Refinancing Mortgages in Switzerland

Karl Frauendorfer and Michael Schürle
Institute for Operations Research, University of St. Gallen
Bodanstrasse 6, 9000 St. Gallen
www.ifu.unisg.ch

Working-Paper 02-09-1
Refinancing Mortgages in Switzerland

Karl Frauendorfer and Michael Schürle
Institute for Operations Research, University of St. Gallen, Switzerland
14.09.2002

Abstract

This paper presents a multistage stochastic programming model for refinancing mortgages with non-contractual maturity under liquidity restrictions in the market. An extension to the management of other products such as savings accounts is straightforward. The evolution of interest rates is modelled by principal components for short-term and a two-factor mean reversion model with long rate and spread for long-term planning. Barycentric approximation provides tight lower and upper bounds for the original problem with relative discretization errors in the order of one per cent.

1 Introduction

Savings accounts, sight deposits and mortgages make up a significant portion of a bank’s assets and liabilities. According to statistics published by the Swiss National Bank, their average share in the balance sheet totals of all financial institutions in Switzerland amounts to 20 % for savings and 53 % for mortgages in 2000. Typically, a large percentage of these products consists of so-called “non-maturing accounts” that can be characterized as follows: First, there is no contractual maturity, i.e., clients may withdraw their investments or repay their mortgages, respectively, at any point in time at no penalty. Second, the customer rate is not indexed to money or capital market rates but may be fixed by the bank as a matter of policy (in contrast to “adjustable rate mortgages” as they are known in the US). As a consequence, the volume of these positions may fluctuate heavily as clients react to changes in the customer rate, the relative attractiveness of alternative investment or financing opportunities due to rising or falling interest rates etc.

1.1 Specific problems of non-maturing accounts

In case of variable-rate (non-fixed) mortgages, the change in volume is positively correlated with interest rates. When the latter are low, a sharp drop
Figure 1: Negative and positive correlation between client rate and volume in demand can be observed since clients switch to fix-rate mortgages in order to hedge themselves against a future interest increase (prepayment risk). On the contrary, the volume of savings accounts grows since their yields become relatively attractive compared to short-term securities, and even some institutional investors like pension funds prefer these deposits to direct investments in money market instruments.

During a period of high interest rates, investors shift their assets from savings accounts to bonds with long maturities (withdrawal risk). This results in a negative correlation between yields and volume change. At the same time, homeowners’ demand for variable-rate mortgages rises significantly. For example, while non-fixed mortgages represented approximately 90% of the entire mortgage volume of the largest Swiss cantonal bank in 1985, their percentage decreased with falling interest rates to less than 60% in 1989. Three years later, a new high of 80% was reached after an increase in market rates.

As the numbers above indicate, a bank is usually unable to refinance its mortgages directly by savings deposits since in general the latter have a much smaller percentage of the balance sheet total. Moreover, the volumes of both products fluctuate “out of phase” as a result of the different correlations with the level of interest rates (see Figure 1). Therefore, mortgages must be refinanced on the money and capital market at higher rates. The challenge for the management is to find a mix of fixed-income instruments which minimizes not only the funding costs but also takes into account the prepayment risk that a significant portion of the volume under management is withdrawn since this would result in a surplus of liabilities over the assets (i.e., the portfolio of instruments with fixed maturities raised in the money and capital market vs. variable mortgages). As an additional complication, there is a political cap on the mortgage rate in Switzerland, and numerous banks were not able to refinance their mortgages at a positive margin during the early 1990s.
1.2 Static vs. dynamic approaches

The current practice in the management of non-maturing accounts is the determination of a so-called replicating portfolio that mimics the behavior of the underlying position. To this end, one constructs a portfolio of fixed income securities whose return replicates the client rate of the relevant asset or liability position plus margin. Transaction costs remain low since liquid money market instruments and swaps are used that are held until maturity to avoid rebalancing (“buy and hold”).

While maturing funds are always renewed at the same maturity, instruments are bought or sold at constant proportions whenever the volume of the position under management increases or decreases due to changes in customer demand. These weights are derived by minimizing the tracking error (i.e., the difference between the average portfolio rate plus margin and the client rate) over a historic sample period under the constraint that the volume of the replicating portfolio matches that of the target account at all points in time. Therefore, prepayment and withdrawal risks are implicitly taken into account.

In this way, uncertain cash flows are transformed into (apparently) certain ones, allowing the bank to manage them like normal positions with fixed maturities. However, the considerable risk of an incorrect transformation remains. For example, different historic sample periods for the determination of a replicating portfolio may result in substantially different portfolio weights. While the minimization of the tracking error aims at the stabilization of the margin over time, it does not provide minimal funding costs or even guarantee a positive margin at all.

The approach is also static in the sense that it does not incorporate the possibility of future changes in random data such as interest rates and volume and their impact on the optimal portfolio composition. Therefore, the question arises if a dynamic policy with active reactions to changes in market rates and customer demand might increase the bank’s profit or reduce its refinancing costs, respectively. In particular, it remains to be clarified if the correlation between interest rates and volume can be exploited more appropriately to manage the different risks associated with non-maturing account positions.

Multistage stochastic programs take all these aspects into account. Based on assumptions about the (joint) dynamics of relevant risk factors that are usually described by stochastic processes, representative scenarios for their future outcomes are generated. In the context of asset & liability management (ALM) problems, the latter quantify the impact of changes in the risk factors on the return of investment or refinancing strategies. Transactions
may take place at discrete points in time over some finite planning horizon for the correction of an initial policy, given new observations of random data in each scenarios. Furthermore, various constraints can be taken into account. For example, this allows the incorporation of liquidity restrictions in the market or limits for the risk exposure with respect to certain positions.

In general, stochastic optimization models result in large-scale programs since they have to include a large number of scenarios to reflect the entire universe of possible future outcomes of risk factors and cash flows. Since multistage programs suffer from an exponential growth in problem size with respect to the number of periods under consideration, the first models for financial planning that appeared in the early 80s (e.g., see Kallberg et al. [25] or Kusy and Ziemba [26]) were restricted to a two-stage structure due to limitations of the computational resources available at this time.

Today, the dramatic improvement of powerful hardware as well as the development of efficient algorithms, in particular if they exploit the special structure and high sparsity inherent to stochastic programs, provide the basis for the solution of problems with even some million scenarios like in the pension fund model of Gondzio and Kouwenberg [20]. Moreover, the inclusion of new theoretical models from the financial literature and related empirical evidence allows a more appropriate modelling of the complex dynamics of risk factors.

Among the first successful commercial multistage applications are the Towers Perrin-Tillinghast ALM system of Mulvey [28], the fixed-income portfolio management models of Zenios et al. [1, 32] or the well-known Russell-Yasuda Kasai financial planning tool of Cariño, Ziemba et al. [3, 4, 5], to mention just a few. Numerous related models have been derived from the latter like the InnoALM system [19] that exploits sophisticated econometric models and scenario dependent correlation matrices to specifically consider extreme market events. For an extensive survey on financial applications, see the book edited by Ziemba and Mulvey [33] or the monograph in Part I of this volume.

Motivated by the difficulties in the management of non-maturing account positions and given their significant importance for most banks in Switzerland, we developed a stochastic optimization model for non-fixed mortgages that has been in use by a major Swiss bank since 1995 and was extended to savings accounts two years later (cf. [16]). Other versions have been implemented, e.g., for pension funds or cash management at an insurance company where one must deal with the seasonal behavior of premium payments, uncertain investment yields and stochastic liabilities due to claims.

In the next section, we introduce a simplified formulation of the stochastic optimization model for refinancing variable mortgages. After a discussion of
relevant characteristics of the underlying risk factors, i.e., interest rates and volume, section 3 describes two term structure models for different planning horizons and an autoregressive process for the volume dynamics. Barycentric approximation as a special scenario generation method that is particularly well suited for the given problem structure is outlined in section 4. Practical experience and results from a case study are reported in section 5. The final section 6 summarizes the most relevant features of the model and gives an outlook to possible directions for its improvement.

