4.6 CONCLUSIONS..... 17
- 5 NELSON-SIEGEL CURVE FITTING AND SCENARIO GENERATION ..... 18**
  - 5.1 INTRODUCTION ..... 18
  - 5.2 THE NELSON-SIEGEL MODEL ..... 18
  - 5.3 FITTING METHODOLOGY ..... 19
  - 5.4 MODELING METHODOLOGY ..... 21
  - 5.5 SCENARIO GENERATION METHODOLOGY..... 21
  - 5.6 NELSON-SIEGEL EMPIRICAL RESULTS ..... 21

<b>6</b>	<b>MODELING PREPAYMENTS BASED ON HISTORICAL YIELD CURVES.....</b>	<b>22</b>
6.1	SCOPE .....	22
6.2	THE OLD PREPAYMENT CURVE .....	22
6.3	THE NEW PREPAYMENT CURVE .....	22
6.4	THE SIMPLIFIED PREPAYMENT MODEL.....	22
6.5	THE TOTAL PREPAYMENT CURVE .....	22
<b>7</b>	<b>A SIMPLE FRAMEWORK FOR REPLICATION OF THE MORTGAGE PORTFOLIO.....</b>	<b>23</b>
7.1	ESTIMATION OF THE MORTGAGE CASH-FLOWS .....	23
7.2	THE PORTFOLIO REPLICATION PROCEDURE .....	23
7.3	THE DISCOUNTED CASH-FLOWS REPLICATION PROCEDURE – EMPIRICAL RESULTS.....	23
7.4	ECONOMIC CAPITAL .....	23
<b>8</b>	<b>THE IMPROVED PREPAYMENT MODEL.....</b>	<b>24</b>
8.1	THE IMPROVEMENT .....	24
8.2	REVISED MODELING OF THE REFINANCINGS – EMPIRICAL RESULTS .....	24
8.3	TESTING THE MODEL OUTPUT AND COMPARISON WITH THE CURRENT PHM .....	24
<b>9</b>	<b>CONCLUSION.....</b>	<b>25</b>
	<b>REFERENCES .....</b>	<b>28</b>
	<b>LIST OF ACRONYMS .....</b>	<b>30</b>
	<b>APPENDIX .....</b>	<b>ERROR! BOOKMARK NOT DEFINED.</b>



# 1 Introduction

With the current development and size of the mortgage markets, mortgage prepayments are a leading risk factor for all financial institutions that operate in the field, including *SNS REAAL*, which is one of the major lenders in the Netherlands. This thesis is written on behalf of *SNS Bank*, within its *GRM* department. *Group Risk Management (GRM)* is a staff department with the following most important tasks:

- measuring and managing risks for the bank and insurance divisions within *SNS REAAL*;
- Capital management;
- Liquidity management;
- Giving the board of directors of the bank, insurance divisions and the group advice on a framework for optimal value creation.

Prepayments have impact on duration gaps and the estimation of economic capital. This thesis thoroughly describes two major issues at *SNS REAAL*, on which we have focused our research during the last 6 months – forecasting mortgage prepayments and the estimation of economic capital for those future prepayments.

The rest of the paper is organized as follows: In chapter 2 we present a literature study on prepayment model types. In chapter 3 we provide a thorough analysis on the mortgage prepayments in the last 5 years. In chapter 4 we describe the models used by 3 different departments of *SNS Group Risk Management (Pricing, ECAP and ALM)* and how the predicted prepayment curve is used by those models, as the purpose is to define the requirements for a new prepayment model. Chapter 5 presents the 3-factor Nelson-Siegel yield curve fitting methodology and the generation of yield curve scenarios for the future by means of simple one-lag autoregressive model. In chapter 6 we present the yield-curve-based prepayment model. The scenario generation and the prepayment model are both used in chapter 7 where we describe the replication procedure and *ECAP* estimation. In chapter 8 we show the improved prepayment model for *ALM* and test it. Chapter 9 concludes and gives recommendations for future improvements.

## 2 Prepayment Models in the Literature

Prepayment models predict unscheduled early principal payments of mortgages. Prepayment rates tend to fluctuate with interest (mortgage) rates and other economic variables. They also depend on the mortgages characteristics like coupon or fixed interest period. Other relevant factors are seasonal effects which in general reflect seasonal variations in housing turnover. The net benefit of the mortgagor due to prepayment (if any) is given by the equation:

$$\left[ (CouponRate - RefinancingRate) \times OutstandingBalance - RefinancingCosts; 0 \right]^+$$

A serious amount of literature is available about modeling mortgage prepayments and most of it is developed for the US market. The market in the Netherlands is quite different, but those models give us thorough insights and a good basis of research. Traditionally the models for valuation of mortgages (and/or mortgage-backed securities) are divided into two groups, based on their type. In recent years also hybrid models (that combine both types) are developed, but the sources are not that extensive. The model types are given below.

