Consistent Re-Calibration of the Discrete Time Multifactor Vasiček Model

Unfinished version, please do not distribute!

Philipp Harms, David Stefanovits, Josef Teichmann, Mario V. Wüthrich

September 21, 2015

Abstract

The discrete time multifactor Vasiček model is a tractable Gaussian spot rate model. Typically, two- or three-factor versions allow to capture the dependence structure between yields with different times to maturity in an appropriate way. In practice, re-calibration of the model to the prevailing market conditions leads to model parameters which change over time. Therefore, the model parameters should be understood as being time-dependent, or even stochastic. Following the consistent re-calibration (CRC) approach of [3] we construct models as concatenations of yield curve increments of Hull-White extended multifactor Vasiček models with different parameters. The CRC approach provides attractive tractable models that preserve the no-arbitrage premise. As a numerical example we fit Swiss interest rates using CRC multifactor Vasiček models.

1 Introduction

The tractability of affine models, such as the Vasiček [7] and the Cox-Ingersoll-Ross [2] models, has made them appealing for term structure modelling. Affine term structure models are based on a (multidimensional) factor process which in turn describes the evolution of the spot rate and the bank account processes. No arbitrage arguments then provide the corresponding zero-coupon bond prices, yield curves and forward rates. Prices in these models are calculated under an equivalent martingale measure for known static model parameters. However, model parameters typically vary over time as financial market conditions change. They may for instance be of regime switching nature and need to be permanently re-calibrated to the actual financial market conditions. In practice, this re-calibration is done on a regular basis (as new information is available). This implies that model parameters are not static and, henceforth, may also be understood as stochastic processes. The re-calibration should preserve the no-arbitrage condition, which provides side constraints in the re-calibration. The aim of this work is to discuss these side constraints with the help of the multifactor Vasiček interest rate model which is tractable

but also flexible model. We show that re-calibration of spot rate models, which satisfies the side constraints, naturally leads to Heath-Jarrow-Morton [4] models with stochastic parameters.

Organisation of the paper. In Section 2 we introduce Hull-White extended multifactor Vasiček models which are the building blocks for consistent re-calibration (CRC) in this work. We define CRC of the Hull-White extended multifactor Vasiček model in Section 3. Section 4 specifies the market price of risk assumptions used to model the factor process under the real world probability measure and the equivalent martingale measure, respectively. In Section 5 we focus on estimation using real market data. Finally, in Section 6 we study the performance of the model on Swiss interest rate data. All proofs are presented in Appendix A, and all figures are presented in Appendix B of this paper.

2 Discrete time Vasiček model and Hull-White extension

Choose a fixed grid size $\Delta > 0$ and consider the discrete time grid $\{0, \Delta, 2\Delta, 3\Delta, \ldots\} = \Delta\mathbb{N}_0$. Choose a (sufficiently rich) filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P}^*)$ with discrete time filtration $\mathbb{F} = (\mathcal{F}(t))_{t\in\mathbb{N}_0}$, where $t \in \mathbb{N}_0$ refers to time point Δt. Assume that $\mathbb{P}^*$ denotes an equivalent martingale measure for a (strictly positive) bank account numeraire $(B(t))_{t\in\mathbb{N}_0}$. $B(t)$ denotes the value at time Δt of an investment of one unit of currency at time 0 into the bank account (i.e. risk-free rollover relative to Δ).

2.1 Discrete time multifactor Vasiček model

We choose $n \in \mathbb{N}$ fixed and introduce the discrete time n-factor Vasiček model.

Notation. Subscript indices refer to elements of vectors and matrices. Argument indices refer to the time points. We denote the $n \times n$ identity matrix by $\mathbb{1} \in \mathbb{R}^{n\times n}$. We also introduce the vectors $\mathbf{1} = (1, \ldots, 1)^\top \in \mathbb{R}^n$ and $\boldsymbol{e}_1 = (1, 0, \ldots, 0)^\top \in \mathbb{R}^n$.

We consider the n-dimensional $\mathbb{F}$-adapted factor process

$$\boldsymbol{X} = (\boldsymbol{X}(t))_{t\in\mathbb{N}_0} = (X_1(t), \ldots, X_n(t))^\top_{t\in\mathbb{N}_0},$$

which generates the spot rate and bank account processes as follows

$$r(t) = \mathbf{1}^\top \boldsymbol{X}(t) \quad \text{and} \quad B(t) = \exp\left\{\Delta \sum_{s=0}^{t-1} r(s)\right\}, \tag{2.1}$$

where $t \in \mathbb{N}_0$; empty sums are set equal to zero. The factor process $\boldsymbol{X}$ is assumed to evolve under $\mathbb{P}^*$ according to

$$\boldsymbol{X}(t) = \boldsymbol{b} + \beta \boldsymbol{X}(t-1) + \Sigma^{\frac{1}{2}} \boldsymbol{\varepsilon}^*(t), \quad t > 0, \tag{2.2}$$

with initial factor $\boldsymbol{X}(0) \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^n$, $\beta \in \mathbb{R}^{n\times n}$, $\Sigma^{\frac{1}{2}} \in \mathbb{R}^{n\times n}$ and $(\boldsymbol{\varepsilon}^*(t))_{t\in\mathbb{N}} = (\varepsilon_1^*(t), \ldots, \varepsilon_n^*(t))^\top_{t\in\mathbb{N}}$ being $\mathbb{F}$-adapted. The following assumptions are in place throughout the paper.

Assumptions 2.1 *We assume that the spectrum of matrix β is a subset of $(-1,1)^n$ and that matrix $\Sigma^{\frac{1}{2}}$ is non-singular. Moreover, for each $t \in \mathbb{N}_0$, we assume that $\boldsymbol{\varepsilon}^*(t)$ is independent of $\mathcal{F}(t-1)$ under $\mathbb{P}^*$ with $\boldsymbol{\varepsilon}^*(t) \overset{\mathbb{P}^*}{\sim} \mathcal{N}(\mathbf{0}, \mathbb{1})$.*

Remark. In Assumptions 2.1 the condition on matrix β ensures that $\mathbb{1} - \beta$ is invertible and that the geometric series generated by β converges. The condition on $\Sigma^{\frac{1}{2}}$ ensures that $\Sigma = \Sigma^{\frac{1}{2}}(\Sigma^{\frac{1}{2}})^\top$ is symmetric positive definite. Under Assumptions 2.1 equation (2.2) defines a stationary process, see [1, Section 11.3].

The model defined by equations (2.1)-(2.2) is called discrete time multifactor Vasiček model. Under the model assumptions we have for $m > t$

$$\boldsymbol{X}(m)|\mathcal{F}(t) \overset{\mathbb{P}^*}{\sim} \mathcal{N}\left((\mathbb{1} - \beta)^{-1} \left(\mathbb{1} - \beta^{m-t}\right) \boldsymbol{b} + \beta^{m-t} \boldsymbol{X}(t), \sum_{s=0}^{m-t-1} \beta^s \Sigma (\beta^\top)^s \right). \tag{2.3}$$

Remark. For $m > t$ the conditional distribution of $\boldsymbol{X}(m)$, given $\mathcal{F}(t)$, depends only on the value $\boldsymbol{X}(t)$ at time Δt and on lag $m-t$. In other words, the factor process (2.2) is a time-homogeneous Markov process.

At time Δt the price of the zero-coupon bond (ZCB) with maturity date $\Delta m > \Delta t$ with respect to filtration $\mathbb{F}$ and equivalent martingale measure $\mathbb{P}^*$ is given by

$$P(t,m) = \mathbb{E}^* \left[\frac{B(t)}{B(m)} \middle| \mathcal{F}(t) \right] = \mathbb{E}^* \left[\exp\left\{ -\Delta \sum_{s=t}^{m-1} \mathbf{1}^\top \boldsymbol{X}(s) \right\} \middle| \mathcal{F}(t) \right].$$

For the proof of the following result see Appendix A.

Theorem 2.2 *The ZCB prices in the discrete time multifactor Vasiček model* (2.1)-(2.2) *with respect to filtration $\mathbb{F}$ and equivalent martingale measure $\mathbb{P}^*$ have affine term structure*

$$P(t,m) = e^{A(t,m) - \boldsymbol{B}(t,m)^\top \boldsymbol{X}(t)}, \quad m > t,$$

with $A(m-1,m) = 0$, $\boldsymbol{B}(m-1,m) = \Delta \mathbf{1}$ and for $m-1 > t \geq 0$

$$\begin{aligned} A(t,m) &= A(t+1,m) - \boldsymbol{B}(t+1,m)^\top \boldsymbol{b} + \frac{1}{2} \boldsymbol{B}(t+1,m)^\top \Sigma \boldsymbol{B}(t+1,m), \\ \boldsymbol{B}(t,m) &= \Delta \left(\mathbb{1} - \beta^\top\right)^{-1} \left(\mathbb{1} - (\beta^\top)^{m-t}\right) \mathbf{1}. \end{aligned}$$

In discrete time multifactor Vasiček model (2.1)-(2.2) the term structure of interest rates (yield curve) takes the following form at time Δt for maturity dates $\Delta m > \Delta t$

$$Y(t,m) = -\frac{1}{\Delta(m-t)} \log P(t,m) = -\frac{A(t,m)}{\Delta(m-t)} + \frac{\boldsymbol{B}(t,m)^\top \boldsymbol{X}(t)}{\Delta(m-t)}, \tag{2.4}$$

with spot rate at time Δt given by $Y(t,t+1) = \mathbf{1}^\top \boldsymbol{X}(t) = r(t)$.

Remark. All models are formulated relative to the filtration $\mathbb{F}$. However, the underlying factor process $\boldsymbol{X}$ is in general not observable in the market. For this reason the ZCB prices in Theorem 2.2 with respect to filtration $\mathbb{F}$ usually cannot be calculated. In the market we only have information about the yield curves. Hence, for parameter estimation we may consider the filtration generated by the term structure of interest rates

$$\mathcal{F}^Y(t) = \sigma\left(\{Y(s,m) \mid s < m, s \leq t\}\right) \subset \mathcal{F}(t), \quad t \geq 0,$$

i.e. we assume that the yield curve process is observable but not necessarily the factor process $\boldsymbol{X}$. Kalman filter techniques are used to infer the unobservable $\boldsymbol{X}$-dynamics from yield curve observations, see [8, Section 3.6.3] and Section 5.2 below.

2.2 Hull-White extended discrete time multifactor Vasiček model

The possible shapes of the Vasiček yield curve (2.4) are restricted by the choice of the parameters $\boldsymbol{b} \in \mathbb{R}^n$, $\beta \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{n\times n}$. These parameters are not sufficient to exactly calibrate the model to an arbitrarily observed initial yield curve. Therefore, we consider the Hull-White extended version (see [5]) of the discrete time multifactor Vasiček model. We replace the factor process defined in (2.2) as follows. For fixed $k \in \mathbb{N}_0$ let $\boldsymbol{X}^{(k)}$ satisfy

$$\boldsymbol{X}^{(k)}(t) = \boldsymbol{b} + \theta(t-k)\boldsymbol{e}_1 + \beta \boldsymbol{X}^{(k)}(t-1) + \Sigma^{\frac{1}{2}}\boldsymbol{\varepsilon}^*(t), \quad t > k, \tag{2.5}$$

with starting factor $\boldsymbol{X}^{(k)}(k) \in \mathbb{R}^n$, $\boldsymbol{e}_1 = (1,0,\ldots,0)^\top \in \mathbb{R}^n$ and function $\theta : \mathbb{N} \to \mathbb{R}$. Model assumption (2.5) corresponds to (2.2) where the first component of $\boldsymbol{b}$ is replaced by the time dependent coefficient $(b_1+\theta(i))_{i\in\mathbb{N}}$ and all other terms ceteris paribus. Without loss of generality we choose the first component for this replacement. Note that parameter b_1 is redundant in the model specification, but for didactical reasons it is used below. The time dependent coefficient θ is called Hull-White extension and it is used to calibrate the model to a given yield curve at a given time point Δk. The upper index $^{(k)}$ denotes that time point and corresponds to the time shift we apply to the Hull-White extension θ in model (2.5). The factor process $\boldsymbol{X}^{(k)}$ generates the spot rate process and the bank account process as in (2.1).
The model defined by assumptions (2.1, 2.5) is called Hull-White extended discrete time multifactor Vasiček model. Under these model assumptions we have for $m > t \geq k$

$$\boldsymbol{X}^{(k)}(m)|\mathcal{F}(t) \overset{\mathbb{P}^*}{\sim} \mathcal{N}\left(\sum_{s=0}^{m-t-1} \beta^s\left(\boldsymbol{b} + \theta(m-s-k)\boldsymbol{e}_1\right) + \beta^{m-t}\boldsymbol{X}^{(k)}(t), \sum_{s=0}^{m-t-1} \beta^s \Sigma (\beta^\top)^s\right).$$

Remark. For $m > t \geq k$ the conditional distribution of $\boldsymbol{X}^{(k)}(m)$, given $\mathcal{F}(t)$, depends only on the factor $\boldsymbol{X}^{(k)}(t)$ at time Δt, and on lags $m-t$ and $m-k$. In other words, factor process (2.5) is a time-inhomogeneous Markov process in this case. Note that the upper index $^{(k)}$ in the notation is important since the conditional distribution depends explicitly on the lag $m-k$.

Theorem 2.3 *The ZCB prices in the Hull-White extended discrete time multifactor Vasiček model (2.1, 2.5) with respect to filtration $\mathbb{F}$ and equivalent martingale measure $\mathbb{P}^*$ have affine term structure*

$$P^{(k)}(t,m) = e^{A^{(k)}(t,m) - \boldsymbol{B}(t,m)^\top \boldsymbol{X}(t)}, \quad m > t \geq k,$$

with $\boldsymbol{B}(t,m)$ as in Theorem 2.2, $A(m-1,m) = 0$ and for $m-1 > t \geq k$

$$\begin{aligned} A^{(k)}(t,m) = A^{(k)}(t+1,m) &- \boldsymbol{B}(t+1,m)^\top \left(\boldsymbol{b} + \theta(t+1-k)\boldsymbol{e}_1\right) \\ &+ \frac{1}{2}\boldsymbol{B}(t+1,m)^\top \Sigma \boldsymbol{B}(t+1,m). \end{aligned}$$

Proof of Theorem 2.3. Proof by induction as the proof of Theorem 2.2. □

In the Hull-White extended discrete time multifactor Vasiček model (2.1, 2.5) the yield curve takes the following form at time Δt for maturity dates $\Delta m > \Delta t \geq \Delta k$

$$Y^{(k)}(t,m) = -\frac{1}{\Delta(m-t)} \log P^{(k)}(t,m) = -\frac{A^{(k)}(t,m)}{\Delta(m-t)} + \frac{\boldsymbol{B}(t,m)^\top \boldsymbol{X}^{(k)}(t)}{\Delta(m-t)}, \tag{2.6}$$

with spot rate at time Δt given by $Y^{(k)}(t,t+1) = \mathbf{1}^\top \boldsymbol{X}^{(k)}(t)$.

Remark. Note that the coefficient $\boldsymbol{B}(t,m)$ in Theorem 2.3 is not affected by the Hull-White extension θ and depends solely on $m-t$, whereas the coefficient $A^{(k)}(t,m)$ depends explicitly on θ, and on lags $m-t$ and $t-k$.

2.3 Calibration of the Hull-White extended model

We consider the term structure model defined by the Hull-White extended factor process $\boldsymbol{X}^{(k)}$ and calibrate the function $\theta \in \mathbb{R}^{\mathbb{N}}$ to a given yield curve at time point Δk. We explicitly introduce the time index k in model (2.5) because the CRC algorithm is a concatenation of multiple Hull-White extended models which are calibrated at different time points Δk, see Section 3 below.

Assume that there is a fixed final time to maturity ΔM and that we observe at time Δk a yield curve $\widehat{\boldsymbol{y}}(k) \in \mathbb{R}^M$ for maturity dates $\Delta(k+1), \ldots, \Delta(k+M)$. For these maturity dates the Hull-White extended discrete time multifactor Vasiček yield curve at time Δk is given by Theorem 2.3, and reads as

$$\boldsymbol{y}^{(k)}(k) = \left(-\frac{1}{\Delta i} A^{(k)}(k,k+i) + \frac{1}{\Delta i}\boldsymbol{B}(k,k+i)^\top \boldsymbol{X}^{(k)}(k)\right)^\top_{i=1,\ldots,M} \in \mathbb{R}^M.$$

For given starting factor $\boldsymbol{X}^{(k)}(k) \in \mathbb{R}^n$, and parameters $\boldsymbol{b} \in \mathbb{R}^n$, $\beta \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{n\times n}$ our aim is to choose the Hull-White extension $\theta \in \mathbb{R}^{\mathbb{N}}$ such that we get an exact fit at time Δk

$$\boldsymbol{y}^{(k)}(k) = \widehat{\boldsymbol{y}}(k). \tag{2.7}$$

The following theorem provides an equivalent condition to (2.7) which allows to calculate the Hull-White extension $\theta \in \mathbb{R}^{\mathbb{N}}$ explicitly.

Theorem 2.4 *Denote by $\boldsymbol{y}^{(k)}(k)$ the yield curve at time Δk obtained from the Hull-White extended discrete time multifactor Vasiček model (2.1, 2.5) for given starting factor $\boldsymbol{X}^{(k)}(k) = \boldsymbol{x} \in \mathbb{R}^n$, parameters $\boldsymbol{b} \in \mathbb{R}^n$, $\beta \in \mathbb{R}^{n\times n}$ and $\Sigma \in \mathbb{R}^{n\times n}$, and Hull-White extension $\theta \in \mathbb{R}^{\mathbb{N}}$. For given $\boldsymbol{y} \in \mathbb{R}^M$ identity $\boldsymbol{y}^{(k)}(k) = \boldsymbol{y}$ holds if and only if the Hull-White extension $\boldsymbol{\theta}$ fulfils*

$$\boldsymbol{\theta} = \mathcal{C}(\beta)^{-1}\boldsymbol{z}\left(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{x}, \boldsymbol{y}\right), \tag{2.8}$$

where $\boldsymbol{\theta} = (\theta_i)^\top_{i=1,\dots,M-1} \in \mathbb{R}^{M-1}$, $\boldsymbol{z}\left(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{x}, \boldsymbol{y}\right) = \left(z_i\left(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{x}, \boldsymbol{y}\right)\right)^\top_{i=1,\dots,M-1} \in \mathbb{R}^{M-1}$ and $\mathcal{C}(\beta) = \left(\mathcal{C}_{ij}(\beta)\right)_{i,j=1,\dots,M-1} \in \mathbb{R}^{(M-1)\times(M-1)}$ are defined by

$$\begin{aligned}
\theta_i &= \theta(i),\\
\mathcal{C}_{ij}(\beta) &= B_1(k+j, k+i+1)\ 1_{\{j\le i\}},\\
z_i\left(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{x}, \boldsymbol{y}\right) &= \sum_{s=k+1}^{k+i}\left(\frac{1}{2}\boldsymbol{B}(s,k+i+1)^\top\Sigma\boldsymbol{B}(s,k+i+1) - \boldsymbol{B}(s,k+i+1)^\top\boldsymbol{b}\right)\\
&\qquad - \Delta\mathbf{1}^\top\left(\mathbb{1} - \beta^{i+1}\right)\left(\mathbb{1}-\beta\right)^{-1}\boldsymbol{x} + \Delta(i+1)y_{i+1}(k),
\end{aligned}$$

with $i, j = 1, \dots, M-1$ and $\boldsymbol{B}(\cdot,\cdot) = \left(B_1(\cdot,\cdot), \cdots, B_n(\cdot,\cdot)\right)^\top$ given by Theorem 2.2.

Theorem 2.4 shows that the Hull-White extension can be calculated inverting the $(M-1)\times(M-1)$ lower triangular positive definite matrix $\mathcal{C}(\beta)$.

3 Consistent re-calibration

The crucial extension is the following: we let parameters $\boldsymbol{b}$, β and Σ vary over time and we re-calibrate the Hull-White extension in a consistent way at each time point, that is, according to the actual choice of the parameter values (using Theorem 2.4). Below we show that this naturally leads to a Heath-Jarrow-Morton [4] (HJM) point of view for term structure modelling.

3.1 Consistent re-calibration algorithm

Assume that $(\boldsymbol{b}(k))_{k\in\mathbb{N}_0}$, $(\beta(k))_{k\in\mathbb{N}_0}$ and $(\Sigma(k))_{k\in\mathbb{N}_0}$ are $\mathbb{F}$-adapted parameter processes with $\beta(k)$ and $\Sigma(k)$ satisfying Assumptions 2.1, $\mathbb{P}^*$-a.s., for all $k \in \mathbb{N}_0$. Based on these parameter processes we will define the n-dimensional $\mathbb{F}$-adapted CRC factor process $\boldsymbol{\mathcal{X}}$ which evolves according to steps (i)-(iv) of the CRC algorithm described below. Thus, factor process $\boldsymbol{\mathcal{X}}$ will define a spot rate model similar to (2.1).
The CRC algorithm works as follows.

(i) Initialisation $k = 0$. Assume that the initial yield curve observation at time 0 is given by $\widehat{\boldsymbol{y}}(0) \in \mathbb{R}^M$. Let $\theta^{(0)} \in \mathbb{R}^{\mathbb{N}}$ be an $\mathcal{F}(0)$-measurable Hull-White extension such that condition (2.7) is satisfied at time 0 for initial factor $\boldsymbol{\mathcal{X}}(0) \in \mathbb{R}^n$, and parameters $\boldsymbol{b}(0)$, $\beta(0)$ and $\Sigma(0)$. By Theorem 2.4 the values $\boldsymbol{\theta}^{(0)} = (\theta^{(0)}(i))_{i=1,\dots,M-1} \in \mathbb{R}^{M-1}$ are given by

$$\boldsymbol{\theta}^{(0)} = \mathcal{C}\left(\beta(0)\right)^{-1}\boldsymbol{z}\left(\boldsymbol{b}(0), \beta(0), \Sigma(0), \boldsymbol{\mathcal{X}}(0), \widehat{\boldsymbol{y}}(0)\right).$$

This provides Hull-White extended Vasiček yield curve $\boldsymbol{y}^{(0)}(0)$ identically equal to $\widehat{\boldsymbol{y}}(0)$ for given initial factor $\boldsymbol{\mathcal{X}}(0)$ and parameters $\boldsymbol{b}(0)$, $\beta(0)$, $\Sigma(0)$.

