# The Clean Thesis Style

Ricardo Langner

*August 26, 2015*

Version: My First Draft

Clean Thesis Style University

CleanThesis

Department of Clean Thesis Style
Institut for Clean Thesis Dev
Clean Thesis Group (CTG)

Documentation

# The Clean Thesis Style

Ricardo Langner

*1. Reviewer* Jane Doe
Department of Clean Thesis Style
Clean Thesis Style University

*2. Reviewer* John Doe
Department of Clean Thesis Style
Clean Thesis Style University

*Supervisors* Jane Doe and John Smith

August 26, 2015

**Ricardo Langner**
*The Clean Thesis Style*
Documentation, August 26, 2015
Reviewers: Jane Doe and John Doe
Supervisors: Jane Doe and John Smith

**Clean Thesis Style University**
*Clean Thesis Group (CTG)*
Institut for Clean Thesis Dev
Department of Clean Thesis Style
Street address
Postal Code and City

# Contents

# 1 A swap curve for insurance risk management, based on no arbitrage short-rate models

## 1.1 Context

The new Solvency II directive defines the calculation of European insurers' technical provisions as the sum of two components, the Best Estimate Liabilities (BEL) and the Risk Margin (RM). The Best Estimate Liabilities (BEL) are defined as the average discounted value of the insurer's future cash-flows, weighted by their probability of occurrence. The Risk Margin is a supplemental amount required for covering the non-hedgeable risks, by involving a capital lockup.

In order to discount the cash-flows relevant in the calculation of the BEL and Risk Margin, an *appropriate* term structure of discount factors is needed. From the no arbitrage pricing theory developed by [HP81], and widely used in insurance *market consistent* pricing of liabilities, the zero rates related to the stochastic discount factors have to be *risk-free*. That is, free from any counterparty credit risk.

There is no easy answer to the question of defining such a *risk-free* rate for insurance liabilities. It could be related the insurer's own assets return, where the liabilities are *perfectly* backed by the assets. But in a *market consistent* approach as required in Solvency II, since not every liability is perfectly backed by the assets, a more fundamental *risk-free* rate also needs to be derived. Deriving such a *common risk-free* rate from market-quoted instruments is also aimed at increasing transparency and comparability of balance sheets across European countries.

For years, in banking, the construction of a term structure of *risk-free* discount factors was based on the assumption that banks are not subject to counterparty credit risk when lending to each other, and liquidity was not an issue. In this context, interbank rates (loosely called LIBOR hereafter) were seen as the best proxies for *risk-free* rates.

From the 2007-2008 financial crisis onwards, the spreads between swaps rates with different tenors started to widen, partly due to the increased reticence of banks to lend to each other. Today, LIBOR is no longer considered as a proxy for *risk-free* rates, and market operators have increasingly started to use Overnight Interest Swaps (OIS) discounting (see [HW12] for example).

Comparatively in the European Insurance market, throughout the quantitative impact studies (the QIS) leading to Solvency II, the questions of *risk-free* term structure construction for valuation have been tackled for years by the CEIOPS and later by the EIOPA (see [CC10] for example). The difficulty in defining a fundamental *risk-free* rate for the insurance market, mainly arises from the fact that a pure market *risk-free* rate could introduce a lot of unwanted market volatility into the insurer's balance sheet. Hence, this discount curve has been adjusted with different spreads through the QIS, and until its most recent specification, making it somewhat, less *consistent* with the market.

As of June 2015 (see [EIO15]), the term structure of discount factors for insurers' liability cash-flows is indeed derived from LIBOR EUR swap (IRS hereafter) rates, as the market for vanilla swaps is considered as 'Deep, Liquid, and Transparent' (the DLT assumption). A credit risk adjustment (CRA) is prescribed by the directive, consisting in a parallel shift applied to LIBOR swap rates. The parallel shift shall not be lower than -35bps or greater than -10 bps. Furthermore, a matching adjustment and a volatility adjustment are other optional parallel shifts which could be applied to the constructed curve.

The volatility adjustment is designed to be used in case of a crisis, causing the widening of sovereign or corporate bonds spreads. On the other hand, the matching adjustment is used in cases where the liabilities are predictable, that is, almost *perfectly* backed. In this paper, we focus on discount curve construction. The matching premium and the volatility adjustment are not further discussed.

Beyond the data and curve adjustments concerns, and considering curve construction methods, [AB13] distinguish between two types of methods: *best fit* methods, and *exact fit* methods. Best fit methods, such as [NS87] and [Sve94] are widely used by central banks. Exact fit methods such as cubic splines methods on the other hand, generally have at least as much parameters as input market products, and provide an exact fit to market data.

While the latter type of methods would be adapted for no arbitrage pricing and trading, the former type are useful for forecasting the yield curve in real world probability (see [DL06] for example). They fit the curve parsimoniously with a few parameters; in an attempt to mimic the factors explaining the variance of the yield curve changes (see [Lit+91] for details). There is another class of models, which combine the idea of using a factors structure, which is the absence of dynamic arbitrages in the curve diffusion, see [Chr+11] for example.

The extrapolation of the constructed curve is also an important subject matter for insurers and pension funds. Indeed, some of their liability cash-flows may have very long maturities, spanning beyond the longest liquid maturities available for market-quoted instruments. The question is, how would spot rates for such long maturities be determined?

As of 2016 in Solvency II, the construction and extrapolation of the swap curve is made by using the Smith-Wilson method described in
[SW01] and in the technical specifications [EIO15]. The Smith-Wilson method constructs the swap curve by exactly fitting the market IRS rates adjusted from a CRA. After a chosen maturity - the last liquid point (LLP), equal to 20 years -, the forward rate is forced by regulatory rules, to converge at an exogenously specified speed to a fixed long term level called the Ultimate Forward Rate (UFR). The UFR is derived as the sum of expected Euro inflation and expected real rates. As of 2016, it is equal to $4.2\%$.

For discount curve construction and extrapolation, we propose a method which relies on closed-form formulas for discount factors available in exogenous (or no arbitrage) short-rate model. It could be both an *exact fit* and *best fit* method, depending on the data at hand, and on how the curve is calibrated to these data. In this framework, the time-varying function ensuring an exact fit to market implied discount factors in exogenous short-rate models is considered to be a piecewise constant function, whose steps become model's parameters. The interpolation of the curve at dates comprised between quoted maturities directly comes from the properties of the model. Pseudo-discount curves can also be constructed in a dual curve environment, by using our method, along with the techniques described for example in [Whi12] and [AB13].

The static discount curve calibrated to market data can then be extrapolated to longer, unobserved maturities, with the forward rates converging to an *ultimate forward rate*. Extrapolation is done by using the same model that the one used for interpolation. On this particular point, our model is hence closer to the [SW01] model than to models which use different methods for interpolation and extrapolation (such as cubic splines for interpolation, and a modified version of [NS87] for extrapolation). We describe ways to either derive an UFR from the data, or to constraint the model to converge to a given UFR.

When it comes to forecasting and/or simulation, if one is interested in no arbitrage pricing, then she can use simulations under a risk neutral probability of the corresponding, consistent (in the sense of [BC99]) exogenous short-rate model. Otherwise, forecasts of the yield curve under the historical probability can be obtained by making use of a functional principal components analysis on the model parameters. Functional principal components analysis is described in [RD91] and [RS05]. It has been applied to forecasting mortality rates by [HU07], and in finance, it has been applied for example in [**Benko2007Functional**].

The advantage of the model presented in this paper, is that, it reconciles in some sense models like [DL06] or [SW01]. Its direct link with an exogenous short-rate model (consistency in the sense of [BC99]) and the possibility of achieving an *exact fit* to swap data means that it could be used as an input for pricing in a risk neutral probability. In addition, although it could be less interpretable than models like [DL06] (in terms of level, slope, and curvature), it could also be used for forecasting the yield curve parsimoniously in historical probability, as demonstrated in sections 1.4.3 and 1.4.4.

In the next sections, we describe the model proposed for discount curve construction and extrapolation, and explain how it could be calibrated to market data. Then, we explain how to obtain forecast of the discount curve, by using the model's parameters. To finish, some numerical examples based on [HW06], [And07], [AP10], [AB13] are presented.

## 1.2 Curve construction and extrapolation

The class of models proposed for discount curve construction and extrapolation relies on short-rate models with a time-varying mean-reversion parameter: exogenous short-rate models. In this section, we provide details on how they are derived.

In the sequel, let $Y$ denote a Lévy process and $W$ a standard brownian motion. We assume that all introduced processes are defined with respect to a stochastic basis $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{Q})$. For every considered Lévy process $Y$, its cumulant function is denoted by $\kappa$, i.e., $\kappa(\theta) = \log \mathbb{E}\left[e^{\theta Y_1}\right]$. As a matter of example, some cumulant functions are given in Table 1.1 for the Brownian motion and for two class of Lévy subordinators parametrized by a single variable $\lambda$ which inversely controls the jump size of the Lévy process. We refer the reader to [**Cont2003**] for more details on Lévy processes. We assume that, under a risk-neutral probability measure $\mathbb{Q}$, the short-term interest

| | Lévy measure | Cumulant |
|---|---|---|
| Brownian motion | $\rho(dx) = 0$ | $\kappa(\theta) = \frac{\theta^2}{2}$ |
| Gamma process | $\rho(dx) = \frac{e^{-\lambda x}}{x} 1_{x>0} dx$ | $\kappa(\theta) = -\log\left(1 - \frac{\theta}{\lambda}\right)$ |
| Inverse Gaussian process | $\rho(dx) = \frac{1}{\sqrt{2\pi x^3}} \exp\left(-\frac{1}{2}\lambda^2 x\right) 1_{x>0} dx$ | $\kappa(\theta) = \lambda - \sqrt{\lambda^2 - 2\theta}$ |

**Tab. 1.1:** Examples of Lévy measures and cumulants

rate is either governed by an extended Lévy-driven Ornstein-Uhlenbeck process (Lévy-driven OU)

$$dX_t = a(b(t) - X_t)dt + \sigma dY_{ct}, \tag{1.1}$$

or an extended CIR process

$$dX_t = a(b(t) - X_t)dt + \sigma\sqrt{X_t}dW_t, \tag{1.2}$$

where the long-term mean parameter $b$ is assumed to be a deterministic function of time, $a$ is a positive parameter which controls the speed of mean-reversion and $\sigma$ is a positive volatility parameter. Concerning the Lévy-driven OU specification 1.1, we use an additional positive parameter $c$ which appears as an increasing change of time $t \to ct$. This parameter can also be interpreted as a volatility parameter but, contrarily to $\sigma$, it controls jump frequency (an increase of $c$ makes the underlying Lévy process jumps more frequently). Let $X_0$ be the value at time $t_0$ of the process $X$. The use of Lévy processes as a driver of short rate or default intensity dynamics stems from the fact that processes driven by some Lévy processes could provide better fit on time series of bond returns than when driven by a Brownian motion. For more details on term structure modeling with Lévy processes, the reader is referred for instance to [**Cariboni2004**; **Crepey2012**; **Eberlein1999**; **Kluge2005**].

**Remark:** Specification 1.1 corresponds to a Lévy Hull-White extended Vasicek model. However, in the seminal Hull-White approach (see [HW90]), the initial term-structure is given as a model input and the function $b$ is defined in such a way that the input term-structure is reproduced by the model. In our approach, contrarily to the Hull and White framework, the deterministic function $b$ is *directly* calibrated on market quotes of interest-rate products.

Under the previous short-rate models, arbitrage-free prices of zero-coupon bonds can be expressed analytically.

### 1.2.1 Curve explicit analytical expressions

We rely on a standard pricing framework where, in absence of arbitrage opportunity, the value at time $t_0$ of a default-free zero-coupon bond with maturity time $t$ is given by

$$P(t_0, t) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_{t_0}^{t} X_u du\right) \mid \mathcal{F}_{t_0}\right], \tag{1.3}$$

where $\mathcal{F}$ is the natural filtration of the short-rate process $X$.