### 2.1 Structural (Option-Based) Approaches

These models (continuous or discrete) are based on option-pricing theory and prepayment is linked to the mortgagor's decision to exercise the outstanding embedded prepayment (call) or default (put) options. Usually there are two sources of uncertainty that are incorporated – changes in interest rates and changes in property (housing) prices (the latter is even not incorporated in some approaches). Those models have the problem that they neglect the borrower-specific characteristics (the irrational<sup>1</sup> behavior of the mortgagors) and are consequently called rational models. To overcome this problem in recent years hybrid models are developed. Those mixed approaches basically split prepayments and model incentives by optimal-call method, but behavior by some empirical model (or by incorporating transaction costs, frictions, etc.). Example is given by Archer and Link (1993) and also by Kau and Slawson (2002).

A “must” in the field is the work of Stanton (1995). He models prepayment based on rational decisions of mortgagors and explicitly modeled heterogeneous transaction costs. The model consistently links prepayment and valuation in a single framework and accounts for structural

---

<sup>1</sup> Throughout the document by “irrational” prepayments we mean exogenous (non-optimal) prepayments

shifts in the economy. The model also accounts for burnout dependence without using an exogenous burnout factor. The model fits the data better than the recent empirical model of Schwartz and Torous (1989).

Archer and Ling (1993) are among the first to incorporate heterogeneous transaction costs in their discrete lattice approach. The latter (preferred by the authors instead of continuous-time option-theoretic model) incorporates endogenous call behavior, at the same time allowing multiple levels of borrowers' transaction costs (rational explanation for burnout). As a result the method is flexible in recognizing exogenous prepayments as well (thus it fits also to the "hybrid" models category). A drawback is that even with heterogeneous transaction costs there still exists a single moment for each transaction cost when interest rates hit some critical level. Then all mortgagors with that transaction cost (or lower) will immediately prepay. If a drop in interest rates occurs again after that up to this same level, no prepayment will be observed (all mortgagors have already done so).

A mixed (hybrid) model is presented also by Kau and Slawson (2002). It is a frictions-adjustable theoretical option-pricing model that accounts for the effects of non-financial factors on borrower decisions while simultaneously maintains optimality (pure financial decisions remain the foundation of the model). The model is a strict form of two-factor option-pricing model, based on both house prices and the spot interest rate, with incorporated frictions (including fixed or variable transaction costs, sub-optimal termination, sub-optimal non-termination, etc.) that explain irrationality of mortgagors. Other frictions may be used as well in that same framework.

Kalotay et al. (2004) discuss the shortcomings of previous option-based models and propose an improved one that values mortgages from the homeowner's perspective.

The approach has two distinguishing features:

- The primary focus is on understanding the market value of a mortgage, in contrast with standard models that strive (often unsuccessfully) to predict future cash flows;
- The authors use two separate yield curves, one for discounting mortgage cash flows and the other for MBS cash flows.

A recent PhD thesis of Sharp (2006) presents a new parsimonious option-theoretic model for borrower's irrational prepayment behavior, i.e. results can be obtained outside the scope of simple rational models. That is accomplished by allowing mortgage values to exceed par and by modeling lags in prepayments.

**The Dutch Market:** Prepayment of Dutch mortgages is restricted by laws to a fixed amount (10% or 20% of the remaining notional) per calendar year, which makes the valuation of those mortgages more complicated. Currently the most wide-spread mortgages in the Netherlands are the interest-only mortgages (followed by the savings mortgages). A paper about prepayment of that kind of mortgages is developed by Kuijpers and Schotman (2006). In contrast to the US, optimal exercise of the Dutch prepayment option is not described so far, although there is rich empirical analysis of the behavior of Dutch mortgagors (see van Bussel (1998), Charlier and Bussel (2003), Alink (2002) and Kuijpers (2004)). The reason that US models cannot be directly applied in the Netherlands is path dependence – direct backward valuation is not applicable in case of partial prepayments (due to the fact that the prepayment policy depends on the rate and number of prepayments in the previous years as well). Thus, as a result we need a non-recombining tree to describe and value the embedded prepayment option. For such a tree the number of nodes increases exponentially in time, making optimal prepayment strategy very difficult (for example a binomial lattice approach). Kuijpers and Schotman (2006) also account for the December prepayment peaks, which are typical for the Dutch mortgage market.