(ii) Increments of the factor process from $k \to k+1$. Assume factor $\boldsymbol{\mathcal{X}}(k)$, parameters $\boldsymbol{b}(k), \beta(k)$ and $\Sigma(k)$, and Hull-White extension $\theta^{(k)}$ are given. Define the Hull-White extended model $\boldsymbol{X}^{(k)} = (\boldsymbol{X}^{(k)}(t))_{t\geq k}$ by

$$\boldsymbol{X}^{(k)}(t) = \boldsymbol{b}(k) + \theta^{(k)}(t-k)\boldsymbol{e}_1 + \beta(k)\boldsymbol{X}^{(k)}(t-1) + \Sigma(k)\boldsymbol{\varepsilon}^*(t), \quad t > k, \tag{3.1}$$

with starting value $\boldsymbol{X}^{(k)}(k) = \boldsymbol{\mathcal{X}}(k)$, $\mathcal{F}(k)$-measurable parameters $\boldsymbol{b}(k)$, $\beta(k)$ and $\Sigma(k)$, and Hull-White extension $\theta^{(k)}$. We update the factor process $\boldsymbol{\mathcal{X}}$ at time $\Delta(k+1)$ according to the $\boldsymbol{X}^{(k)}$-dynamics, that is, we set

$$\boldsymbol{\mathcal{X}}(k+1) = \boldsymbol{X}^{(k)}(k+1).$$

This provides $\mathcal{F}(k+1)$-measurable yield curve at time $\Delta(k+1)$ for maturity dates $\Delta m > \Delta(k+1)$

$$Y^{(k)}(k+1,m) = -\frac{A^{(k)}(k+1,m)}{\Delta(m-(k+1))} + \frac{\boldsymbol{B}^{(k)}(k+1,m)^\top \boldsymbol{\mathcal{X}}(k+1)}{\Delta(m-(k+1))},$$

with $A^{(k)}(m-1,m) = 0$ and $\boldsymbol{B}^{(k)}(m-1,m) = \Delta\mathbf{1}$, and recursively for $m-1 > t \geq k$

$$\begin{aligned} A^{(k)}(t,m) &= A^{(k)}(t+1,m) - \boldsymbol{B}^{(k)}(t+1,m)^\top \left(\boldsymbol{b}(k) + \theta^{(k)}(t+1-k)\boldsymbol{e}_1\right) \\ &\quad + \frac{1}{2}\boldsymbol{B}^{(k)}(t+1,m)^\top \Sigma(k)\boldsymbol{B}^{(k)}(t+1,m), \\ \boldsymbol{B}^{(k)}(t,m) &= \Delta\left(\mathbb{1} - \beta^\top(k)\right)^{-1}\left(\mathbb{1} - \beta^\top(k)^{m-t}\right)\mathbf{1}. \end{aligned}$$

This is exactly the no-arbitrage price under $\mathbb{P}^*$ if the parameters $\boldsymbol{b}(k)$, $\beta(k)$ and $\Sigma(k)$ and the Hull-White extension $\theta^{(k)}$ remain constant for all $t > k$.

(iii) Parameter update and re-calibration at $k+1$. Assume that at time $\Delta(k+1)$ the parameters $(\boldsymbol{b}(k), \beta(k), \Sigma(k))$ are updated to $(\boldsymbol{b}(k+1), \beta(k+1), \Sigma(k+1))$. We may think of this parameter update as a consequence of model selection after we observe a new yield curve at time $\Delta(k+1)$. This is discussed in more detail in Section 5 below. The no-arbitrage yield curve at time $\Delta(k+1)$ from the model with parameters $(\boldsymbol{b}(k), \beta(k), \Sigma(k))$, and Hull-White extension $\theta^{(k)}$ is given by

$$\boldsymbol{y}^{(k)}(k+1) = \left(Y^{(k)}(k+1,k+2), \ldots, Y^{(k)}(k+1,k+1+M)\right)^\top \in \mathbb{R}^M.$$

The parameter update $(\boldsymbol{b}(k), \beta(k), \Sigma(k)) \mapsto (\boldsymbol{b}(k+1), \beta(k+1), \Sigma(k+1))$ requires re-calibration of the Hull-White extension, otherwise arbitrage is introduced into the model. This re-calibration provides $\mathcal{F}(k+1)$-measurable Hull-White extension $\theta^{(k+1)} \in \mathbb{R}^{\mathbb{N}}$ at time $\Delta(k+1)$. The values $\boldsymbol{\theta}^{(k+1)} = (\theta^{(k+1)}(i))_{i=1,\ldots,M-1} \in \mathbb{R}^{M-1}$ are given by, see Theorem 2.4,

$$\boldsymbol{\theta}^{(k+1)} = \mathcal{C}\left(\beta(k+1)\right)^{-1} \boldsymbol{z}\left(\boldsymbol{b}(k+1), \beta(k+1), \Sigma(k+1), \boldsymbol{\mathcal{X}}(k+1), \boldsymbol{y}^{(k)}(k+1)\right). \tag{3.2}$$

and the resulting yield curve $\boldsymbol{y}^{(k+1)}(k+1)$ under the updated parameters is identically equal to $\boldsymbol{y}^{(k)}(k+1)$. Note that this CRC makes the upper index (k) in the yield curve superfluous because the Hull-White extension is re-calibrated to the new parameters such that the resulting yield curve remains unchanged. Therefore, we write $\mathcal{Y}(k,\cdot)$ in the sequel for the CRC yield curve with factor $\boldsymbol{\mathcal{X}}(k)$, parameters $\boldsymbol{b}(k), \beta(k), \Sigma(k)$, and Hull-White extension $\theta^{(k)}$.

(iv) Iteration. Iterate items *(ii)-(iii)* for $k \geq 0$. □

Remark. For the implementation of the above algorithm we need to consider the following issue. Assume we start the algorithm at time 0 with initial yield curve $\widehat{\boldsymbol{y}}(0) \in \mathbb{R}^M$. At times Δk, for $k > 0$, calibration of $\boldsymbol{\theta}^{(k)} \in \mathbb{R}^{M-1}$ requires yields with times to maturity beyond ΔM. Either yields for these times to maturity are observable and the length of $\boldsymbol{\theta}^{(k)}$ is reduced in every step of the CRC algorithm or an appropriate extrapolation method beyond the latest available maturity date is applied in every step.

3.2 Heath-Jarrow-Morton representation

We analyse the yield curve dynamics $(\mathcal{Y}(k,\cdot))_{k\in\mathbb{N}_0}$ obtained by the CRC algorithm of Section 3.1. Due to re-calibration (3.2) the yield curve fulfils the following identity for $m > k+1$

$$\begin{aligned}\mathcal{Y}(k+1,m) &= -\frac{A^{(k)}(k+1,m)}{\Delta(m-(k+1))} + \frac{\boldsymbol{B}^{(k)}(k+1,m)^\top \boldsymbol{\mathcal{X}}(k+1)}{\Delta(m-(k+1))} \\ &= -\frac{A^{(k+1)}(k+1,m)}{\Delta(m-(k+1))} + \frac{\boldsymbol{B}^{(k+1)}(k+1,m)^\top \boldsymbol{\mathcal{X}}(k+1)}{\Delta(m-(k+1))},\end{aligned} \tag{3.3}$$

where the first line is based on the $\mathcal{F}(k)$-measurable parameters $(\boldsymbol{b}(k), \beta(k), \Sigma(k))$ and Hull-White extension $\theta^{(k)}$, and the second line is based on the $\mathcal{F}(k+1)$-measurable parameters and Hull-White extension $(\boldsymbol{b}(k+1), \beta(k+1), \Sigma(k+1), \theta^{(k+1)})$ after the CRC step. Note that in the re-calibration only $(\boldsymbol{b}(k+1), \beta(k+1), \Sigma(k+1))$ can be chosen exogenously and the Hull-White extension $\theta^{(k+1)}$ is used for consistency property (3.2). Our aim is to express $\mathcal{Y}(k+1,m)$ as a function of $\boldsymbol{\mathcal{X}}(k)$ and $\mathcal{Y}(k,m)$. Using equations (3.1) and (3.3) we have for $m > k+1$

$$\begin{aligned}\Delta\left(m-(k+1)\right)\mathcal{Y}(k+1,m) &= -A^{(k)}(k+1,m) \\ &\quad + \boldsymbol{B}^{(k)}(k+1,m)^\top \left(\boldsymbol{b}(k) + \theta^{(k)}(1)\boldsymbol{e}_1 + \beta(k)\boldsymbol{\mathcal{X}}(k) + \Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}^*(k+1)\right).\end{aligned} \tag{3.4}$$

This provides the following theorem, see Appendix A for the proof.

Theorem 3.1 *Under equivalent martingale measure $\mathbb{P}^*$ the yield curve dynamics $(\mathcal{Y}(k,\cdot))_{k\in\mathbb{N}_0}$ obtained by the CRC algorithm of Section 3.1 has the following HJM representation for $m > k+1$*

$$\begin{aligned}\Delta(m-(k+1))\mathcal{Y}(k+1,m) &= \Delta(m-k)\mathcal{Y}(k,m) - \Delta\mathcal{Y}(k,k+1) \\ &\quad + \frac{1}{2}\boldsymbol{B}^{(k)}(k+1,m)^\top \Sigma(k)\boldsymbol{B}^{(k)}(k+1,m) \\ &\quad + \boldsymbol{B}^{(k)}(k+1,m)^\top \Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}^*(k+1),\end{aligned}$$

with $\boldsymbol{B}^{(k)}(k+1,m) = \Delta\left(\mathbb{1} - \beta^\top(k)\right)^{-1}\left(\mathbb{1} - \beta^\top(k)^{m-k-1}\right)\mathbf{1}$.

Key observation. Observe that in Theorem 3.1 a remarkable simplification happens. Simulating CRC algorithm (3.1)-(3.2) to future time points $\Delta k > 0$ does not require the calculation of the Hull-White extensions $(\theta^{(k)})_{k\in\mathbb{N}_0}$ according to (3.2), but the knowledge of the parameter process $(\boldsymbol{b}(k), \beta(k), \Sigma(k))_{k\in\mathbb{N}_0}$ is sufficient. The Hull-White extensions are fully encoded in the yield curve process $(\mathcal{Y}(k,\cdot))_{k\in\mathbb{N}_0}$, and we can avoid inversion of (potentially) high dimensional matrices $\mathcal{C}(\beta(k))_{k\in\mathbb{N}_0}$.

Further remarks.

- CRC of the multifactor Vasiček spot rate model can be defined directly in the HJM framework assuming stochastic dynamics for the parameters. However, solely from the HJM representation, one cannot see that the yield curve dynamics is obtained combining well understood Hull-White extended multifactor Vasiček spot rate models using the CRC algorithm of Section 3, that is, the Hull-White extended multifactor Vasiček model gives a functional form to the HJM representation.
- The CRC algorithm of Section 3 does not rely directly on $(\boldsymbol{\varepsilon}^*(t))_{t\in\mathbb{N}}$ having independent and Gaussian components. The CRC algorithm is feasible as long as explicit formulas for ZCB prices in the Hull-White extended model are available. Therefore, one may replace the Gaussian innovations by more general distribution assumptions such as normal variance mixtures. This replacement is feasible as long as conditional exponential moments can be calculated under the new innovation assumption. Under non-Gaussian innovations, it will no longer be the case that the HJM representation does not depend on the Hull-White extension $\theta^{(k)} \in \mathbb{R}^{\mathbb{N}}$. However, the HJM representation will only depend on the the first term $\theta^{(k)}(1)$ of the Hull-White extension.
- Interpretation of the parameter processes will be given in Section 5, below.

4 Real world dynamics and market-price of risk

All previous derivations were done under an equivalent martingale measure $\mathbb{P}^*$ for the bank account numeraire $(B(t))_{t\in\mathbb{N}_0}$. In order to fit market data using the multifactor Vasiček model we also need to track the processes under the real world measure denoted by $\mathbb{P}$. We present a specific change of measure which provides tractable spot rate dynamics also under the real world measure. For simplicity we choose a sufficiently large finite time horizon $T \in \mathbb{N}$ and we assume that $\mathcal{F}_T = \mathcal{F}$. For $\boldsymbol{x} \in \mathbb{R}^n$ set

$$\boldsymbol{\lambda}(\boldsymbol{x}) = \boldsymbol{\lambda} + \Lambda \boldsymbol{x},$$

where $\boldsymbol{\lambda} \in \mathbb{R}^n$ and $\Lambda \in \mathbb{R}^{n\times n}$. For fixed $k \in \mathbb{N}_0$ let the factor process $\boldsymbol{X}^{(k)}$ be given by (3.1). The n-dimensional $\mathbb{F}$-adapted process $(\boldsymbol{\lambda}(\boldsymbol{X}^{(k)}(t)))_{t\geq k}$ describes the market-price of risk. We define the following $\mathbb{P}^*$-density process $(\xi(t))_{t\geq k}$

$$\xi(t) = \exp\left\{-\frac{1}{2}\sum_{s=k+1}^{t}\left\|\boldsymbol{\lambda}\left(\boldsymbol{X}^{(k)}(s-1)\right)\right\|_2^2 + \sum_{s=k+1}^{t}\boldsymbol{\lambda}\left(\boldsymbol{X}^{(k)}(s-1)\right)^\top \boldsymbol{\varepsilon}^*(s)\right\}, \quad t > k.$$

The real world probability measure $\mathbb{P}$ is then obtained by the Radon-Nikodym derivative

$$\frac{d\mathbb{P}}{d\mathbb{P}^*} = \xi(T). \tag{4.1}$$

An immediate consequence is that for $t > k$

$$\boldsymbol{\varepsilon}(t) = \boldsymbol{\lambda} + \Lambda \boldsymbol{X}^{(k)}(t-1) + \boldsymbol{\varepsilon}^*(t),$$

has a standard Gaussian distribution under $\mathbb{P}$, conditionally given $\mathcal{F}(t-1)$. This implies that under the real world measure $\mathbb{P}$ the factor process $(\boldsymbol{X}^{(k)}(t))_{t\geq k}$ is described for $t > k$ by

$$\boldsymbol{X}^{(k)}(t) = \boldsymbol{a}^{(k)}(t) + \alpha(k)\boldsymbol{X}^{(k)}(t-1) + \Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}(t), \tag{4.2}$$

where we define

$$\boldsymbol{a}^{(k)}(t) = \boldsymbol{b}(k) + \theta^{(k)}(t-k)\boldsymbol{e}_1 - \Sigma(k)^{\frac{1}{2}}\boldsymbol{\lambda} \quad \text{and} \quad \alpha(k) = \beta(k) - \Sigma(k)^{\frac{1}{2}}\Lambda, \tag{4.3}$$

with $\boldsymbol{a}^{(k)}(t)$ and $\alpha(k)$ being $\mathcal{F}(k)$-measurable. Formula (4.2) describes the spot rate process under the real world measure $\mathbb{P}$. As in Assumptions 2.1 we require Λ to be such that the spectrum of $\alpha(k)$ is a subset of $(-1,1)^n$. This ensures that $\boldsymbol{X}^{(k)}$ is a stationary process also under the real world measure $\mathbb{P}$.

Theorem 3.1 provides the HJM view of the re-calibrated Hull-White extended discrete time multifactor Vasiček model under equivalent martingale measure $\mathbb{P}^*$. In order to predict yield curves we need to study the dynamics under the real world measures $\mathbb{P}$. Let $(\boldsymbol{\lambda}(k))_{k\in\mathbb{N}_0}$ and $(\Lambda(k))_{k\in\mathbb{N}_0}$ be $\mathbb{F}$-adapted stochastic processes describing the market price of risk. Theorems 3.1 and equations (4.2)-(4.3) provide the following corollary.

Corollary 4.1 *Under real world measure $\mathbb{P}$ satisfying* (4.1) *the yield curve dynamics $(\mathcal{Y}(k,\cdot))_{k\in\mathbb{N}_0}$ obtained by the CRC algorithm of Section 3.1 has the following HJM representation for $m > k+1$*

$$\begin{aligned}\Delta\left(m-(k+1)\right)\mathcal{Y}(k+1,m) = {} & \Delta(m-k)\mathcal{Y}(k,m) - \Delta\mathcal{Y}(k,k+1) \\ & + \frac{1}{2}\boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)\boldsymbol{B}^{(k)}(k+1,m) \\ & - \boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)^{\frac{1}{2}}\boldsymbol{\lambda}(k) \\ & - \boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)^{\frac{1}{2}}\Lambda(k)\boldsymbol{\mathcal{X}}(k) \\ & + \boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}(k+1),\end{aligned}$$

with $\boldsymbol{B}^{(k)}(k+1,m) = \Delta\left(\mathbb{1} - \beta^\top(k)\right)^{-1}\left(\mathbb{1} - \beta^\top(k)^{m-k-1}\right)\mathbf{1}$.

We obtain additional drift terms $-\boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)^{\frac{1}{2}}\boldsymbol{\lambda}(k)$ and $-\boldsymbol{B}^{(k)}(k+1,m)^\top\Sigma(k)^{\frac{1}{2}}\Lambda(k)\boldsymbol{X}(k)$ which are characterized by the market-price of risk parameters $\boldsymbol{\lambda}(k) \in \mathbb{R}^n$ and $\Lambda(k) \in \mathbb{R}^{n\times n}$.

5 Choice of parameter process

The CRC algorithm of Section 3 requires exogenous specification of the models for the parameter process of multifactor Vasiček model (2.1)-(2.2) and the market price of risk process, i.e. we need to model under $\mathbb{P}$ the process

$$\left(\boldsymbol{b}(t), \beta(t), \Sigma(t), \boldsymbol{\lambda}(t), \Lambda(t)\right)_{t\in\mathbb{N}_0}. \tag{5.1}$$

By equation (3.1) the one-step ahead development of the CRC factor process $\boldsymbol{\mathcal{X}}$ under $\mathbb{P}$ reads as

$$\boldsymbol{\mathcal{X}}(t+1) = \boldsymbol{b}(t) + \theta^{(t)}(1)\boldsymbol{e}_1 - \Sigma(t)^{\frac{1}{2}}\boldsymbol{\lambda}(t) + \left(\beta(t) - \Sigma(t)^{\frac{1}{2}}\Lambda(t)\right)\boldsymbol{\mathcal{X}}(t) + \Sigma(t)^{\frac{1}{2}}\boldsymbol{\varepsilon}(t+1), \qquad (5.2)$$

with $\mathcal{F}(t)$-measurable parameters $\boldsymbol{b}(t)$, $\beta(t)$ and $\Sigma(t)$, and Hull-White extension $\theta^{(t)}$. Thus, on the one hand, the factor process $(\boldsymbol{\mathcal{X}}(t))_{t\in\mathbb{N}_0}$ evolves according to (5.2), and on the other hand parameters $(\boldsymbol{b}(t), \beta(t), \Sigma(t), \boldsymbol{\lambda}(t), \Lambda(t))_{t\in\mathbb{N}_0}$ evolve according to the financial market conditions. Note that the process $(\theta^{(t)})_{t\in\mathbb{N}_0}$ of Hull-White extensions is fully determined through CRC by (3.2). In order to distinguish the evolutions of $(\boldsymbol{\mathcal{X}}(t))_{t\in\mathbb{N}_0}$ and $(\boldsymbol{b}(t), \beta(t), \Sigma(t), \boldsymbol{\lambda}(t), \Lambda(t))_{t\in\mathbb{N}_0}$, respectively, we assume that process (5.1) changes on a slower pace than the factor process and, therefore, parameters can be assumed to be constant over a short time window. This assumption motivates the following approach to specify a model for process (5.1). For each time point Δt we fit multifactor Vasiček model (2.1)-(2.2) with fixed parameters $(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{\lambda}, \Lambda)$ on observations from a short time window $\{t-K+1, \ldots, t\}$ of length K. For estimation we assume that we have yield curve observations $(\widehat{\boldsymbol{y}}(k))_{k=t-K+1,\ldots,t} = ((\widehat{y}_1(k), \ldots, \widehat{y}_M(k)))_{k=t-K+1,\ldots,t}$ for times to maturity $\Delta\tau_1, \ldots, \Delta\tau_M$. In this way we obtain time series for the parameters from historical data. Finally, we fit a stochastic model to these parameter time series. In the following we discuss the interpretation of the parameters and present two different estimation procedures.

5.1 Interpretation of parameters

▷ *Level and speed of mean reversion.* By equation (2.3) we have under $\mathbb{P}^*$ for $m > t$

$$\begin{aligned}\mathbb{E}^*\left[\boldsymbol{X}(m)|\mathcal{F}(t)\right] &= (\mathbb{1}-\beta)^{-1}\left(\mathbb{1}-\beta^{m-t}\right)\boldsymbol{b} + \beta^{m-t}\boldsymbol{X}(t),\\ \mathbb{E}^*\left[r(m)|\mathcal{F}(t)\right] &= \mathbf{1}^\top(\mathbb{1}-\beta)^{-1}\left(\mathbb{1}-\beta^{m-t}\right)\boldsymbol{b} + \mathbf{1}^\top\beta^{m-t}\boldsymbol{X}(t).\end{aligned}$$

Thus, β determines the speed at which the factor process $(\boldsymbol{X}(t))_{t\in\mathbb{N}_0}$ and the spot rate process $(r(t))_{t\in\mathbb{N}_0}$ return to their long term means

$$\lim_{m\to\infty}\mathbb{E}^*\left[\boldsymbol{X}(m)|\mathcal{F}(t)\right] = (\mathbb{1}-\beta)^{-1}\boldsymbol{b} \quad \text{and} \quad \lim_{m\to\infty}\mathbb{E}^*\left[r(m)|\mathcal{F}(t)\right] = \mathbf{1}^\top(\mathbb{1}-\beta)^{-1}\boldsymbol{b}.$$

A sensible choice of $(\beta(t))_{t\in\mathbb{N}_0}$ adapts the speed of mean reversion to the prevailing financial market conditions at each time point Δt.

▷ *Instantaneous variance.* By equation (2.3) we have under $\mathbb{P}^*$ for $t > 0$

$$\mathrm{Cov}^*\left[\boldsymbol{X}(t)|\mathcal{F}(t-1)\right] = \Sigma, \quad \text{and} \quad \mathrm{Var}^*\left[r(t)|\mathcal{F}(t-1)\right] = \mathbf{1}^\top\Sigma\mathbf{1}.$$

Thus, matrix Σ plays the role of the instantaneous covariance matrix of $\boldsymbol{X}$, and it describes the instantaneous spot rate volatility.

5.2 Maximum likelihood estimation and Kalman filter

The log-likelihood function of $\boldsymbol{X}$ cannot be computed directly since we do not observe the factor process. A way out is to use time series of cross sectional yields $(\widehat{\boldsymbol{y}}(k))_{k=t-K+1,\ldots,t}$ and fit the

model parameters of the unobservable process $\boldsymbol{X}$ dynamically using the Kalman filter, see [8, Section 3.6.3].

▷ *Transition system.* The evolution of the unobservable process $\boldsymbol{X}$ under $\mathbb{P}$ is assumed to be given by

$$\boldsymbol{X}(k) = \boldsymbol{a} + \alpha \boldsymbol{X}(k-1) + \Sigma^{\frac{1}{2}} \boldsymbol{\varepsilon}(k), \quad k \in \{t-K+1, \dots, t\},$$

with initial factor $\boldsymbol{X}(t-K) \in \mathbb{R}^n$, and parameters $\boldsymbol{a} = \boldsymbol{b} - \Sigma^{\frac{1}{2}} \boldsymbol{\lambda}$ and $\alpha = \beta - \Sigma^{\frac{1}{2}} \Lambda$. The initial factor $\boldsymbol{X}(t-K)$ is updated according to the output of the Kalman filter for the previous observation window $\{t-K, \dots, t-1\}$. The initial factor is set to zero for the first observation window available.