2 Formulation of the optimization model

2.1 Framework for decision making under uncertainty

We assume that the evolution of random data is described by a discrete-time stochastic process \( \omega := (\omega_t; t = 1, \ldots, T) \) defined in a probability space \((\Omega, \mathcal{F}, P)\). The sample space can be decomposed with respect to time as \( \Omega := \Omega_1 \times \ldots \times \Omega_T \) with \( \Omega_t \subset \mathbb{R}^{M_t} \). The filtration \( \mathcal{F} := \{\mathcal{F}_t; t = 1, \ldots, T\} \) specifies the information structure and satisfies \( \mathcal{F}_t \subset \mathcal{F}_{t+1} \). Each \( \sigma \)-field \( \mathcal{F}_t := \sigma(\omega^t) \) is generated by observations \( \omega^t := (\omega_1, \ldots, \omega_t) \) of the data process that are known at time \( t \), and \( P_t : \mathcal{F}_t \to [0,1] \) denotes the corresponding conditional probability measure.

Restricting ourselves to the linear case, the underlying decision problem can be outlined as follows: At any point in time \( t = 0, \ldots, T \), we have to decide upon a portfolio composition \( x_t \) in order to meet an uncertain target \( h_{t+1} \) in the future (here: the volume of variable mortgages). The value of the portfolio after one period is given by \( T_t \cdot x_t \) where the elements of the matrix \( T_t \) represent, e.g., prices of individual positions that may be affected by the outcomes of some risk factors \( \omega_{t+1} \in \Omega_{t+1} \). Since it is likely that the portfolio fails to meet the target, a correction \( x_{t+1} \) may be required after a realization of \( \omega_{t+1} \) has been observed to compensate the discrepancy \( h_{t+1} - T_t \cdot x_t \). Such a recourse action \( x_{t+1} \) is penalized by the fixed matrix \( W_{t+1} \).

The objective is to minimize the cost \( c'_t \cdot x_t \) in the current period \( t \) plus the expected cost \( E_{t+1} \phi_{t+1}(x^t, \omega^{t+1}) = \int_{\Omega_{t+1}} \phi_{t+1}(x^t, \omega^{t+1})dP_{t+1} \) for future transactions that depends on the sequence of observations \( \omega^{t+1} := (\omega^t, \omega_{t+1}) \) and earlier decisions \( x^t := (x_0, \ldots, x_t) \). This can be stated recursively for \( t = T, \ldots, 1 \) in terms of value functions as

\[
\phi_t(x^{t-1}, \omega^t) := \min_{\omega'} c'_t(\omega') \cdot x_t(\omega') + \int_{\Omega_{t+1}} \phi_{t+1}(x^t, \omega^{t+1})dP_{t+1}(\omega_{t+1} | \omega^t) \\
\text{s.t. } W_t \cdot x_t(\omega^t) = h_t(\omega^t) - T_{t-1}(\omega^t) \cdot x_{t-1}(\omega^{t-1}), x_t(\omega^t) \geq 0
\]
with the boundary condition \( \phi_{T+1}(\cdot) := 0 \) whereas the problem for the first stage is

\[
\phi_0 := \min c_0' \cdot x_0 + \int_{\Omega} \phi_1(x_0, \omega_1) dP_1(\omega_1) \\
\text{s.t. } W_0 \cdot x_0 = h_0, \ x_0 \geq 0.
\]  

(2)

A decision \( x_t \) may depend only on information \( \omega^t \) available at time \( t \) and not on future outcomes of random data \( \omega_{t+1}, \ldots, \omega_T \). This property is known as nonanticipativity. For the simplicity of notation, we do not stress the dependency of \( x_t \) on \( \omega_t \) in the sequel but assume implicitly that \( x_t \equiv \{x_t|\mathcal{F}_t\} \) is adapted to the filtration \( \mathcal{F}_t \).

Due to linearity of the objective function and constraints, the multistage stochastic program given by (1) and (2) is convex. Its properties are widely discussed in the literature (e.g., see Part I of this volume or the textbook by Birge and Louveaux [2]). Here, we focus on the special case where the vectors \( \omega_t \in \Omega_t \subset \mathbb{R}^{M_t} \) of random data can be decomposed into two components: \( \eta_t \in \Theta_t \subset \mathbb{R}^{K_t} \) affects the coefficients in the objective, \( \xi_t \in \Xi_t \subset \mathbb{R}^{L_t} \) influences the demand on the right-hand-side of constraints with \( M_t = K_t + L_t \). Then, the value function given by (1) is a saddle function for all \( t = 1, \ldots, T \), and the corresponding optimization problem can be solved if the following conditions hold:

(i) The support of \( \omega_t = (\eta_t, \xi_t) \) is covered by compact and convex sets \( \Omega_t = \Theta_t \times \Xi_t \).

(ii) \( c_t(\eta^t), h_t(\xi^t) \) are linear affine in their respective random vectors \( \eta^t := (\eta_1, \ldots, \eta_t), \xi^t := (\xi_1, \ldots, \xi_t) \), and the matrices \( T_{t-1} \) are deterministic.

In other words, the value function (1) takes on the form

\[
\phi_t(x^{t-1}, \omega^t) := \min c_t'(\eta^t) \cdot x_t + \int_{\Omega_{t+1}} \phi_{t+1}(x^t, \omega^{t+1}) dP_{t+1}(\omega_{t+1}|\omega^t) \\
\text{s.t. } W_t \cdot x_t + T_{t-1} \cdot x_{t-1} = h_t(\xi^t), \ x_t \geq 0
\]  

(1')

Since \( \phi_{T+1}(\cdot) = 0 \), this implies that the problem for the final stage \( T \) is convex in \( (x^{T-1}, \xi^T) \) and concave in \( \eta^T \). When calculating the expectations \( E_t \phi_{t+1} \) in the remaining stages \( T-1, \ldots, 1 \), the probability measures \( P_{t+1} \) depend on \( \omega^t \). As a consequence, the saddle property of the value function in \( T \) may not be “inherited” to the previous stages due to the integration with respect to \( P_{t+1}(\cdot|\omega^t) \). However, if the corresponding distribution functions may be represented, e.g., in the form \( Q_{t+1}(\omega_{t+1} + H_{t+1}(\omega^t)) \), where \( Q_{t+1} \) is
independent of $\omega^t$ and $H_{t+1}$ is a linear mapping, then it can be shown that the relevant expectation functionals $E_{t+1}\phi_{t+1}(x^t, \omega^{t+1})$ are continuous saddle functions (cf. [13, 16]). This leads to the requirement:

(iii) The distribution function of $P_t(\cdot|\omega^{t-1})$ depends linearly on the past.

In particular, this covers the case that the distribution of risk factors is independent of prior outcomes. If conditions (i)–(iii) are fulfilled, the value functions $\phi_t(x^{t-1}, \eta^t, \xi^t)$ are convex in $(x^{t-1}, \xi^t)$ and concave in $\eta^t$ for $t = 1, \ldots, T$ which is called the entire convex case (see Figure 2 for the case of $K_t = L_t = 1$). These conditions are sufficient, and the saddle property may also be given for more general problem types (cf. [13] for details).

2.2 Specification of the funding problem

Given the description of the mortgage funding problem in section 1.1, the formulation of the optimization model is straightforward: Let $D = \{1, 2, \ldots, D\}$ denote the set of maturity dates for fixed-income instruments held in the current portfolio from decisions in the past ($D$ is the longest available maturity). Standard maturities traded in the money and capital market which can be used for refinancing are given by the set $D^S \subseteq D$. Due to liquidity restrictions in the Swiss interbank market which must be observed by large banks, transaction costs (bid-ask spreads) increase when a certain volume is exceeded. To this end, the amount to be refinanced in each maturity is split into several tranches. The maximum number of possible tranches for maturity $d$ is given by $I^d$, $I^d := \{1, \ldots, I^d\}$ is a corresponding index set, and $x^d_{i,t}$ denotes the transaction volume in tranche $i \in I^d, d \in D^S$. 
The coefficient $\phi_{d,i,t}$ quantifies the costs for raising $1$ in the $i$-th tranche of maturity $d$ at time $t$. It contains the net present value of all interest payments until maturity of the instrument, transaction costs as well as “penalty spreads” for exceeding the tranche limits. Since the planning horizon must be truncated at time $T$ while the bank’s business is (hopefully) continued beyond that date, the present value of all cash flows which occur after the terminal stage, i.e., outstanding interest payments and repayment of the face value discounted at the yield curve in $T$, is also included to incorporate end effects.