## 2.2 Empirical (Econometric) Approaches

In these models the prepayment function is empirically estimated (by means of regression analysis or often within hazard frameworks). In general those models are used to forecast prepayment cash flows. Those models are still preferred in practice for the reason that they account for irrational behavior of mortgagors (due to the fact that they explain the future by the realized historical data). In general simulated (forecasted) interest rates are used as explanatory variables for cash flow projection. Recently the advances in credit risk modeling also motivate a number of papers on prepayment of mortgages. Nakagawa and Shouda (2005) use a structural approach – they define an unobservable prepayment cost process and compare it to the firm value process (used in the default risk modeling literature). Goncharov (2005) uses an intensity-based approach as in reduced-form credit risk models in order to value mortgages. Kau et al. (2004) use a reduced-form intensity-based approach to model prepayment and default behavior for individual mortgages in an explicitly defined proportional hazard framework. They also validate it empirically by calibrating the model to a large data set of historical mortgage market prices.

An important question arises considering econometric models – can we assume that all mortgagors in a pool of same loan characteristics feature independent prepayment and default behavior? Many earlier models use pooled data, but since 1986 the models are

calibrated by using individual loans. It is well known that past refinancing incentives due to low mortgage refinancing rates play a huge role on prepayment speeds at the pool level in both the present and the future – this effect is known as “burnout” (the higher the fraction already prepaid, the less likely will a future prepayment occur at any IR level). To incorporate burnout in the model, it should be added as an explanatory variable (like did Schwartz and Torous (1989) for example). By doing that, they claim they do not need a separate assumption about the pool heterogeneity.

Most empirical models use either past prepayment rates or (and) other endogenous variables like burnout, seasonality etc. in order to explain current prepayment. The purpose is to fit the shape of observed prepayment data unrestricted by many theoretical considerations. In general those models are reduced-form heuristic representations for some underlying process. Thus, it is not clear at all how those models will perform in a different economic environment. For example if the interest rate process changes (or say the mortgage contract terms change) the prepayment process should change as well, but purely empirical models cannot catch the magnitude of that change. Nevertheless many major Wall Street companies have developed their own econometric models. Most of them are based on SMM (Single Month Mortality) that reflects seasonal and age prepayment variations, and housing turnovers. Articles on those models are available, but they do not provide sufficient details about the models themselves (just about the variables used).

Kolbe and Zagst (2006) introduce a prepayment-risk-neutral valuation model for mortgage-backed-securities, based on the proportional hazard model by Kau et al (2004). The new thing is that the general economic environment is especially accounted for in the prepayment process by an additional factor, which fits to the quarterly GDP growth rate. That is why we refer to that model as a hybrid model too.

**The Dutch Market:** Two publications are available and provide thorough and sufficiently explanatory empirical analysis of the mortgage market in the Netherlands (as mentioned already those are Charlier and Bussel (2001), and Alink (2002)). Alink (2002) investigates the variables that have explanatory power on Dutch mortgage prepayments and develops an empirical regression model that uses them. Several regression methods are compared, but the final choice is a logistic regression due to the fact that no aggregation of data is needed. Charlier and Bussel (2001) develop two separate models for interest-only and savings mortgages on a loan-by-loan basis, based on SMM and maximization of log-likelihood. The models allow for exogenous prepayments.

In Hayre (2003) an empirical model based on pool-level data is proposed. The modelling framework involves a separate submodel for each one of the four prepayment causes:

- Home sales (housing turnover)
- Refinancings (old for new mortgage)
- Defaults (foreclosure and subsequent liquidation of the mortgage)
- Curtailments (partial prepayments) & full payoffs

The projections of the four submodels are then summed and to obtain the total projected prepayment rate. Compared to Alink (2002) the presented model is not that statistical, but uses economic and logical arguments instead of data analysis to evaluate the impact of the used variables. Important about the robustness (in economics, “robustness” defines the ability of a financial trading system to remain effective under different markets and market conditions) of the model is the following citation by Hayre (2003): “Therefore, the traditional static statistical model will not work well over time. Instead, to handle changes in the environment or the infrastructure that determine present prepayment behavior, we use time-dependent parameters as necessary”.

### 3 Analysis of the Interest-Only Mortgage Portfolio of SNS REAAL

In this chapter we provide thorough analysis of the prepayments of interest-only mortgages of SNS REAAL in the last five years.

#### 3.1 The Data

Considering the mortgage portfolio of SNS REAAL for mortgages initiated after January 1993 and full prepayments for the period January 2004 to March 2009 (due to the fact that the database has no records of prepayment reasons for earlier prepayments), we are interested in interest-only mortgages. By full prepayments we mean all prepayments that lead to termination of the mortgage contract. Later on we show that partial prepayments are not a significant part of the total prepayments and for that reason we do not include them in the definition of prepayments (except those that are partial, but also full prepayments at the same time).