Remark. Parameters $(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{\lambda}, \Lambda)$ are assumed to be constant over the estimation window $\{t-K+1, \dots, t\}$. Thus, we drop the index k compared to equations (4.2)-(4.3). For estimation we assume that the factor process evolves according to the time-homogeneous multifactor Vasiček model (2.1)-(2.2). The Hull-White extension is calibrated to the yield curve at time Δt given the estimated parameter values of the time-homogeneous model.

▷ *Measurement system.* The evolution of the yield curve process under $\mathbb{P}$ is assumed to be given by

$$\boldsymbol{Y}(k) = \boldsymbol{d} + D\boldsymbol{X}(k) + S^{\frac{1}{2}} \boldsymbol{\eta}(k), \quad k \in \{t-K, \dots, t\},$$

where

$$\begin{aligned}
\boldsymbol{Y}(k) &= (Y(k, k+\tau_1), \dots, Y(k, k+\tau_M))^\top \in \mathbb{R}^M, \\
\boldsymbol{d} &= \left(-(\Delta\tau_1)^{-1} A(k, k+\tau_1), \dots, -(\Delta\tau_M)^{-1} A(k, k+\tau_M)\right)^\top \in \mathbb{R}^M, \\
D_{ij} &= (\Delta\tau_i)^{-1} B_j(k, k+\tau_i), \quad 1 \le i \le M, \quad 1 \le j \le n,
\end{aligned}$$

with $A(\cdot,\cdot)$ and $\boldsymbol{B}(\cdot,\cdot) = (B_1(\cdot,\cdot), \dots, B_n(\cdot,\cdot))^\top$ given by Theorem 2.2. The M-dimensional $\mathcal{F}(k)$-measurable noise term $S^{\frac{1}{2}} \boldsymbol{\eta}(k)$ for non-singular $S^{\frac{1}{2}} \in \mathbb{R}^{M \times M}$ indicates that zero-coupon yields are itself estimated from financial market data. We assume that, $\boldsymbol{\eta}(k)$ is independent of $\mathcal{F}(k-1)$ and $\boldsymbol{\varepsilon}(k)$ under $\mathbb{P}$, and that $\boldsymbol{\eta}(k) \overset{\mathbb{P}}{\sim} \mathcal{N}(\boldsymbol{0}, \mathbb{1})$.

Given parameter and market price of risk values $(\boldsymbol{b}, \beta, \Sigma, \boldsymbol{\lambda}, \Lambda)$ one can estimate the factor using the following iterative procedure. For $k \in \{t-K, \dots, t\}$ and fixed t we consider the σ-field $\mathcal{F}^{\boldsymbol{Y}}(k) = \sigma\left(\boldsymbol{Y}(s) \mid t-K \le s \le k\right) \subset \mathcal{F}(k)$ and describe the estimation procedure in this state space model.

▷ *Anchoring.* Fix initial factor $\boldsymbol{X}(t-K) = \boldsymbol{x}(t-K|t-K-1)$ and initialise

$$\begin{aligned}
\boldsymbol{x}(t-K+1|t-K) &= \mathbb{E}\left[\boldsymbol{X}(t-K+1) \middle| \mathcal{F}^{\boldsymbol{Y}}(t-K)\right] = \boldsymbol{a} + \alpha \boldsymbol{x}(t-K|t-K-1), \\
\Sigma(t-K+1|t-K) &= \mathrm{Cov}\left(\boldsymbol{X}(t-K+1) \middle| \mathcal{F}^{\boldsymbol{Y}}(t-K)\right) = \Sigma.
\end{aligned}$$

▷ *Forecasting the measurement system.* At time $k \in \{t-K+1, \dots, t\}$ we have

$$\begin{aligned} \boldsymbol{y}(k|k-1) &= \mathbb{E}\left[\boldsymbol{Y}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k-1)\right] = \boldsymbol{d} + D\boldsymbol{x}(k|k-1), \\ F(k) &= \mathrm{Cov}\left(\boldsymbol{Y}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k-1)\right) = D\Sigma(k|k-1)D^\top + S, \\ \boldsymbol{\zeta}(k) &= \widehat{\boldsymbol{y}}(k) - \boldsymbol{y}(k|k-1). \end{aligned}$$

▷ *Bayesian inference in the transition system.* The prediction error $\boldsymbol{\zeta}(k)$ is used to update the unobservable factors.

$$\begin{aligned} \boldsymbol{x}(k|k) &= \mathbb{E}\left[\boldsymbol{X}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k)\right] = \boldsymbol{x}(k|k-1) + K(k)\boldsymbol{\zeta}(k), \\ \Sigma(k|k) &= \mathrm{Cov}\left(\boldsymbol{X}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k)\right) = (\mathbb{1} - K(k)D)\,\Sigma(k|k-1), \end{aligned}$$

where $K(k)$ denotes the Kalman gain matrix given by

$$K(k) = \mathrm{Cov}\left(\boldsymbol{X}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k-1)\right) D^\top \mathrm{Cov}\left(\boldsymbol{Y}(k)\middle|\mathcal{F}^{\boldsymbol{Y}}(k-1)\right)^{-1} = \Sigma(k|k-1)D^\top F(k)^{-1}.$$

▷ *Forecasting the transition system.* For the unobservable factor process we have the following forecast

$$\begin{aligned} \boldsymbol{x}(k+1|k) &= \mathbb{E}\left[\boldsymbol{X}(k+1)\middle|\mathcal{F}^{\boldsymbol{Y}}(k)\right] = \boldsymbol{a} + \alpha\boldsymbol{x}(k|k), \\ \Sigma(k+1|k) &= \mathrm{Cov}\left(\boldsymbol{X}(k+1)\middle|\mathcal{F}^{\boldsymbol{Y}}(k)\right) = \alpha\Sigma(k|k)\alpha^\top + \Sigma. \end{aligned}$$

The Kalman filter procedure above allows to infer factor $\boldsymbol{X}$ given parameter and market price of risk values. Of course, in this section, we are interested in estimating these values in the first place. For this purpose the procedure above can be used in conjunction with maximum likelihood estimation as explained in the following.

▷ *Likelihood function.* For the underlying parameters $\boldsymbol{\Theta} = (\boldsymbol{b}, \beta, \Sigma, \boldsymbol{a}, \alpha)$ we have the following likelihood function given the observations $\mathcal{F}^{\boldsymbol{Y}}(t)$

$$\mathcal{L}_{\mathcal{F}^{\boldsymbol{Y}}(t)}(\boldsymbol{\Theta}) = \prod_{k=t-K+1}^{t} \frac{\exp\left(-\frac{1}{2}\boldsymbol{\zeta}(k)^\top F(k)^{-1}\boldsymbol{\zeta}(k)\right)}{(2\pi)^{\frac{M}{2}} \det F(k)^{\frac{1}{2}}}. \tag{5.3}$$

The maximum likelihood estimator (MLE) $\widehat{\boldsymbol{\Theta}}^{\mathrm{MLE}} = (\widehat{\boldsymbol{b}}^{\mathrm{MLE}}, \widehat{\beta}^{\mathrm{MLE}}, \widehat{\Sigma}^{\mathrm{MLE}}, \widehat{\boldsymbol{a}}^{\mathrm{MLE}}, \widehat{\alpha}^{\mathrm{MLE}})$ is found by maximizing the likelihood function $\mathcal{L}_{\mathcal{F}^{\boldsymbol{Y}}(t)}(\boldsymbol{\Theta})$ over $\boldsymbol{\Theta}$, given the data $\mathcal{F}^{\boldsymbol{Y}}(t)$. In order to perform optimisation we need to specify the noise term matrix S. This noise term matrix can also be seen as the degree of information contained in the observations, that is, the smaller the diagonal entries of S the more informative are the observations. In the numerical example of Section 6 we analyse the sensitivity in S.

5.3 Estimation motivated by continuous time modelling

5.3.1 Rescaling the time grid

Assume factor process $(\boldsymbol{X}(t))_{t\in\mathbb{N}_0}$ is given under $\mathbb{P}$ by $\boldsymbol{X}(0) \in \mathbb{R}^n$ and for $t > 0$

$$\boldsymbol{X}(t) = \boldsymbol{a} + \alpha\boldsymbol{X}(t-1) + \Sigma^{\frac{1}{2}}\boldsymbol{\varepsilon}(t),$$

where $\boldsymbol{a} = \boldsymbol{b} - \Sigma^{\frac{1}{2}}\boldsymbol{\lambda}$ and $\alpha = \beta - \Sigma^{\frac{1}{2}}\Lambda$. Furthermore, assume that α is a diagonalisable matrix with $\alpha = TDT^{-1}$ for $T \in \mathbb{R}^{n\times n}$ and diagonal matrix $D \in (-1,1)^{n\times n}$. Then, the transformed process $\boldsymbol{Z} = (T^{-1}\boldsymbol{X}(t))_{t\in\mathbb{N}_0}$ evolves according to

$$\boldsymbol{Z}(t) = \boldsymbol{c} + D\boldsymbol{Z}(t-1) + \Psi^{\frac{1}{2}}\boldsymbol{\varepsilon}(t), \quad t > 0,$$

where $\boldsymbol{c} = T^{-1}\boldsymbol{a}$ and $\Psi = T^{-1}\Sigma(T^{-1})^\top$. For $d > 0$ the d-step ahead conditional distribution of $\boldsymbol{Z}$ under $\mathbb{P}$ is given by

$$\boldsymbol{Z}(t+d)|\mathcal{F}(t) \overset{\mathbb{P}}{\sim} \mathcal{N}\left(\boldsymbol{\mu} + \gamma\boldsymbol{Z}(t), \Gamma\right), \quad t \geq 0,$$

where $\boldsymbol{\mu} = (\mathbb{1} - D)^{-1}\left(\mathbb{1} - D^d\right)\boldsymbol{c}$, $\gamma = D^d$ and $\Gamma = \sum_{s=0}^{d-1} D^s\Psi D^s$. Suppose we have estimated $\boldsymbol{\mu} \in \mathbb{R}^n$, the diagonal matrix $\gamma \in (-1,1)^n$ and $\Gamma \in \mathbb{R}^{n\times}$ on the time grid with size $d\Delta$, for instance, using MLE estimation as explained in Section 5.2. We are interested in recovering the parameters $\boldsymbol{c}$, D and Ψ of the dynamics on the refined time grid with size Δ from $\boldsymbol{\mu}$, γ and Γ. The diagonal matrix D and $\boldsymbol{c}$ are reconstructed from the diagonal matrix γ as follows.

$$D = \gamma^{\frac{1}{d}} = \mathbb{1} + \frac{1}{d}\log(\gamma) + o\left(\frac{1}{d}\right), \quad \text{as } d \to \infty,$$
$$\boldsymbol{c} = (\mathbb{1} - \gamma)^{-1}\left(\mathbb{1} - \gamma^{\frac{1}{d}}\right)\boldsymbol{\mu} = \frac{1}{d}(\mathbb{1} - \gamma)^{-1}\log\left(\gamma^{-1}\right)\boldsymbol{\mu} + o\left(\frac{1}{d}\right), \quad \text{as } d \to \infty,$$

where logarithmic and power functions applied to diagonal matrices are defined on their diagonal elements. Note that for $i, j = 1, \ldots, n$ we have

$$\Gamma_{ij} = \sum_{s=0}^{d-1} \gamma_{ii}^{\frac{s}{d}}\Psi_{ij}\gamma_{jj}^{\frac{s}{d}} = \Psi_{ij}\sum_{s=0}^{d-1}\left(\gamma_{ii}^{\frac{1}{d}}\gamma_{jj}^{\frac{1}{d}}\right)^s = \Psi_{ij}\frac{1 - \gamma_{ii}\gamma_{jj}}{1 - (\gamma_{ii}\gamma_{jj})^{\frac{1}{d}}}.$$

Therefore, we recover Ψ from γ and Γ as follows.

$$\Psi = \frac{1}{d}\upsilon\Gamma + o\left(\frac{1}{d}\right), \quad \text{as } d \to \infty,$$

where $\upsilon = (\log(\gamma_{ii}\gamma_{jj})(1 - \gamma_{ii}\gamma_{jj})^{-1})_{i,j=1,\ldots,n} \in \mathbb{R}^{n\times n}$. From the formulas for $\boldsymbol{c}$, D and Ψ we observe that the drift term and the random term in the $\boldsymbol{Z}$-dynamics live on different scales as $d \to \infty$.

Under $\mathbb{P}$ for $t > 0$ we have

$$\begin{aligned}
&\mathbb{E}\left[\mathcal{D}_t\boldsymbol{Z}\left(\mathcal{D}_t\boldsymbol{Z}\right)^\top\middle|\mathcal{F}_{t-1}\right] = \mathrm{Cov}\left[\mathcal{D}_t\boldsymbol{Z}, \mathcal{D}_t\boldsymbol{Z}|\mathcal{F}_{t-1}\right] + \mathbb{E}\left[\mathcal{D}_t\boldsymbol{Z}|\mathcal{F}_{t-1}\right]\mathbb{E}\left[\mathcal{D}_t\boldsymbol{Z}|\mathcal{F}_{t-1}\right]^\top \\
&= \mathrm{Cov}\left[\boldsymbol{Z}(t), \boldsymbol{Z}(t)|\mathcal{F}_{t-1}\right] + \left(\mathbb{E}\left[\boldsymbol{Z}(t)|\mathcal{F}_{t-1}\right] - Z(t-1)\right)\left(\mathbb{E}\left[\boldsymbol{Z}(t)|\mathcal{F}_{t-1}\right] - Z(t-1)\right)^\top \\
&= \Psi + \left(\boldsymbol{c} + (D - \mathbb{1})\boldsymbol{Z}(t-1)\right)\left(\boldsymbol{c} + (D - \mathbb{1})\boldsymbol{Z}(t-1)\right)^\top,
\end{aligned}$$

where $\mathcal{D}_t\boldsymbol{Z} = \boldsymbol{Z}(t) - \boldsymbol{Z}(t-1)$. Since

$$\boldsymbol{c} + (D - \mathbb{1})\boldsymbol{Z}(t-1) = \frac{1}{d}\left((\mathbb{1} - \gamma)^{-1}\log\left(\gamma^{-1}\right) + \log(\gamma)\boldsymbol{Z}(t-1)\right) + o\left(\frac{1}{d}\right), \quad \text{as } d \to \infty,$$

we obtain as $d \to \infty$

$$\mathbb{E}\left[\mathcal{D}_t\boldsymbol{X}\left(\mathcal{D}_t\boldsymbol{X}\right)^\top\middle|\mathcal{F}_{t-1}\right] = T^{-1}\mathbb{E}\left[\mathcal{D}_t\boldsymbol{Z}\left(\mathcal{D}_t\boldsymbol{Z}\right)^\top\middle|\mathcal{F}_{t-1}\right]\left(T^{-1}\right)^\top = \Sigma + o\left(\frac{1}{d}\right), \tag{5.4}$$

where $\Delta\boldsymbol{X}(t) = \boldsymbol{X}(t) - \boldsymbol{X}(t-1)$.

5.3.2 Longitudinal realised covariations of yields

We consider the yield curve increments within the discrete-time multifactor Vasiček model (2.1)-(2.2). The innovations of the yield process $(Y(t,t+\tau))_{t\in\mathbb{N}_0}$ for fixed time to maturity $\Delta\tau > 0$ are given by

$$\begin{aligned}\mathcal{D}_{t,\tau}Y &= Y(t,t+\tau) - Y(t-1,t-1+\tau)\\ &= \frac{1}{\Delta\tau}\boldsymbol{B}(t,t+\tau)^\top(\boldsymbol{X}(t)-\boldsymbol{X}(t-1)) = \frac{1}{\Delta\tau}\boldsymbol{B}(t,t+\tau)^\top\Delta\boldsymbol{X}(t),\end{aligned}$$

where $\Delta\boldsymbol{X}(t)|\mathcal{F}(t-1) \overset{\mathbb{P}}{\sim} \mathcal{N}(\boldsymbol{a}+(\alpha-\mathbb{1})\boldsymbol{X}(t-1),\Sigma)$. For times to maturity $\Delta\tau_1,\Delta\tau_2 > 0$ we get under $\mathbb{P}$

$$\mathbb{E}\left[\mathcal{D}_{t,\tau_1}Y\mathcal{D}_{t,\tau_2}Y|\mathcal{F}_{t-1}\right] = \frac{1}{\Delta^2\tau_1\tau_2}\boldsymbol{B}(t,t+\tau_1)^\top\mathbb{E}\left[\Delta\boldsymbol{X}(t)\Delta\boldsymbol{X}(t)^\top\middle|\mathcal{F}_{t-1}\right]\boldsymbol{B}(t,t+\tau_2).$$

By equation (5.4) for small grid size Δ we estimate the last expression by

$$\mathbb{E}\left[\mathcal{D}_{t,\tau_1}Y\mathcal{D}_{t,\tau_2}Y|\mathcal{F}_{t-1}\right] \approx \frac{1}{\tau_1\tau_2}\mathbf{1}^\top(\mathbb{1}-\beta^{\tau_1})(\mathbb{1}-\beta)^{-1}\Sigma\left(\mathbb{1}-\beta^\top\right)^{-1}\left(\mathbb{1}-\left(\beta^\top\right)^{\tau_2}\right)\mathbf{1}. \tag{5.5}$$

The latter is very interesting because of the following reasons.

- Formula (5.5) does not depend on the unobservable factors $\boldsymbol{X}$.
- Formula (5.5) allows for direct cross-sectional estimation of β and Σ. That is, β and Σ can directly be estimated from market observations (without knowing the market-price of risk).
- Formula (5.5) is helpful to determine the number of factors needed to fit the model to market yield curve increments.

Let $\widehat{y}_1(k)$ and $\widehat{y}_2(k)$ be market observations for times to maturity $\Delta\tau_1$ and $\Delta\tau_2$, and at times $k \in \{t-K+1,\dots,t\}$. Then, the expectation on the left hand side of (5.5) can be estimated by the realised covariation

$$\widehat{\mathrm{RCov}}(t,\tau_1,\tau_2) = \frac{1}{K}\sum_{k=t-K+1}^{t}\left(\widehat{y}_1(k)-\widehat{y}_1(k-1)\right)\left(\widehat{y}_2(k)-\widehat{y}_2(k-1)\right). \tag{5.6}$$

5.3.3 Cross-sectional estimation of β and Σ

The realised covariation estimators $\widehat{\mathrm{RCov}}(t,\tau_i,\tau_j)$ for $i,j = 1,\dots,M$ allow to estimate the parameters β and Σ at time Δt in the following way. For given symmetric weights $w_{ij} = w_{ji} \geq 0$ we solve the least squares problem

$$\begin{aligned}\left(\widehat{\beta}^{\mathrm{RCov}},\widehat{\Sigma}^{\mathrm{RCov}}\right) = \arg\min_{\beta,\Sigma}\Bigg\{\sum_{i,j=1}^{M} w_{ij}\Bigg(&\widehat{\mathrm{RCov}}(t,\tau_i,\tau_j)\\ &-\frac{1}{\tau_i\tau_j}\mathbf{1}^\top(\mathbb{1}-\beta^{\tau_i})(\mathbb{1}-\beta)^{-1}\Sigma\left(\mathbb{1}-\beta^\top\right)^{-1}\left(I-\left(\beta^\top\right)^{\tau_j}\right)\mathbf{1}\Bigg)^2\Bigg\},\end{aligned} \tag{5.7}$$

where we optimise over β and Σ satisfying Assumptions 2.1.

5.3.4 Inference on market-price of risk

Next we aim at determining the market price parameters $\boldsymbol{\lambda}$ and Λ under the density process assumption of Section 4. For this purpose we combine the MLE estimation of Section 5.2 with the estimation based on the realised covariations of yields of Section 5.3.3. First, we estimate β and Σ by $\widehat{\beta}^{\mathrm{RCov}}$ and $\widehat{\Sigma}^{\mathrm{RCov}}$ as explained in Section 5.3.3. Secondly, we estimate $\boldsymbol{a}$, $\boldsymbol{b}$ and α maximising (5.3) for fixed β and Σ, i.e.

$$\left(\widehat{\boldsymbol{b}}^{\mathrm{MLE}}, \widehat{\boldsymbol{a}}^{\mathrm{MLE}}, \widehat{\alpha}^{\mathrm{MLE}}\right) = \arg\max_{\boldsymbol{b},\boldsymbol{a},\alpha} \mathcal{L}_{\mathcal{F}\boldsymbol{Y}(t)}\left(\boldsymbol{b}, \widehat{\beta}^{\mathrm{RCov}}, \widehat{\Sigma}^{\mathrm{RCov}}, \boldsymbol{a}, \alpha\right), \tag{5.8}$$

where we optimise over $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{a} \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}^{n\times n}$ with spectrum in $(-1,1)^n$. The constraint on the matrix α ensures that the factor process is also stationary under the real world measure $\mathbb{P}$.

From equation (4.3) we have $\boldsymbol{\lambda} = \Sigma^{-\frac{1}{2}}(\boldsymbol{b} - \boldsymbol{a})$ and $\Lambda = \Sigma^{-\frac{1}{2}}(\beta - \alpha)$. This motivates inference of $\boldsymbol{\lambda}$ by

$$\widehat{\boldsymbol{\lambda}} = \left(\widehat{\Sigma}^{\mathrm{RCov}}\right)^{-\frac{1}{2}}\left(\widehat{\boldsymbol{b}}^{\mathrm{MLE}} - \widehat{\boldsymbol{a}}^{\mathrm{MLE}}\right), \tag{5.9}$$

and inference of Λ by

$$\widehat{\Lambda}(k) = \left(\widehat{\Sigma}^{\mathrm{RCov}}\right)^{-\frac{1}{2}}\left(\widehat{\beta}^{\mathrm{RCov}} - \widehat{\alpha}^{\mathrm{MLE}}\right). \tag{5.10}$$

5.4 Conclusions

Formulas (5.7), (5.8), (5.9), (5.10) and (3.2) fully specify the inference of *historical* parameters, market-prices of risk and Hull-White extensions. The only free variables we still have in the inference procedure are the choice of the dimension of the multifactor Vasiček model and of window length K (which will determine the viscosity of the parameters).

Going forward we will need to model the parameter process and the market-price of risk process in order to predict future yield curves. An alternative way that avoids explicit modelling of these processes is to push the yield curve period by period through Corollary 4.1 and then make path-wise inference of the parameters as described above. This reflects an empirical non-parametric modelling approach where parameters are fully encoded in the new yield curve observations.

6 Numerical example for Swiss interest rates

We choose $\Delta = 1/252$ which corresponds to a daily time grid (assuming that a financial year has 252 business days). For the Swiss currency (CHF) we consider as yield observations the Swiss Average Rate (SAR), the London Interbank Offered Rate (LIBOR) and the Swiss Confederation Bond (SWCNB). See Figures B.1-B.21.

