INTEREST RATE THEORY AND CONSISTENCY PROBLEMS

by

Florina Halasan

M.Sc., Al.I.Cuza University, Romania, 2001

AN ESSAY SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES

(Department of Mathematics)

We accept this essay as conforming
to the required standard

The University of British Columbia

December 2003

Abstract

We present three different approaches to interest rate modelling, namely spot, forward and market models. Afterwards we derive general necessary and sufficient conditions for the mutual consistency of a given parameterized family of forward rate curves and the dynamics of a given interest rate model. Consistency in this context means that the interest rate model will produce forward rate curves belonging to the parameterized family. As an application, we look at Nelson–Siegel, Svensson and exponential-polynomial families and their consistency with some short rate models. The majority of these results can be found in [4].

Contents

Acknowledgements

This essay was written under the supervision of Dr. Ulrich Haussmann. I would like to thank him for his guidance, insight and support throughout.

Florina Halasan

The University of British Columbia
December 2003

Chapter 1

Introduction

1.1 Interest Rate Modelling

The modelling of the term structure of interest rates has produced a variety of approaches since the advent of arbitrage-free pricing theory and it continues to occupy the efforts of both academics and practitioners.

Unlike other asset classes (equities, foreign exchange), where the lognormal Black-Scholes framework is universally accepted, no such agreement exists with regard to interest rate modelling. One reason for this is that the phenomenon we are attempting to model - the random fluctuation of the whole yield curve - is much more complex than the movements of a single stock or index price. One can intuitively relate this to the difference in the dynamics of a scalar variable (in the case of an index) and a vector (representing the yield curve).

A second reason, that is perhaps more fundamental from a market perspective, relates to the nature of the vanilla market in interest rate derivatives. This consists of caps/floors and swaptions, which the market prices using the Black framework where the respective forward Libor and swap rate underlyings are lognormal but the discount factors are non-stochastic. Thus the market standard for the purposes of hedging must regard these vanilla instruments to be independent, where the volatility matrix for swaption prices has for the most part no bearing on the volatility curve associated with the cap/floor market. Moreover, the assumption of simultaneous lognormal behaviour in the Libor and swap rates is not mathematically easy to reconcile. Nevertheless, the goal of interest rate modelling is to provide a framework under which a large class of interest rate sensitive securities can be priced in a consistent manner.[10]

The term structure of interest rates or the yield curve can be described in variety of different ways, which are equivalent:

• Zero Coupon (or discount) bond prices $P(t,T)$: $P(t,T)$ is the value at time t of a discount bond paying one unit at time T.

• Yields: $Y(t,T) = -\log P(t,T)/(T-t)$ i.e. $P(t,T) = \exp(-Y(t,T)(T-t))$.

• Instantaneous forward rates: $f(t,T) = -\frac{\partial \log P(t,T)}{\partial T}$.

• Discrete Libor rates: $L(t,T) = \frac{(\frac{P(0,t)}{P(0,T)}-1)}{(T-t)}$.

• Spot rates: $r(t) = f(t,t)$.

However, beyond these basic identities there remains the freedom to choose in terms of "microscopic" detail (eg, specifying dynamics, volatilities and number of factors) and more importantly in terms of the bond pricing "framework". The latter is determined partly by the actual variable used to describe the model, and can be categorized into three families: spot rate, forward rate and market models. Although all three of these prescriptions are mathematically consistent (by definition of a term structure model), each approach leads to distinct development, implementation and calibration issues. Moreover, the resulting intuition gained from using the model, especially important for relating the pricing and hedging of products based on the model, is different in each case. Before the characteristic of each framework is elaborated and comparisons made between them, it is useful to first discuss what is required in any term structure model (see [10]).

Spot models (pioneered by Vasicek) attempt to describe the bond dynamics through directly modelling the short-term interest rate. Heath Jarrow and Morton (HJM) established a general framework and formulated the interest rate dynamics explicitly in terms of the continuously compounded forward rate. Market models are a class of models within the HJM framework that describe variables directly observed in the market, such as the discretely compounding Libor and swap rates.

We will consider all our term structure models defined on a filtered probability space $(\Omega, \mathcal{F}, Q, \{\mathcal{F}_t\}_{t\geq 0})$, and that our models are free from arbitrage in the sense that the probability measure Q is a martingale measure.

1.2 Spot Rate Models

The first generation of models developed were generally spot rate-based. This choice was due to a combination of mathematical convenience and tractability, or numerical ease of implementation. Furthermore, the most widely used of these models are one-factor models, in which the entire yield curve is specified by a single stochastic state variable, in this case the spot or short-term rate.

In the literature there are a large number of proposals on how to specify the Q-dynamics for the short rate, r. We present, as summarized in [3], a (far from complete) list of the most popular models. If a parameter is time dependent, this is written out explicitly. Otherwise all parameters are constant.