For $t > 0$ these coefficients depend on future interest rates and, hence, are uncertain. It is assumed that the $\phi_{d,i,t}$ for all $d \in D^S$ at time $t$ are functions of a random vector $\eta_t \in \mathbb{R}^{K_t}$. As discussed in the sequel, the elements of $\eta_t$ are the state variables of a term structure model which describes the evolution of the yield curve by a $K_t$-dimensional stochastic process. A formal specification of the functional relationship between the interest rate risk factors $\eta_t$ and $\phi_{d,i,t}$ is omitted here because its notation becomes rather cumbersome.

Analogously, the mortgage volume $\upsilon_t$ depends on the random vector $\xi_t \in \mathbb{R}^{L_t}$. It may be correlated with $\eta_t$ to reflect a dependency between interest rates and volume. With the objective to minimize the expected discounted refinancing costs over the planning horizon $T$, the corresponding multistage stochastic program is:

$$\min \int \Omega \left\{ \sum_{t=0}^{T} \sum_{d \in D^S} \sum_{i \in I_d} \phi_{d,i,t}(\eta_t) \cdot x_{d,i,t} \right\} dP(\omega) \quad (3)$$

subject to

$$x_{d,t} - x_{d,t+1} = 0 \quad t = 0, \ldots, T; \forall d \notin D^S; \text{ a.s.} \quad (3.1)$$

$$x_{d,t} - x_{d,t-1} - \sum_{i \in I_d} x_{d,i,t} = 0 \quad t = 0, \ldots, T; \forall d \in D^S; \text{ a.s.} \quad (3.2)$$

$$\sum_{d \in D} x_{d,t} = \upsilon_t(\xi_t) \quad t = 0, \ldots, T; \text{ a.s.} \quad (3.3)$$

$$\ell_{i,t}^d \leq x_{i,t}^d \leq u_{i,t}^d \quad t = 0, \ldots, T; \forall i \in I_d^d; \forall d \in D^S; \text{ a.s.} \quad (3.4)$$

$$x_{i,t}^d \geq 0; \quad x_{i,t}^d \equiv \{x_{i,t}^d|\mathcal{F}_t\} \quad t = 0, \ldots, T; \forall i \in I_d^d; \forall d \in D^S; \text{ a.s.} \quad (3.5)$$

$$x_{d,t} \in \mathbb{R}; \quad x_{d,t} \equiv \{x_{d,t}^d|\mathcal{F}_t\} \quad t = 0, \ldots, T; \forall d \in D; \text{ a.s.} \quad (3.6)$$

The budget constraint (3.1) ensures that the position $x_{d,t}^d$ maturing after $d$ periods equals the corresponding value in the previous period for non-traded maturity dates while (3.2) corrects it by the new transactions in $t$ for
traded ones. Herein, \( x^d_{-1} \) denotes the amount in the initial portfolio held in maturity \( d \). Constraint (3.3) requires that the complete portfolio matches the stochastic mortgage volume \( \nu_t \) at all points in time.

Lower and upper limits \( \ell^d_{l,t} \) and \( \ell^u_{u,t} \) for tranches with different penalty spreads to reflect liquidity restrictions are given by (3.4). Obviously, the absolute amount that can be refinanced in maturity \( d \) can be controlled by a corresponding number of tranches \( i^d \). All tranches with low indices are “filled” to their limits first in the optimal solution since the spreads are strictly increasing. Nonanticipativity constraints (3.5) and (3.6) require that decisions are adapted to the filtration \( \mathcal{F}_t \). Moreover, the restrictions (3.1)–(3.6) at stage \( t = 1, \ldots, T \) must hold almost surely, i.e., for all observations of the uncertain interest rate and volume risk factors \( \omega_t := (\eta_t, \xi_t) \) with positive probability.

Since we distinguish between stochastic factors \( \eta_t \) that affect only coefficients in the objective and those \( \xi_t \) that occur solely in the constraints with deterministic left-hand-sides, it can easily be seen that the mortgage funding problem above has the same structure as the special type of multi-stage stochastic programs given by (1') and (2). This is a particularly useful property for the solution method introduced in section 4.

3 Modelling the dynamics of risk factors

Since the optimization problem (3) contains uncertain coefficients, we must specify stochastic processes that describe the dynamics of the relevant risk factors, i.e., interest rates of traded maturities and the volume of the mortgage position under management. The corresponding process parameters are then estimated by suitable statistical procedures. Alternative models should be compared by means of econometric tests which allow the rejection of inappropriate specifications. In this way, one can identify the model that provides the best explanation of the empirically observed data among different alternatives.

3.1 Term structure models for interest rate scenarios

Interest rates exhibit a number of characteristic features that should be taken into account to model their dynamics realistically. For example, principal component analysis reveals that three stochastic factors are sufficient to explain more than 95% of interest rate volatility. These factors can be associated with changes in the level, steepness and curvature of the yield curve (see Figure 3 for Swiss Franc Euromarket rates, cf. Schürle [31]). Moreover, one
can observe that interest rates fluctuate around a long-term mean during an economic cycle. This *mean reversion* property is incorporated in most term structure models that have been introduced in the financial literature over the last two decades (e.g., see James and Webber [23] for a comprehensive introduction).

The relevance of these different aspects depends on the underlying application and/or the desired planning horizon. For example, when a portfolio of short-term securities is managed for a horizon of up to one year like in cash management problems, it should be taken into account that the short end of the yield curve often shows complex and quickly varying shapes. This can be reflected well using three principal components as risk factors that are standard normally distributed and orthogonal by construction while the mean reversion property may be neglected. The sensitivities of relevant interest rates with respect to the principal components can easily be derived by means of most statistical software tools.

On the other hand, this approach implies a non-stationary distribution of interest rates. As a consequence, the model generates values that are outside of the “usual” range or may even become negative when the planning horizon is extended. A simple way of modelling mean reversion of a risk factor $\eta_t$ is the introduction of a drift term

$$\frac{d\eta_t}{dt} = \kappa(\theta - \eta_t), \quad \kappa > 0. \tag{4}$$

If $\eta_t < \theta$ ($\eta_t > \theta$) at time $t$, then the left-hand-side of this equation is positive (negative) which will cause $\eta_t$ to increase (decrease) towards its mean reversion level $\theta$ at a speed controlled by parameter $\kappa$. Rewriting (4) and adding a volatility term where noise is a function of a Wiener process with increment $dz_t$ as usual in modelling financial data yields the (continuous-time) process

$$d\eta_t = \kappa(\theta - \eta_t)dt + \sigma\eta_t^\gamma dz_t. \tag{5}$$
The instantaneous volatility $\sigma \eta_t^\gamma$ may depend on the current level of the factor if $\gamma > 0$ to reflect a possible heteroscedasticity of interest rates. Moreover, this prevents the process from becoming negative since in continuous time the volatility term becomes zero when $\eta_t = 0$ and only the positive drift remains (see [23] for details).

Given $K$ stochastic processes of type (5) for one or more risk factors in a “classical” term structure model, Ito’s lemma can be used to derive a process for the price of a discount bound. By construction of a hedge portfolio consisting of $K+1$ instruments with different maturities, one obtains a partial differential equation (PDE) for the term structure using a no-arbitrage argument. However, the latter can be solved analytically only for some special cases. Example are models of the affine type where (i) all $K$ factors follow a process of the form (5) with identical exponent $\gamma \in \{0, 0.5\}$, (ii) are orthogonal and (iii) sum up to the instantaneous rate $r_t$, i.e., the yield for an infinitely short holding period (see Frauendorfer and Schürle [17]).