▷ *Short times to maturity.* The SAR is an ongoing volume weighted average rate calculated by the Swiss National Bank based on repo transactions between financial institutions. It is used for short times to maturity of at most 3 months. For SAR we have the Over-Night SARON that corresponds to a time to maturity of Δ (one business day) and the SAR Tomorrow-Next

(SARTN) for time to maturity 2Δ (two business days). The latter is not completely correct because SARON is a collateral over-night rate and tomorrow-next is a call money rate for receiving money tomorrow which has to be paid back the next business day. Moreover, we have the SAR for times to maturity of 1 week (SAR1W), 2 weeks (SAR2W), 1 month (SAR1M) and 3 months (SAR3M), see also [6].

▷ *Short to medium times to maturity.* The LIBOR reflects times to maturity which correspond to 1 month (LIBOR1M), 3 months (LIBOR3M), 6 months (LIBOR6M) and 12 months (LIBOR12M) in the interbank market.

▷ *Medium to long times to maturity.* The SWCNB is based on Swiss government bonds and it is used for times to maturity which correspond to 2 years (SWCNB2Y), 3 years (SWCNB3Y), 4 years (SWCNB4Y), 5 years (SWCNB5Y), 7 years (SWCNB7Y), 10 years (SWCNB10Y), 20 years (SWCNB20Y) and 30 years (SWCNB30Y).

This data is available from December 8, 1999, and we set September 15, 2014 to be the last observation date. Of course, SAR, LIBOR and SWCNB do not exactly model risk-free zero-coupon bonds and these different classes are not completely consistent because prices are determined slightly differently for the different classes. In particular, this can be seen during the 2008-2009 financial crisis. However, this data is in many cases the best approximation to CHF risk-free zero-coupon yields that is available. For the longest times to maturity of SWCNB one may also raise issues about liquidity of these instruments.
In Figures B.22-B.28 we compute the realised covariations of yields as defined in (5.6). In Figures B.14, B.27 and B.28 we observe that SAR fits SWCNB better than LIBOR. For this reason we drop LIBOR and build daily yield curve from SAR and SWCNB.
In this numerical example we restrict ourselves to a multifactor Vasiček model with β and α having the following form

$$\beta = \operatorname{diag}(\beta_{11}, \ldots, \beta_{nn}), \quad \text{and} \quad \alpha = \operatorname{diag}(\alpha_{11}, \ldots, \alpha_{nn}),$$

where $-1 \leq \beta_{11}, \ldots, \beta_{nn}, \alpha_{11}, \ldots, \alpha_{nn} \leq 1$. In the following we explain exactly how to perform the very delicate task of parameter estimation in multifactor Vasiček (2.1)-(2.2) using the procedure explained in Section 5. Since the factor process $\boldsymbol{X}$ describes the spot rate dynamics we choose to use short times to maturity (SAR) to estimate the parameters $\boldsymbol{b}$, β, Σ, $\boldsymbol{a}$ and α which describe the dynamics of the factor process in the multifactor Vasiček model with constant parameter. Then, we determine the Hull-White extension θ in order to match the prevailing yield curve given by SAR and SWCNB.
First, we determine an appropriate number of factors n to use. The more factors we use the better we can fit the model to the data, however, the dimensionality of the estimation problem increases quadratically in the number of factors. Therefore, we look for a trade off between accuracy of the fit and the number of parameters in the model. In Figures B.29-B.31 we fit $\beta_{11}, \ldots, \beta_{nn}$ and Σ solving optimisation (5.7) for $n = 1, 2, 3$. In the 1-factor case we observe

that the fit to market realised volatilities is rather poor. In particular, we cannot capture the dependence on τ accurately. In the 2-factor model market realised volatilities can be fitted much more accurately than in the 1-factor case. In the 3-factor model we observe a slight improvement compared to the 2-factor case. Considering these results we restrict ourselves from now on to 2-factor Vasiček models with parameters

$$\boldsymbol{b} = (b_1, b_2)^\top, \quad \beta = \operatorname{diag}(\beta_{11}, \beta_{22}), \quad \Sigma^{\frac{1}{2}} = \begin{pmatrix} \sqrt{\Sigma_{11}} & 0 \\ \rho_{21}\sqrt{\Sigma_{22}} & \sqrt{\Sigma_{22}}\sqrt{1-\rho_{21}^2} \end{pmatrix},$$
$$\boldsymbol{a} = (a_1, a_2)^\top \quad \text{and} \quad \alpha = \operatorname{diag}(\alpha_{22}, \alpha_{22}),$$

where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^2$, $-1 \leq \beta_{11}, \beta_{22}, \alpha_{11}, \alpha_{22} \leq 1$, $\Sigma_{11}, \Sigma_{22} > 0$ and $-1 < \rho < 1$.
From formulas (5.7), (5.8), (5.9), (5.10) we obtain time series for the model parameters and market prices of risk, see Figures B.32-B.43. Finally, we calibrate the Hull-White extension by equation (3.2) to the full yield curve (SAR and SWCNB), see Figures B.44-B.46.
TODO: discuss the results so far before going on with this case study.

A Proofs

Proof of Theorem 2.2. We prove Theorem 2.2 by induction as in [8, Theorem 3.16] where ZCB prices are derived under the assumption that β and Σ are diagonal matrices. We have $P(m-1,m) = \exp\left(-\Delta \mathbf{1}^\top \boldsymbol{X}(m-1)\right)$, which proves the claim for $t = m-1$. Assume that Theorem 2.2 holds for $t+1 \in \{2, \ldots, m-1\}$. We verify that it also holds for $t \in \{1, \ldots, m-2\}$. Under equivalent martingale measure $\mathbb{P}^*$ we have using the tower property for conditional expectations and the induction assumption

$$\begin{aligned}
P(t,m) &= \exp\left\{-\Delta \mathbf{1}^\top \boldsymbol{X}(t)\right\} \mathbb{E}^* \left[\mathbb{E}^* \left[\exp\left\{-\Delta \sum_{s=t+1}^{m-1} \mathbf{1}^\top \boldsymbol{X}(s)\right\} \middle| \mathcal{F}(t+1)\right] \middle| \mathcal{F}(t)\right] \\
&= \exp\left\{-\Delta \mathbf{1}^\top \boldsymbol{X}(t)\right\} \mathbb{E}^* \left[P(t+1,m) | \mathcal{F}(t)\right] \\
&= \exp\left\{-\Delta \mathbf{1}^\top \boldsymbol{X}(t)\right\} \mathbb{E}^* \left[\exp\left\{A(t+1,m) - \boldsymbol{B}(t+1,m)^\top \boldsymbol{X}(t+1)\right\} \middle| \mathcal{F}(t)\right] \\
&= \exp\left\{-\Delta \mathbf{1}^\top \boldsymbol{X}(t) + A(t+1,m) - \boldsymbol{B}(t+1,m)^\top \left(\boldsymbol{b} + \beta \boldsymbol{X}(t)\right) + \frac{1}{2} \boldsymbol{B}(t+1,m)^\top \Sigma \boldsymbol{B}(t+1,m)\right\} \\
&= \exp\left\{A(t+1,m) - \boldsymbol{B}(t+1,m)^\top \boldsymbol{b} + \frac{1}{2} \boldsymbol{B}(t+1,m)^\top \Sigma \boldsymbol{B}(t+1,m) - \left(\boldsymbol{B}(t+1,m)^\top \beta + \Delta \mathbf{1}^\top\right) \boldsymbol{X}(t)\right\}.
\end{aligned}$$

This proves the following recursive formula for $m-1 > t \geq 0$

$$\begin{aligned}
A(t,m) &= A(t+1,m) - \boldsymbol{B}(t+1,m)^\top \boldsymbol{b} + \frac{1}{2} \boldsymbol{B}(t+1,m)^\top \Sigma \boldsymbol{B}(t+1,m), \\
\boldsymbol{B}(t,m) &= \beta^\top \boldsymbol{B}(t+1,m) + \Delta \mathbf{1}.
\end{aligned}$$

Finally, note that the recursive formula for $\boldsymbol{B}(\cdot,\cdot)$ implies

$$\boldsymbol{B}(t,m) = \Delta \sum_{s=0}^{m-t-1} \left(\beta^\top\right)^s \mathbf{1} = \Delta \left(\mathbb{1} - \beta^\top\right)^{-1} \left(\mathbb{1} - \left(\beta^\top\right)^{m-t}\right) \mathbf{1}.$$

This concludes the proof. □

Proof of Theorem 2.4. First, observe that the condition $\boldsymbol{y}^{(k)}(k) = \boldsymbol{y}$ imposes conditions only on the values $\theta(1), \ldots, \theta(M-1)$. Secondly, note that the vector $\boldsymbol{\theta}$ such that the condition is satisfied can be calculated recursively in the following way.

(i) First component θ_1. We have $A^{(k)}(k+1,k+2)=0$, $\boldsymbol{B}(k+1,k+2)=\Delta\mathbf{1}$ and

$$A^{(k)}(k,k+2)=-\Delta\mathbf{1}^\top\boldsymbol{b}-\Delta\theta(1)+\frac{\Delta^2}{2}\mathbf{1}^\top\Sigma\mathbf{1},$$

see Theorem 2.3. Solving the last equation for θ_1 we have

$$\theta_1=\frac{\Delta}{2}\mathbf{1}^\top\Sigma\mathbf{1}-\mathbf{1}^\top\boldsymbol{b}-\frac{1}{\Delta}A^{(k)}(k,k+2).$$

From (5.5) we obtain

$$A^{(k)}(k,k+2)=\Delta\mathbf{1}^\top\left(\mathbb{1}-\beta^2\right)(\mathbb{1}-\beta)^{-1}\boldsymbol{x}-2\Delta y_2.$$

This is equivalent to

$$\theta_1=\frac{\Delta}{2}\mathbf{1}^\top\Sigma\mathbf{1}-\mathbf{1}^\top\boldsymbol{b}-\mathbf{1}^\top\left(\mathbb{1}-\beta^2\right)(\mathbb{1}-\beta)^{-1}\boldsymbol{x}+2y_2. \tag{A.1}$$

(ii) Recursion $i\to i+1$. Assume we have determined $\theta_1,\dots,\theta_i$ for $i=1,\dots,M-2$. We want to determine θ_{i+1}. We have $A^{(k)}(k+i+1,k+i+2)=0$ and iteration implies

$$A^{(k)}(k,k+i+2)=-\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\left(\boldsymbol{b}+\theta(s-k)\boldsymbol{e}_1\right)+\frac{1}{2}\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\Sigma\boldsymbol{B}(s,k+i+2).$$

Solving the last equation for θ_{i+1} and using $\boldsymbol{B}(k+i+1,k+i+2)=\Delta\mathbf{1}$ we have

$$\begin{aligned}\theta_{i+1}=&-\frac{1}{\Delta}A^{(k)}(k,k+i+2)-\frac{1}{\Delta}\sum_{s=k+1}^{k+i}\boldsymbol{B}(s,k+i+2)^\top\left(\boldsymbol{b}+\theta(s-k)\boldsymbol{e}_1\right)-\mathbf{1}^\top\boldsymbol{b}\\&+\frac{1}{2\Delta}\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\Sigma\boldsymbol{B}(s,k+i+2).\end{aligned}$$

From (5.5) we obtain

$$A^{(k)}(k,k+i+2)=\Delta\mathbf{1}^\top\left(\mathbb{1}-\beta^{i+2}\right)(\mathbb{1}-\beta)^{-1}\boldsymbol{x}-\Delta(i+2)y_{i+2}.$$

This is equivalent to

$$\begin{aligned}\theta_{i+1}&=(i+2)y_{i+2}-\mathbf{1}^\top\left(\mathbb{1}-\beta^{i+2}\right)(\mathbb{1}-\beta)^{-1}\boldsymbol{x}-\frac{1}{\Delta}\sum_{s=k+1}^{k+i}\boldsymbol{B}(s,k+i+2)^\top\left(\boldsymbol{b}+\theta_{s-k}\boldsymbol{e}_1\right)\\&\quad-\mathbf{1}^\top\boldsymbol{b}+\frac{1}{2\Delta}\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\Sigma\boldsymbol{B}(s,k+i+2)\\&=(i+2)y_{i+2}-\mathbf{1}^\top\left(\mathbb{1}-\beta^{i+2}\right)(\mathbb{1}-\beta)^{-1}\boldsymbol{x}-\frac{1}{\Delta}\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\boldsymbol{b}\\&\quad-\frac{1}{\Delta}\sum_{s=k+1}^{k+i}B_1(s,k+i+2)\theta_{s-k}+\frac{1}{2\Delta}\sum_{s=k+1}^{k+i+1}\boldsymbol{B}(s,k+i+2)^\top\Sigma\boldsymbol{B}(s,k+i+2).\end{aligned} \tag{A.2}$$

This recursion allows to determine the components of $\boldsymbol{\theta}$. Note that equation (A.2) can be written as

$$\left(\mathcal{C}(\beta)\boldsymbol{\theta}\right)_{i+1}=z_{i+1}\left(\boldsymbol{b},\beta,\Sigma,\boldsymbol{x},\boldsymbol{y}\right),\quad i=1,\dots,M-2.$$

Observe that the lower triangular matrix $\mathcal{C}(\beta)$ is invertible since $\det\mathcal{C}(\beta)=\Delta^{M-1}>0$. Hence, equations (A.1) and (A.2) prove (2.8). □

Proof of Theorem 3.1. We add and subtract $-A^{(k)}(k,m)+\boldsymbol{B}^{(k)}(k,m)^\top\boldsymbol{\mathcal{X}}(k)$ to equation (3.4) and obtain

$$\begin{aligned}\Delta\left(m-(k+1)\right)\mathcal{Y}(k+1,m)=&A^{(k)}(k,m)-A^{(k)}(k+1,m)-A^{(k)}(k,m)\\&+\boldsymbol{B}^{(k)}(k,m)^\top\boldsymbol{\mathcal{X}}(k)-\boldsymbol{B}^{(k)}(k,m)^\top\boldsymbol{\mathcal{X}}(k)\\&+\boldsymbol{B}^{(k)}(k+1,m)^\top\left(\boldsymbol{b}(k)+\theta^{(k)}(k+1)\boldsymbol{e}_1+\beta(k)\boldsymbol{X}(k)+\Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}^*(k+1)\right).\end{aligned} \tag{A.3}$$

We have the two identities

$$-A^{(k)}(k,m) + \boldsymbol{B}^{(k)}(k,m)^\top \boldsymbol{\mathcal{X}}(k) = \Delta(m-k)\mathcal{Y}(k,m),$$

$$A^{(k)}(k,m) - A^{(k)}(k+1,m) = -\boldsymbol{B}^{(k)}(k,m)^\top \left(\boldsymbol{b}(k) + \theta^{(k)}(k+1)\boldsymbol{e}_1\right) + \frac{1}{2}\boldsymbol{B}^{(k)}(k+1,m)^\top \Sigma(k)\boldsymbol{B}^{(k)}(k+1,m).$$

Therefore, the right-hand side of equality (A.3) is rewritten as

$$\begin{aligned}\Delta(m-(k+1))\mathcal{Y}(k+1,m) = \Delta(m-k)\mathcal{Y}(k,m) &+ \left(\boldsymbol{B}(k+1,m)^\top \beta(k) - \boldsymbol{B}(k,m)^\top\right)\boldsymbol{\mathcal{X}}(k) \\ &+ \frac{1}{2}\boldsymbol{B}^{(k)}(k+1,m)^\top \Sigma(k)\boldsymbol{B}^{(k)}(k+1,m) + \boldsymbol{B}^{(k)}(k+1,m)^\top \Sigma(k)^{\frac{1}{2}}\boldsymbol{\varepsilon}^*(k+1).\end{aligned}$$

Observe that

$$\boldsymbol{B}^{(k)}(k+1,m)^\top \beta(k) = \Delta\left(\sum_{s=0}^{m-k-2}\left(\beta^\top(k)\right)^s \mathbf{1}\right)^\top \beta(k) = \Delta\mathbf{1}^\top \sum_{s=1}^{m-k-1}\beta(k)^s = \boldsymbol{B}(k,m)^\top - \Delta\mathbf{1}^\top,$$

and that $Y(k,k+1) = \mathbf{1}^\top \boldsymbol{X}(k)$. This proves the claim. □

B Figures

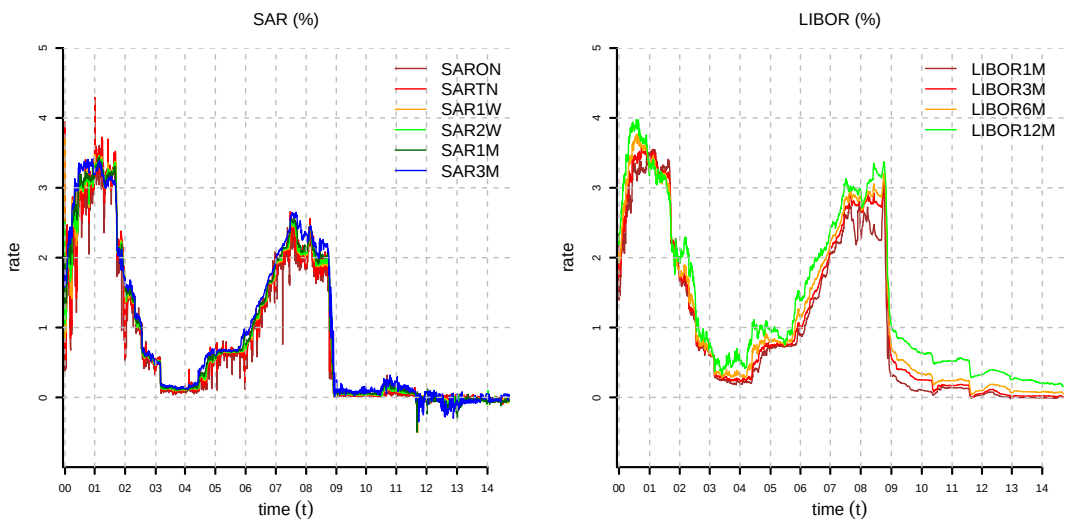

Figure B.1: (lhs) Swiss Average Rate (SAR); (rhs) London Interbank Offered Rate (LIBOR) from December 8, 1999, until September 15, 2014.

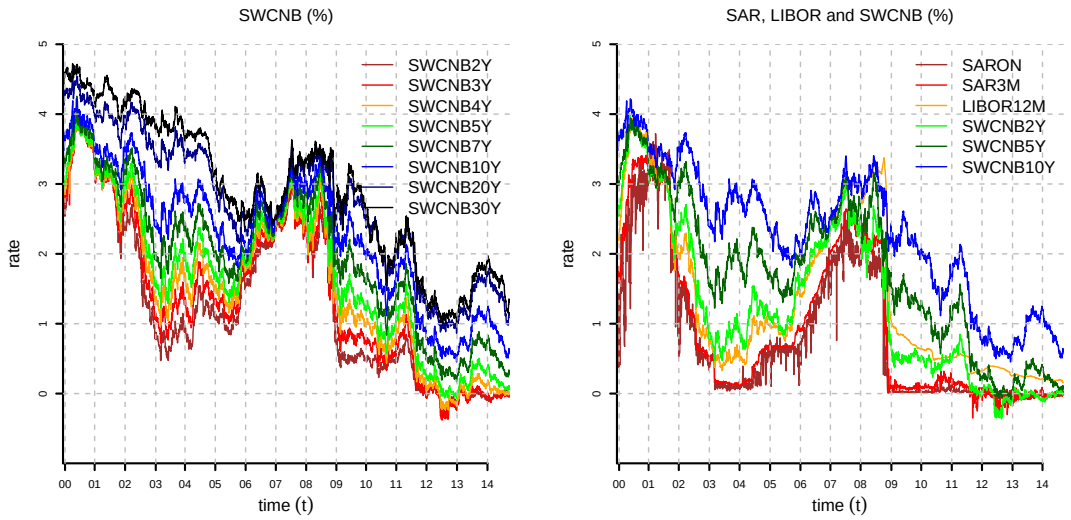

Figure B.2: (lhs) Swiss Confederation Bond (SWCNB); (rhs) a selection of SAR, LIBOR and SWCNB from December 8, 1999, until September 15, 2014.

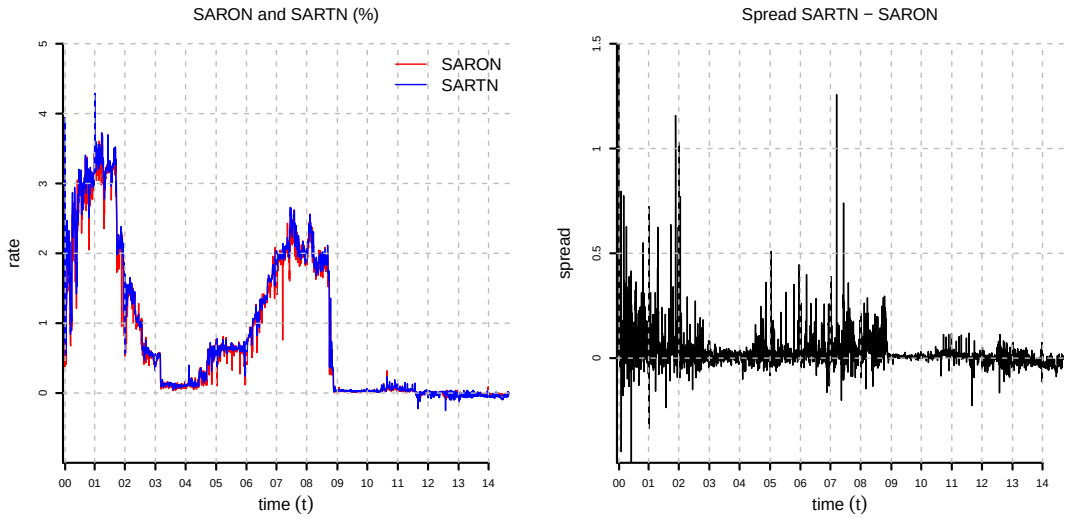

Figure B.3: (lhs) SARON and SARTN from December 8, 1999, until September 15, 2014; (rhs) spread between SARTN and SARON.