#### 3.2 General Results<sup>2</sup>

We use a sample of 26936 fixed-rate mortgages (variable-rate mortgages are filtered out, because they are considered riskless, thus the final number). We then categorize the prepayments in four groups, according to the prepayment reason, as recorded in the database (the record used is *aflos\_reden\_cd* and contains the following values – A, E, I, O). Thus we distinguish between the following categories (prepayment reasons):

1. 'A': External refinancing (with another bank);
2. 'E': Transfer of the property (selling the property due to moving to different one);
3. 'I': Internal refinancing (again with SNS REAAL);
4. 'O': General prepayments (for other reasons).

Internal and external refinancings are considered separately by the used MATLAB program, but are reported together (averaged) throughout the report, because for now we make no distinction between both.

For each of the above categories we separate the mortgages in subgroups, depending on the type of penalty paid. For that reason we use the database record *decode\_omschr1*. Thus for each main category we distinguish between the following types of penalties:

---

<sup>2</sup> The MATLAB program output is in part 1.1 of the appendix, exact numbers may be found there.

1. Full penalty [Volledige boeteberekening] (penalty on the whole original notional amount of the mortgage);
2. No penalty [Geen boeteberekening];
3. Latency penalty<sup>3</sup> [Alleen extra rente i.v.m. te late aankondiging] (penalty due to late notification of the bank – less than 30 days in advance);
4. Partial penalty [Alleen boete over niet-vrij aflosbaar] (penalty paid on the difference between the original notional and the fine-free part of the notional, which in general is either 10% or 20% of the first).

### **3.3 Refinancings (Categories 1 & 3)**

### **3.4 Property Transfers (Category 2)**

### **3.5 General Prepayments (Category 4)**

### **3.6 Partial Prepayments**

### **3.7 The 3% Cap**

### **3.8 General Conclusions**

---

<sup>3</sup> LatencyFine = MortgageRate x (30d – actual number of days till prepayment) x OriginalNotional.



## **4 Types of Prepayment Curves and their Usage by SNS REAAL**

### **4.1 Introduction**

In order to model the prepayments of SNS REAAL's mortgage portfolio we need to know what the requirements are for such a model. For that purpose in this chapter we discuss the types of prepayments (prepayment curves) used by three departments of SNS REAAL: *PRICING*, *ECAP* (Economic Capital) and *ALM* (Asset & Liability Management). Our scope is to get better understanding of the current models that those departments use for their own purposes and of the requirements they have relative to those prepayments. It is crucial to explain how those different types of prepayments are used for further calculations (mathematical formulae will be provided as well).

The remaining part of the chapter is structured as follows: first we state some common facts about the three models of the three different departments and then in separate parts we describe those models and the role that prepayment rates play in them (occasionally we also share our opinion about the correctness of those models and give recommendations for future improvements if any, based on the analysis and statistics of the mortgage portfolio); in the last part we provide conclusions about the requirements that the modeled prepayments should fulfill and suggestions about how to model them in (hopefully) a very simple way.

### **4.2 Common Features**

### **4.3 ALM Department**

### **4.4 ECAP Department**

### **4.5 PRICING Department**

### **4.6 Conclusions**

## 5 Nelson-Siegel Curve Fitting and Scenario Generation

In this first part we give a proposal for a long-term (until maturity of a mortgage pool) yield curve scenario generation. The methodology presented is the one adopted by Diebold & Li in [9]. Below we briefly describe the method in theory without going into the details.

### 5.1 Introduction

We propose a variation of the Nelson-Siegel exponential components framework that models the entire yield curve, period by period, as a three-dimensional parameter, which evolves dynamically. The three time-varying parameters of the 3-factor Nelson-Siegel model may be interpreted as level, slope and curvature factors. We propose autoregressive models to model those factors separately and produce forecasts, which are more accurate at long horizons than most standard benchmark forecasts.

### 5.2 The Nelson-Siegel Model

The 3-factor Nelson-Siegel yield curve is given by the equation<sup>4</sup>:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t} \left( \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} \right) + \beta_{3,t} \left( \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau) \right),$$

where we interpret the parameters as:

- $\tau$  is the desired maturity;
- $\lambda_t$  governs the exponential decay rate. Small values of the parameter produce slow decay and can better fit the curve at long maturities and vice versa. A way to determine lambda is to estimate it for each month in the past by means of optimization numerical method like Newton-Rhapon for example and then consider simply the average value (we will see that this is convenient indeed in order to use a single lambda only). Thus the fits will be comparable in terms of betas. The lambda estimation methodology is given in the next chapter;
- $\beta_{1,t}$  may be interpreted as a long-term factor (level factor);
- $\beta_{2,t}$  may be considered a short-term factor (slope factor);
- $\beta_{3,t}$  may be considered a medium-term factor (curvature factor).