In order to overcome these restrictions, we implemented several exponential functions for an interpolation of the term structure instead of an arbitrage-free PDE. For the ease of exposition, only the rather simple form

$$R(s_t, l_t, d) = (s_t + \beta_1 \cdot d) \cdot e^{-\beta_2 d} + l_t$$

is depicted here where $R(s_t, l_t, d)$ denotes the spot rate for maturity $d$. While the fix parameters $\beta_1, \beta_2$ control the shape of the yield curve, we model the long rate $l_t$ and the spread $s_t := r_t - l_t$ by the stochastic processes

$$ds_t = \kappa_s(\theta_s - s_t)dt + \sigma_s dz_{1,t}$$
$$dl_t = \kappa_l(\theta_l - l_t)dt + \sigma_l l_t^\gamma dz_{2,t}.$$  \hspace{1cm} (7)

The volatility of the long rate may depend on the current level of $l_t$ when $\gamma > 0$ to incorporate a possible heteroscedasticity. Obviously, the process for the spread must allow for negative values to reflect normal ($s_t < 0$) as well as inverse term structures ($s_t > 0$). Although this does not preclude negative interest rates when the spread becomes sufficiently large, they are extremely unlikely for realistic parameter estimates.

The process specification (7) resembles in some sense the model of Schaeffer and Schwartz [30] that uses the same state variables but assumes $\gamma = 0.5$ and no correlation between both Wiener processes to derive an analytical approximation for the term structure PDE. The advantage of our approach is that we have more flexibility in the specification of yield curve functions than in conventional term structure models. Furthermore, we may choose any value for the correlation $\rho = dz_{1,t} \cdot dz_{2,t}$ between the Wiener processes and the volatility exponent $\gamma$. Using discrete time approximations of (7)
\[ s_{t+1} = s_t + \kappa_s (\theta_s - s_t) \Delta t + \sigma_s \sqrt{\Delta t} \epsilon_{1,t} \]
\[ l_{t+1} = l_t + \kappa_l (\theta_l - l_t) \Delta t + \sigma_l l_t \sqrt{\Delta t} \epsilon_{2,t}, \]

(8)

parameters can easily be estimated for both processes separately with maximum likelihood (under the assumption that the residuals \( \epsilon_{i,t} \sim N(0,1) \), \( i = 1, 2 \), are serially independent and uncorrelated with \( s_t \) and \( l_t \)). Herein, the long rate \( l_t \) is approximated by observations of the 5 year CHF Euromarket rate as the longest available (liquid) maturity and the spread \( s_t \) by the difference between the one-month rate and the latter.

In view of the approximation method introduced in the sequel for the solution of multistage stochastic programs, we restrict ourselves to fixed values \( \gamma \in \{0, 1\} \) when we estimate the parameters of the discrete processes (8). Under these specifications, the saddle property of value functions discussed in section 2.1 will be preserved. An analysis of different historic sample periods reveals that \( \gamma = 1 \) provides higher likelihood values in most cases which supports the assumption of heteroscedastic interest rates. After calibration of the processes, an estimate for \( \rho \) is obtained from the cross-correlation between the residuals. Finally, the parameters of the yield curve interpolation function (here: \( \beta_1, \beta_2 \)) are determined that allow for the best fit of (6) to the observed rates of all maturities in the historic sample by quadratic minimization.

3.2 Specification of the volume process

At first sight, the specification and calibration of stochastic processes for the variable mortgage volume seems to be easier than for interest rates since it is directly observable and, in contrast to the term structure, does not depend on several “latent” factors. However, data for the estimation and assessment of alternative model are often difficult to obtain in practice. The Swiss National Bank publishes only an aggregate of variable and fixed mortgages in its monthly reports, and the percentage of non-fixed mortgages is only available on a yearly basis since 1996. Furthermore, fluctuations in mortgage demand often depend on the type of bank. For example, the clientele of large commercial banks in Switzerland consists mainly of urban population that tends to react more actively to changes in the economy than customers of smaller cooperative banks in the countryside. As a consequence, any volume model has to calibrated individually to the specific bank situation.

According to the problem description in section 1.1, it seems plausible that the demand for non-maturing accounts depends on the level of interest rates. In case of savings accounts, Schürle [31] found that a trend-stationary process with two factors of a term structure model as explanatory variables
provides a good description for real deposit positions. However, when the implementation of the mortgage funding model started in 1993, we did not have a sufficiently large data set at our disposal to support such a hypothesis. Therefore, we simply model the volume by the autoregressive process

\[
\tilde{v}_t = a + \rho \tilde{v}_{t-1} + \xi_t
\]

(9)
to which a deterministic trend is added, i.e., \( v_t = \tilde{v}_t + bt \). The stochastic component (9) has a long-term mean of \( a/(1 - \rho_v) \) if the process is stationary (i.e., \(|\rho_v| < 1\)). After correcting a sample time series by the trend \( bt \), the parameters of the resulting standard AR(1) process can be easily estimated by means of most statistical standard software. A dependency on interest rates is taken into account by the correlation between the stochastic factor \( \xi_t \) and the residuals of the discrete-time interest rate processes (8).

4 Solution of multistage stochastic programs

The numerical difficulty in solving a problem of type (1) and (2) lies in the nested minimization and multi-dimensional integration of value functions. Since the latter are given only implicitly as the solution of stochastic programs with respect to the remaining stages, this integration cannot performed analytically, and numerical techniques are required that can broadly be classified into simulation-based methods and bound-based approximations.

In the former case, random samples are drawn from the underlying probability distributions, e.g., to derive stochastic quasi-gradients or for the application of stochastic decomposition algorithms. While the computational effort is independent of the dimension of random data, these approaches provide only probabilistic bounds for the discretization error. Loosely speaking, this is the error that results from replacing the entire universe of possible outcomes of random data by a relatively small set of scenarios.

In contrast to this, bound-based approximation methods partition the domain of random data into cells and determine representative points within them. The accuracy may be improved by adding new scenarios that result from partitioning the initial cells into smaller ones (refinement). For multistage problems, a careful control of the refinement process is necessary since the number of scenarios grows exponentially with the dimension size and the desired accuracy (curse of dimensionality). However, generalizations of the well-known inequalities of Jensen [24] and Edmundson/Madansky [27] that exploit the saddle property of value functions discussed in section 2.1 provide exact lower and upper bounds to the original problem (bounds based on
these inequalities and their refinements in the context of two-stage stochastic programming were introduced, e.g., by Huang et al. [21]). This allows a deliberate refinement for those scenarios and stages where the largest discretization error is observed to reduce the growth in problem size.

Since the evolution of interest rates and volume is modelled by low-dimensional processes, the underlying saddle property motivates the application of a bound-based scenario generation approach for the solution of the mortgage funding problem (3). In order to deal with the observed correlations between the risk factors, approximation schemes must take into account cross moment information in a numerically efficient way. Such techniques have been developed, e.g., by Edirisinghe and Ziemba [6, 8, 9, 10] or Frauendorfer [11, 13, 14] (see also Edirisinghe [7] for a survey on bound-based approximations).

4.1 Barycentric approximation

In the sequel, we concentrate on the barycentric approximation technique that was originally introduced in Frauendorfer [11] for two-stage stochastic programs and extend it to the multistage case. The basic idea is to replace the implicitly given value functions of type (1′) by two bilinear functions that can easily be integrated. Then, the best points where the original value function must be supported by its bilinear approximations in order to minimize the discretization error are the so-called generalized barycenters. These are determined with respect to cross-simplices (or briefly: ×-simplices), i.e., the Cartesian product of regular simplices, that cover the support of random data.