When the mean-reverting level $b$ is a deterministic function of time, the following proposition, which is a classical result in the theory of affine term-structure models, gives an analytical expression for $P(t_0, t)$ in the class of Lévy-driven OU models.

**Proposition:** In the Lévy-driven OU model (1.1), the value at time $t_0$ of a default-free zero-coupon bond with maturity $t$ is given by

$$P(t_0, t) = \exp\left(-\phi(t - t_0)X_0 - a\int_{t_0}^{t} b(u)\phi(t-u)du - c\psi(t-t_0)\right) \tag{1.4}$$

where the functions $\phi$ and $\psi$ are defined by

$$\phi(s) := \frac{1}{a}\left(1 - e^{-as}\right), \tag{1.5}$$

$$\psi(s) := -\int_{0}^{s} \kappa\left(-\sigma\phi(s-\theta)\right) d\theta. \tag{1.6}$$

***Proof.*** Using Itô's lemma, the Lévy-driven OU process is such that, for any $t > t_0$

$$X_t = e^{-a(t-t_0)}X_0 + a\int_{t_0}^{t} b(\theta)e^{-a(t-\theta)}d\theta + \sigma\int_{t_0}^{t} e^{-a(t-\theta)}dY_{c\theta}. \tag{1.7}$$

and, using (1.1) and (1.7), the integral $\int_{t_0}^{t} X_u du$ can be reformulated as

$$\int_{t_0}^{t} X_u du = \phi(t-t_0)X_0 + a\int_{t_0}^{t} b(u)\phi(t-u)du + \sigma\int_{t_0}^{t}\phi(t-u)dY_{cu}. \tag{1.8}$$

Expression (1.4) is obtained from (1.3) and (1.8) and by using Lemma 3.1 in [**Eberlein1999**]. □

**Remark:** Note that the function $\phi$ does not depend on the Lévy process specification. Moreover, for most Lévy processes, the integral of the cumulant transform in (1.6) has no simple closed-form solution but can be easily computed numerically. The reader is referred to [**Hainaut2007**] for examples of Lévy processes for which the function $\psi$ defined by (1.6) admits a closed-form expression.

Similar analytical expressions are available under model specification 1.2 where the underlying short-rate process follows an extended CIR process with deterministic long-term mean parameter $b$.**proposition** In the extended CIR model (1.2), the value at time $t_0$ of a zero-coupon bond with maturity $t$ is given by

$$P(t_0, t) = \exp\left(-X_0\varphi(t-t_0) - a\int_{t_0}^{t}\varphi(t-u)b(u)du\right) \tag{1.9}$$

where $\varphi$ is given by

$$\varphi(s) := \frac{2(1-e^{-hs})}{h+a+(h-a)e^{-hs}} \tag{1.10}$$

and $h := \sqrt{a^2+2\sigma^2}$.

***Proof.*** For any maturity date $t$, thanks to the Feynman-Kac formula, the function $\tilde{P}$ defined for any $u$ such that $t_0 \leq u \leq t$ by

$$\tilde{P}(u,x) := \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_u^t X_u du\right) \mid X_u = x\right]$$

is solution of the following PDE

$$\frac{\partial \tilde{P}(u,x)}{\partial u} + a\left(b(u)-x\right)\frac{\partial \tilde{P}(u,t)}{\partial x} + \frac{1}{2}\sigma^2 x \frac{\partial^2 \tilde{P}(u,t)}{\partial x^2} - \tilde{P}(u,t)x = 0, \tag{1.11}$$

with the final condition $\tilde{P}(t,x) = 1$, for all $x$. It is straightforward to check that the function $\tilde{P}$ defined by

$$\tilde{P}(u,x) = \exp\left(-x\varphi(t-u) - a\int_u^t \varphi(t-s)b(s)ds\right)$$

with $\varphi$ given by (1.10) is solution of PDE (1.11).

□

Depending on the chosen term-structure model, Proposition 1.2.1 or Proposition 1.2.1 can be used to compute present values of market instruments involved in the curve construction.

### 1.2.2 Piecewise-constant long-term mean parameter $b(t)$

Typically, in order to obtain an *exact* fit in the Hull-White extended Vasicek model (that is, a model from class 1.1 with $c = 1$ and a Brownian motion as Lévy driver), we have to choose:

$$b(t) = \frac{1}{a}\frac{df^M}{dt}(0,t) - f^M(0,t) + \frac{1}{2}\frac{\sigma^2}{a^3}\left(1-e^{-2at}\right) \tag{1.12}$$

where $t \mapsto f^M(0,t)$ are the market implied instantaneous forward rates. In our framework, the market-implied discount curve $P^M(0,t)$ at date $t_0 = 0$ is constructed by considering a piecewise constant function $t \mapsto b(t)$, whose steps are derived from vanilla (IRS) or overnight swaps (OIS) cash-flows. We let $T_1, \dots, T_n$, be the maturities of market quoted IRS, with Credit Risk Adjustment (CRA), or OIS. We assume that the function $t \mapsto b(t)$ is piecewise-constant, with:

$$b(t) = b_i, \quad \text{for } T_{i-1} \le t < T_i, \quad i = 1, \dots, n \tag{1.13}$$

$$b(t) = b_{n+1}, \quad \text{for } t \ge T_n \tag{1.14}$$

and $T_0 = t_0 = 0$.

Under this specification of the long-term mean parameter, closed-form formulas can be obtained for the discount factors. Under model class 1.1 and given that $t_0 = 0$, the integral term in equation (1.4) becomes

$$I_{n+1}(t) = \sum_{k=1}^{n} b_k \left(\xi(t - T_{k-1} \wedge t) - \xi(t - T_k \wedge t)\right) \tag{1.15}$$

for any $t \le T_n$ and

$$I_{n+1}(t) = \sum_{k=1}^{n} b_k \left(\xi(t - T_{k-1}) - \xi(t - T_k)\right) + b_{n+1}\xi(t - T_n) \tag{1.16}$$

for $t > T_n$ where $\xi$ is defined as

$$\xi(s) := s - \phi(s), \; s \ge 0. \tag{1.17}$$

Under model class 1.2 and given that $t_0 = 0$, the integral term in equation (1.9) becomes

$$I_{n+1}(t) = \sum_{k=1}^{n} b_k \left(\eta(t - T_{k-1} \wedge t) - \eta(t - T_k \wedge t)\right) \tag{1.18}$$

for any $t \le T_n$ and

$$I_{n+1}(t) = \sum_{k=1}^{n} b_k \left(\eta(t - T_{k-1}) - \eta(t - T_k)\right) + b_{n+1}\eta(t - T_n) \tag{1.19}$$

for $t > T_n$ where $\eta$ is defined as

$$\eta(s) := 2\left[\frac{s}{h+a} + \frac{1}{\sigma^2}\log\frac{h + a + (h-a)e^{-hs}}{2h}\right]. \tag{1.20}$$

The previous result can also be found in [**Bielecki2014**] under a more general form. [**Schlogl2000**] also consider an extended CIR model with piecewise-constant parameter in order to construct initial yield-curves but prices of zero-coupon bonds are given under a recursive form.

Recall that our aim is to construct a discount curve by fitting the mean-reversion function $b$ on quoted swaps observed for different standard maturities. The interpolation of the curve at intermediary dates between quoted swaps maturities directly comes from the properties of the model. Contrarily to the Hull-White approach where an exogenous term-structure is given as a model input, in our approach, the fitted risk-neutral short-rate model is, by construction, consistent in the sense of [BC99].

In the next section, we explain how model parameters can be calibrated in different situations, for the construction of OIS and IRS (with CRA) discount curves. We focus our presentation on the particular extended Vasicek short-rate model

$$dX_t = a(b(t) - X_t)dt + \sigma dW_t. \tag{1.21}$$

This model belongs to class 1.1 with $c = 1$ and with a Brownian motion as Lévy driver. However, the presented calibration method can easily be adapted to other mean-reverting models of class 1.1 and 1.2. Pseudo-discount curves could also be constructed in a dual curve environment, by using this method along with the techniques described for example in [AB13] and [Whi12].

### 1.2.3 Calibration of the model

This section is about the calibration of our model. Section 1.2.3 discusses the calibration of the liquid part of the curve to swaps, 1.2.3 is about calibration by using swaps, caps and swaptions, and 1.2.3 describes curve extrapolation. Section 1.3 describes another way of choosing the parameters, by applying cross-validation to forecasts under the historical measure.

#### Calibration of the liquid part

Considering that there are $N$ quoted swaps used for constructing the discount curve as of today, and at most $M$ coupon payment dates for all the swaps, we let $V$ be the vector of current values for the market swaps with length equal to $N$. $C$ is the $N \times M$ matrix containing in each row, the swaps' coupon payments. $P$ is the vector of discount factors that we are trying to derive, having a length equal to $M$.

Three methods might be envisaged for calibrating the model, depending on the data at hand:

- A method to be used **if an exact fit is required**, and can be found. That is, if we require $V = CP$
- A method to be used **if an exact fit cannot necessarily be found**, but approximated
- A method to be used **when the dataset is noisy, and a smooth curve is required**

These three methods are described hereafter, and numerical examples can be found in section 1.4.

- **If an exact fit is required**, then it is possible to guess *reasonable* values for $a$ and $\sigma$ (say, $a$ between $0.05$ and $1$, and $\sigma$ between $1\%$ and $5\%$), and use an iterative curve calibration (also known as *bootstrapping*, but different from statistical bootstrap resampling) to solve $V = CP$. On figure 1.1, we can observe that the discount rates obtained for 1000 values of $a \in [0.1, 10]$ and $\sigma \in [0, 0.1]$ do not exhibit particular differences at quoted swaps maturities. The corresponding forward rates in figure 1.2 exhibit more differences.

  This type of method was used for any vanilla swap before the 2007 crisis, no matter its tenor. It is relevant only for extracting discount factors from OIS which are considered to be perfectly collateralized, or in Solvency II context. As of 2016, single curve construction in Solvency II, is applied to IRS, along with a parallel CRA, comprised between 10bps and 35bps.

  In order to describe the curve's calibration procedure, we will use a formulation similar to the one in [AP10]. We let $T_1 < \ldots < T_n$ be the maturity dates of OIS or IRS minus CRA, with the same currency on both legs. The swap payment dates occur at dates $t_j$, with a frequency belonging to $\{1\,month,\ 3\,months,\ 6\,months,\ 1\,year\}$.

  The single curve construction, in the specific Hull & White-consistent case treated in this paper, is made as follows:

  1. Guess $a$ and $\sigma$: any *reasonable* values for $a$ and $\sigma$ will produce an *exact* fit for discount factors and discount rates (cf. figure 1.1)

2. *Loop on i:* At each step $T_i$ corresponding to the $i^{th}$ input swap maturity, suppose that the discount factors and $b_j$s are known for any $t_j < T_i$

3. Make a guess for $b_i$

4. Use results from section 1.2.2, to derive the discount factors at intermediate swap payment dates: $T_{i-1} \leq t_j \leq T_i$. No interpolation is required.

5. Calculate $V_i$, the value of the $i^{th}$ swap. While $V_i \neq 0$ return to point 2. Typically, the points 3 to 5 are solved iteratively with a root search algorithm.

- **If no solution is available for equation** $V = CP$ by iterative curve calibration, then similarly to [And07], it is possible to search for $P$ minimizing:

$$\frac{1}{2N}\left(V - CP\right)^T W^2 \left(V - CP\right) \tag{1.22}$$

$W$, a diagonal matrix of weights is used. These weights are based on inverse duration, such as proposed by [Bli97]; with elements:

$$w_j = \frac{1/d_j}{\sum_{j=1}^{N} 1/d_j} \tag{1.23}$$

Weights such as $w_j$'s are commonly used to give more importance to the short end of the curve, which is hence fitted more accurately. But other weighting schemes might be envisaged.