---

<sup>4</sup> The factorization provided is different than the original one proposed by Nelson & Siegel in [21], but is preferable (for reasons see [8]).

### 5.3 Fitting Methodology

To fit the 3-factor model and estimate the historical values of the model parameters we use a MATLAB fitting function. To estimate the beta for each month in the past we need to first estimate the parameter lambda. To solve for lambda we may use for example Newton-Rhapon optimization method to solve the non-linear least squares (non LLS) problem:

$$\min_{\lambda_t, \beta_t} \sum_{i=1}^N (r_{i,t} - y_t(\tau_i))^2,$$

where:

- $r_{i,t}$  are the historical rates at month  $t$  with maturities  $\tau_i$  as given in part 1 of the appendix;
- $\beta_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t})^T$  is the vector of beta parameters;
- $N$  is the number of rates we try to fit (number of maturities).

A weakness of this approach is the existence of multitude of local minima for the optimization problem. Thus it is oversensitive to the initial guess of the parameters lambda and beta.

We see that for a given lambda the curve equation is actually linear in each  $\beta_i$ . Therefore for a given lambda we can solve the linear least squares problem for beta and estimate the unique minimum:

$$\min_{\beta_t} \sum_{i=1}^{21} (r_{i,t} - y_t(\tau_i))^2$$

through the orthogonal projection on the observation vector  $r_t = (r_1, \dots, r_{21})^T$ , because we need:

$$\hat{\beta}_t = (X_{\lambda_t}^T X_{\lambda_t})^{-1} X_{\lambda_t}^T r_t,$$

where:

$$\bullet \quad X_{\lambda_t} = \begin{pmatrix} \phi_{\lambda_t,1}(\tau_1) & \phi_{\lambda_t,2}(\tau_1) & \phi_{\lambda_t,3}(\tau_1) \\ \phi_{\lambda_t,1}(\tau_2) & \phi_{\lambda_t,2}(\tau_2) & \phi_{\lambda_t,3}(\tau_2) \\ \dots & \dots & \dots \\ \phi_{\lambda_t,1}(\tau_{21}) & \phi_{\lambda_t,2}(\tau_{21}) & \phi_{\lambda_t,3}(\tau_{21}) \end{pmatrix} \text{ with}$$

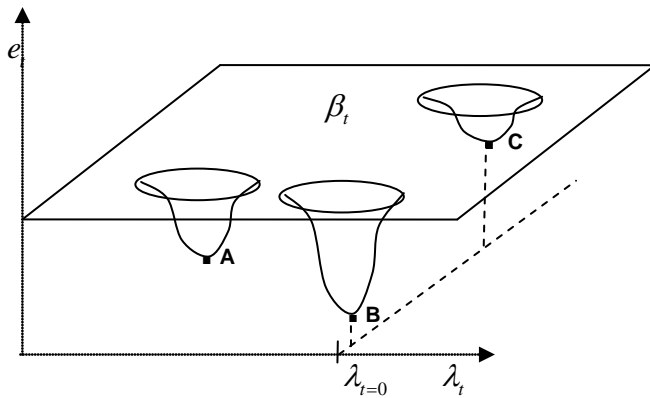
$$\phi_{\lambda_t}(\tau) = (\phi_{\lambda_t,1}(\tau), \phi_{\lambda_t,2}(\tau), \phi_{\lambda_t,3}(\tau)) = \left( 1, \left[ \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} \right], \left[ \frac{1 - \exp(-\lambda_t \tau)}{\lambda_t \tau} - \exp(-\lambda_t \tau) \right] \right).$$

Thus we transform the above non LLS problem into a new problem in just one dimension (for lambda):

$$\min_{\lambda_t} \sum_{i=1}^{21} (r_{i,t} - y_t(\tau_i, \beta_t = (X_{\lambda_t}^T X_{\lambda_t})^{-1} X_{\lambda_t}^T r_{i,t}))^2 .$$

What we have done is to translate the general 4-dimensional optimization problem into a 1-dimensional (for lambda) one, which is shown on the figure below and:

- $e_t = \sum_{i=1}^{21} (r_{i,t} - y_t(\tau_i))^2$  is the error the problem with three minima A, B and C;
- $\lambda_{t=0}$  is the value of lambda for which we have the minima B and C (C is only a local extremum and B is the global one). The generic search algorithm will be easily fooled to converge to C instead to B, when the initial guess of  $\lambda_t$  is closer to C.