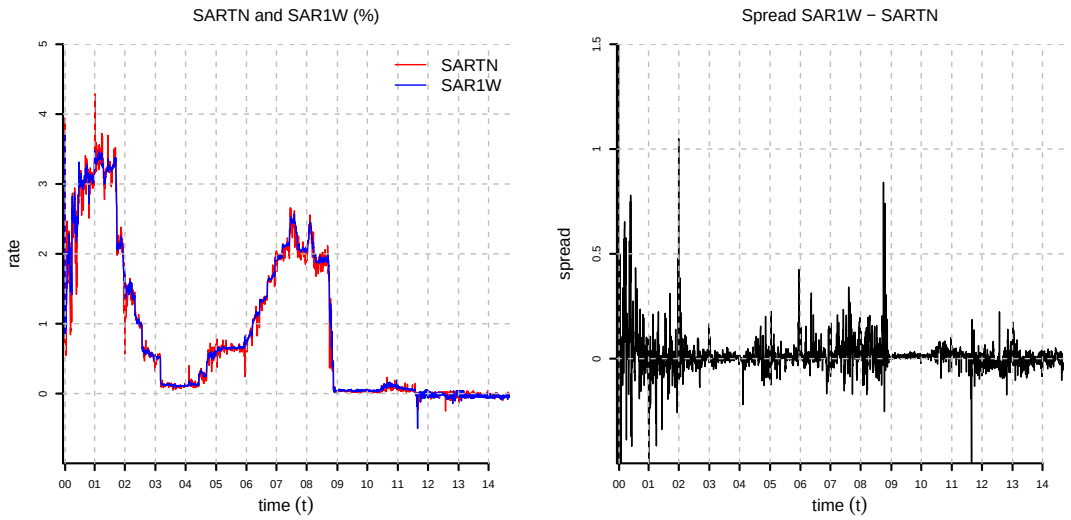

Figure B.4: (lhs) SARTN and SAR1W from December 8, 1999, until September 15, 2014; (rhs) spread between SAR1W and SARTN.

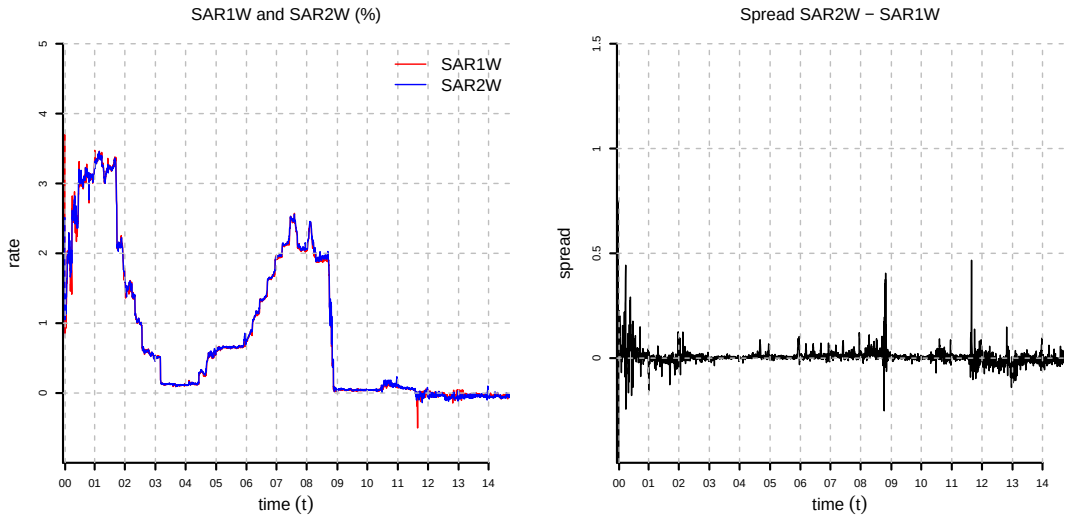

Figure B.5: (lhs) SAR1W and SAR2W from December 8, 1999, until September 15, 2014; (rhs) spread between SAR2W and SAR1W.

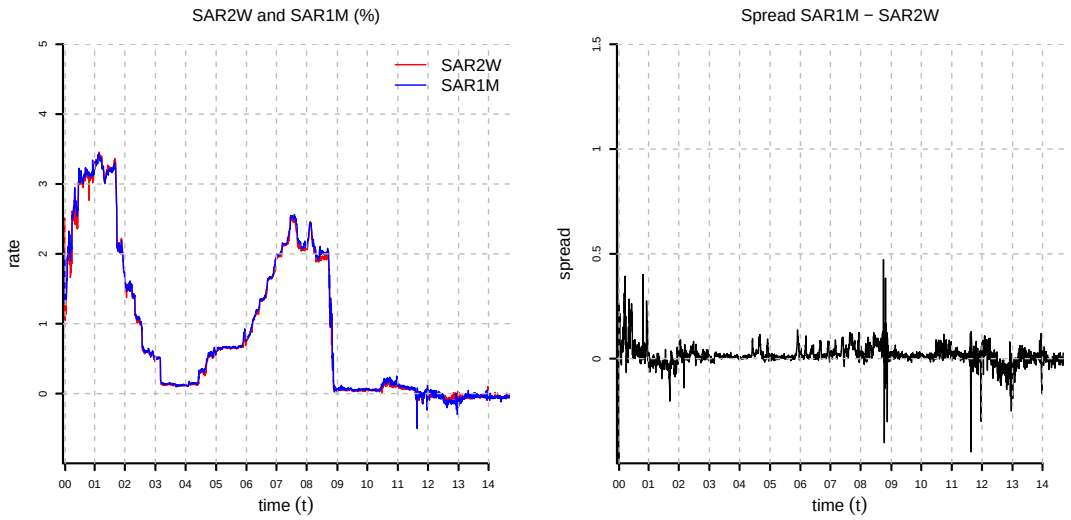

Figure B.6: (lhs) SAR2W and SAR1M from December 8, 1999, until September 15, 2014; (rhs) spread between SAR1M and SAR2W.

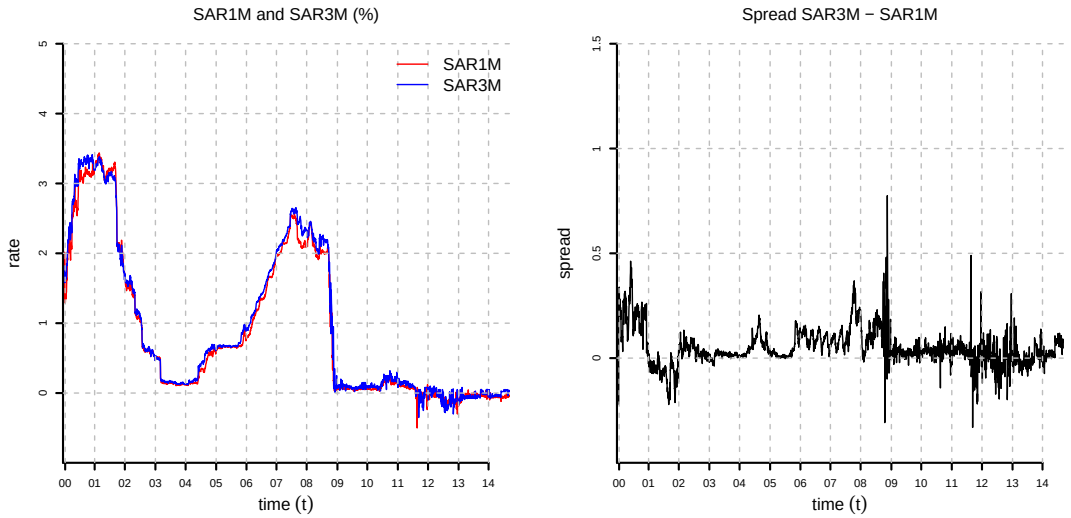

Figure B.7: (lhs) SAR1M and SAR3M from December 8, 1999, until September 15, 2014; (rhs) spread between SAR3M and SAR1M.

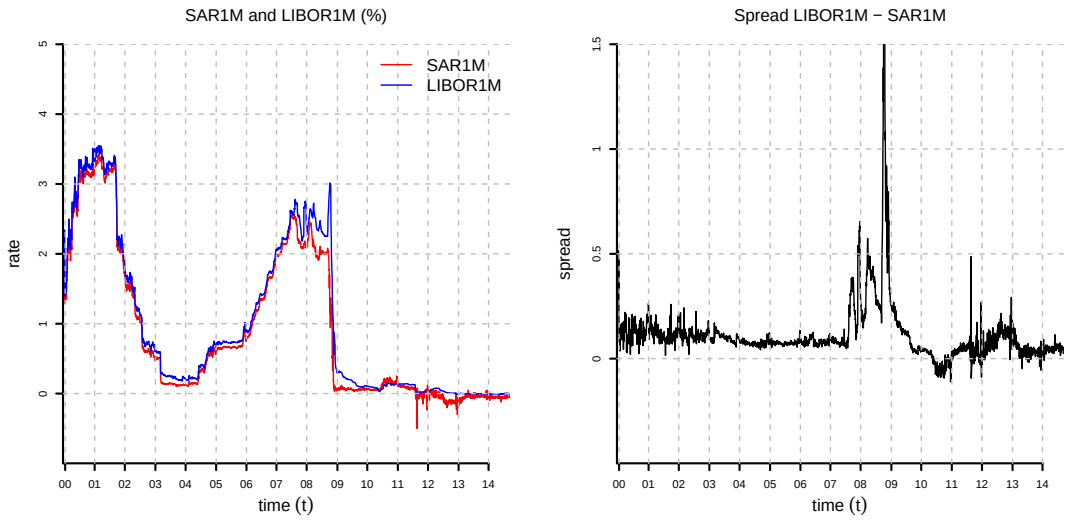

Figure B.8: (lhs) SAR1M and LIBOR1M from December 8, 1999, until September 15, 2014; (rhs) spread between LIBOR1M and SAR1M. Observe that these instruments have the same time to maturity of 1 month. However, SAR1M is secured whereas LIBOR1M reflects conditions at the unsecured money market with a resulting positive credit spread.

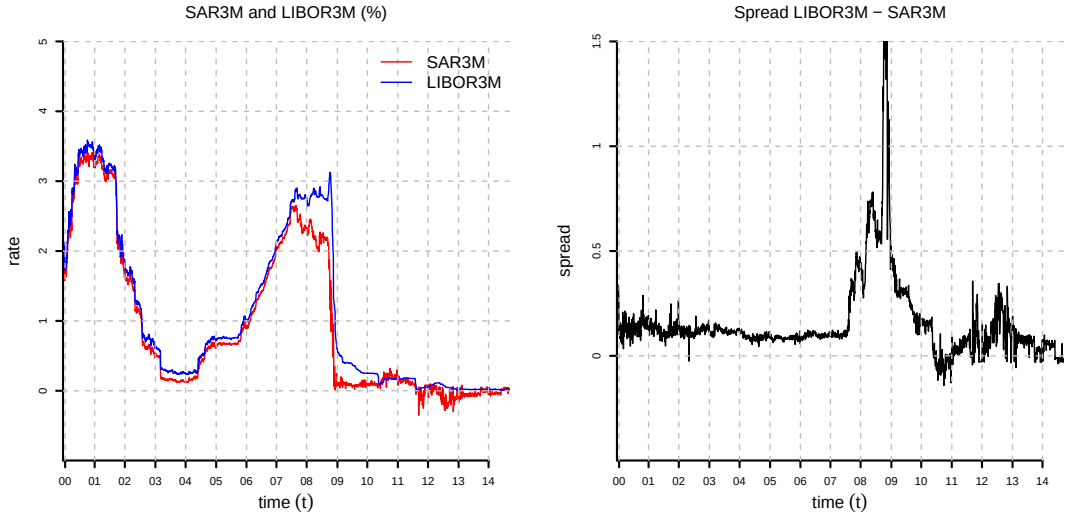

Figure B.9: (lhs) SAR3M and LIBOR3M from December 8, 1999, until September 15, 2014; (rhs) spread between LIBOR3M and SAR3M. Observe that these instruments have the same time to maturity of 3 months. However, SAR3M is secured whereas LIBOR3M reflects conditions at the unsecured money market with a resulting positive credit spread.

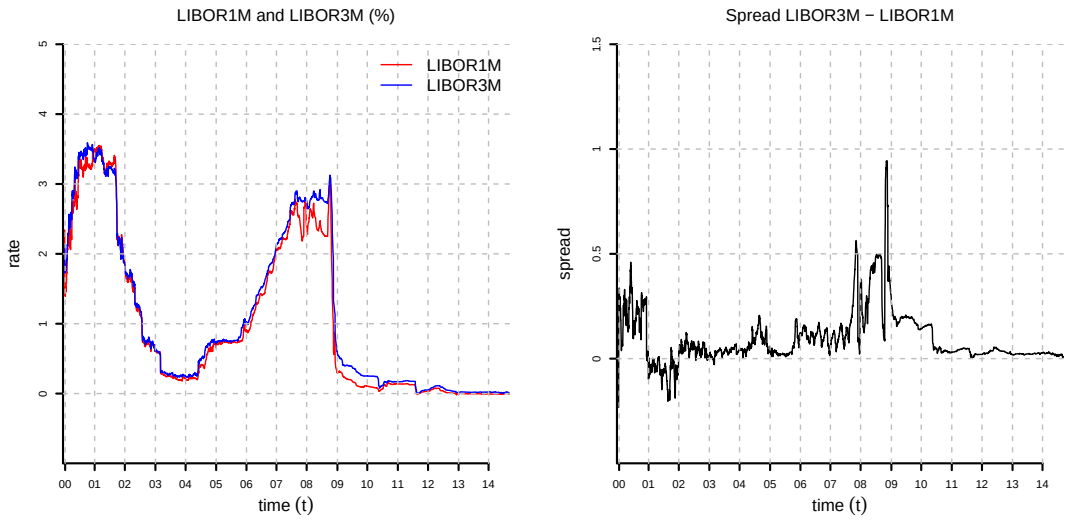

Figure B.10: (lhs) LIBOR1M and LIBOR3M from December 8, 1999, until September 15, 2014; (rhs) spread between LIBOR3M and LIBOR1M.

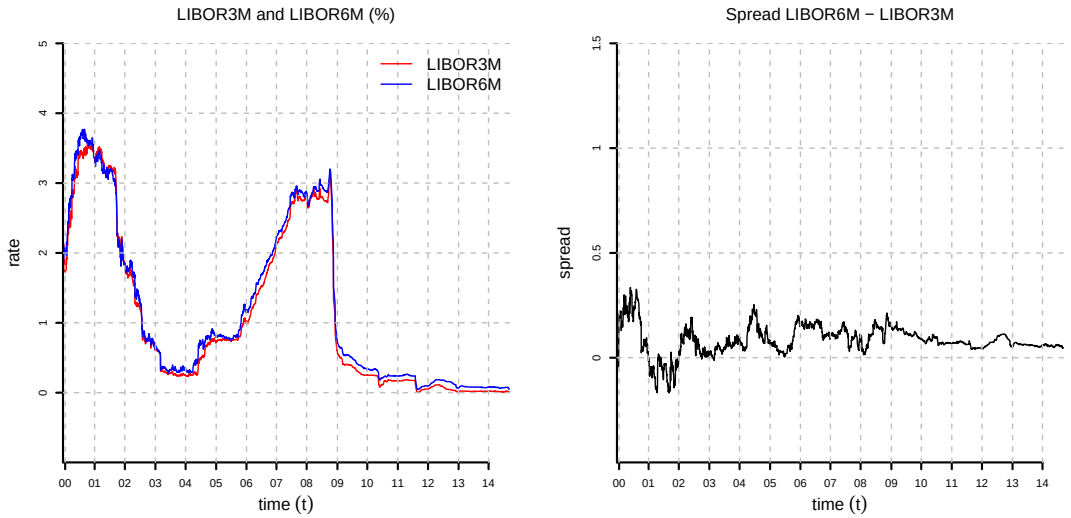

Figure B.11: (lhs) LIBOR3M and LIBOR6M from December 8, 1999, until September 15, 2014; (rhs) spread between LIBOR6M and LIBOR3M.

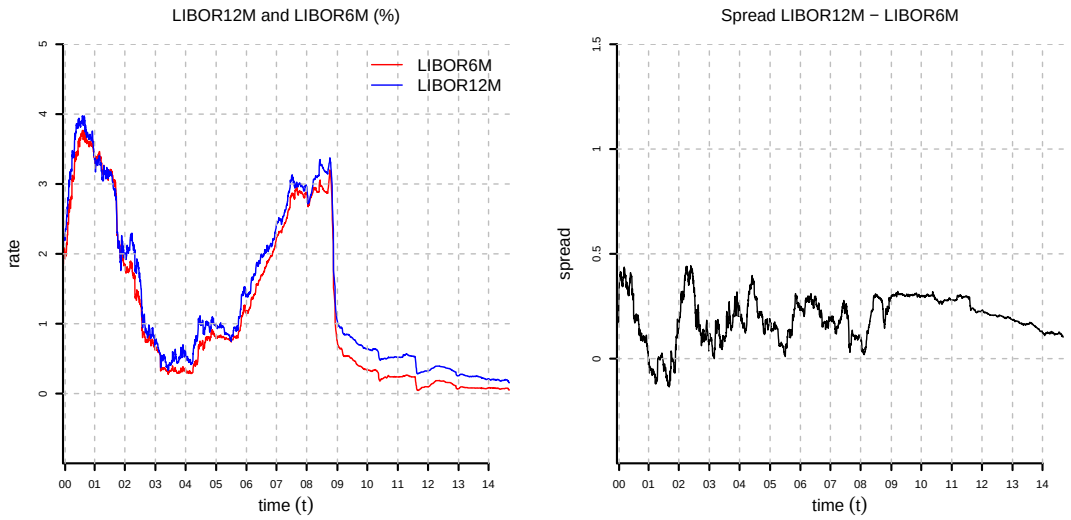

Figure B.12: (lhs) LIBOR6M and LIBOR12M from December 8, 1999, until September 15, 2014; (rhs) spread between LIBOR12M and LIBOR6M.

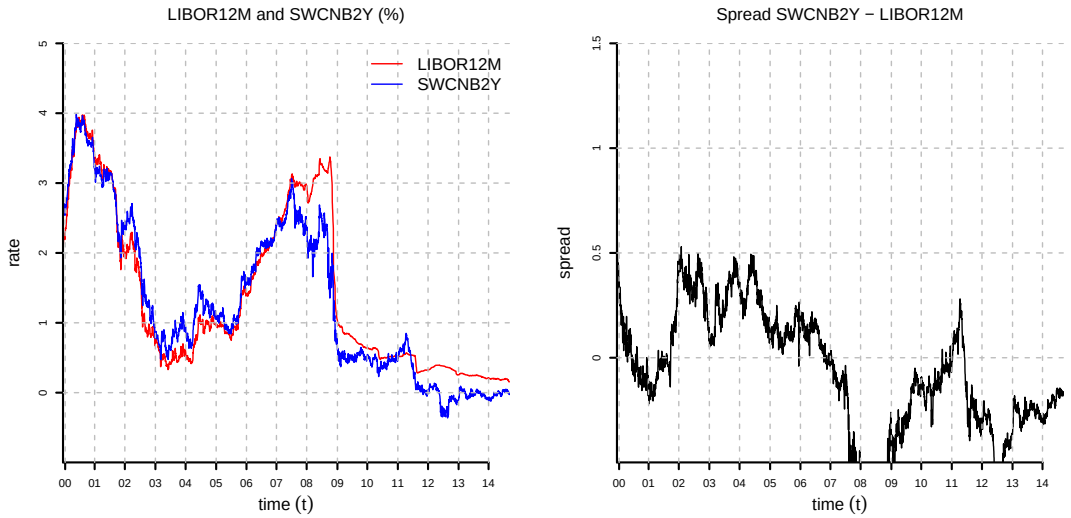

Figure B.13: (lhs) LIBOR12M and SWCNB2Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB2Y and LIBOR12M. This picture indicates that LIBOR for medium times to maturity and SWCNB for long times to maturity are not completely consistent.

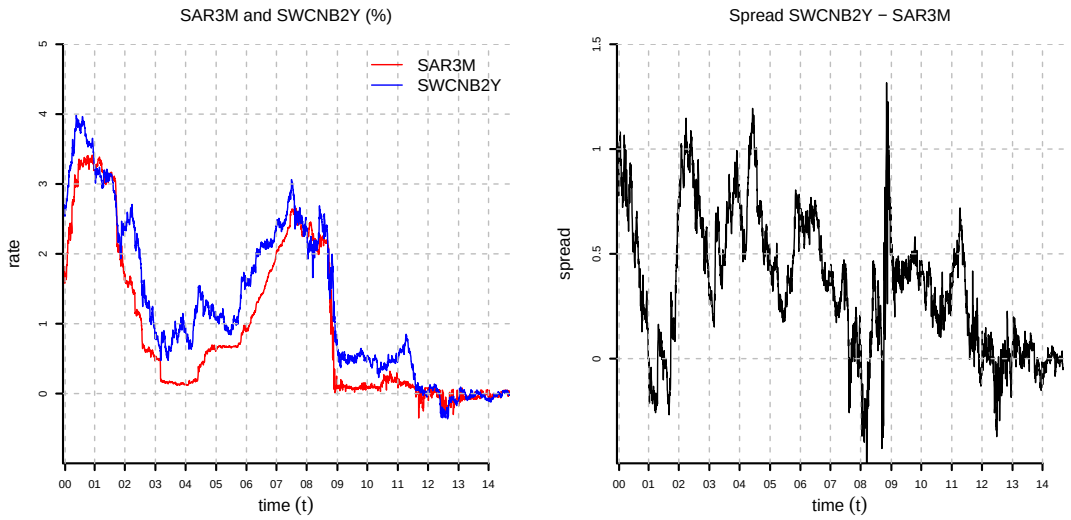

Figure B.14: (lhs) SAR3M and SWCNB2Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB2Y and SAR3M. This picture indicates that SAR fits better to SWCNB than LIBOR and, therefore, we may drop LIBOR.

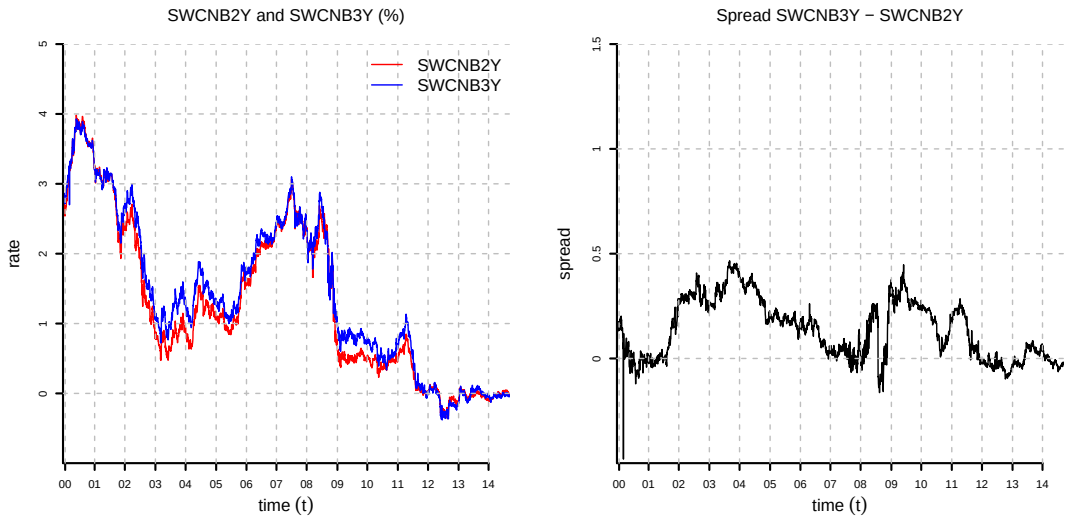

Figure B.15: (lhs) SWCNB2Y and SWCNB3Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB3Y and SWCNB2Y.