For the ease of exposition, we assume that the sets Θ_{t}(ω^{t-1}) ⊂ R^{K_{t}} and Ξ_{t}(ω^{t-1}) ⊂ R^{L_{t}} that cover the support of η_{t} and ξ_{t} are regular simplices. Note that both may depend on prior observations ω^{t-1} although this is not stressed in the notation for simplicity. Let the vertices of Θ_{t} and Ξ_{t} be denoted by u_{ν_{t}}, ν_{t} = 0, . . . , K_{t}, and v_{μ_{t}}, μ_{t} = 0, . . . , L_{t}. The barycentric weights λ_{t}(η_{t}) = (λ_{t,0}(η_{t}), . . . , λ_{t,K_{t}}(η_{t}))′ of η_{t} with respect to Θ_{t} are those nonnegative weights that allow the representation of η_{t} as a linear combination of the vertices u_{ν_{t}} and sum up to one:

\[ \lambda_{t,0} + \lambda_{t,1} + \ldots + \lambda_{t,K_{t}} = 1 \]
\[ u_{t,0}\lambda_{t,0} + u_{t,1}\lambda_{t,1} + \ldots + u_{t,K_{t}}\lambda_{t,K_{t}} = \eta_{t}. \]

The barycentric weights τ_{t}(ξ_{t}) = (τ_{t,0}(ξ_{t}), . . . , τ_{t,L_{t}}(ξ_{t}))′ of ξ_{t} with respect to Ξ_{t} are defined analogously. λ_{t}(\cdot), τ_{t}(\cdot) can be obtained as the solution of
and provides the barycenters taking into account the distance from each sample point to the edges, providing the generalized barycenters \( \xi_t \) of (the volume risk factor) and \( C \) are equivalent to \( u \) respectively. These mass distributions add up to one, i.e.,

\[
\sum_{\nu} M_{\nu_\mu}(\Theta_t \times \{v_{\mu_\nu}\}) = 1, \quad \nu = 0, \ldots, K_t
\]

with respect to the \( \times \)-simplex \( \Theta_t \times \Xi_t \) where

\[
M_{\nu_\mu}(\Theta_t \times \{v_{\mu_\nu}\}) = \int \lambda_{\nu_\mu}(\eta_t, \xi_t | \omega^{t-1}) dP_t(\eta_t, \xi_t | \omega^{t-1}), \quad \mu = 0, \ldots, L_t
\]

are the mass distributions induced by the probability measure \( P_t \) on the \( K_t \)-dimensional simplices \( \Theta_t \times \{v_{\mu_\nu}\} \) and the \( L_t \)-dimensional simplices \( \{u_{\nu_\mu}\} \times \Xi_t \), respectively. These mass distributions add up to one, i.e.,

\[
\sum_{\nu_\mu=0}^{K_t} M_{\nu_\mu}(\{u_{\nu_\mu}\} \times \Xi_t) = 1 \quad \text{and} \quad \sum_{\nu_\mu=0}^{L_t} M_{\nu_\mu}(\Theta_t \times \{v_{\mu_\nu}\}) = 1.
\]

Therefore, we may interpret (14) and (15) as probabilities assigned to the points \( (u_{\nu_\mu}, \xi_{\nu_\mu}) \) with probability \( M_{\nu_\mu}(\{u_{\nu_\mu}\} \times \Xi_t) \) for \( \nu = 0, \ldots, K_t \) and \( (\eta_{\mu_\nu}, v_{\mu_\nu}) \) with probability \( M_{\mu_\nu}(\Theta_t \times \{v_{\mu_\nu}\}) \) for \( \mu = 0, \ldots, L_t \).

An illustration is given in Figure 4 where the samples represent the joint distribution of \( \eta \) and \( \xi \) for \( K = L = 1 \) (the time index is omitted for simplicity), indicating a negative correlation between the random data. In the one-dimensional case, the simplices covering the support of \( \eta \) and \( \xi \) are intervals and, thus, the resulting \( \times \)-simplex is a rectangle. For instance, in Figure 4 (a) the edges \( AB \) and \( CD \) cover the support of \( \eta \) (here: the factor of a term structure model), i.e., \( A \) and \( D \) correspond to vertex \( u_0 \) while \( B \) and \( C \) are equivalent to \( u_1 \). Analogously, \( AD \) and \( BC \) represent the domain of (the volume risk factor) \( \xi \) and correspond to an interval with vertices \( v_0 \) and \( v_1 \).

Projecting the distribution mass onto \( AB \) and \( CD \) as in Figure 4 (b), taking into account the distance from each sample point to the edges, provides the barycenters \( \eta_0 \) and \( \eta_1 \). For each simplex, they are determined as the
“center of gravity” of the projected mass, and their probabilities are equivalent to the proportion of the latter to a total mass of one. Analogously, the barycenters $\xi_0$ and $\xi_1$ in Figure 4 (c) result from a projection of the mass onto $AD$ and $BC$, respectively.

An advantageous feature from a computational point of view is that the generalized barycenters in (12), (13) and their probabilities (14), (15) are completely derived from the first moments of $\eta_t$ and $\xi_t$ as well as the bilinear cross moments $E_t(\eta_{\nu_t} \cdot \xi_{\mu_t})$, $\nu_t = 0, \ldots, K_t$, $\mu_t = 0, \ldots, L_t$. Since the covariance of two random variables is determined by the first moments and the corresponding cross moments, the measures $Q^u_t$ and $Q^l_t$ incorporate implicitly the correlation between $\eta_t$ and $\xi_t$ as indicated by the different coordinates of the corresponding barycenters in Figure 4 (b) and (c). However, cross moments (or covariances, respectively) between different elements of $\eta_t$ are not taken into account (the same holds for the components of $\xi_t$). Therefore, no assumptions of independence between the random variables are required.

### 4.2 Lower and upper bounds for value functions

By application of (12)–(15) at each stage $t$, the original probability measure $P_t$ is approximated by two discrete measures $Q^l_t$ and $Q^u_t$ with supports

$$\text{supp} Q^l_t = \{(u_{\nu_t}, \xi_{\nu_t}) | \nu_t = 0, \ldots, K_t\}$$

$$\text{supp} Q^u_t = \{(\eta_{\mu_t}, v_{\mu_t}) | \mu_t = 0, \ldots, L_t\}. \quad (16)$$

The corresponding probabilities are given by $q^l_t(u_{\nu_t}, \xi_{\nu_t}) := \mathcal{M}_{\nu_t}(\{u_{\nu_t}\} \times \Xi_t)$ and $q^u_t(\eta_{\mu_t}, v_{\mu_t}) := \mathcal{M}_{\mu_t}(\Theta_t \times \{v_{\mu_t}\})$. Substituting $P_t$ in (1)’ by these approximate measures provides the new value functions

$$\psi_t(x^{t-1}, \omega^t) := \min c'_t(\eta^t) \cdot x_t + \int_{\Omega_{t+1}} \psi_{t+1}(x^{t'}, \omega^{t+1}) dQ^l_{t+1}(\omega_{t+1} | \omega^t)$$

$$\text{s.t.} \quad W_t \cdot x_t + T_{t-1} \cdot x_{t-1} = h_t(\xi^t), \quad x_t \geq 0 \quad (18)$$
\[ \Psi_t(x^{t-1}, \omega^t) := \min_{c} c_t'(\eta_t) \cdot x_t + \int_{\Omega_{t+1}} \Psi_{t+1}(x^t, \omega^t+1) dQ_{t+1}(\omega_{t+1}|\omega^t) \]

s.t. \( W_t \cdot x_t + T_{t-1} \cdot x_{t-1} = h_t(\xi_t), \ x_t \geq 0 \) \hspace{1cm} (19)

for \( t = 1, \ldots, T \) with \( \psi_{T+1}(\cdot) = \Psi_{T+1}(\cdot) := 0 \). Both \( \psi_t \) and \( \Psi_t \) are bilinear functions since the integrand \( \lambda_{\nu}(\eta_t) \cdot \tau_{\mu}(\xi_t) \) in (12) and (13) is bilinear in \( (\eta_t, \xi_t) \). It is shown in [13] that the following relation holds:

\[ \psi_t(x^{t-1}, \omega^t) \leq \phi_t(x^{t-1}, \omega^t) \leq \Psi_t(x^{t-1}, \omega^t), \] \hspace{1cm} (20)