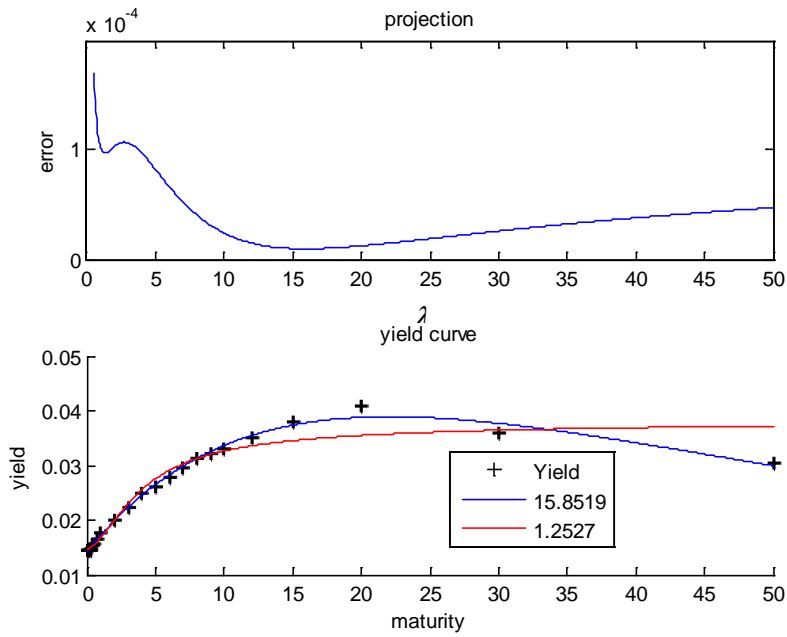


To avoid this issue we translate the problem into 1-dimensional setting, actually projecting the 4-dimensional error function  $e_t$  onto the  $(e_t, \lambda_t)$  plane, thus having for each lambda the unique minimizing vector beta. See an example figure below (note that maturities are in

years instead of months as we will transform them later). Two things are to be noticed:

- The projection shows two optimal lambdas for this particular yield curve (each of them has its corresponding beta of course);
- The fitted curves for both lambdas are given in the plot and we see that the best fit is the blue one for lambda equal to 15,85 (in case of the picture only).

To illustrate the relation to the previous figure, take for example  $\lambda_{t=0}$  there and notice that the improved method would never choose C (the local minimum) instead of B (the global one), because the projection of B on the  $(e_t, \lambda_t)$  plane is lower than that of C. Important issue then is of course again the initial guess of lambda. This we can estimate according to the data under consideration and is given in the empirical results. In case of the above figures a good starting lambda is  $\lambda_t = \lambda$ .



Fitting will provide us with one value of  $\beta_t = (\beta_{1,t}, \beta_{2,t}, \beta_{3,t})^T$  for each month of the past. According to [9] we know: “Clearly the 3-factor model is capable of replicating a variety of yield curve shapes: upward sloping, downward sloping, humped and inverted humped.”

#### 5.4 Modeling Methodology

#### 5.5 Scenario Generation Methodology

#### 5.6 Nelson-Siegel Empirical Results

## **6 Modeling Prepayments Based on Historical Yield Curves**

### **6.1 Scope**

It is known that the prepayment curve is an important input for the models of three separate departments of *SNS Bank*, namely *PRICING*, *ECAP* and *ALM*. The specific requirements for the types of prepayment curves and the way those are used by these same three departments are described in chapter 4. Exactly these requirements will we try to meet when modeling the mortgage portfolio prepayments.

### **6.2 The Old Prepayment Curve**

We specify here what the main issues with the usage of the current prepayment PHM (Proportional Hazard Model) are:

### **6.3 The New Prepayment Curve**

### **6.4 The Simplified Prepayment Model**

### **6.5 The Total Prepayment Curve**

## **7 A Simple Framework for Replication of the Mortgage Portfolio**

### **7.1 Estimation of the Mortgage Cash-Flows**

### **7.2 The Portfolio Replication Procedure**

### **7.3 The Discounted Cash-Flows Replication Procedure – Empirical Results**

### **7.4 Economic Capital**

## **8 The Improved Prepayment Model**

### **8.1 The Improvement**

### **8.2 Revised Modeling of the Refinancings – Empirical Results**

### **8.3 Testing the Model Output and Comparison with the Current PHM**



## 9 Conclusion

In this thesis we used a 3-factor Nelson-Siegel model to interpret yield curves as a combination of level, slope and curvature parameters. Those we used to explain past mortgage refinancings and to forecast expected future ones. We also used those parameters to generate yield curve scenarios for the future. Based on these future scenarios and on the simple prepayment model presented, we developed a mortgage pool replication technique that defined the estimation of economic capital for prepayments in a basic framework.

As the major purpose of this graduation internship was to develop a simple and easy to interpret, but yet accurate prepayment model for the *ALM* department of *SNS REAAL*, we also introduced an improved methodology, this time based on historical mortgage curves. We do not change the methodology, we again use Nelson-Siegel to fit the curves from the past and use the three parameters to explain past refinancings and forecast future ones.