- **If the swaps data are noisy, or if one is interested in fitting smoothly noisy bonds data** a third method could be envisaged. It consists in penalizing the possibly large changes in forward rates' (approximate) second derivatives and/or in $b_i$s. The objective function to be minimized is:

$$\frac{1}{2N}\left(V - CP\right)^T W^2 \left(V - CP\right) + \lambda_1 \sum_{i=1}^{N} (f_i^{''} - f_{i-1}^{''})^2 + \lambda_2 \sum_{i=1}^{N} (b_i - b_{i-1})^2 \tag{1.24}$$

where $f_i^{''}$ is the approximate second derivative (using finite differences) of the discrete forward rates at time $T_i$. Typically, $\lambda_1$ and $\lambda_2$ can be found by cross-validation. An example can be found in section 1.4.

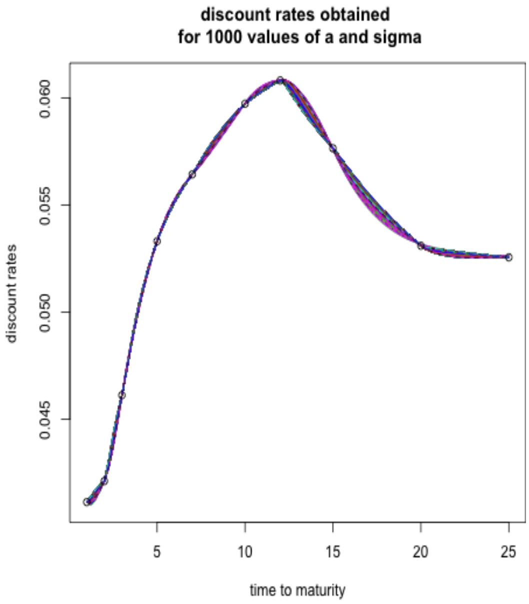


Fig. 1.1: Discount rates obtained for 1000 values of $a \in [0.1, 10]$ and $\sigma \in [0, 0.1]$

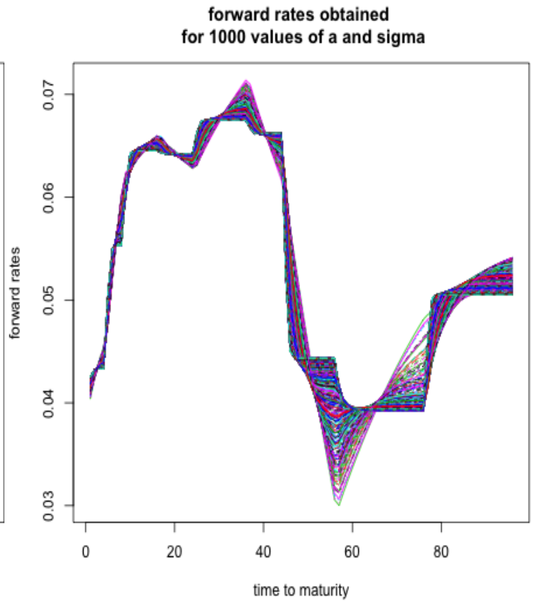


Fig. 1.2: Forward rates obtained for 1000 values of $a \in [0.1, 10]$ and $\sigma \in [0, 0.1]$

### Calibration using derivatives

Another way for picking $a$ and $\sigma$ might be to calibrate the underlying short-rate model to a set of caps and swaptions. The optimization procedure would involve the following steps: choosing $a$ and $\sigma$, construct the initial curve with an *exact fit* using the results described in section 1.2.3; use it as an input for theoretical caps and swaptions prices formulas implied by the underlying short-rate model, until $a$ and $\sigma$ which minimize the difference between theoretical and market prices for caps and swaptions are found.

### Curve extrapolation

Using the Hull and White extended Vasicek model, it is possible to derive the instantaneous forward rates from the discount factors formula. We can write:

$$f(0,t) = -\frac{\partial log(P(0,t)}{\partial t} = X_0 e^{-at} + a \int_0^t e^{-a(t-u)} b(u) du - \frac{\sigma^2}{2} \phi^2(t) \qquad (1.25)$$

Hence, in our framework, using the fact that $t \mapsto b(t)$ is piecewise constant, we can also write:

$$f^M(0,t) = X_0 e^{-at} + a \sum_{i=1}^{n} b_i \left[\phi(t - T_{i-1} \wedge t) - \phi(t - T_i \wedge t)\right] + a b_{n+1} \phi(t - T_n \wedge t) - \frac{\sigma^2}{2}\phi^2(t) \tag{1.26}$$

This formula directly provides an input for the simulation of Hull & White short-rate, with parameters $a$, $\sigma$ and $b_1, \ldots, b_n$ previously calibrated to market data.

Hence, let $t$ grow to $\infty$, we have:

$$f^M(0,\infty) = b_{n+1} - \frac{\sigma^2}{2a^2} \tag{1.27}$$

If we assume that the UFR is exogenously chosen, and denote it by $f_\infty$, we are able to derive the parameter $b_{n+1}$ as:

$$b_{n+1} = f_\infty + \frac{\sigma^2}{2a^2} \tag{1.28}$$

This enables to re-write equation (1.19), when extrapolation is required, as:

$$I_{n+1}(t) = \sum_{k=1}^{n} b_k \left(\xi(t - T_{k-1} \wedge t) - \xi(t - T_k \wedge t)\right) + \left(f_\infty + \frac{\sigma^2}{2a^2}\right)\xi(t - T_n \wedge t) \tag{1.29}$$

If **a fixed *ultimate forward rate* (UFR) is defined exogenously**, one can increase or decrease the parameter $a$, to achieve a convergence of $f^M(0,t)$ to $f_\infty$ at a pre-specified maturity. A period of convergence $\tau_{cv}$ after the *Last Liquid Point* (LLP) is defined. Starting from a low value such as $a = 0.1$, $a$ is increased until:

$$f^M(0, LLP + \tau_{cv}) = f_\infty$$

or

$$|f^M(0, LLP + \tau_{cv}) - f_\infty| < tol$$

for a given $\sigma$, and a given numerical tolerance $tol$.

Otherwise, an ***ultimate forward rate* (UFR) can be derived from market data**. A static discount curve is fitted to a fraction of the quoted swaps available, called the *training* set. After the construction of the curve on this fraction of the data, we evaluate how well, when extrapolated to a given exogenous UFR, it would price the remaining swaps in a *test* set.

The LLP provided by the prudential authority (as of 2016, a maturity 20 years), could be used to define the frontier between the *training* and *test* set. Otherwise, one can define a percentage of the swaps data to be used as a *training* dataset, for example 80% or 90% of the available swaps.

Both of these methods for curve extrapolation are applied in the numerical examples, in section 1.4.

## 1.3 Forecasting with Functional PCA

The idea that a few principal components explain a major part of the changes in bonds returns originates from [Lit+91]. This idea is now well accepted and applied to yield curve forecasting; the interested reader could refer to [DL06] or [Chr+11] for example.

We use a similar rationale, but apply it somewhat differently. The changes in the swap curve over time, are explained by the changes observed in the calibrated parameters $b_i$s over time. Considering the fact that our model for fitting each cross section of yields is already *overparametrized* (as it uses at least as much parameters as swap rates available in the input dataset), the use of models such as an unrestricted Vector Autoregressive (VAR) to predict the $b_i$s could lead to poor forecasts, with high variance.

Functional Principal Components Analysis in the spirit of [RD91] and [RS05], and more precisely Functional Principal Components Regression, was hence seen as one of the most immediate candidate to achieve a reduction of the problem's dimension. This method is used for example by [HU07] for forecasting log mortality rates. It has also been applied in finance, for example in [**Benko2007Functional**].

We consider functional data of the form:

$$b_x^{a,\sigma}(t) \tag{1.30}$$

These are the parameters $b_i$s obtained by fitting each cross section of swap rates; observed at increasing times $t \in \{t_1, \ldots, t_N\}$, for increasing maturities $x \in \{x_1, \ldots, x_p\}$.

The calibration method is the one described in section 1.2.3, with $a$ and $\sigma$ kept fixed over time.

### Finding the Functional Principal Components

Using the approach described in details in [RS05], we let $\mathbf{B}$ be the matrix containing at line $i$ and column $j$:

$$\mathbf{B}_{i,j} = b_{x_j}^{a,\sigma}(t_i) \tag{1.31}$$

With $i = 1, \ldots, N$ and $j = 1, \ldots, n$, $n > p$. For each cross section of $b_i$s calibrated at time $t_i$, a cubic spline interpolation is applied to $x \mapsto b_x^{a,\sigma}(t_i)$, so that the $b_i$s values are now equally spaced on a larger grid of maturities spanning $[x_1, x_p]$. Let $w$ be the fixed interpolation step applied to $x \mapsto b_x^{a,\sigma}(t_i)$ on $[x_1, x_p]$, and:

$$\mathbf{V} = \frac{1}{N}\mathbf{B}^T\mathbf{B} \tag{1.32}$$

$\mathbf{V}$ is the covariance matrix of the $b_i$'s, when we consider that the columns of $\mathbf{B}$ have been centered. We are then looking for the vectors $\xi^{a,\sigma}$, the (approximate) functional principal components, verifying:

$$w\mathbf{V}\xi^{a,\sigma} = \rho\xi^{a,\sigma} \tag{1.33}$$

This is equivalent to searching the eigenvalues and eigenvectors of $\mathbf{V}$, so that:

$$\mathbf{V}u = \lambda u \tag{1.34}$$

and $\rho = w\lambda$. This problem of finding eigenvalues and eigenvectors of $\mathbf{V}$ is typically solved by using the Singular Value Decomposition (SVD) of $\mathbf{B}$, and taking the normalized right singular vectors as functional principal components. The interested reader can refer to [**jolliffe2002principal**] and [RS05] for details. Another interesting resource on Functional Principal Component Analysis is [**shang2014survey**].

### Forecasting using Principal Components regression

Having obtained the functional principal components, a least squares regression of the cross sections of $b_i$s is carried out. The $b_i$s are expressed as a linear combination of the previously constructed functional principal components, plus an error term:

$$\forall t \in \{t_1, \ldots, t_N\}, \; b_x^{a,\sigma}(t) = \beta_{t,0} + \sum_{k=1}^{K} \beta_{t,k}\xi_k^{a,\sigma}(x) + \epsilon_t(x) \tag{1.35}$$

$K$ is the number of functional principal components. These functional principal components are not highly correlated by construction, so that we can use univariate time series forecasts for each of the $K+1$ time series, and h-step ahead forecasts of the $b_i$s as:

$$\hat{b}_x^{a,\sigma}(t+h) = \hat{\beta}_{t+h|t,0} + \sum_{k=1}^{K} \hat{\beta}_{t+h|t,k} \xi_k^{a,\sigma}(x) \tag{1.36}$$

Once the forecasts $\hat{b}_x^{a,\sigma}(t+h)$ are obtained, they can be plugged into formulae from section 1.2.2 to deduce h-step ahead forecasts for the discount factors and discount rates.

For choosing *good* values for $a$, $\sigma$ and $K$, we typically used a cross-validation on grids of values for these three parameters, and rolling origin estimation/forecasting, as described in section 1.4.

## 1.4 Numerical examples

In order to illustrate how the methods described in the previous sections work, we use IRS and OIS data from [And07], [AP10], [AB13], an example of bonds data from [HW06]; *a curve where all cubic splines produce negative forward rates*. For forecasting the curves, we use market EUR 6M IRS data, (from which we give detailed summaries) with a CRA adjustment equal to 10bps.

For the data from [AP10], we assume that the swaps cash-flows payments occur on an annual basis as for OIS. From [AB13], we consider mid quotes from Eonia OIS and 6-month Euribor IRS as of December 11, 2012. These data sets are all reproduced in the appendices.