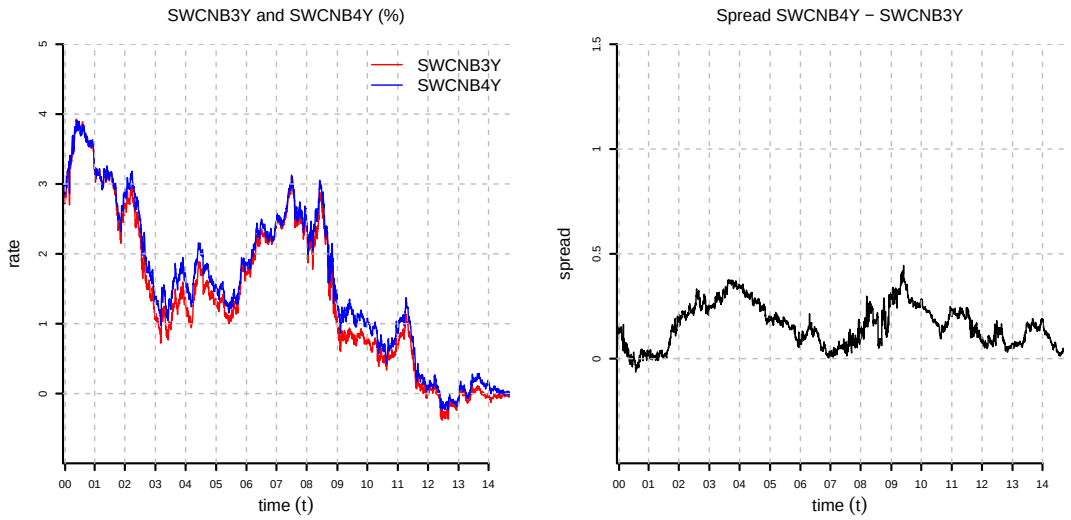

Figure B.16: (lhs) SWCNB3Y and SWCNB4Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB4Y and SWCNB3Y.

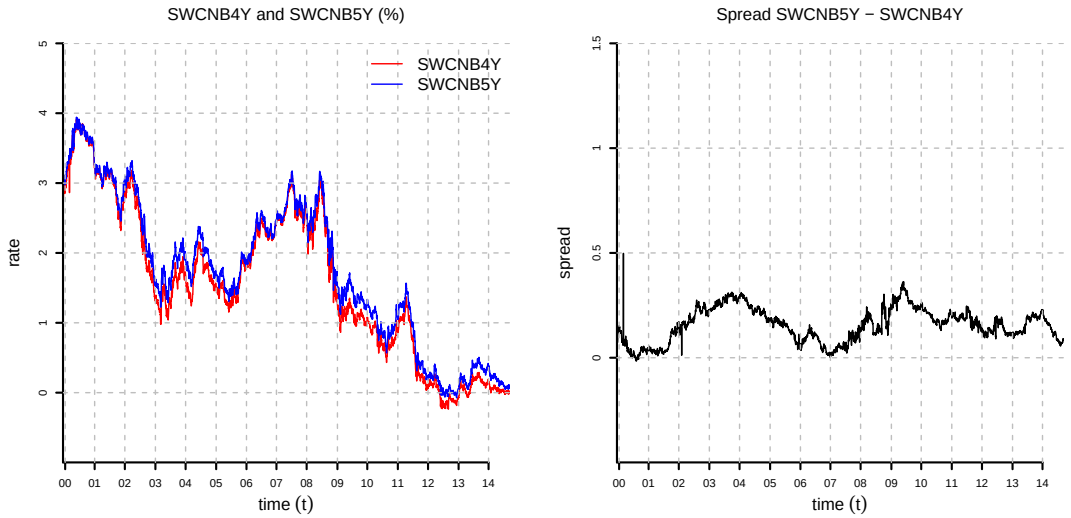

Figure B.17: (lhs) SWCNB4Y and SWCNB5Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB5Y and SWCNB4Y.

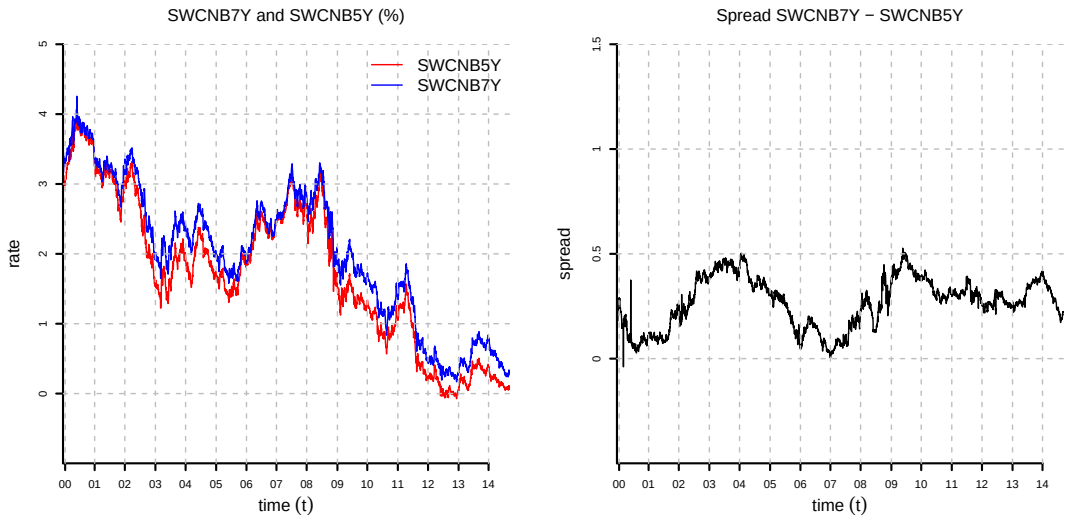

Figure B.18: (lhs) SWCNB5Y and SWCNB7Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB7Y and SWCNB5Y.

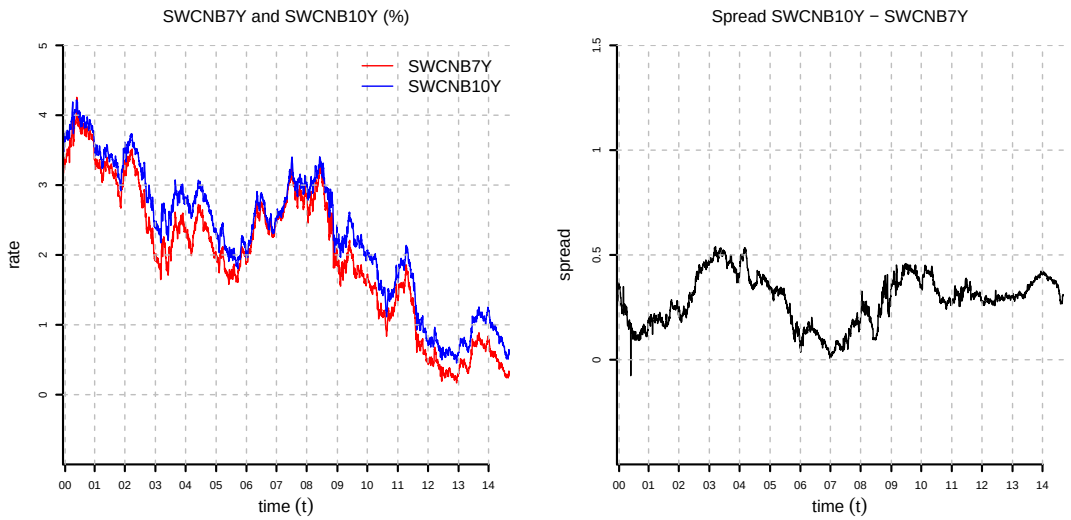

Figure B.19: (lhs) SWCNB7Y and SWCNB10Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB10Y and SWCNB7Y.

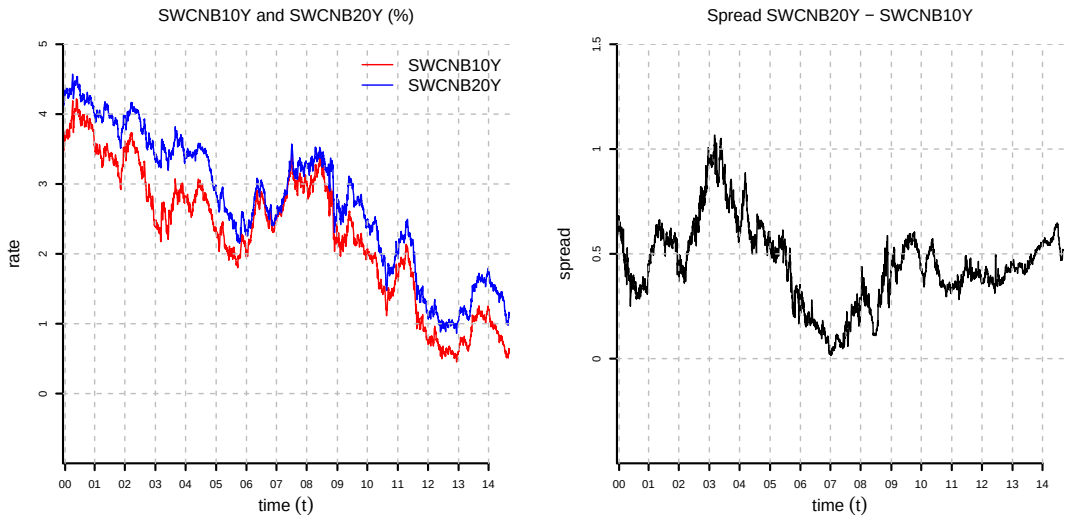

Figure B.20: (lhs) SWCNB10Y and SWCNB20Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB20Y and SWCNB10Y.

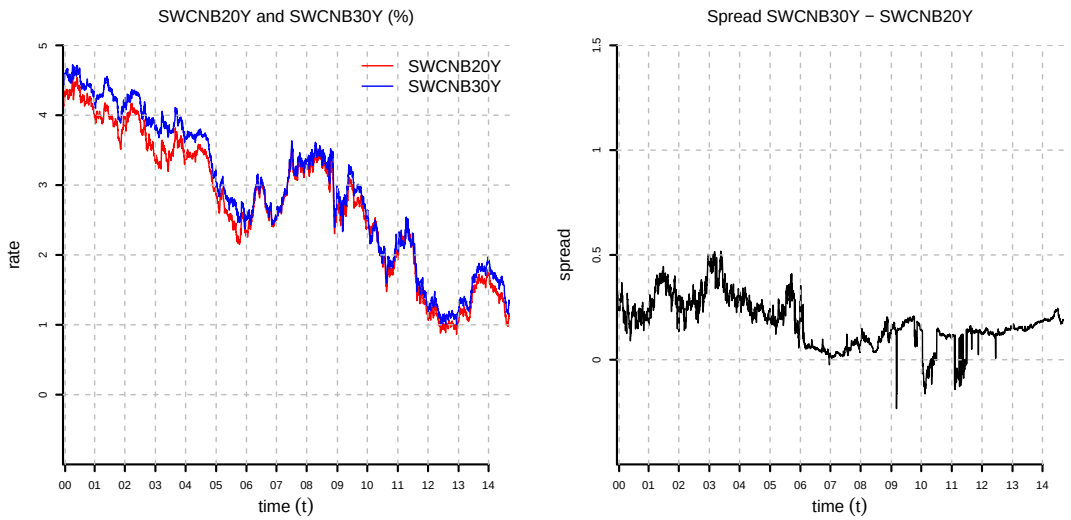

Figure B.21: (lhs) SWCNB20Y and SWCNB30Y from December 8, 1999, until September 15, 2014; (rhs) spread between SWCNB30Y and SWCNB20Y.

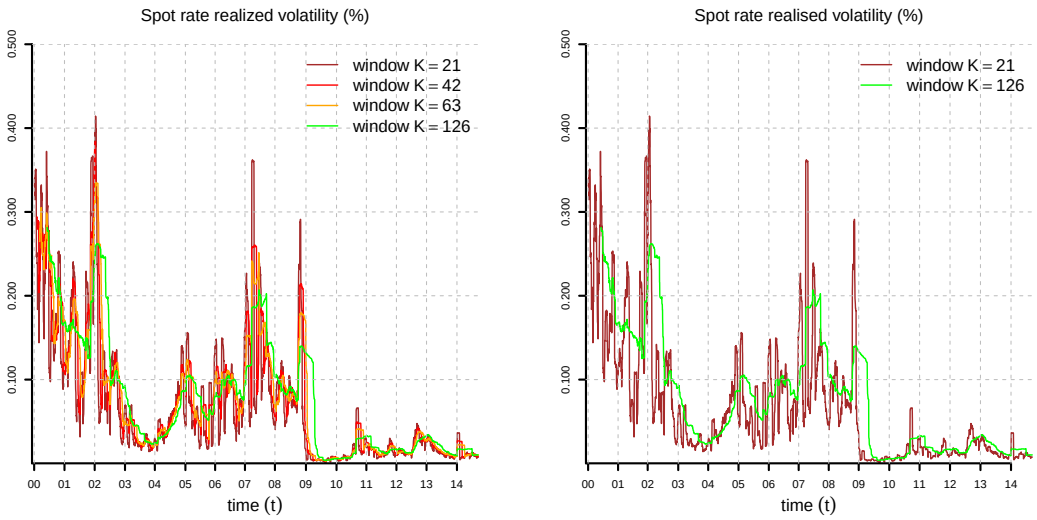

Figure B.22: Spot rate realised volatility $\widehat{\mathrm{RCov}}(t,1,1)^{\frac{1}{2}}$ for window lengths $K = 21, 42, 63, 126$ which becomes more viscous for longer window lengths. See formula (5.6).

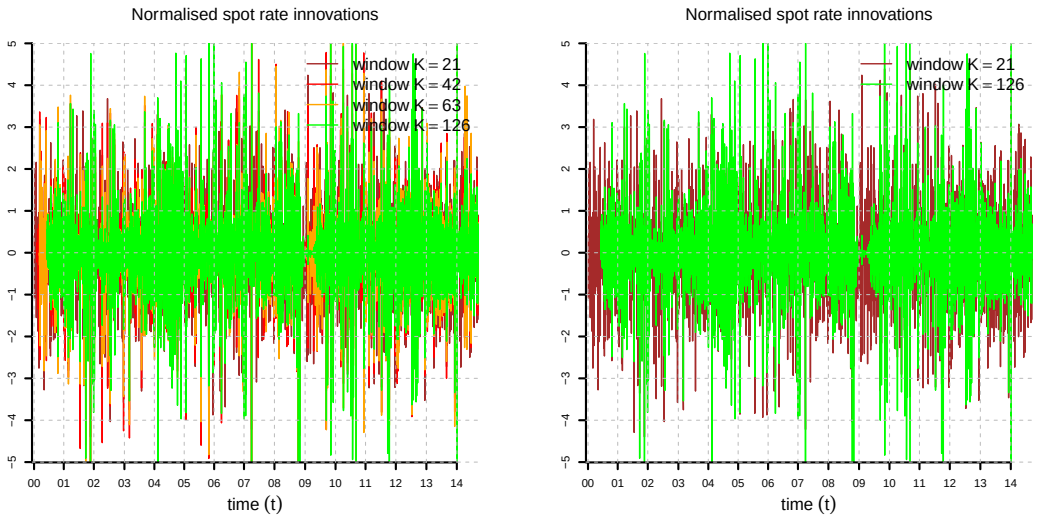

Figure B.23: Normalised spot rate innovations $\widehat{\mathrm{RCov}}(t,1,1)^{-\frac{1}{2}}\Delta r(t)$ having for all window lengths roughly empirical mean 0 and empirical variance 1.

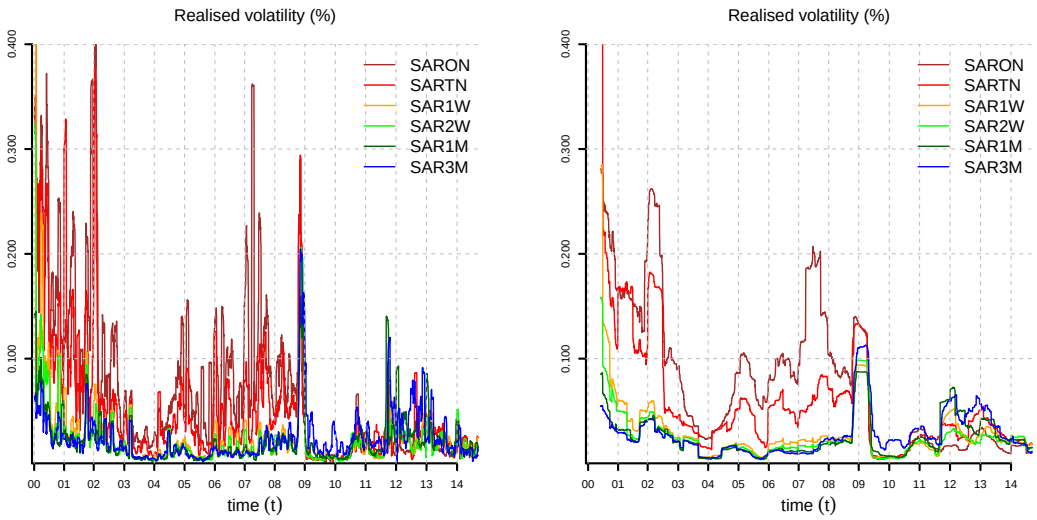

Figure B.24: SAR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)^{\frac{1}{2}}$ for $\tau = 1, 2, 5, 10, 21, 63$, window length $K = 21$ (lhs) and $K = 126$ (rhs).

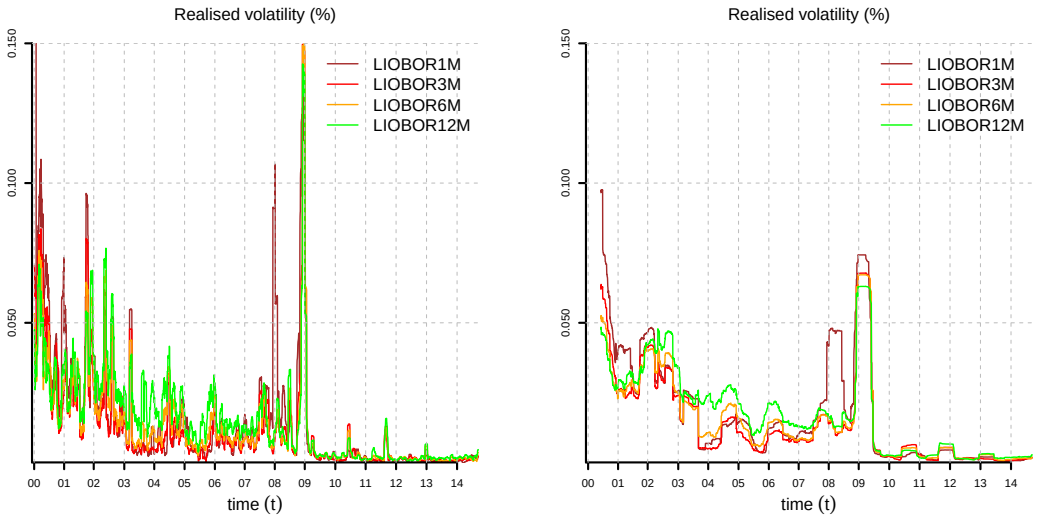

Figure B.25: CHF LIBOR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)^{\frac{1}{2}}$ for $\tau = 21, 63, 126, 252$, window length $K = 21$ (lhs) and $K = 126$ (rhs).

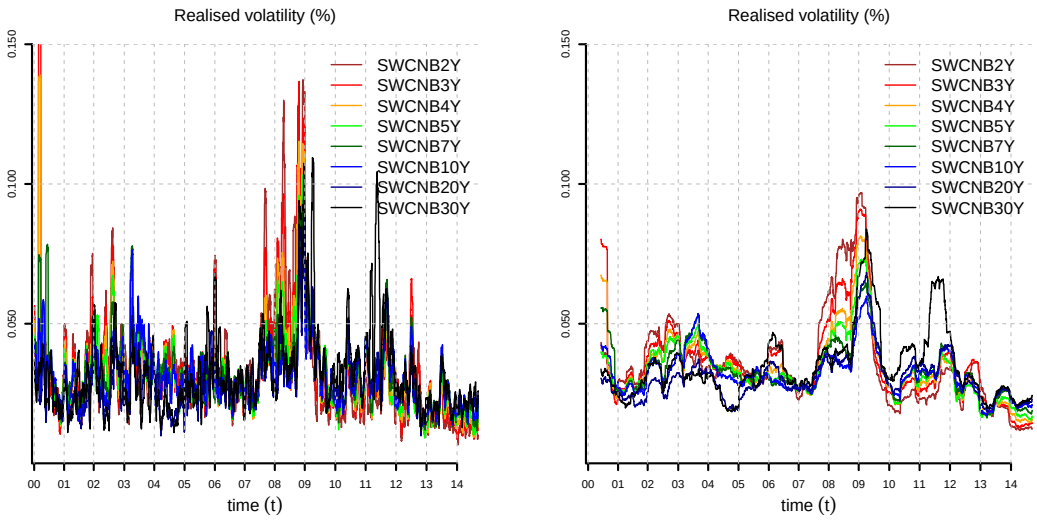

Figure B.26: SWCNB realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)^{\frac{1}{2}}$ for $\tau/252 = 2,3,4,5,7,10,20,30$, window length $K = 21$ (lhs) and $K = 126$ (rhs).

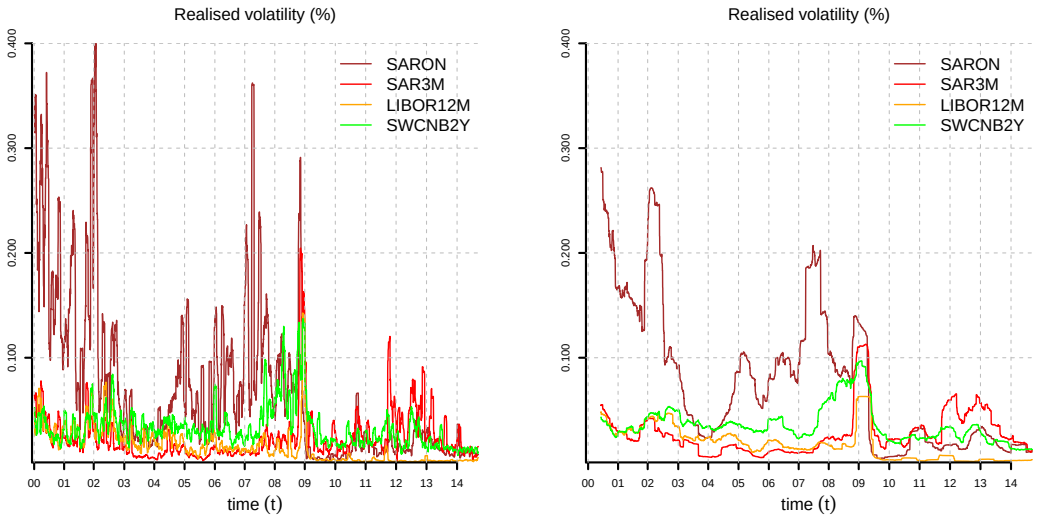

Figure B.27: SAR, CHF LIBOR and SWCNB realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)^{\frac{1}{2}}$ for $\tau = 1,63,252,504$, window length $K = 21$ (lhs) and $K = 126$ (rhs). Note that LIBOR looks rather differently from SAR and SWCNB after the financial crisis of 2008.