An issue that still needs to be further investigated is a way to incorporate the more accurate, improved prepayment model (instead of the yield-curve-based one) in the mortgage pool replication. At this time it is impossible due to the fact that this improved model does not depend on the yield curve (and interest rates in particular), but the replicating portfolio does. For this reason we need to use much more inaccurate prepayment model in the replication procedure (the yield-curve-based one). A possible solution may be to use yield curve Nelson-Siegel parameters in the refinancings regression, but include all separate spreads (put on top of the yield curve to get the mortgage curve) as well. In that way we will keep the relation between prepayments and yield curve scenarios, but we will account for the enormous (unseen in the past) spreads that we observe nowadays, because of the financial crisis. The reason why this is not our adopted methodology is that currently we do not have the historical spreads in the database (the only available spread is the liquidity one). Thus, upon finding better data, the replication procedure described in this thesis may be improved. Other possibility would be to decompose the mortgage curve into yield curve and spread curve and apply Nelson-Siegel on both, then include all parameters in the refinancings regression.

Another improvement of the replication procedure would be the matching of all mortgage cash-flows with those of the replicating instruments (and not only the discounted ones). It should be clear though, that this will add another dimension to our replication and will substantially increase the calculation efforts and time. We already proposed a specific way for replication of all cash-flows (namely the LP method), but genetic algorithms may be used

as well. Another possible improvement is the usage of index amortizing swaps in the replicating portfolio.

We used only 85 observations to model refinancings in the improved methodology (as mortgage rates are available since 12.2001) and for that reason proper backtesting was hard to do. Thus we recommend additional backtesting in the future upon gathering of more historical mortgage curves.

We also recommend the improvement of the incentive calculation from part 2 of the appendix by means of using the later on discovered historical mortgage curves (see chapter 8) instead of the linear interpolation methodology adopted due to absence of better data. That will also improve the output of the prepayment analysis program (chapter 3).

Before we come to the end of this thesis, we would like to make some conclusions about the decisions made throughout the research and the significance of the described results for *SNS REAAL*. In the last number of years the *ALM* department of *SNS Bank* has been using a quite complex empirical prepayment model – the proportional hazard one. The main reasons for which *ALM* is not satisfied with the PHM and for the initiation of this research are three. First, the recalibration of the model is too complex and time-consuming. Second, the model output (forecasted prepayment curve) is hard to explain (interpret), in the sense that the relations between the different explanatory variables, used by the PHM, and the CPR are unknown. Third, two other sub-departments of *Group Risk Management* use separate components (parts) of the total prepayment curve, namely the minimal and movement prepayments, but those are not outputs of the current PHM (all drawbacks of this model we provided in chapter 6). The improved prepayment model described in chapter 8 satisfies all these major requirements: it provides a linear framework and perfect relation between the forecasted prepayments and prepayment drivers; it is recalibrated in just minutes and it may be used by 3 departments instead of *ALM* only. On the other hand the model we provide is not a standard model available in the literature, most of which are very complex, striving to achieve the highest possible accuracy. Just because our requirements were of different nature (unlike accuracy for example), we needed to “tailor” our proposed model to suit them. Thus we developed a simple and practical model with good accuracy which had been accepted by *ALM* with enthusiasm.

Apart from the model itself, the prepayment statistics in the report provide a good basis for analysis of the behavior of mortgagors in the Netherlands. Its MATLAB code may also be used (with slight adjustments) to provide the same analysis for each specific mortgage type

the bank sells, which offers us a plausible way to create expectations for the future development of these specific mortgages<sup>5</sup>.

Chapter 4 gives an overview of the models used by three different departments of *SNS REAAL* – *ECAP*, *PRICING* and *ALM*. The description is detailed, but also easy to understand due to the fact that we keep the same notation throughout the whole chapter, which makes it a perfect tool for people that need to familiarize themselves with more than one of the models. Also the yield curve scenario generation provided in chapter 6 may be used separately in all kind of market risk models for analysis of worst case scenarios and stress testing. The *ECAP* estimation methodology provides a basis for comparison between its output and the output of the current *ECAP* model for prepayments, which brings further security for the bank.

Thus, the research underlying this thesis was a sequence of accomplishing separate tasks that were logically linked. Each of them is significant for *SNS REAAL* in its own way, but also in combination with the others.