In section 1.4.1, four calibration methods are tested to illustrate section 1.2.3. The method proposed in this paper [1] is denoted by CMN. It is compared to two iterative curve calibration methods, with linear (LIN) and natural cubic splines (SPL) interpolation on missing dates, and the [SW01] method (SW). Section 1.4.1 also illustrates 1.2.3. We use a dataset from [And07]; a direct *bootstrapping* without regularization produces wiggly spot and forward rates. The effects of the regularization of approximate second derivative for forward rates and calibrated $b_i$'s is illustrated. Such a regularization could also be applied to noisy bonds data.

[1] Actually applied to Hull & White model, but which can be applied to other short-rate models.

In section 1.4.1, the interpolation method is tested on a *curve where all cubic methods produce negative forward rates*, from [HW06]. Section 1.4.2 illustrates the possible extrapolation methods described in section 1.2.3. 1.4.3 illustrates the curves' forecasting method introduced in section 1.3.

The discount factors usually display no particular subtleties, so they are deliberately omitted. We present discount rates and discrete forwards instead, and the discrete forwards are taken to be 3-month forward rates.

### 1.4.1 Curve calibration

#### On swaps data from [AP10]

Below on figures 1.3, 1.4, 1.5, 1.6, are the discount rates and discrete forwards obtained for the four methods described in the previous section; two *bootstrapping* methods, with linear (LIN) and natural cubic splines (SPL) interpolation on missing dates, the [SW01] method (SW), and the method described in 1.2.3 with an exact fit, denoted as CMN. The discount rates are presented as a dashed line, and the forward rates as a plain colored line.

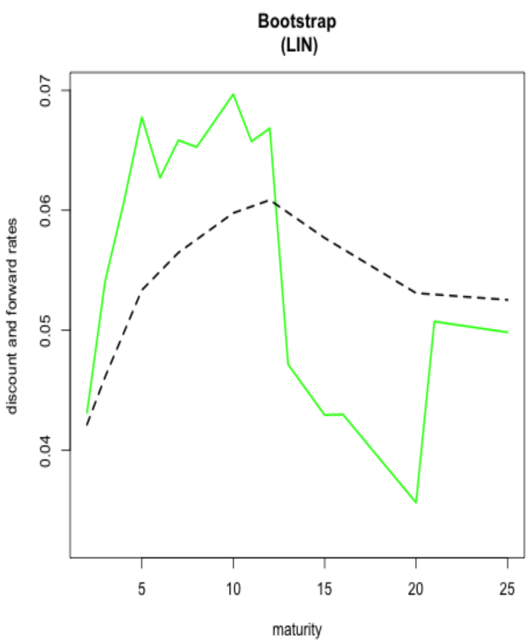


**Fig. 1.3:** *Bootstrapping* with linear interpolation

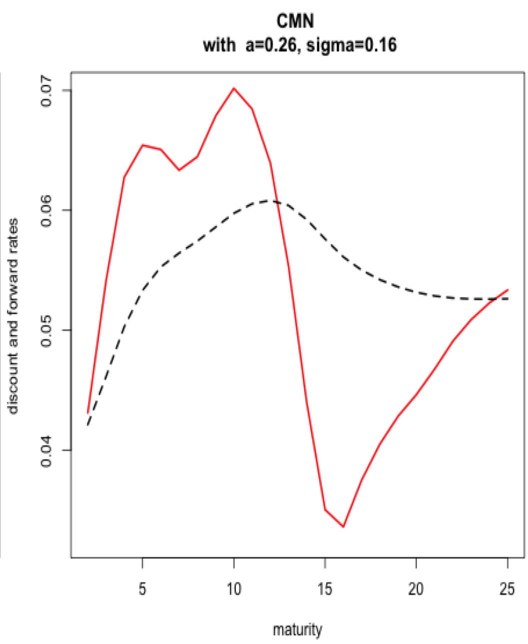


**Fig. 1.4:** CMN applied to Hull and White extended Vasicek

As demonstrated on figures 1.3, 1.4, 1.5 and 1.6, the discount rates produced by the four methods are quite similar. The discrete forward rates better exhibit the differences between them.

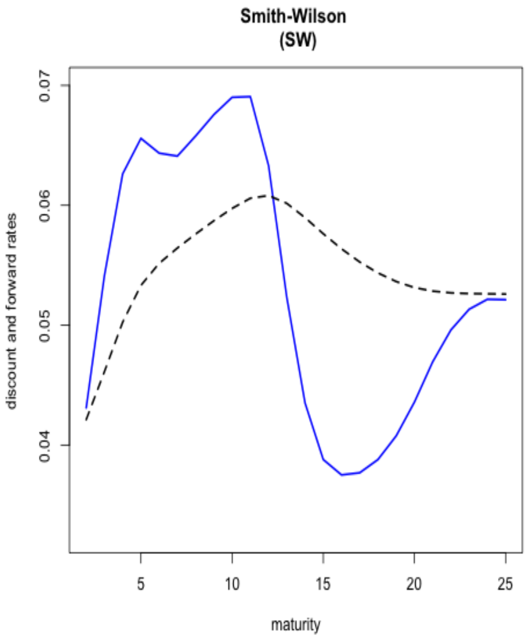


Fig. 1.5: Smith-Wilson method

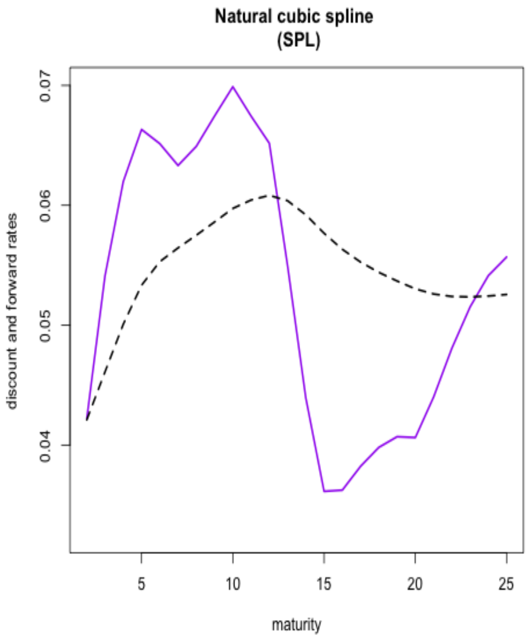


Fig. 1.6: Natural cubic spline

Curve construction with linear interpolation between quoted swaps maturities (on figure 1.3), produces a saw-tooth like forward curve, which might not be desirable, and the other methods produce more regular forward curves.

For the method described in this paper - denoted as CMN on figure 1.4 - and applied to the extended Vasicek model, the discrete forwards (with an exact fit required, as described in section 1.2.3) reflect the fact that the discount factors' construction relies on a piece-wise constant function, with slight changes in first derivatives at quoted swap maturities. This effect remains very reasonable however, as the discrete forward curve is highly similar to those produced by the other models, and doesn't exhibit large changes at quoted swap maturities.

With $a$=0.2557, $\sigma$=0.1636, the parameters $b_i$s from table 1.2 are obtained. They are presented along with the parameters $\xi_i$s obtained by the method in [SW01], with $a = 0.1$. $a = 0.1$ is actually given as default parameter by Solvency II's technical specifications, and using the notations from QIS5 technical specifications.

### On noisy swaps data

This section illustrates what may happen if the method from section 1.2.3 is applied directly to noisy data, without regularization of the parameters. We use data from [And07].

Tab. 1.2: Parameters obtained for CMN and Smith-Wilson

| Maturity | $b_i$ | $\xi_i$ |
|---|---|---|
| 1 | 0.0661 | -16.680 |
| 2 | 0.1894 | 23.556 |
| 3 | 0.2523 | -0.8413 |
| 5 | 0.2523 | -8.9116 |
| 7 | 0.2523 | 3.3552 |
| 10 | 0.2806 | 7.9600 |
| 12 | 0.2523 | -14.098 |
| 15 | 0.2089 | 3.9119 |
| 20 | 0.2553 | 3.4828 |
| 25 | 0.2616 | -1.9497 |

Figure 1.7 on the left describes the discount and forward rates obtained without regularization, with $a = 0.3655$ and $\sigma = 0.0037$. On the right, figure 1.8 describes the discount and forward rates obtained by minimizing the objective function in equation (1.24), and using the parameters $\lambda_1 = 1e-08$ and $\lambda_2 = 1e-05$, $a = 9.8891$ and $\sigma = 0.3957$.

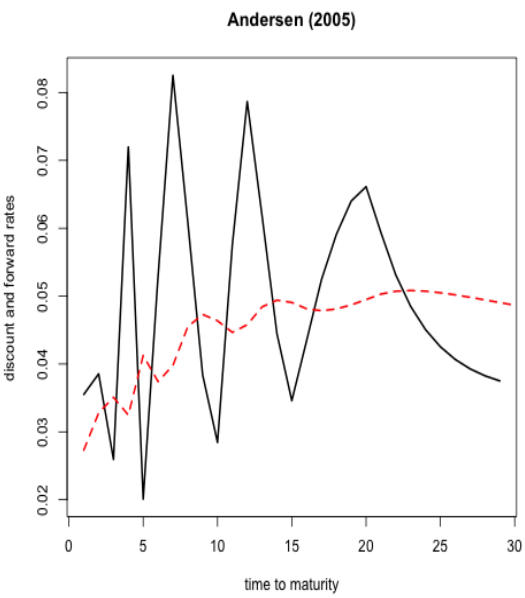


Fig. 1.7: Curve calibration without regularization

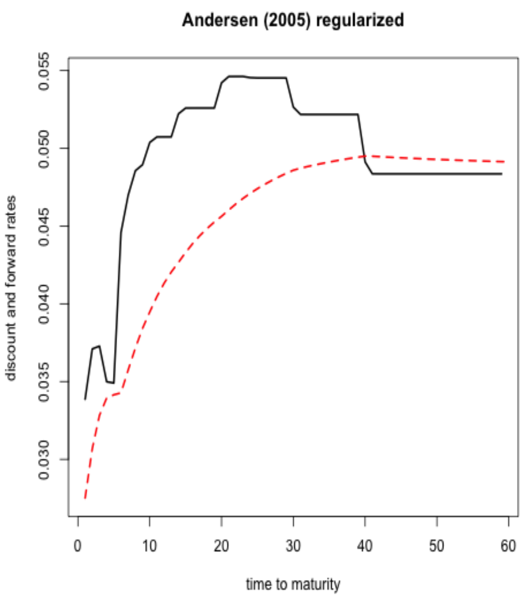


Fig. 1.8: Curve calibration with regularization

In order to pick $\lambda_1$ and $\lambda_2$, we make a grid search on couples $(\lambda_1, \lambda_2)$. For each $(\lambda_1, \lambda_2)$, a minimization based on derivatives is applied, with multiple restarts of the minimization algorithm. Multiple restarts avoid getting trapped into local minima.

Tab. 1.3: Parameters obtained for unregularized and regularized CMN

| Maturity | unregularized $b_i$ | regularized $b_i$ |
|---|---|---|
| 0.5 | 0.0253 | 0.0281 |
| 1 | 0.1100 | 0.0363 |
| 1.5 | -0.0078 | 0.0383 |
| 2 | 0.0929 | 0.0383 |
| 2.5 | -0.0005 | 0.0380 |
| 3 | -0.1360 | 0.0352 |
| 4 | 0.2901 | 0.0358 |
| 5 | 0.1975 | 0.0478 |
| 7 | 0.1654 | 0.0497 |
| 10 | -0.0056 | 0.0515 |
| 12 | 0.1315 | 0.0533 |
| 15 | 0.1392 | 0.0554 |
| 20 | 0.0688 | 0.0553 |
| 30 | 0.039 | 0.0491 |

Table 1.3 contains both the unregularized and regularized $b_i$s. The unregularized ones naturally exhibit a higher variance, because an exact fit to each swap rate in the noisy dataset is required. The regularized $b_i$s exhibit a lower variance, at the expense of a higher bias in the fitting of the data from [And07].