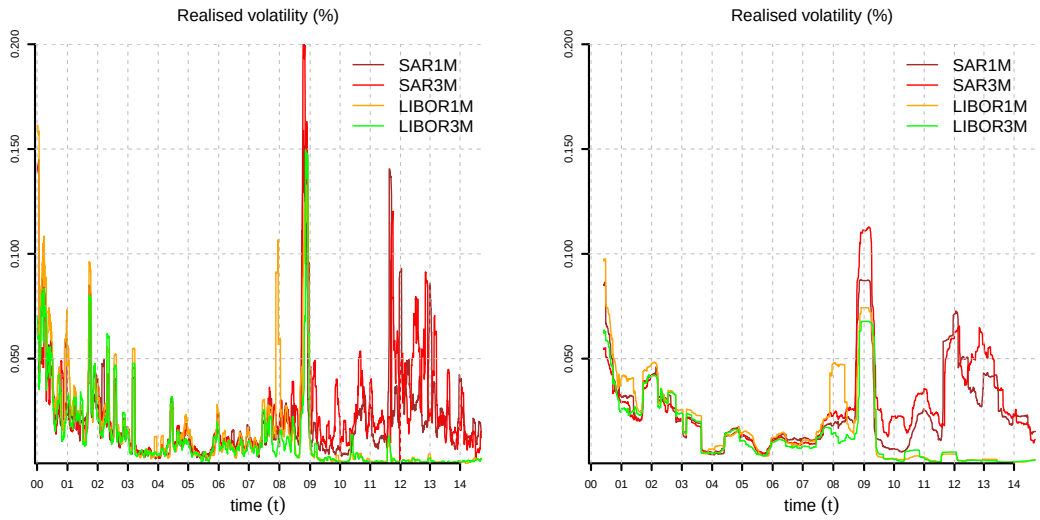

Figure B.28: SAR and CHF LIBOR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)^{\frac{1}{2}}$ for $\tau = 21, 63$, window length $K = 21$ (lhs) and $K = 126$ (rhs). Note that LIBOR looks rather differently from SAR after the financial crisis of 2008.

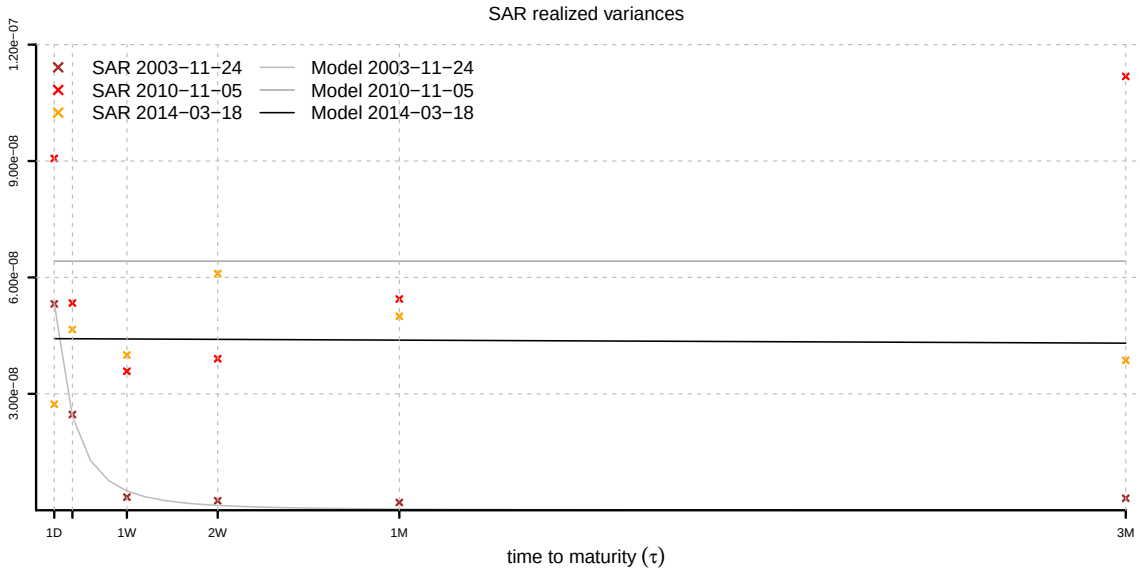

Figure B.29: SAR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)$ for $K = 126$, $\tau = 1, 2, 5, 10, 21, 63$ and three observation dates compared to the realised volatility of the 1-factor Vasiček model fitted by optimisation (5.7) for $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$, $\tau_6 = 63$ and $w_{ij} = 1_{i=j}$. We observe that the 1-factor model is not flexible enough to achieve an accurate fit.

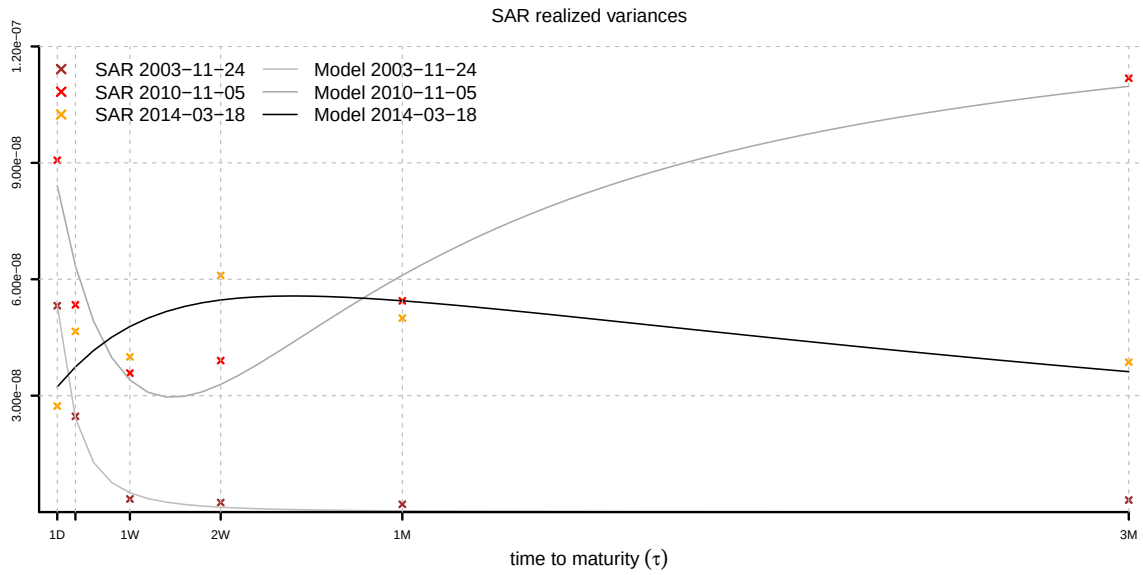

Figure B.30: SAR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)$ for $K = 126$, $\tau = 1, 2, 5, 10, 21, 63$ and three observation dates compared to the realised volatility of the 2-factor Vasiček model fitted by optimisation (5.7) for $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$, $\tau_6 = 63$ and $w_{ij} = 1_{i=j}$. The 2-factor model achieves an suitable fit.

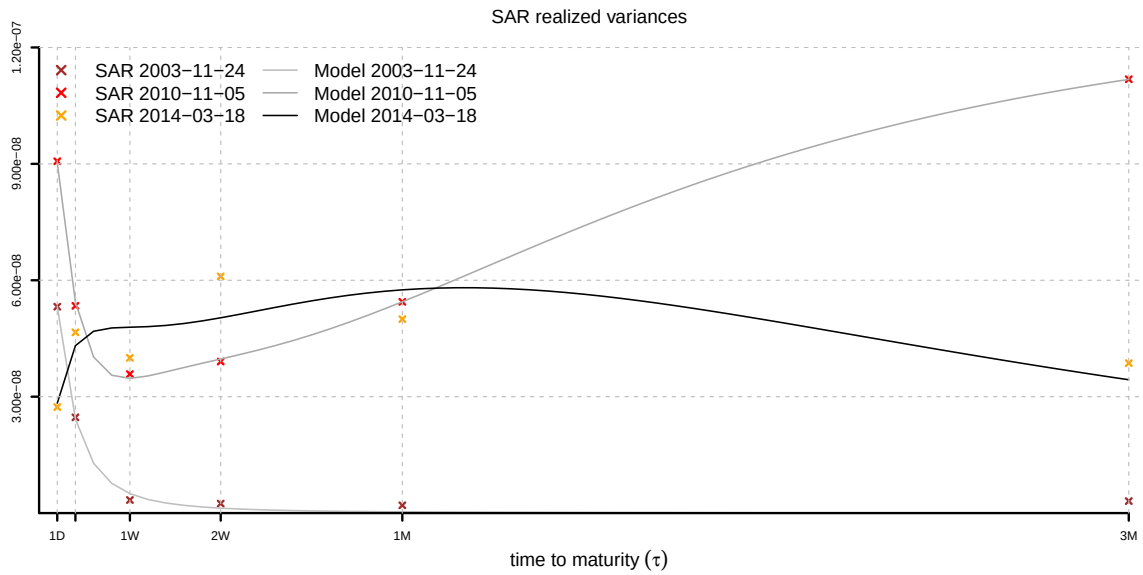

Figure B.31: SAR realised volatility $\widehat{\mathrm{RCov}}(t,\tau,\tau)$ for $K = 126$, $\tau = 1, 2, 5, 10, 21, 63$ and three observation dates compared to the realised volatility of the 3-factor Vasiček model fitted by optimisation (5.7) for $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$, $\tau_6 = 63$ and $w_{ij} = 1_{i=j}$. The 3-factor model achieves a slightly better fit than the 2-factor model.

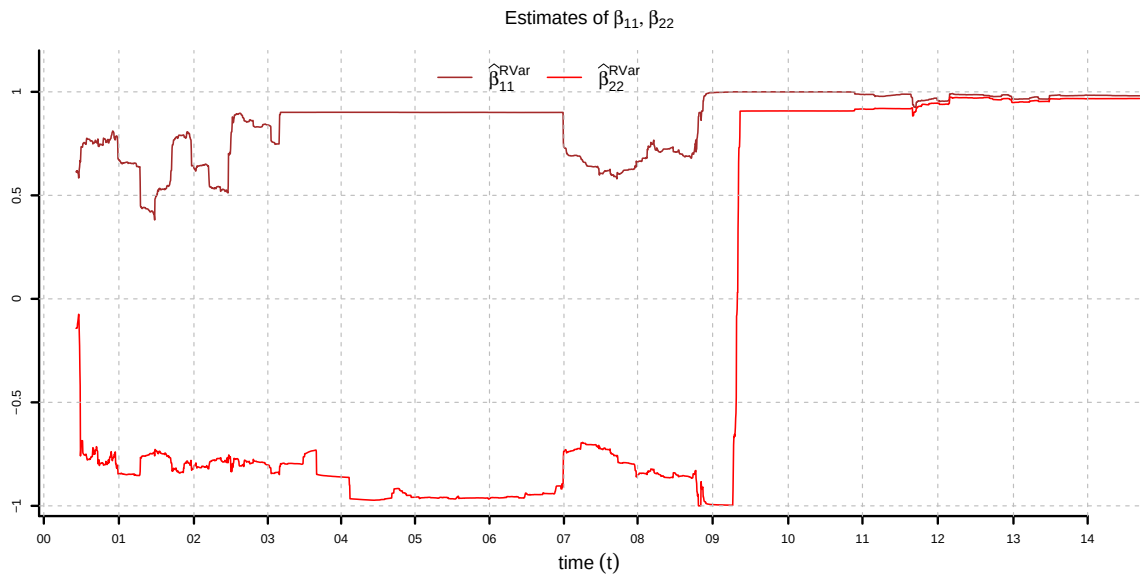

Figure B.32: Estimation of β_{11} and β_{22} by optimisation (5.7) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$.

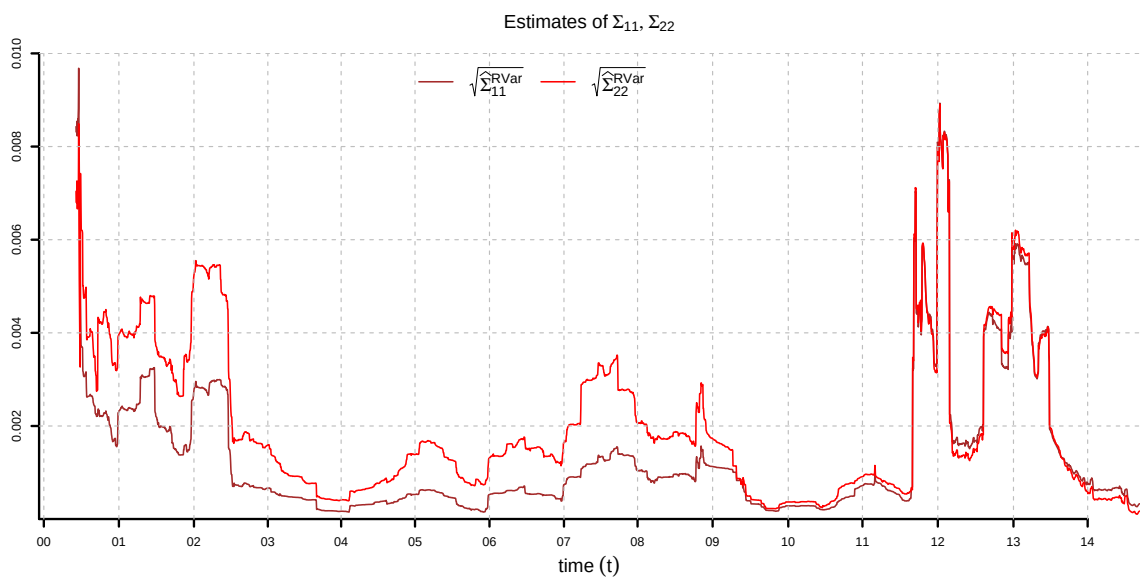

Figure B.33: Estimation of Σ_{11} and Σ_{22} by optimisation (5.7) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$.

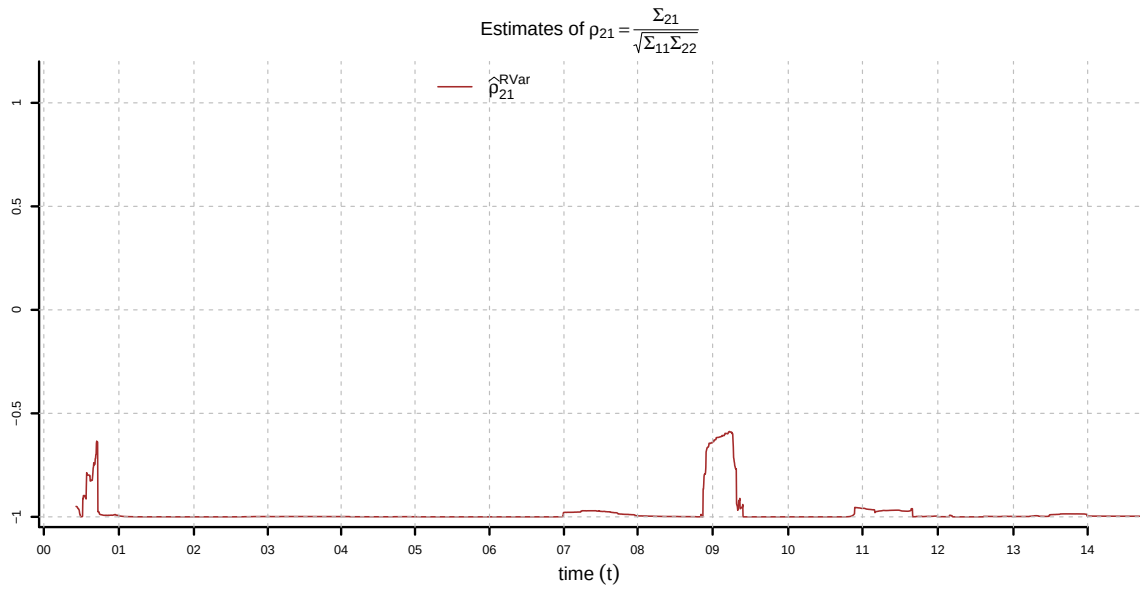

Figure B.34: Estimation of $\rho_{21} = \mathrm{Corr}\,(X_1(t), X_2(t)|\mathcal{F}(t-1))$ by optimisation (5.7) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$.

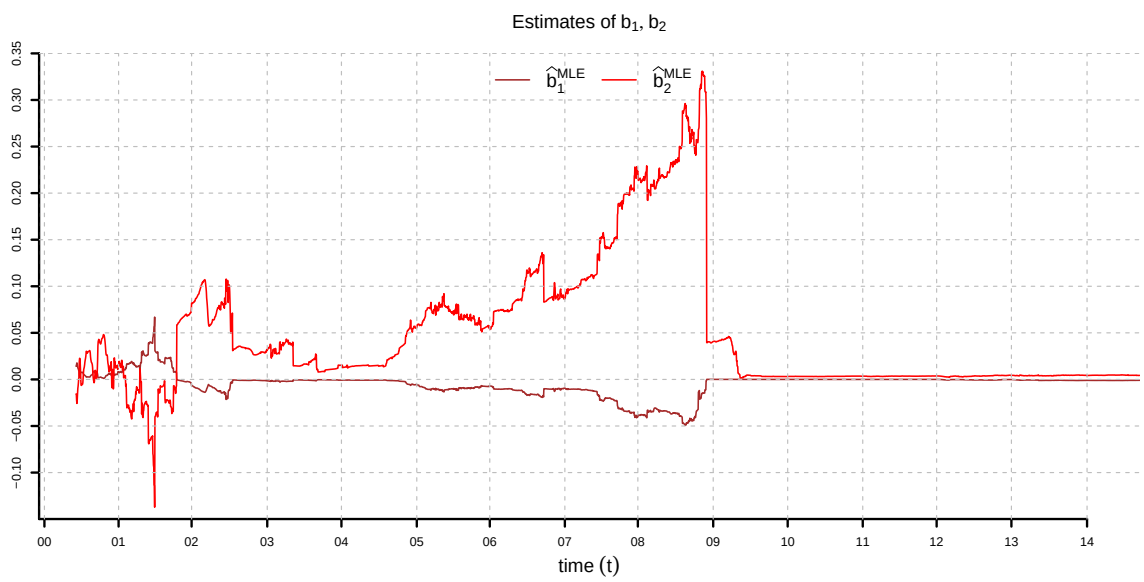

Figure B.35: Estimation of b_1 and b_2 by optimisation (5.8) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$ and $S = 10^{-4} \cdot \mathbb{1}$.

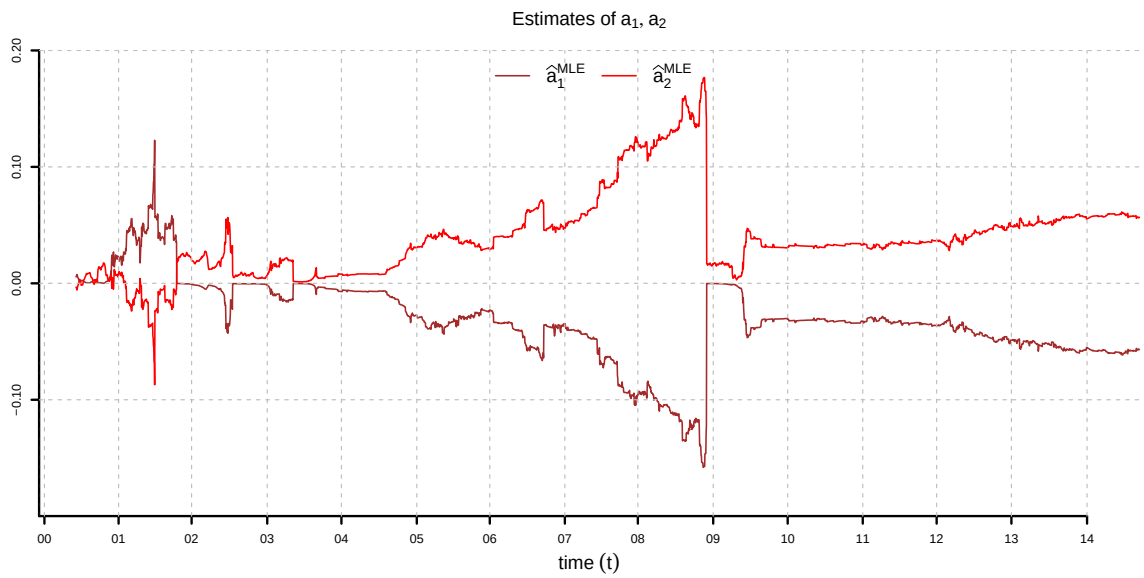

Figure B.36: Estimation of a_1 and a_2 by optimisation (5.8) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$ and $S = 10^{-4} \cdot \mathbb{1}$.

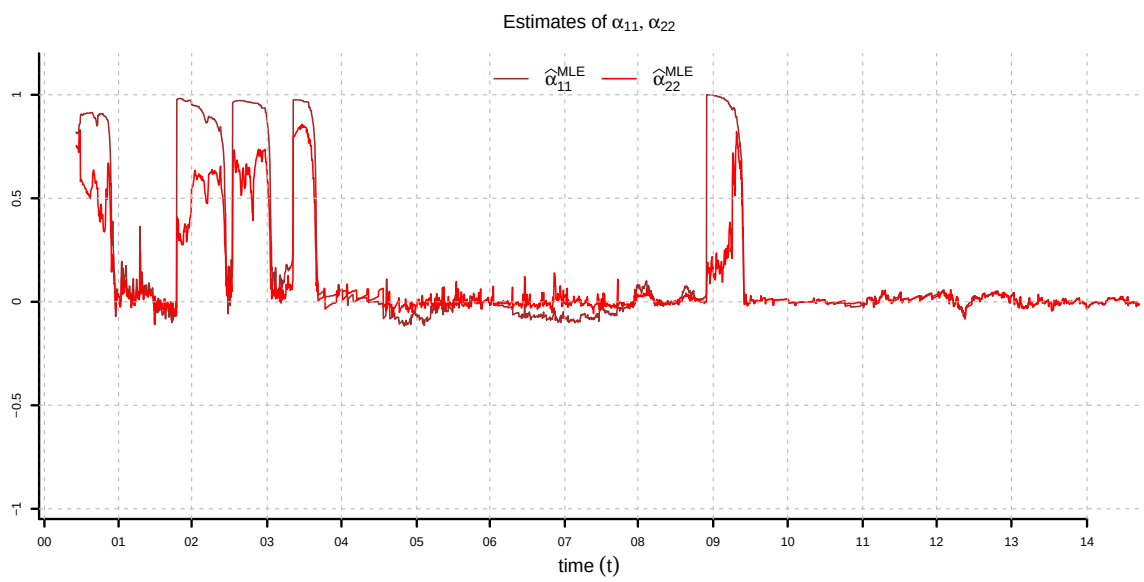

Figure B.37: Estimation of α_{11} and α_{22} by optimisation (5.8) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$ and $S = 10^{-4} \cdot \mathbb{1}$.