The meaning of this inequality is that the value function \( \phi_t(\cdot) \) with respect to the original measure \( P_t \) (which is a saddle function) is supported from below and above by the two bilinear functions \( \psi_t(\cdot) \) and \( \Psi_t(\cdot) \) with respect to the approximate measures \( Q^l_t \) and \( Q^u_t \). The barycenters \( \xi_{\nu}, \nu_t = 0, \ldots, K_t \), are the supporting points for the minorant while the majorant is supported in \( \eta_{\mu}, \mu_t = 0, \ldots, L_t \). This situation is illustrated in Figure 5 (again, the case \( K = L = 1 \) is considered and the time index omitted for simplicity). Obviously, the two bilinear functions can easily be integrated since the calculation of the expectations reduces to the weighted sums

\[ E_t \psi_t(x^{t-1}, \omega^t) = \sum_{\omega_t \in \text{supp} Q^l_t} \psi_t(x^{t-1}, \omega^t) \cdot q^l_t \]

\[ E_t \Psi_t(x^{t-1}, \omega^t) = \sum_{\omega_t \in \text{supp} Q^u_t} \Psi_t(x^{t-1}, \omega^t) \cdot q^u_t \]

which was the intention of the approximation. Therefore, the problems that result from the substitution of the original conditional measures \( P_t \) in (1') and (2) by \( Q^l_t \) and \( Q^u_t \) can be treated as deterministic multistage programs.

Their solutions provide policies \( x^l := (x^l_0, \ldots, x^l_T) \) and \( x^u := (x^u_0, \ldots, x^u_T) \). While the decisions for \( t > 0 \) correspond to outcomes in the barycentric scenarios that may be seen as “representative” rebalancing actions, only the policy for \( t = 0 \) will be implemented and is of interest for the user. However, a situation may occur where the first-stage decisions are not unique, i.e., \( x^l_0 \neq x^u_0 \). Then, more accurate bounds can be achieved when the support of random data at time \( t \) is partitioned into \( \ell_t(\omega^{t-1}) \) sub-\( \times \)-simplices with

\[ \bigcup_{i=1}^{\ell_t} \Omega^l_t = \Omega_t \subset \text{supp} \omega_t, \] \hspace{1cm} (21)

\[ \Omega^l_i \cap \Omega^l_j = \emptyset, \quad i \neq j; i, j = 1, \ldots, \ell_t, \] \hspace{1cm} (22)

\[ \Omega^l_i \text{ are regular } \times \text{-simplices for } i = 1, \ldots, \ell_t. \] \hspace{1cm} (23)
A collection $\mathcal{P}^\ell_t(\omega^{t-1}) := \{\Omega^1_t, \ldots, \Omega^\ell_t\}$ which satisfies (21)-(23) is called a partition of the support of $\omega_t$, and an approximation can be obtained by application of the scheme (12)-(15) to each element $\Omega^i_t$ individually. Note that the probabilities assigned to the individual outcomes must be adjusted according to the percentage of the distribution mass that is covered by the corresponding $\times$-simplex. In case that the accuracy of a bound is not sufficient, one may split one of the (sub-) $\times$-simplices in the initial partition $\mathcal{P}^\ell_t$ (refinement, see Figure 6). Then, the solution of the corresponding approximate problem based on the new partition $\mathcal{P}^{\ell_t+1}$ must be as least as good as the former bound. As $\ell_t \to \infty$ and the sub-$\times$-simplices become arbitrarily small with respect to diameter, the approximate value functions $\psi_t$ and $\Psi_t$ epi-converge to $\phi_t$ (see Frauendorfer [13] for details).

Nevertheless, dividing the elements of a partition without strategy may dramatically increase the number of scenarios and, hence, the computational complexity of the corresponding deterministic optimization problems. For example, one may refine the partition with the largest discretization error $\epsilon_t(\omega^i) := \Psi_t(u^{t-1}, \omega^i) - \psi_t(u^{t-1}, \omega^i)$ until the desired accuracy is achieved. If $\epsilon_t(\cdot) = 0$, then the approximation of $\phi_t$ is exact, and (further) refinements will not improve the solution. In this sense, the existence and utilization of exact bounds may be seen as one of the most important features of barycentric approximation for the solution of multistage stochastic programs (see Frauendorfer and Marohn [15] for refinement techniques in the context of barycentric approximation).
4.3 Cross-simplicial coverages and convergence

Obviously, the distributions induced by the stochastic processes we considered in sections 3.1 and 3.2 have unbounded support. For the determination of a \( \times \)-simplicial coverage, \( P_t(\Theta_t \times \Xi_t|\eta_t^{-1}, \xi_t^{-1}) \) must therefore be substituted by a normalized truncation so that \( P_t(\Theta_t \times \Xi_t|\eta_t^{t-1}, \xi_t^{t-1}) \geq 1 - \varepsilon \) for some sufficiently small \( \varepsilon > 0 \).

In case that the partition consists of a single \( \times \)-simplex only and all \( K \) risk factors are standard normally distributed and uncorrelated, such a coverage may be constructed as follows: A sphere with radius \( \delta \) around the origin contains a percentage of \( 2\Phi(\delta) - 1 \) of the total mass distribution (\( \Phi \) denotes the c.d.f.). It can be covered by a regular simplex in \( \mathbb{R}^K \) with \( K + 1 \) vertices (note that such a simplex is the polyhedron with the smallest possible number of independent vertices). In the one-dimensional case, this simplex reduces to an interval \( [-\delta, \delta] \) and for \( K = 2 \) to a triangle whose vertices may be chosen, e.g., as \( u_0 = (-\sqrt{3}\delta, \delta)' \), \( u_1 = (\sqrt{3}\delta, \delta)' \) and \( u_2 = (0, -2\delta)' \). For example, with \( \delta = 2 \) the circle contains more than 95% of the probability mass, and since the latter is entirely covered by the simplex, any outcome within a range of two standard deviations will be considered in the approximation.

For arbitrary \( K \)-dimensional normal distributions with expectation \( \mu \) and covariance matrix \( \Sigma \), we make use of the fact that a standard normally distributed random vector \( Z \in \mathbb{R}^K \) can be transformed into another one \( Y \sim \mathcal{N}(\mu, \Sigma) \) using the lower triangular matrix of the Cholesky decomposition \( \Gamma \) of \( \Sigma \), i.e., \( \Sigma = \Gamma \cdot \Gamma' \). Given a simplex with vertices \( u_i^Z \), \( i = 0, \ldots, K \), around the truncated support of an uncorrelated standard normal distribution, the vertices of a simplex that covers the actual correlated distribution are obtained from the transformation \( \mu + \Gamma \cdot u_i^Z \). This procedure is performed separately for the distributions of \( \eta_t \) and \( \xi_t \) at time \( t \) to determine the vertices of the (single) \( \times \)-simplex \( \Theta_t \times \Xi_t \) in the initial partition.

According to results in [16] where we applied successive refinements to initial partitions consisting of only one \( \times \)-simplex, the convergence is rather slow when the (lower) bound based on the vertices of \( \Theta_t \) and barycenters for \( \xi_t \) is refined while the other (upper) bound seems to be stable. The former effect can be explained geometrically by “too extreme” coordinates of the
vertices we selected to cover the support of $\eta_t$. As a consequence, the saddle function exhibits a relatively large degree of concavity with respect to the risk factors $\eta_t$ and cannot be approximated well by bilinear functions.

Since the split of a simplex generates only one new point but the existing vertices remain in the partition, the influence of extreme outcomes in the initial discretization decreases slowly. Furthermore, the split of a $\times$-simplex in the partition for the first stage provides a better improvement than in case of the partition where the largest discretization error was observed. This is not surprising since the approximation of the distribution in $t = 0$ affects all scenarios and, hence, has the highest impact on the solution.

4.4 Disretization of interest rate processes

To achieve tighter bounds, possibly with less refinement steps, we modified the procedure for the determination of initial coverages in combination with the discrete-time version of the two-factor mean reversion model (8). While a triangle covers the truncated support of $\eta_t = (s_t, l_t)'$ with only three vertices which keeps the scenario size moderate, a better approximation of the circle around the mass of the (uncorrelated) standard normal distribution can obviously be obtained using polygons with $K_t' > 3$ vertices (see Figure 7). More precisely, we determine pentagons as approximations of the (unit) circle that are partitioned into three triangles for the first stages of the stochastic program. Combining them with intervals that cover the support of $\xi_t$ results in initial partitions of $\ell_t = 3$ cross-simplices. After the transformation according to the expectation and covariance matrix of the actual distributions at time $t$, generalized barycenters and weights are determined for each of them individually according to (12)–(15).