### On *a curve where all cubic methods produce negative forward rates*, with data from [HW06]

The dataset from this section is used in [HW06], and is described as *a curve where all cubic methods produce negative forward rates*. It is reproduced in the appendices.

Figure 1.9 illustrates the discount rates (dashed line), and discrete forward rates (plain coloured line) obtained with a linear interpolation of the bond yields. The discrete forward remain positive on all maturities, but again exhibit a sawtooth profile. As expected, the natural cubic spline on figure 1.10 produces negative discrete forward rates on this dataset.

Figures 1.11 and 1.12 present the results obtained on data from [HW06]. Figure 1.12 presents the sign of discrete forward rates as a function of $a$ and $\sigma$. We consider that discrete forward rates' sign is negative if a least one discrete forward rate is negative. We observe on both figures 1.11 and 1.12 that a low value of $a$ might produce negative forward rates on maturities comprised between $15$ and $20$. But a high value always produces positive forward rates.

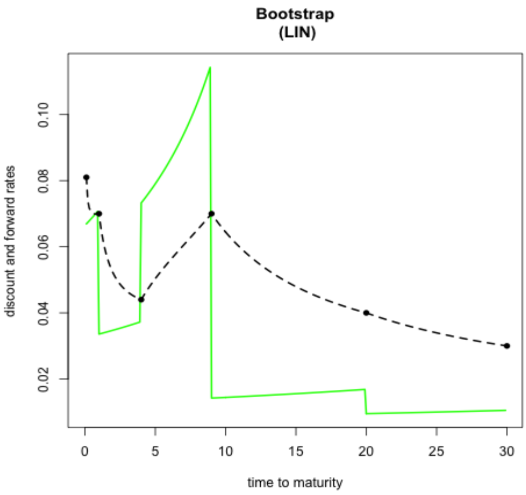


**Fig. 1.9:** Linear interpolation on *a curve where all cubic methods produce negative forward rates*

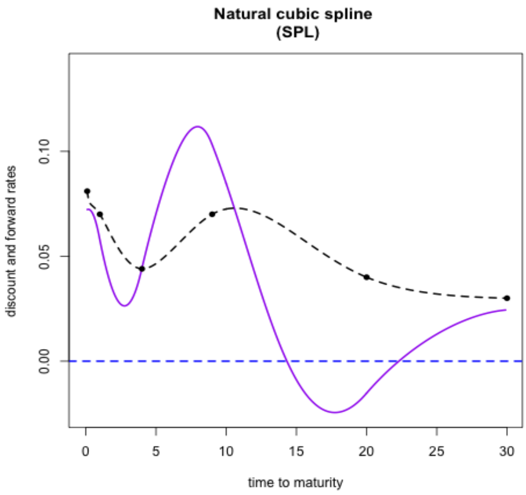


**Fig. 1.10:** Natural cubic spline interpolation on *a curve where all cubic methods produce negative forward rates*

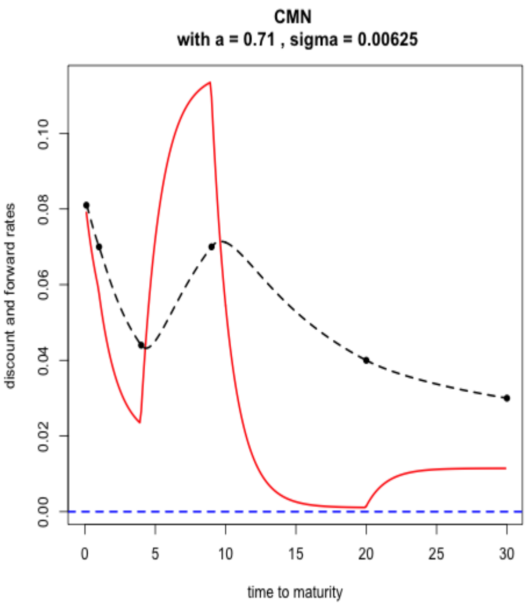


**Fig. 1.11:** CMN interpolation on *a curve where all cubic methods produce negative forward rates*

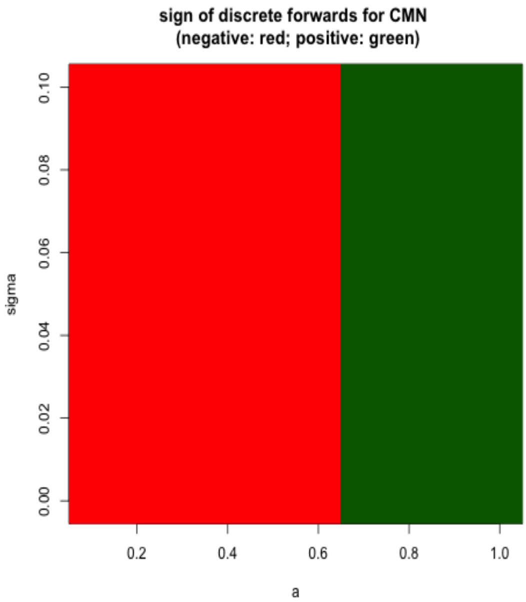


**Fig. 1.12:** Sign of discrete forwards for CMN as function of $a$ and $\sigma$, on *a curve where all cubic methods produce negative forward rates*

This is explained by what we saw in section 1.2.3: in the Hull and White extend Vasicek case, $a$ controls the speed of convergence of forward rates to the UFR: the higher the $a$, the faster the convergence of forward rates to the UFR on long-term maturities. The parameters obtained by CMN interpolation (for producing figure 1.11), with $a = 0.71$ and $\sigma = 0.0062$ are presented in table 1.4.

**Tab. 1.4:** Parameters obtained CMN with $a = 0.71$ and $\sigma = 0.0062$ on [HW06] data

| Maturity | $b_i$ |
|---|---|
| 0.1 | 0.0718 |
| 1 | 0.0351 |
| 4 | 0.0018 |
| 9 | 0.1162 |
| 20 | 0.0011 |
| 30 | 0.0114 |

### 1.4.2 Curve extrapolation on data from [AB13]

In this section, we use the extrapolation methods described in 1.2.3, on OIS and IRS (with CRA adjustment equal to $10$bps) data from [AB13].

#### With Solvency II technical specifications, on IRS + CRA

Extrapolation to a fixed UFR equal to $4.2\%$ is tested, using CMN and the Smith-Wilson method. For both methods, the Last Liquid Point (LLP) is equal to 20 years, and convergence to the UFR is forced to 40 years after the LLP.

For the CMN method, the parameters are $a = 0.174$ and $\sigma = 0.0026$, and for the Smith-Wilson method, $a = 0.125$. The resulting discount and forward curves are presented in figures 1.13 and 1.14, and the parameters $b_i$s and $\xi_i$s in table 1.5.

The discount and forward curves produced by both methods are similar, as seen on figures 1.13 and 1.14. The convergence of the Smith-Wilson method to the UFR seems to be slighty faster. This is caused by the fact that for CMN, we use instantaneous forward rates to assess the convergence to the UFR, whereas for the Smith-Wilson method, we use discrete forwards.

#### With OIS data, and a data driven UFR

For this example, we use OIS data from [AB13] presented in the appendices. A training set containing 14 swap rates ($90\%$ of the dataset) with increasing maturities starting at $1$ and ending at $20$ is made up.

**Tab. 1.5:** Parameters for CMN ($b_i$) and Smith-Wilson ($\xi_i$) extrapolation

| Maturity | $b_i$ | $\xi_i$ |
|---|---|---|
| 1 | 0.0019 | -2.5888 |
| 2 | 0.0112 | 0.7585 |
| 3 | 0.0266 | 0.1415 |
| 4 | 0.0352 | 1.3153 |
| 5 | 0.0438 | 0.4726 |
| 6 | 0.0378 | -0.8809 |
| 7 | 0.0399 | 1.2010 |
| 8 | 0.0387 | -0.8965 |
| 9 | 0.0338 | -0.3536 |
| 10 | 0.0376 | 0.7268 |
| 11 | 0.0363 | -0.1582 |
| 12 | 0.0353 | 1.2852 |
| 13 | 0.0312 | -1.9866 |
| 14 | 0.0239 | 0.4161 |
| 15 | 0.0285 | 0.7056 |
| 16 | 0.0211 | -0.7112 |
| 17 | 0.0208 | -1.7105 |
| 18 | 0.0182 | 1.9922 |
| 19 | 0.0248 | -1.5542 |
| 20 | 0.0172 | 0.5125 |
| 21 | 0.0272 | 1.0148 |
| 22 | 0.0189 | -2.1158 |
| 23 | 0.0025 | 3.4051 |
| 24 | 0.0021 | -3.7822 |
| 25 | 0.0020 | 2.7013 |
| 26 | 0.0239 | -2.8668 |
| 27 | 0.0195 | 2.2513 |
| 28 | 0.0274 | -0.8877 |
| 29 | 0.0202 | -7.1463 |
| 30 | 0.0326 | 8.5322 |

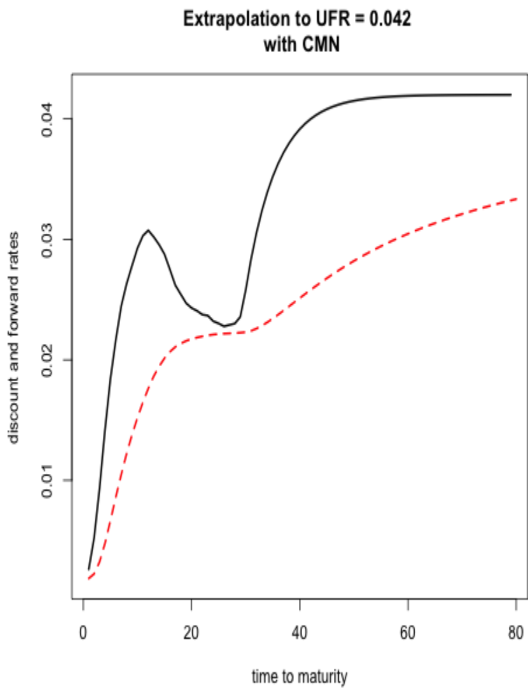


Fig. 1.13: Extrapolation to $UFR = 4.2\%$ with CMN

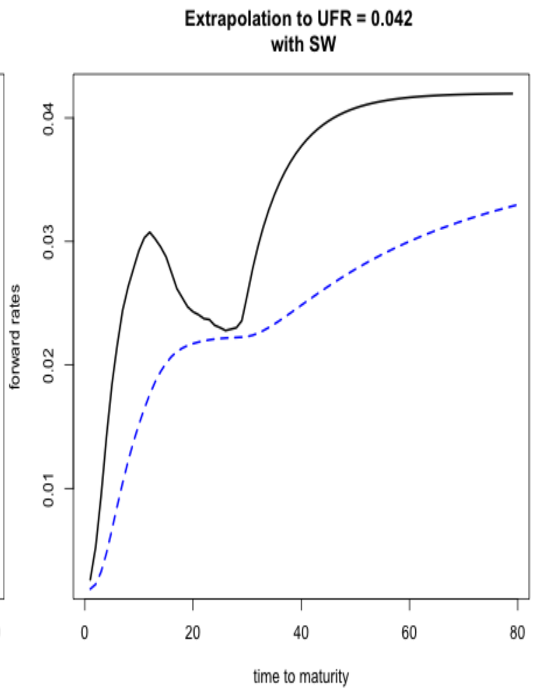


Fig. 1.14: Extrapolation to $UFR = 4.2\%$ with Smith-Wilson

This training set is used to construct the discount curve, which is then extrapolated to 30-year maturity and beyond, using different values for the UFR. The two remaining swaps, with maturities equal to 25 and 30, are placed into the test set.