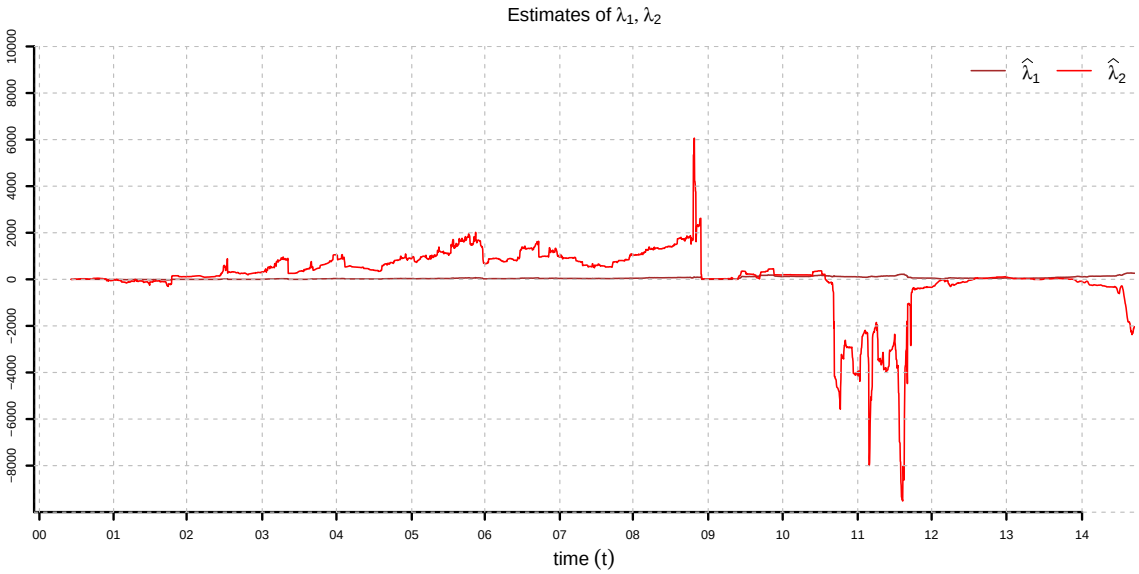

Figure B.38: Estimation of the market price of risk parameter $\boldsymbol{\lambda}$ by (5.9).

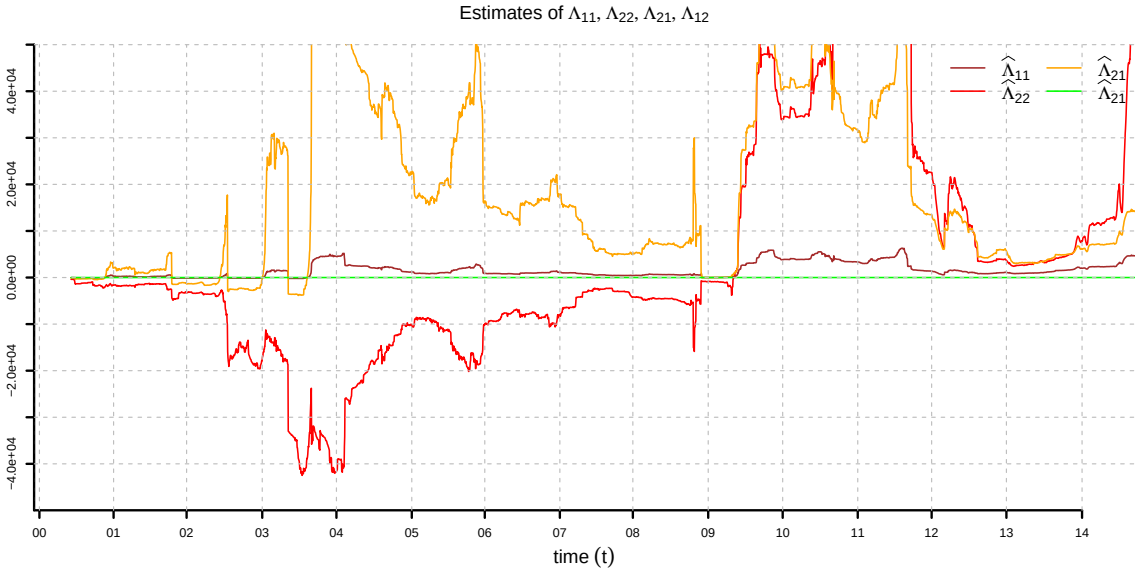

Figure B.39: Estimation of the market price of risk parameter Λ by (5.10).

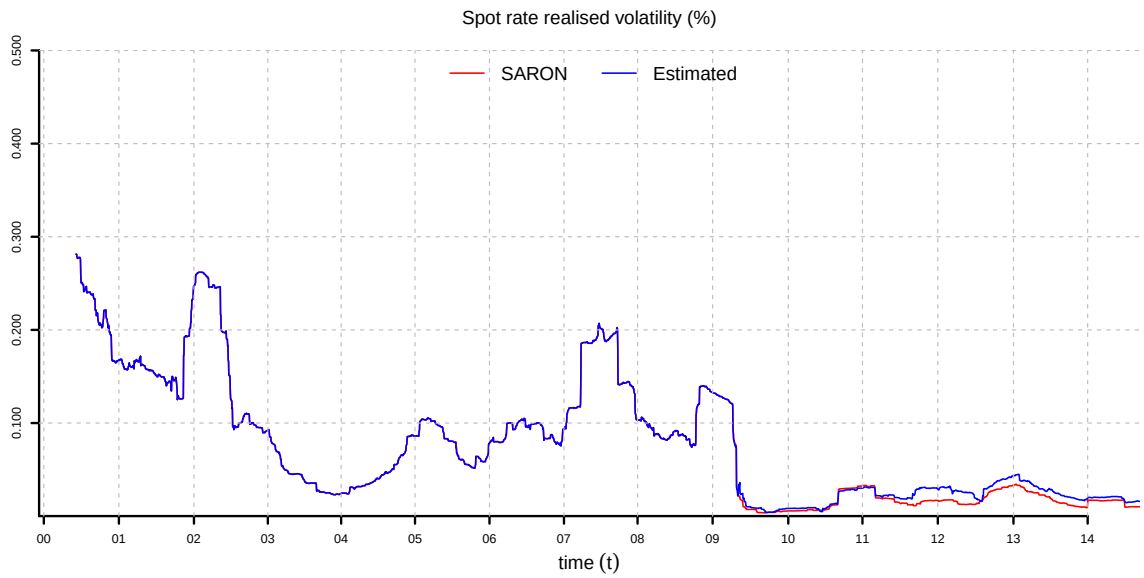

Figure B.40: Comparison of spot rate realised volatility $\widehat{\mathrm{RCov}}(t,1,1)^{\frac{1}{2}}$ and estimated instantaneous standard deviation $\mathbf{1}^\top \Sigma^{\mathrm{RCov}} \mathbf{1}$, where Σ^{RCov} is estimated by optimisation (5.7) in the 2-factor case for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$.

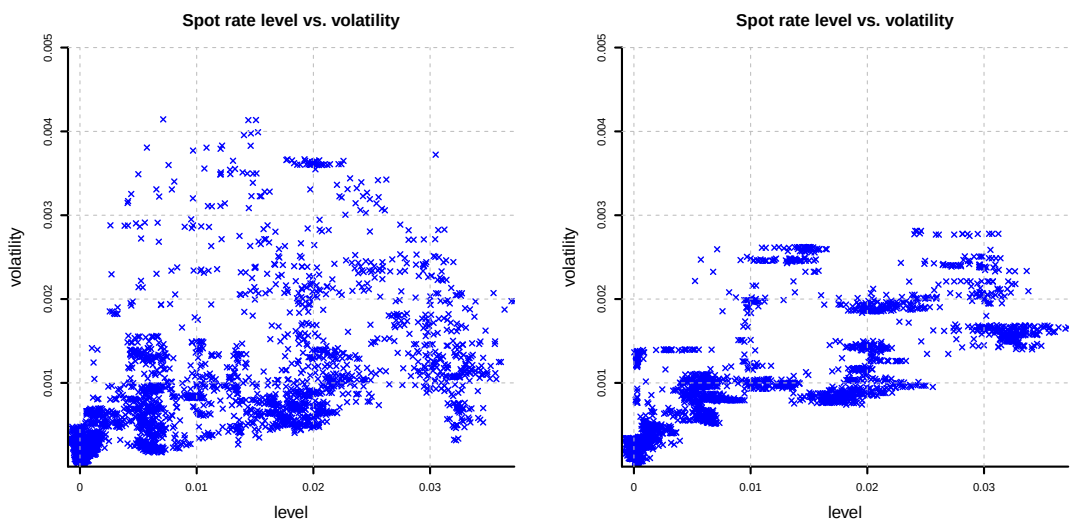

Figure B.41: Level of spot rate $r(t-1)$ versus realised volatility $\mathbf{1}^\top \widehat{\Sigma}^{\mathrm{RCov}}(t)\mathbf{1}$ (lhs) for window length $K = 21$; (rhs) for window length $K = 126$. Outliers in the volatility direction need to be considered with care because there is a lag in estimation of $\widehat{\Sigma}^{\mathrm{RCov}}(t)$ compared to the level of spot rate through the viscosity choice K.

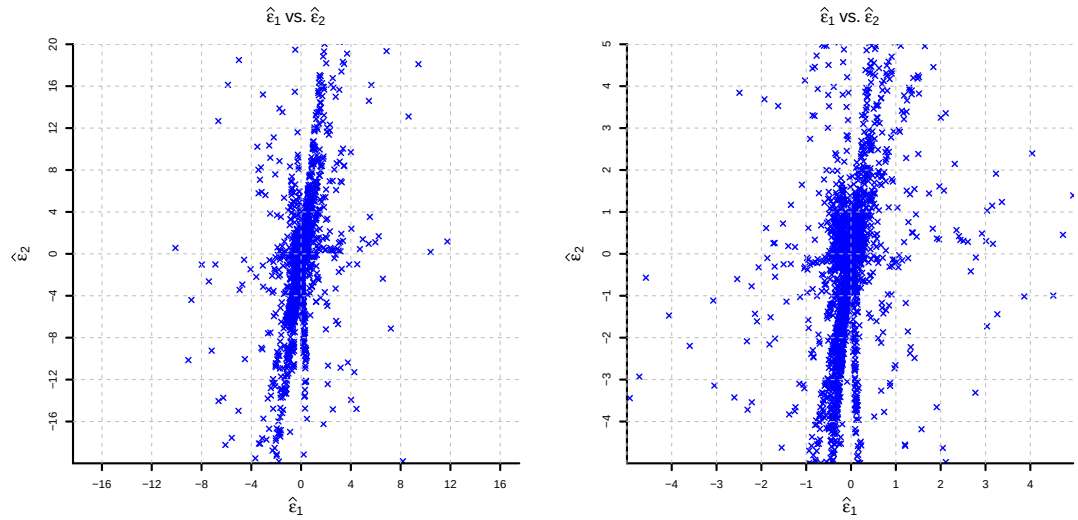

Figure B.42: Estimated residuals $\widehat{\boldsymbol{\varepsilon}}(t+1) = \widehat{\Sigma}^{\mathrm{RCov}}(t)^{-\frac{1}{2}} \Big(\boldsymbol{X}(t+1) - \widehat{\boldsymbol{a}}^{\mathrm{MLE}}(t) - \widehat{\alpha}^{\mathrm{MLE}}(t) \boldsymbol{X}(t) \Big)$, where $\boldsymbol{X}$ is obtained from the Kalman filter of Section 5.2 with window length $K = 126$. Empirical mean is TODO and empirical standard deviation is TODO.

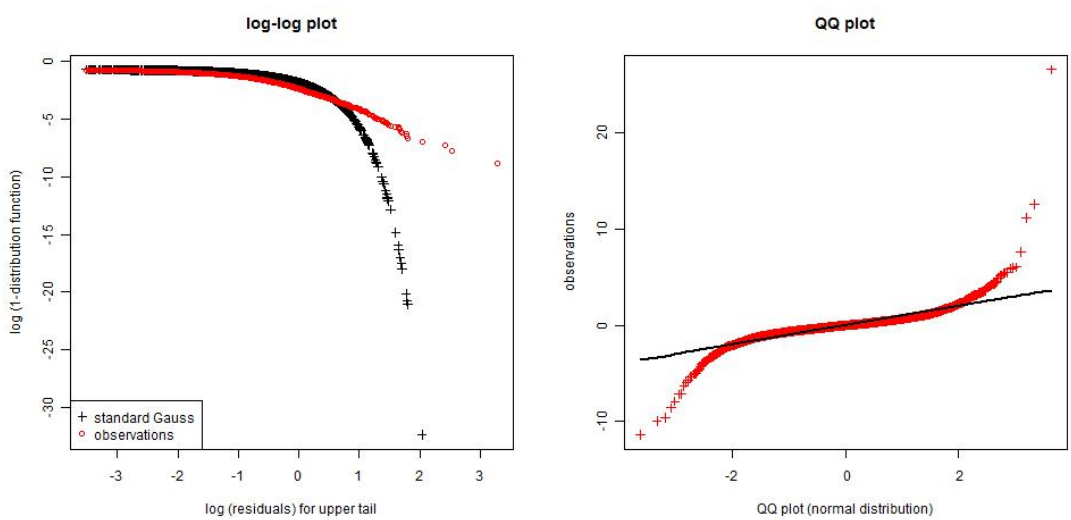

Figure B.43: Estimated residuals $\widehat{\boldsymbol{\varepsilon}}(t+1) = \widehat{\Sigma}^{\mathrm{RCov}}(t)^{-\frac{1}{2}} \Big(\boldsymbol{X}(t+1) - \widehat{\boldsymbol{a}}^{\mathrm{MLE}}(t) - \widehat{\alpha}^{\mathrm{MLE}}(t) \boldsymbol{X}(t) \Big)$ for a window length of $K = 126$. (lhs) log-density-plot $\widehat{\boldsymbol{\varepsilon}}_1(t+1)$; (rhs) QQ-plot $\widehat{\boldsymbol{\varepsilon}}_1(t+1)$. We observe that these residuals are more heavy-tailed than Gaussian distributions which is also suggested by the empirical standard deviation of TODO.

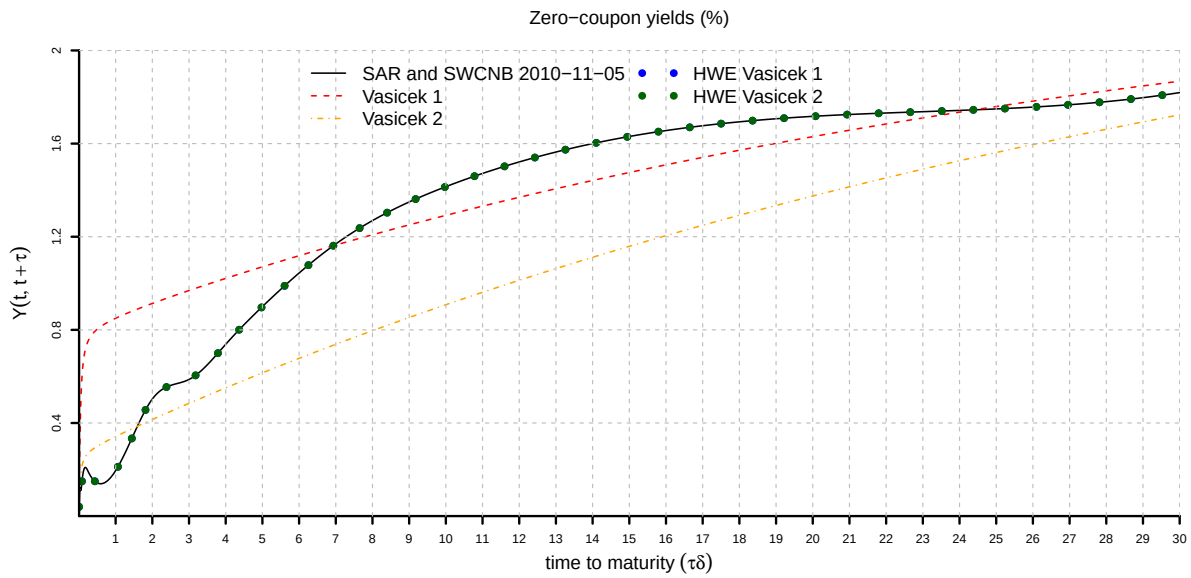

Figure B.44: Calibration of 2-factor Vasiček models and Hull-White extensions as of 3 November, 2010. Vasiček 1 is obtained in the following way. First, β and Σ are estimated by optimisation (5.7) for $K = 126$, $M = 6$, $\tau_1 = 1$, $\tau_2 = 2$, $\tau_3 = 5$, $\tau_4 = 10$, $\tau_5 = 21$ and $\tau_6 = 63$. Secondly, $\boldsymbol{b}$ is calibrated to the prevailing yield curve. Vasiček 2 is obtained from optimisations (5.7) and (5.8). The Hull-White extended models match the market yield curve exactly.

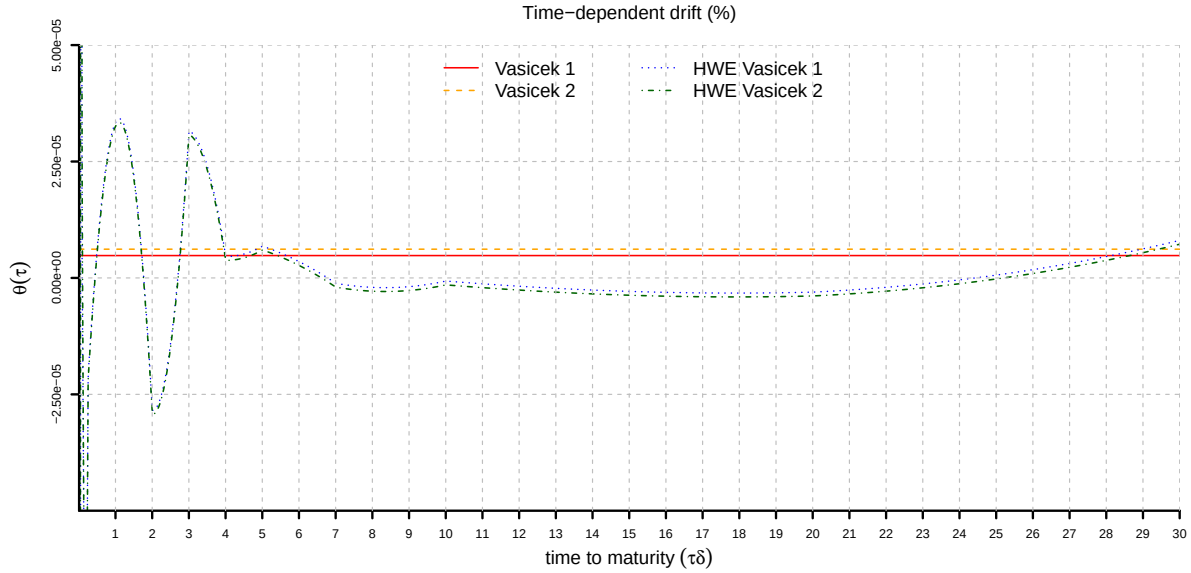

Figure B.45: The Hull-White extension θ corresponding to the models of Figure B.44.

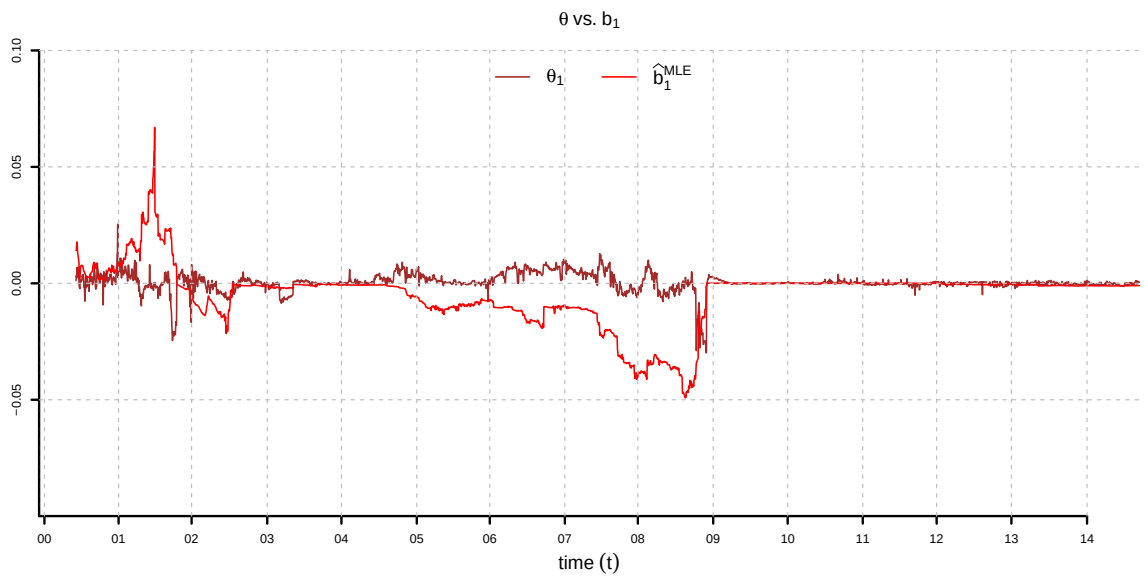

Figure B.46: θ_1 given by (A.1) for $\widehat{\boldsymbol{b}}^{\mathrm{MLE}}$, $\widehat{\beta}^{\mathrm{RCov}}$ and $\widehat{\Sigma}^{\mathrm{RCov}}$.

References

[1] Brockwell, P.J., Davis, R.A. (1991). Time Series: Theory and Methods. Springer.

[2] Cox, J.C., Ingersoll, J.E., Ross, S.A. (1985). A theory of the term structure of interest rates. *Econometrica* **53/2**, 385-407.

[3] Harms, P., Stefanovits, D., Teichmann, J., Wüthrich, M.V. (2015). Consistent recalibration of yield curve models. *arXiv*:1502.02926

[4] Heath, D., Jarrow, R., Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claim valuation. *Econometrica* **60/1**, 77-105.

[5] Hull, J., White, A. (1994). Branching out. *Risk* **7**, 34-37.

[6] Jordan, T.J. (2009). SARON - an innovation for the financial markets. Launch event for Swiss Reference Rates, Zurich, August 25, 2009.

[7] Vasiček, O. (1977). An equilibrium characterization of the term structure. *Journal Financial Economics* **5/2**, 177-188.

[8] Wüthrich, M.V., Merz, M. (2013). Financial Modeling, Actuarial Valuation and Solvency in Insurance. Springer.