Because this discretization yields 9 outcomes for the lower bound (corresponding to the number of vertices in the partition), we use squares partitioned into two simplices ($\ell_t = 2$) and then single triangles ($\ell_t = 1$) at later stages where the approximation has less impact on the solution of the complete stochastic program in order to reduce the growth in problem size. The specific size $\ell_t$ of the initial partitions for all $t = 1, \ldots, T$ depends on the number of periods that must be achieved. The total number of scenarios at the $t$-th stage is given by

---

1General polyhedra to obtain tighter bounds are also exploited in Gassmann and Ziemba [18]. However, they develop an upper bound for the case of single polyhedrons with arbitrary number of vertices while we divide the initial polyhedron that covers the support of $\eta_t$ in regular sub-simplices to apply the barycentric approximation scheme.
Figure 7: Coverages for the uncorrelated and correlated case with barycenters

\[
s^l_t := \prod_{\tau=1}^{t} (K_\tau + 1) \cdot \ell_\tau = 3^t \cdot \prod_{\tau=1}^{t} \ell_\tau
\]

\[
s^u_t := \prod_{\tau=1}^{t} (L_\tau + 1) \cdot \ell_\tau = 2^t \cdot \prod_{\tau=1}^{t} \ell_\tau
\]

for the lower and upper bound, respectively. In our experience, we obtain a better accuracy for a similar problem size compared to sequential refinements of partitions where we start with a single \(\times\)-simplex for the discretization. For example, in case of an 8-stage quarterly planning problem (\(D^S = \{3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y\}\)) with \(\ell_1 = 3, \ell_2 = 2, \ell_3 = \ldots = \ell_7 = 1\) and \(\delta = 2\), we have 13122 scenarios in the lower (larger) approximation. The objective function values are 5818.6 for the lower bound and 5888.4 for the upper bound\(^2\), equivalent to a relative difference of only 1.20%.

In case of the principal component model for short-term planning, we make use of the fact that the \(K_t\) components of \(\eta_t\) are orthogonal (uncorrelated) by construction. This allows to represent the \(\Theta_{t,1}^{K_t}\), themselves as a \(\times\)-simplices, i.e., \(\Theta_t = \Theta^{K_1}_t \times \ldots \times \Theta^{K_t}_t\). Each \(\Theta^{K_i}_i = [-\delta, \delta] \subset \mathbb{R}\) is an interval that covers the truncated support of the \(i\)-th principal component \(\eta^{K_i}_i\), \(i = 1, \ldots, K_t\). The required modified formulae for the barycentric approximation scheme can be found in [12] for the two-stage case. For example, to obtain the generalized barycenters \(\eta_{\mu_t} = (\eta^{1}_{\mu_t}, \ldots, \eta^{K_t}_{\mu_t})\) and corresponding probabilities, we have to replace \(\lambda_{\nu_t}(\eta_t)\) in (12)–(15) by \(\prod_{i=1}^{K_t} \lambda_{\nu_t^{i}}(\eta^{i}_t)\) with \(\nu_t = (\nu^{1}_t, \ldots, \nu^{K_t}_t), \nu^{i}_t = 0, \ldots, K_t\). Note that each principal component \(\eta^{i}_t\) may still be correlated with the volume risk factor \(\xi_t\).

The \(\times\)-simplex \(\Theta_t\) as the Cartesian product of \(K_t\) intervals is a multi-dimensional rectangle with \(2^{K_t}\) vertices, and the number of scenarios at stage \(t\) is given by \(s^l_t := \prod_{\tau=1}^{t} 2^{K_\tau}\) for the lower bound. The growth in problem

\(^2\)Process parameters: \(\kappa_s = 0.8687, \theta_s = -0.0061, \sigma_s = 0.0223, \kappa_t = 0.1940, \theta_t = 0.0507, \sigma_t = 0.1532, \rho = -0.3610, \phi_1 = -0.002689, \phi_2 = 0.4293\) for the term structure model (\(\Delta t = 0.25\)) and \(a = 83.82, b = 0, \rho_\psi = 0.9835, \sigma_\xi = 166.0\) for the volume process.
size for the upper approximation with \( s^* := \prod_{t=1}^{T} 2^{L_t} \) is less dramatic since we assumed \( L_t = L = 1 \). Thus, with \( K_1 = K = 3 \) relevant factors, the size of the scenario set in the former case is multiplied by 8 with each additional stage and no refinements.

Since this increase may be too restrictive for the number of periods that can be taken into account, the growth in problem size may be reduced by considering all principal components only in the first stages while the third and then the second factor are ignored at later points in time. This is motivated by the empirical observation that they contribute only 3% and 19%, respectively, to the dynamics of interest rates (cf. [31]). As a consequence, the scenarios incorporate complex tilt and humped movements at the beginning but may reflect only shifts of the yield curve towards the end of the planning horizon where the outcomes have less impact on the first-stage decision. In this way, we achieve a lower and upper bound of 12830.9 and 12873.3, i.e., a relative difference of 0.33% for a 7-stage monthly planning problem with maturities \( D^S = \{1M, 2M, 3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y\} \) and dimension size \( K_1 = K_2 = 3, K_3 = \ldots = K_6 = 2 \) which results in 16384 scenarios.

5 Application to the funding problem

Scenarios generated by means of any discretization method may be represented by a tree (e.g., see [14] for a formal description). Nodes of this scenario tree with depth \( t \) correspond to outcomes of random data at time \( t \). Using the barycentric approximation scheme described above results in a total number of nodes \( \sum_{t=0}^{T} s^t_1 \) for the lower bounding problem which is the larger one since we have at least as many interest rate than volume risk factors (\( K_t \geq L_t \forall t \)). The funding model (3) has \( D + \sum_{d \in D^S} I^d \) variables at each stage. For instance, with traded maturities \( D^S = \{3M, 6M, 1Y, 2Y, 3Y, 4Y, 5Y\} \) and 8 tranches for each of them, we have 76 variables which must be duplicated for all nodes. Thus, the corresponding deterministic program for the example above with 13122 scenarios already consists of more than 1.5 million variables in 19666 nodes. The problem generation and solution with standard optimization tools such as Cplex requires up to a few hours on a medium-size

\[ 3 \text{We estimated the factor sensitivities for CHF Euromarket rates as shown in Figure 3 with principal component analysis, and the correlations between the factor scores and the volume risk factor } \xi_t \text{ are } 0.34, -0.18 \text{ and } -0.03. \text{ The different magnitudes of the objective values obtained with the mean reversion model above results from the fact that only a fraction of the mortgage volume could be included in the optimization there due to quarterly planning, i.e., the actual portfolio had to be split into four components since the period length was extended to three months there.} \]
workstation (Sun Ultra 10, 1 GB RAM). This may be seen as tolerable since
the determination of refinancing policies is performed only once a month.

As a consequence of the exponential growth in problem size, we can deal
with a limited number of stages only although the planning horizon of the
refinancing problem is actually infinite. Our experience from various case
studies is that an increase in the number of stages (even at the cost of a
reduced accuracy of the approximation) in general leads to a higher per-
formance. This can be explained by the fact that the model has a greater
flexibility for corrections of the initial portfolio. Moreover, an extension of
the horizon allows a better consideration of the impact of future changes in
the risk factors (e.g., a sharp drop in volume that might lead to a surplus of
liabilities over the mortgage position) for the first-stage decision. However, in
practice we are not able to solve problems with more than 10 stages with the
model formulation and solution techniques described before. If we wanted to
include the whole spectrum of traded instruments in the interbank market
from 1 month to 5 years and since the period length is given by the shortest
maturity, this would correspond to a planning horizon of less than one year
which we consider as insufficient for the long-term model.