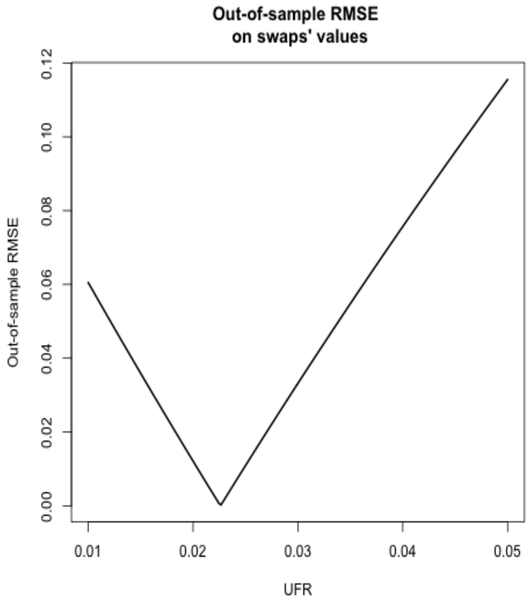


Fig. 1.15: Out-of-sample RMSE on swap values, as a function of UFR

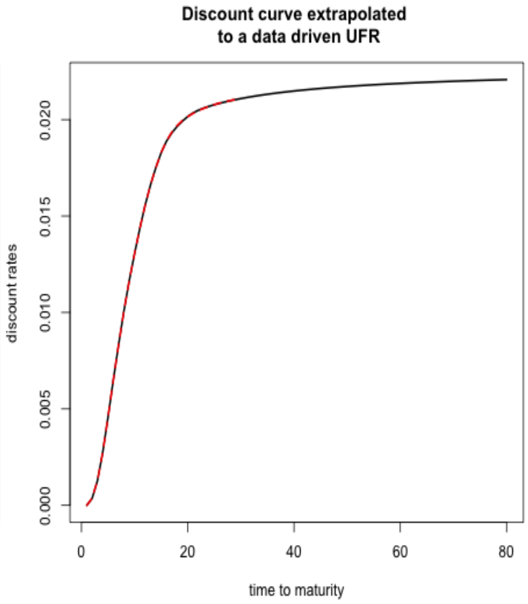


Fig. 1.16: Extrapolation of OIS curve to a data driven $UFR = 0.0226$

Figure 1.15 presents the out-of-sample RMSE obtained on swaps values from the test set, as a function of UFR. This error decreases until $UFR = 0.0226$ (notice that this value would depend on the step chosen on the grid of UFRs), and then, starts to increase again. Figure 1.16 displays the discount curve constructed on the training set, extrapolated to a 80-year maturity with an UFR equal to $0.0226$ (the one minimizing the out-of-sample RMSE on the chosen grid of UFRs) is presented.

### 1.4.3 12-months ahead forecast on historical IRS + CRA

In this section, we apply ideas from section 1.3 to real world IRS data observed monthly from december 2013 to april 2016, adjusted from a CRA equal to $10$bps.

Figure 1.17 and table 1.6 are to be read together. They contain the informations on the spot rates derived from the IRS data adjusted from a CRA, using CMN with $a = 0.3655$ and $\sigma = 0.0037$ (other values than $a = 0.3655$ and $\sigma = 0.0037$ would produce the same results as the fitting is exact for many different values of these parameters).

The static curves are generally upward sloping, and as time passes, lower and lower spot rates are encountered. In addition, negative rates are observed in table 1.6; which is coherent with the current context.

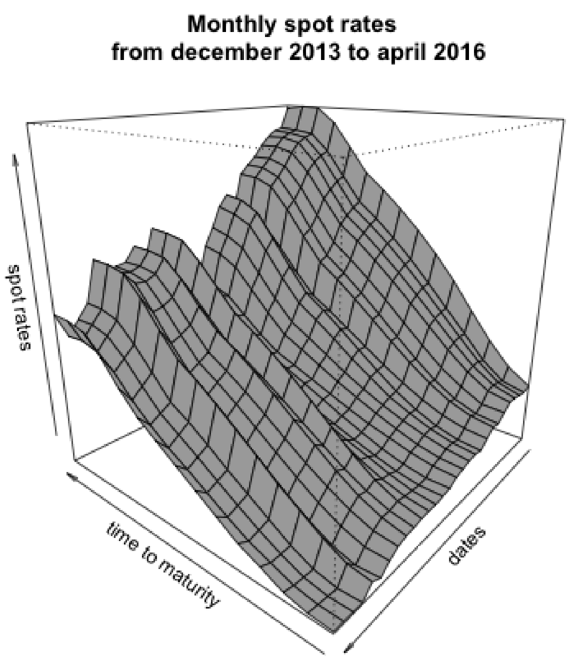


Fig. 1.17: Spot rates observed from december 2013 to april 2016

**Tab. 1.6:** Descriptive statistics for the spot rates observed from december 2013 to april 2016

| Maturity | Min. | 1st Qrt | Median | Mean | 3rd Qrt | Max. |
|---|---|---|---|---|---|---|
| 1 | -0.0026 | -0.0008 | 0.0000 | 0.0003 | 0.0019 | 0.0031 |
| 3 | -0.0023 | 0.0002 | 0.0009 | 0.0013 | 0.0028 | 0.0065 |
| 5 | -0.0008 | 0.0017 | 0.0030 | 0.0037 | 0.0056 | 0.0117 |
| 10 | 0.0046 | 0.0059 | 0.0090 | 0.0101 | 0.0132 | 0.0211 |
| 15 | 0.0063 | 0.0097 | 0.0127 | 0.0141 | 0.0179 | 0.0258 |
| 20 | 0.0069 | 0.0113 | 0.0144 | 0.0157 | 0.0199 | 0.0272 |
| 30 | 0.0071 | 0.0118 | 0.0155 | 0.0164 | 0.0208 | 0.0270 |

**Tab. 1.7:** Average out-of-sample error on real world IRS data + CRA

| Method | Parameters | Avg. OOS error |
|---|---|---|
| CMN - `auto.arima` | $K = 5, a = 1, \sigma = 0.1555$ | 0.0031 |
| CMN - `ets` | $K = 5, a = 1, \sigma = 0.2$ | 0.0037 |
| NS - `auto.arima` | $\lambda = 1.8889$ | 0.0031 |
| NS - `ets` | $\lambda = 1.8889$ | 0.0035 |
| NSS - `auto.arima` | $\lambda_1 = 21, \lambda_2 = 21$ | 0.0027 |
| NSS - `ets` | $\lambda_1 = 7, \lambda_2 = 3$ | 0.0035 |

### Benchmarking the model

Benchmarks are subjective. The one presented in this section does not aim at showing that one method is always superior to the other. It aims at showing that the method presented in this paper produces forecasts which are (more than) reasonable, and actually close to other well-known methods forecasts (on this given dataset).

Forecasts from the model presented in section 1.3 are hence compared to those of two other models constructed in the spirit of by the [DL06]. The cross sections of yields described by figure 1.17 and table 1.6 are fitted by the [NS87] model (NS), and its extension by [Sve94] (NSS). The formulas for the spot rates from these models are respectively:

$$R^M(t,T) = \beta_{t,1} + \beta_{t,2}\left[\frac{1-e^{-T/\lambda}}{T/\lambda}\right] + \beta_{t,3}\left[\frac{1-e^{-T/\lambda}}{T/\lambda} - e^{-T/\lambda}\right] \tag{1.37}$$

and

$$R^M(t,T) = \beta_{t,1} + \beta_{t,2}\left[\frac{1-e^{-T/\lambda_1}}{T/\lambda_1}\right] + \beta_{t,3}\left[\frac{1-e^{-T/\lambda_1}}{T/\lambda_1} - e^{-T/\lambda_1}\right] \quad (1.38)$$

$$+ \beta_{t,4}\left[\frac{1-e^{-T/\lambda_2}}{T/\lambda_2} - e^{-T/\lambda_2}\right] \quad (1.39)$$

Forecasts $\hat{R}^M(t+h,T)$ are obtained by fitting univariate time series to the parameters $\beta_{t,i}, i = 1, \ldots, 4$ with automatic ARIMA (`auto.arima`) and exponential smoothing (`ets`) models from [HK08]. This automatic selection is done only for the sake of the benchmarking exercise, and in order to conduct the experience in fairly similar conditions for all the methods. In practice, a visual inspection and an actual study of the univariate time series would of course be required.

For all the methods the six methods, CMN, NS, NSS with `auto.arima` and `ets`, we obtain 12-months ahead forecasts, from rolling estimation windows of a fixed 6 months length, starting in december 2013. That is, the models are trained on 6 months data, and predictions are made on 12 months data; successively. The average out-of-sample RMSE are then calculated for each method, on the whole surface of observed and forecasted yields.

The *best* parameters for CMN are obtained by cross-validation, with $K \in \{2,3,4,5,6\}$, 5 values of $a$ comprised between $0.9$ and $1$, and $10$ values of $\sigma$ comprised between $0$ and $0.2$. For NS and NSS, $\lambda_1$ and $\lambda_2$ are chosen by cross-validation, using the rolling estimation/forecasting we have just described.

### Bootstrap simulation of 12-months ahead spot rates

In this section, we use the last 12 months of the dataset to construct the functional principal components. Using 12 months as the length of the fixed window for estimation, we get an average out-of-sample RMSE of $0.0026$ (on a smaller number of testing samples than the 6 months estimation window, of course).

An AR(1) is fitted to the observed univariate time series $(\beta_{t,i})_t$, $i = 0, \ldots, K$, with $a = 1$, $\sigma = 0.0089$, and $K = 3$ chosen by cross-validation. The three functional principal components' characteristics are summarized in table 1.8. We notice that the first functional principal component explains already $99.2415\%$ of the changes in $b_i$s, and the first three functional principal components selected by cross-validation explain $99.9220\%$.

Tab. 1.8: Importance of Principal components

| Indicator | PC1 | PC2 | PC3 |
|---|---|---|---|
| Standard deviation | 0.1286 | 0.2461 | 0.2246 |
| Proportion of variance (in %) | 99.2415 | 0.5489 | 0.1315 |
| Cumulative Proportion (in %) | 99.2415 | 99.7904 | 99.9220 |

Figure 1.18 presents the autocorrelation functions of the residuals of AR(1) fitted to $(\beta_{t,i})$, $i = 0, \ldots, 3$ from april 2015 to april 2016. The residuals from AR(1) fitted to $(\beta_{t,i})_t$, $i = 1, \ldots, 3$ could be considered as stationary, but those from the AR(1) fitted to $(\beta_{t,0})_t$ seems to be closer to an AR(4).

We denote these residuals by $(\epsilon_{t,i})_t$, $i = 0, \ldots, 3$. In order to obtain simulations for the $(\beta_{t,i})_t$, $i = 0, \ldots, 3$, it is possible to use a Gaussian hypothesis on the residuals. We choose to create one thousand bootstrap resamples with replacement of the $(\epsilon_{t,i})_t$, $i = 0, \ldots, 3$ [2], denoted as $(\epsilon^*_{t,i})_t$, $i = 0, \ldots, 3$, and create new pseudo values for $(\beta_{t,i})$, $i = 0, \ldots, 3$:

$$\beta^*_{t,i} = \beta_{t,i} + \epsilon^*_{t,i},\ i = 0, \ldots, 3$$

Having done this, AR(1) forecasts $\beta^*_{t+h|t,i}$ can be obtained, in order to construct:

$$\hat{b}^{a,\sigma,*}_x(t+h) = \hat{\beta}^*_{t+h|t,0} + \sum_{k=1}^{K} \hat{\beta}^*_{t+h|t,k} \xi^{a,\sigma}_k(x) \tag{1.40}$$

The $\hat{b}^{a,\sigma,*}_x(t+h)$ can then be plugged into formulae **??** and 1.19 to deduce simulations of h-step ahead forecasts for the discount factors and discount rates.