---

<sup>5</sup> That was even done for the so called “Rentedermpet” and “Plafondrente”

## References

- [1] Alink B. (2002): *“Mortgage Prepayments in the Netherlands”* - Proefschrift Universiteit Twente.
- [2] Archer W. & Ling D. (1993): *“Pricing Mortgage-Backed Securities: Integrating Optimal Call and Empirical Models of Prepayment”*.
- [3] Boshuizen F.: *“CPR en verhuiskans in EKAP en Pricing”* – Internal report.
- [4] Boshuizen F.: *“Beschrijving EKAP embedded opties model – verhuisoctie”* – Internal Report.
- [5] Boshuizen F. & Akker R. van den & Kruk R. van der & Dijk B. van & Veldhuizen D.: *“Optieprijsmodel Hypotheken Pricing Bank”* - Internal report.
- [6] Boshuizen F. & Buursink A. & Ederov B. & Jendrichovsky M. & Kruk R. van der & Meijer P. & Ploegh A.: *“20090730 Eindrapport Embedded opties”*.
- [7] Bussel A. van (1998): *“Valuation and interest rate risk of mortgages in the Netherlands”* - Proefschrift Universiteit Maastricht.
- [8] Charlier E. & van Bussel A. (2001): *“Prepayment Behavior of Dutch Mortgagors: An Empirical Analysis”*.
- [9] Diebold F. & Li C. (2005): *“Forecasting the term structure of government bond yields”* - Journal of Econometrics 130 (2006) 337-364.
- [10] Goncharov Y. (2005): *“An Intensity-Based Approach to the Valuation of Mortgage Contracts and Computation of the Endogenous Mortgage Rate”*.
- [11] Hayre L. (2003): *“Prepayment Modeling and Valuation of Dutch Mortgages”*.
- [12] Hilliard E. & Kau J. & Slawson V. (1998): *“Valuing Prepayment and Default in a Fixed-Rate Mortgage: A Bivariate Binomial Options Pricing Technique”*.
- [13] Jendrichovsky M. (2009): *“The Probability of Movement in the ECAP Prepayments Model”* - Internal report.
- [14] Kalotay A. & Yang D. & Fabozzi F. (2004): *“An Option-Theoretic Prepayment Model for Mortgages and Mortgage-Backed Securities”*.
- [15] Kau J. & Keenan D. & Muller W. & Epperson J. (1992): *“A Generalized Valuation Model for Fixed-Rate Residential Mortgages”*.

- [16] Kau J. & Keenan D. & Smurov A. (2004): “*Reduced-Form Mortgage Valuation*”.
- [17] Kau J. & Slawson C. (2002): “*Frictions, Heterogeneity and Optimality in Mortgage Modeling*”.
- [18] Kocherlakota R. & Rosenbloom E. & Shiu E. (1988): “*Algorithms for Cash-Flow Matching*” – Transactions of Society of Actuaries vol.40 pt.1.
- [19] Kolbe A. & Zagst R. (2006): “*A Hybrid-Form Model for the Prepayment-Risk-Neutral Valuation of Mortgage-Backed-Securities*”.
- [20] Kuijpers B. (2004): “*Mortgage valuation and the term structure of interest rates*” - Proefschrift, Universiteit Maastricht.
- [21] Kuijpers B. & Schotman P. (2006): “*Valuation and Optimal Exercise of Dutch Mortgage Loans with Prepayment Restrictions*”.
- [22] Nakagawa H. & Shouda T. (2005): “*Analyses of Mortgage-Backed Securities Based on Unobservable Prepayment Cost Processes*”.
- [23] Nelson C. & Siegel A. (1987) “*Parsimonious Modeling of Yield Curves*” – Journal of Business 60 473-489.
- [24] Roelofs T.: “*Prepayment Modellen: Hoe nu verder?*” – Internal report.
- [25] Schwartz E. & Torous W. (1989): “*Prepayment and the Valuation of Mortgage-Backed Securities*”.
- [26] Schwartz E. & Torous W. (1992): “*Prepayment, Default and the Valuation of Mortgage Pass-Through Securities*”.
- [27] Sharp N. (2006): “*Advances in Mortgage Valuation: An Option-Theoretic Approach*”.
- [28] Stanton R. (1995): “*Rational Prepayment and the Valuation of Mortgage-Backed Securities*”.
- [29] Veldhuizen D. (2008): “*An Alternative for Hedging Pipeline Risk and Prepayment Risk of Mortgages, Using Forward Starting Swaptions and Index Amortizing Swaps*” – Proefschrift, Vrije Universiteit Amsterdam.
- [30] Werf K. van der: “*Waardering Prepayment-optie hypotheek*” – Internal report.

## List Of Acronyms

ACF – Autocorrelation function

ALM – Asset and liability management

AR – Autoregression

ARMA – Autoregressive moving average

CF – Cash-flow

CPR – Conditional prepayment rate

DCF – Discounted cash-flow

ECAP – Economic capital

MSE – Mean squared error

OLS – Ordinary least square

PACF – Partial autocorrelation function

PHM – Proportional hazard model