Therefore, the determination of refinancing decisions is carried out in two
steps: First, we perform an optimization run with maturities between one and
two years where scenarios are based on the mean reversion model. When the
solution indicates that a short-term policy should be implemented, we use the
principal component model to analyze if the funding costs can be reduced
by shorter maturities than one year. Alternatively, the liability portfolio
may be split into a short- and a long-term component that are optimized
separately. The corresponding volumes depend, e.g., on the current portfolio
composition, limits given by the bank’s internal risk management system, or
forecasts for interest rates and mortgage demand.

Before the stochastic optimization model was applied to real positions, a
case study had been conducted to assess its performance for a period of high
yields. The study was based on monthly money market rates and the volume
of a real mortgage position provided by a Swiss bank (see Figure 8, the lag
between interest rates and volume results from the fact that banks adjust the
relevant client rate with some delay to a new market situation). The level of
interest rates at which mortgages had to be refinanced on the market reached
up to 10% while the client rate that the bank receives never climbed above
7% due to the political cap in Switzerland. As a consequence, refinancing the
mortgage position with a replicating portfolio of 25% six-month, 50% one-
year and 25% three-year instruments which serves as benchmark provides
a (negative) average margin of $-0.21\%$ for the sample period. The static
policy required also to *invest* significant amounts at low yields because the
share of portfolio positions with longer maturities was still too large when the mortgage demand dropped, resulting in a surplus of liabilities.

For simplicity, we did not follow the “2-step procedure” outlined before with separate runs of the long-term and the short-term planning models. Instead, we used the former only with maturities from 3 months to 5 years. With 9 stages, this quarterly planning is equivalent to a time horizon of only two years. According to our experience, the performance of the model improves significantly for longer periods. In order to achieve this number of stages, we did not split the initial partitions.

Liquidity limits for transactions without penalties ranged from 400 mio. for the shortest to 100 mio. for the longest maturity. Transactions costs increased by one basis point (BP) with each additional tranche of 50 mio. except for 3 months where tranches are twice as large. For the initial portfolio composition, it was assumed that the static 6M/1Y/3Y-mix had been implemented in the past, i.e., all positions had to be renewed within the first three years. Parameters for the stochastic processes were updated semi-annually based on observations of the previous five years.

<table>
<thead>
<tr>
<th>method</th>
<th>ref. costs [%]</th>
<th>margin [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>avg.</td>
<td>std.dev.</td>
</tr>
<tr>
<td>SP model</td>
<td>6.13</td>
<td>1.38</td>
</tr>
<tr>
<td>repl. portf.</td>
<td>6.50</td>
<td>1.80</td>
</tr>
</tbody>
</table>

Table 1: Comparison of dynamic policies with the replicating portfolio (RP)

After a refinancing decision had been found, the portfolio positions were updated and a new optimization was started with the next set of interest rates and mortgage volume out of the sample period (“roll-over planning”). The result after all 73 runs is summarized in Table 1. Compared to the replicating portfolio benchmark, the average refinancing costs for the renewed positions
could be reduced from 6.50% to 6.13%. The margin as the difference between the client rate and the refinancing costs increased correspondingly by 37 BP to 0.16%. While this number implies only a small profit at first sight, one must take into account that the static approach was not able to provide a positive margin at all for the specific market situation in the case study which results in significant losses for the bank. Moreover, the standard deviation was also reduced noticeably compared to the replicating portfolio although the latter was constructed to minimize the volatility of the margin.

6 Conclusions

An extended version of the multistage stochastic programming model described in this paper has been in use now by a major Swiss bank since 1995 for refinancing variable mortgages. The modifications are, e.g., additional constraints for the portfolio duration to achieve a target that is frequently defined by the bank’s board of directors. While the two-factor mean reversion model provides a good description of both ends of the term structure due to the selection of state variables, the error of fit is slightly worse for medium-term maturities. Therefore, a third factor has been introduced that controls the curvature of the yield curve. Unfortunately, we cannot reveal all details of this commercial model in publications.

As evidence for the performance of the model since it has been applied in practice, the average funding rate compared to the static mix that was exploited by the bank before as benchmark is shown in Figure 9. With an average margin of approximately 70 BP that could be achieved with the replicating portfolio approach in the long run, an increase in the order of 37 BP represents a significant improvement of the bank’s profits. Before the application of another version of the model to savings accounts started in
1997, another case study with data of a real position was conducted (see Frauendorfer and Schürle [16] for a detailed description). The results were compared to the performance of two replicating portfolios that had been used by the bank for the management of different deposit positions.

Figure 10 shows the evolution of the margin between the return of the invested funds and the client rate for the dynamic policies determined by the stochastic optimization model and the static benchmark policies. According to Table 2, the average margin over the sample period could be improved by 25 BP compared to the better replicating portfolio while volatility is significantly reduced. The increase in performance at lower risk can be explained by the possibility of rebalancing transactions in the dynamic approach. Similar observations have also been reported for other multistage stochastic programming models (e.g., see Cariño and Ziemba [5]).

<table>
<thead>
<tr>
<th>method</th>
<th>weights repl. portf.</th>
<th>margin [%]</th>
<th>mean</th>
<th>std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 y</td>
<td>2 yrs</td>
<td>5 yrs</td>
<td></td>
</tr>
<tr>
<td>SP model</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>2.66</td>
</tr>
<tr>
<td>repl. portf. 1</td>
<td>0.0 0.5 0.5</td>
<td></td>
<td></td>
<td>2.41</td>
</tr>
<tr>
<td>repl. portf. 2</td>
<td>0.35 0.35 0.3</td>
<td></td>
<td></td>
<td>2.40</td>
</tr>
</tbody>
</table>

Table 2: Performance of stochastic optimization model for savings accounts

There are many possible ways to improve the model in the form presented here, e.g., with respect to the volume process specification (9). However, an empirical investigation of alternative models failed because of the lack of publicly available data. Moreover, new mortgage products have recently been introduced in Switzerland that led to a noticeable change in the demand for variable mortgages, and it is difficult to correct the structural break induced by such a “non-stochastic event” in the data. We are also still investigating other term structure models with respect to their practical use for scenario generation.
Some authors derive interest rate scenarios from (binomial) lattice models that have been developed for pricing derivative securities. In a continuous-time framework, this is equivalent to a one-factor specification with time-dependent parameters. Since these models are calibrated only to the term structure observed at present time, we are not confident if such an approach can provide a suitable description of long-term interest rate behavior but should rather be seen as an interpolation technique for the initial yield curve. As Hull and White [22] point out: “It is important to distinguish between the goal of developing a model that adequately describes term-structure movements and the goal of developing a model that adequately values most of the interest-rate-contingent claims that are encountered in practice. It is quite possible that a two- or three-state variable model is necessary to achieve the first goal”.

According to our experience, models for a planning horizon of several years should be calibrated to a historical data set that covers at least an economic cycle. In can also be helpful to assess term structure model by comparing scenarios generated in a simulation study with the characteristics of empirically observed data as suggested by Frauendorfer and Schürle [17]. Interest rate scenarios based on the principal component and the two-factor mean reversion models in section 3.1 may not be free of arbitrage. This aspect should be taken into account in applications with simultaneous investing and refinancing. Otherwise, the model could try to exploit such “spurious” arbitrage opportunities in its decisions although they result only from a miss-specification (e.g., see Pflug [29] for arbitrage-free scenario generation methods).

Using barycentric approximation, we are able to obtain tight bounds for the original problem when the support of random data is partitioned into several $\times$-simplices. Despite the low discretization error, different first-stage decisions for the lower and upper approximate problems may still occur. This requires an improvement of the scenario selection, and the investigation of efficient refinement techniques based on suitable error measures is still ongoing. However, the example of the funding problem presented here and many other approaches described in this book show that multistage stochastic programming models have already passed the state of research and may be applied successfully to financial decision making under uncertainty.

Acknowledgement: We are grateful to W.T. Ziemba for his helpful comments on an earlier version of this paper that helped improve the exposition significantly.
References