The simulations (1000) of 12-months ahead discount rates are presented in figures 1.19 and 1.20.

### 1.4.4 6-months and 36-months ahead forecast on longer historical data

In this second example, we use interest rate swaps data from the Federal Reserve Bank of St Louis website [3] observed monthly, from july 2000 to september 2016, with maturities equal to $1, 2, 3, 4, 5, 7, 10, 30$, and a tenor equal to three months.

[2]Even if for $\epsilon_{t,0}$, considering figure 1.18, this makes a strong stationarity assumption on the residuals.
[3]Available at https://fred.stlouisfed.org/categories/32299

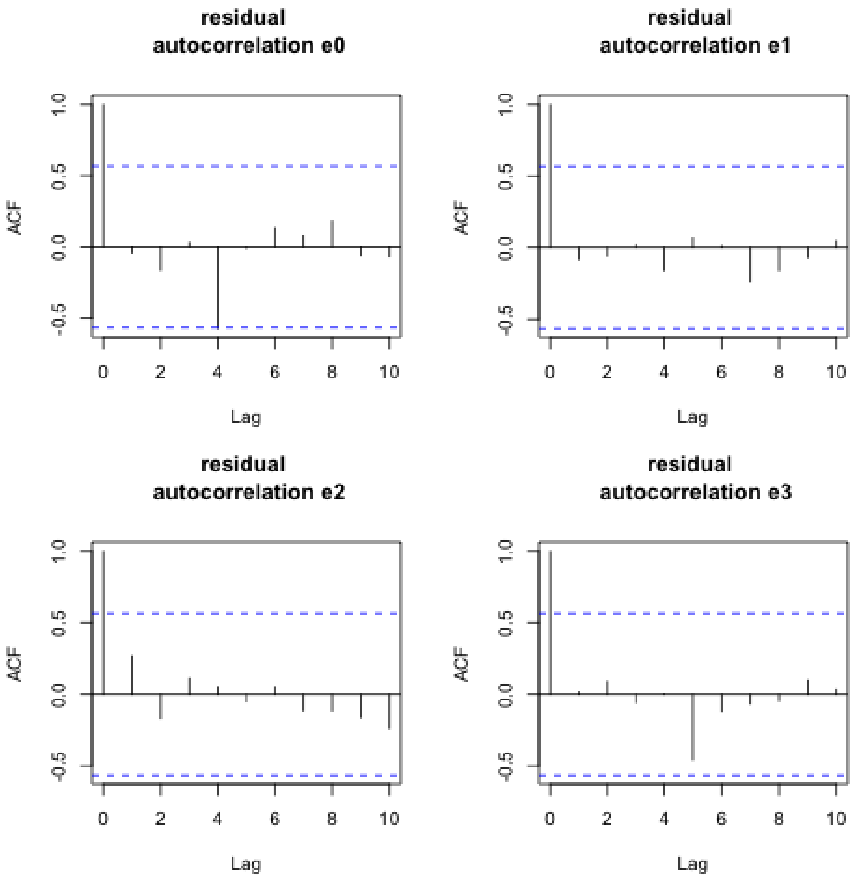


**Fig. 1.18:** Autocorrelation functions for the residuals of univariate time series(AR(1)) on $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$

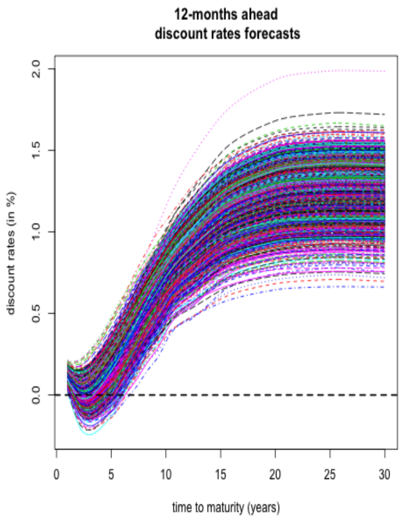


**Fig. 1.19:** Curves simulated with principal components from april 2015 to april 2016, and bootstrap ressampling of the residuals

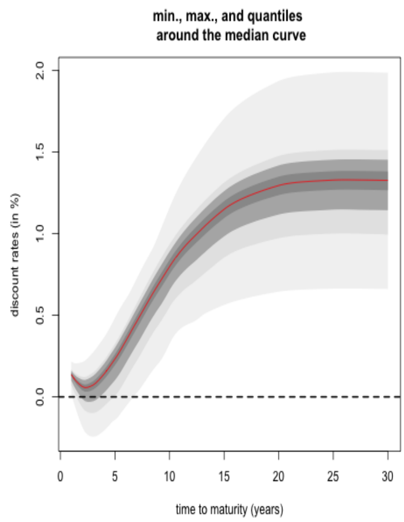


**Fig. 1.20:** Min., Max., and quartiles around the median curve for the simulations

**Tab. 1.9:** Descriptive statistics for fitted parameters $b_i$s from april 2015 to april 2016

| Maturity | Min. | 1st Qrt | Median | Mean | 3rd Qrt | Max. |
|---|---|---|---|---|---|---|
| 1 | -0.0026 | -0.0021 | -0.0010 | -0.0013 | -0.0004 | -0.0003 |
| 3 | 0.0000 | 0.0026 | 0.0040 | 0.0035 | 0.0048 | 0.0058 |
| 5 | 0.0025 | 0.0076 | 0.0030 | 0.0092 | 0.0108 | 0.0143 |
| 10 | 0.0115 | 0.0174 | 0.0090 | 0.0188 | 0.0208 | 0.0230 |
| 15 | 0.0122 | 0.0168 | 0.0127 | 0.0192 | 0.0220 | 0.0228 |
| 20 | 0.0117 | 0.0148 | 0.0144 | 0.0171 | 0.0195 | 0.0211 |
| 30 | 0.0080 | 0.0116 | 0.0155 | 0.0134 | 0.0150 | 0.0178 |

In figure 1.21, we represent the eight time series of swap rates, observed for each maturity $1, 2, 3, 4, 5, 7, 10, 30$, between july 2000 and september 2016. The swap rates for different maturities generally exhibit a decreasing trend, and are nearly equal to 0 by the end of 2016 for the shortest maturities.

Starting in 2006, the spreads between swap rates with different maturities start to narrow, until the end of 2007, and swap rates for short maturities are relatively high. This is the period corresponding to the Liquidity and Credit Crunch 2007-2008. Table 1.10 below presents the descriptive statistics for the data.

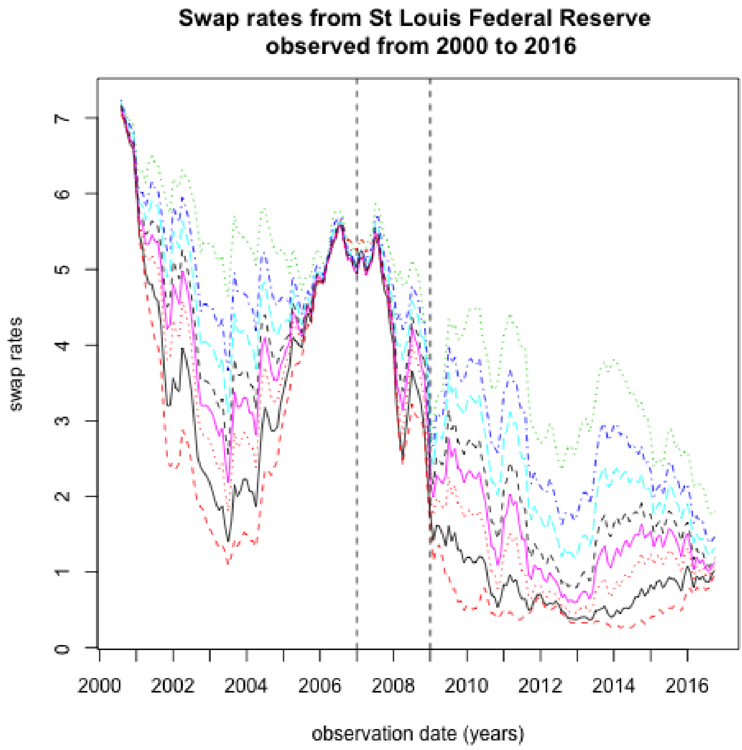


**Fig. 1.21:** Swap rates data (in %) from St Louis Federal Reserve Bank, at maturities $1, 2, 3, 4, 5, 7, 10, 30$

**Tab. 1.10:** Descriptive statistics for St Louis Federal Reserve data

| Maturity | Min. | 1st Qrt | Median | Mean | 3rd Qrt | Max. |
|---|---|---|---|---|---|---|
| 1 | 0.0026 | 0.0050 | 0.0134 | 0.0211 | 0.0336 | 0.0705 |
| 2 | 0.0037 | 0.0078 | 0.0182 | 0.0239 | 0.0390 | 0.0712 |
| 3 | 0.0046 | 0.0108 | 0.0236 | 0.0269 | 0.0422 | 0.0714 |
| 4 | 0.0060 | 0.0134 | 0.0280 | 0.0296 | 0.0439 | 0.0715 |
| 5 | 0.0078 | 0.0167 | 0.0316 | 0.0319 | 0.0456 | 0.0717 |
| 7 | 0.0119 | 0.0215 | 0.0368 | 0.0354 | 0.0483 | 0.0720 |
| 10 | 0.0139 | 0.0261 | 0.0419 | 0.0388 | 0.0502 | 0.0724 |
| 30 | 0.0175 | 0.0327 | 0.0465 | 0.0440 | 0.0537 | 0.0720 |

We transformed these swap rates into zero rates by using a single curve calibration (that is, ignoring the counterparty credit risk) with linear interpolation between the maturities; one of the methods used in section 1.4.1. Then, as in the previous section, NS, NSS, CMN are used for fitting and forecasting the curves, with `auto.arima` applied to the factors.

We obtain 6-months and 36-months ahead forecasts, from rolling training/testing windows (as in the last section) with respectively, a fixed 6 and 36 months length. The average out-of-sample RMSE are then calculated for each method, on the whole set of observed and forecasted yields.

The *best* hyperparameters - associated with the lowest out-of-sample average RMSE - for each model are obtained through a search on a grid of values. For a 6-months horizon, they are (using the notations from section 1.4.3):

- NS: $\lambda = 1.6042$
- NSS: $\lambda_1 = 1.6250$ $\lambda_2 = 1.6250$
- CMN: $a = 177.8279$, $\sigma = 3.9473e - 04$, $K = 6$

and for a 36-months horizon:

- NS: $\lambda = 1.4271$
- NSS: $\lambda_1 = 1.575$ $\lambda_2 = 1.575$
- CMN: $a = 14.6780$, $\sigma = 0.0011$, $K = 4$

The following results are obtained for the out-of-sample average RMSE:

Tab. 1.11: Descriptive statistics for out-of-sample RMSE, for training window = 6 months, and testing window = 6 months

| Method | Min. | 1st Qrt | Median | Mean | 3rd Qrt | Max. | Std. Dev |
|---|---|---|---|---|---|---|---|
| NS | 0.00101 | 0.00269 | 0.00409 | 0.00481 | 0.00595 | 0.01530 | 0.00296 |
| NSS | 0.00102 | 0.00269 | 0.00411 | 0.00481 | 0.00595 | 0.01537 | 0.00296 |
| CMN | 0.00115 | 0.00256 | 0.00396 | 0.00468 | 0.00580 | 0.01600 | 0.00302 |

Tab. 1.12: Descriptive statistics for out-of-sample RMSE, for training window = 36 months, and testing window = 36 months

| Method | Min. | 1st Qrt | Median | Mean | 3rd Qrt | Max. | Std. Dev |
|---|---|---|---|---|---|---|---|
| NS | 0.00356 | 0.00703 | 0.01044 | 0.01489 | 0.01609 | 0.21500 | 0.0213 |
| NSS | 0.00300 | 0.00690 | 0.01114 | 0.01484 | 0.01689 | 0.21570 | 0.0201 |
| CMN | 0.00402 | 0.00945 | 0.01279 | 0.01452 | 0.01917 | 0.03710 | 0.0070 |

Using tables 1.11, 1.12 and figures 1.22 and 1.23, we observe that CMN give results which are close to those from NS and NSS, with a lower average out-of-sample RMSE in both cases. For a training window equal to six months, and testing window of six months, the results obtained by the three methods are pretty similar, and the same performance is observed for the three during the financial crisis. For a training and testing window of thirty six months length, CMN has a lower mean and standard deviation for out-of-sample RMSE overall, but doesn't perform the best in the period of financial crisis 2007-2009.

## 1.5 Conclusion

In this paper, we introduced a method for swap discount curve construction and extrapolation. This method relies on the closed form formulas for discount factors available in exogenous short-rate models. We presented different ways to calibrate and extrapolate the model on different data sets from the existing literature. Moreover, we showed that the model's parameters contain a certain predictive power, enabling to obtain swap curves' forecasts, with predictive distribution.

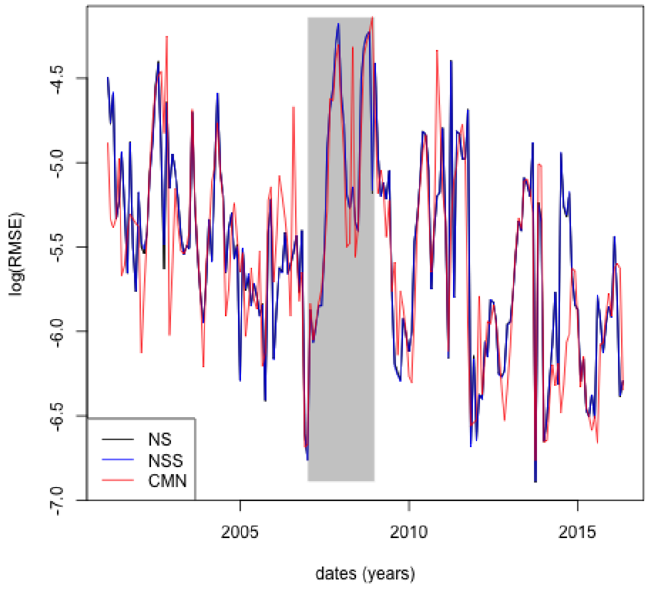


**Fig. 1.22:** log(out-of-sample RMSE) for training window = 6 months, and testing window = 6 months

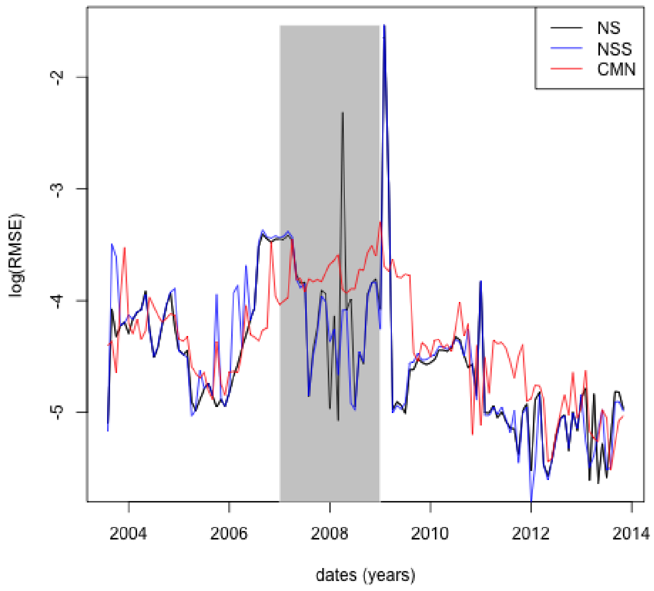


**Fig. 1.23:** log(out-of-sample RMSE) for training window = 36 months, and testing window = 36 months

## 1.6 Appendix

### 1.6.1 Data from [AP10]

| Maturity | Swap Par Rate |
|---|---|
| 1 | 4.20% |
| 2 | 4.30% |
| 3 | 4.70% |
| 5 | 5.40% |
| 7 | 5.70% |
| 10 | 6.00% |
| 12 | 6.10% |
| 15 | 5.90% |
| 20 | 5.60% |
| 25 | 5.55% |

### 1.6.2 Data from [And07]

| Maturity | Swap Par Rate |
|---|---|
| 0.5 | 2.75% |
| 1 | 3.10% |
| 1.5 | 3.30% |
| 2 | 3.43% |
| 2.5 | 3.53% |
| 3 | 3.30% |
| 4 | 3.78% |
| 5 | 3.95% |
| 7 | 4.25% |
| 10 | 4.50% |
| 12 | 4.65% |
| 15 | 4.78% |
| 20 | 4.88% |
| 30 | 4.85% |

### 1.6.3 Data from [HW06]

| Maturity | Continuous yield |
|---|---|
| 0.1 | 8.10% |
| 1 | 7.00% |
| 4 | 4.40% |
| 9 | 7.00% |
| 20 | 4.00% |
| 30 | 3.00% |

### 1.6.4 Data from [AB13]

| Maturity | EUR6M IRS | Eonia OIS |
|---|---|---|
| 1 | 0.286% | 0.000% |
| 2 | 0.324% | 0.036% |
| 3 | 0.424% | 0.127% |
| 4 | 0.576% | 0.274% |
| 5 | 0.762% | 0.456% |
| 6 | 0.954% | 0.647% |
| 7 | 1.135% | 0.827% |
| 8 | 1.303% | 0.996% |
| 9 | 1.452% | 1.147% |
| 10 | 1.584% | 1.280% |
| 11 | 1.703% | 1.404% |
| 12 | 1.809% | 1.516% |
| 13 | 1.901% | - |
| 14 | 1.976% | - |
| 15 | 2.037% | 1.764% |
| 16 | 2.086% | - |
| 17 | 2.123% | - |
| 18 | 2.150% | - |
| 19 | 2.171% | - |
| 20 | 2.187% | 1.939% |
| 21 | 2.200% | - |
| 22 | 2.211% | - |
| 23 | 2.220% | - |
| 24 | 2.228% | - |
| 25 | 2.234% | 2.003% |
| 26 | 2.239% | - |
| 27 | 2.243% | - |
| 28 | 2.247% | - |
| 29 | 2.251% | - |
| 30 | 2.256% | 2.038% |
| 35 | 2.295% | - |
| 40 | 2.348% | - |
| 50 | 2.421% | - |
| 60 | 2.463% | - |

# Bibliography


[AB13] Ferdinando M Ametrano and Marco Bianchetti. „Everything you always wanted to know about multiple interest rate curve bootstrapping but were afraid to ask“. In: *Available at SSRN 2219548* (2013) (cit. on pp. 2–4, 9, 16, 22, 36).

[And07] Leif Andersen. „Discount curve construction with tension splines“. In: *Review of Derivatives Research* 10.3 (2007), pp. 227–267 (cit. on pp. 4, 11, 16, 18, 20, 34).

[AP10] Leif BG Andersen and Vladimir V Piterbarg. *Interest rate modeling*. Atlantic Financial Press, 2010 (cit. on pp. 4, 10, 16, 17, 34).

[BC99] Tomas Björk and Bent Jesper Christensen. „Interest rate dynamics and consistent forward rate curves“. In: *Mathematical Finance* 9 (1999), pp. 323–348 (cit. on pp. 4, 9).

[Bli97] Robert R Bliss. „Testing Term Structure Estimation Methods“. In: *Advances in Futures and Options Research* 9 (1997), pp. 197–231 (cit. on p. 11).

[CC10] CFOForum and CROForum. „QIS 5 Technical Specification Risk-free interest rates“. In: (2010) (cit. on p. 2).

[Chr+11] Jens HE Christensen, Francis X Diebold, and Glenn D Rudebusch. „The affine arbitrage-free class of Nelson–Siegel term structure models“. In: *Journal of Econometrics* 164.1 (2011), pp. 4–20 (cit. on pp. 3, 14).

[DL06] Francis X Diebold and Canlin Li. „Forecasting the term structure of government bond yields“. In: *Journal of econometrics* 130.2 (2006), pp. 337–364 (cit. on pp. 2, 4, 14, 26).

[EIO15] EIOPA. „Technical documentation of the methodology to derive EIOPA's risk- free interest rate term structures“. In: *https://eiopa.europa.eu/Publications* (2015) (cit. on pp. 2, 3).

[HK08] Robin Hyndman and Yeasmin Khandakar. „Automatic time series forecasting: The forecast Package for R“. In: *Journal of Statistical Software* 27.3 (2008), pp. 1–22 (cit. on p. 27).

[HP81] J Michael Harrison and Stanley R Pliska. „Martingales and stochastic integrals in the theory of continuous trading“. In: *Stochastic processes and their applications* 11.3 (1981), pp. 215–260 (cit. on p. 1).

[HU07] Rob J Hyndman and Md Shahid Ullah. „Robust forecasting of mortality and fertility rates: a functional data approach“. In: *Computational Statistics & Data Analysis* 51.10 (2007), pp. 4942–4956 (cit. on pp. 4, 14).

[HW06] Patrick S Hagan and Graeme West. „Interpolation methods for curve construction". In: *Applied Mathematical Finance* 13.2 (2006), pp. 89–129 (cit. on pp. 4, 16, 17, 20, 22, 35).

[HW12] John Hull and Alan White. „The FVA debate". In: *Risk magazine* 25.8 (2012), pp. 83–85 (cit. on p. 2).

[HW90] J. Hull and A. White. „Pricing interest-rate derivative securities". In: *The Review of Financial Studies* 3.4 (1990), pp. 573–592 (cit. on p. 5).

[Lit+91] Robert B Litterman, José Scheinkman, and Laurence Weiss. „Volatility and the yield curve". In: *The Journal of Fixed Income* 1.1 (1991), pp. 49–53 (cit. on pp. 2, 14).

[NS87] Charles R Nelson and Andrew F Siegel. „Parsimonious modeling of yield curves". In: *Journal of business* (1987), pp. 473–489 (cit. on pp. 2, 3, 26).

[RD91] James O Ramsay and CJ Dalzell. „Some tools for functional data analysis". In: *Journal of the Royal Statistical Society. Series B (Methodological)* (1991), pp. 539–572 (cit. on pp. 4, 14).

[RS05] JO Ramsay and BW Silverman. „Springer Series in Statistics". In: *Functional data analysis*. Springer, 2005 (cit. on pp. 4, 14, 15).

[Sve94] Lars EO Svensson. *Estimating and interpreting forward interest rates: Sweden 1992-1994*. Tech. rep. National Bureau of Economic Research, 1994 (cit. on pp. 2, 26).

[SW01] A Smith and T Wilson. „Fitting yield curves with long term constraints". In: *London: Bacon and Woodrow* (2001) (cit. on pp. 3, 4, 16–18).

[Whi12] Richard White. „Multiple Curve Construction". In: *Quantitative research, OpenGamma* (2012) (cit. on pp. 3, 9).

# List of Figures

# List of Tables

## Colophon

This thesis was typeset with $\LaTeX\,2_\varepsilon$. It uses the *Clean Thesis* style developed by Ricardo Langner. The design of the *Clean Thesis* style is inspired by user guide documents from Apple Inc.

Download the *Clean Thesis* style at `http://cleanthesis.der-ric.de/`.

# Declaration

You can put your declaration here, to declare that you have completed your work solely and only with the help of the references you mentioned.

*City, August 26, 2015*

Ricardo Langner