1. Vasicek [14]

$$dr = (b - ar)dt + \sigma dW, \quad (a > 0), \tag{1.1}$$

2. Cox-Ingersoll-Ross (CIR) [9]

$$dr = a(b - r)dt + \sigma\sqrt{r}dW, \tag{1.2}$$

3. Dothan

$$dr = ardt + \sigma rdW, \tag{1.3}$$

4. Black-Derman-Toy

$$dr = \Theta(t)rdt + \sigma(t)rdW, \tag{1.4}$$

5. Ho-Lee

$$dr = \Theta(t)dt + \sigma dW, \tag{1.5}$$

6. Hull-White (extended Vasicek)

$$dr = (\Theta(t) - a(t)r)dt + \sigma(t)dW, \quad (a(t) > 0), \tag{1.6}$$

7. Hull-White (extended CIR)

$$dr = (\Theta(t) - a(t)r)dt + \sigma\sqrt{r}dW. \quad (a(t) > 0) \tag{1.7}$$

These models are distinguished by the exact specification of the spot rate dynamics through time, in particular the form of the diffusion process, and hence the underlying distribution of the spot rate.

1.3 Forward Rate Models

An alternative approach to modelling the term structure was offered by the Heath, Jarrow and Morton [6] structure. In contrast to the spot rate approach, they model the entire yield curve as a state variable, providing conditions in a general framework that incorporates all the principles of arbitrage-free pricing and discount bond dynamics. The HJM methodology uses as the driving stochastic variable the instantaneous forward rates, the evolution of which is dependent on a specific (usually deterministic) volatility function.

Because of the relationship between the spot rate and the forward rate, $r(t) = f(t,t)$, any spot rate model is also an HJM model. In fact, any interest rate model that satisfies the principles of arbitrage-free bond dynamics must be within the HJM framework.

We now turn to the specification of the Heath–Jarrow–Morton (HJM) framework. We specify everything under the martingale measure Q.

Assumption. We assume that, for every fixed $T > 0$, the forward rate $f(\cdot, T)$ has a stochastic differential which under the equivalent martingale measure is given by

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t), \tag{1.8}$$

$$f(0,T) = f^*(0,T), \tag{1.9}$$

where W is a (d-dimensional) Q-Wienner process whereas $\alpha(\cdot, T)$ and $\sigma(\cdot, T)$ are adapted processes.

Note that we use the observed forward rate curve $\{f^*(0,T); T \geq 0\}$ as the initial condition. This will automatically give us a perfect fit between observed and theoretical bond prices at $t = 0$.

HJM drift condtition. *Under the martingale measure Q, the processes α and σ must satisfy the following relation, for every t and every $T \geq t$.*

$$\alpha(t,T) = \sigma(t,T) \int_t^T \sigma(t,s)' ds. \tag{1.10}$$

where $'$ denotes transpose.

1.3.1 The Musiela Parametrization

In many practical applications it is more natural to use time to maturity, rather than time of maturity, to parameterize bonds and forward rates. If we denote running time by t, time of maturity by T, and time to maturity by x, then we have $x = T - t$, and in terms of x the forward rates are defined as follows, [11].

Definition 1.1 For all $x \geq 0$ the forward rates $r(t, x)$ are defined by the relation

$$r(t, x) = f(t, t + x). \tag{1.11}$$

Suppose that we have now the standard HJM–type model for the forward rates under the martingale measure Q. The question is to find the Q–dynamics for $r(t, x)$, and we have the following result, known as the Musiela equation.

Proposition 1.1 (The Musiela equation) *Assume that the forward rate dynamics under Q are given by* (1.8)*. Then*

$$\begin{aligned} dr(t, x) = \Big\{ \frac{\partial}{\partial x} r(t, x) + \sigma_0(t, x) \int_0^x \sigma_0(t, u)' du \Big\} dt + \\ + \sigma_0(t, x) dW(t) \end{aligned} \tag{1.12}$$

where

$$\sigma_0(t, x) = \sigma(t, t + x).$$

Proof Using a slight variation of the Ito formula we have

$$dr(t, x) = df(t, t + x) + \frac{\partial f}{\partial T}(t, t + x)dt,$$

where the differential in the term $df(t, t + x)$ only operates on the first t. We thus obtain

$$dr(t, x) = \alpha(t, t + x)dt + \sigma(t, t + x)dW(t) + \frac{\partial}{\partial x} r(t, x)dt,$$

and, using the HJM drift condition, we obtain our result.□

The point of the Musiela parametrization is that it highlights equation (1.12) as an infinite dimensional SDE. It has become an indispensable tool of modern interest rate theory.

1.4 Market Models

The motivation for the development of market models arose from the fact that, although the HJM framework is appealing theoretically, its standard formulation is based on continuously compounded rates and is therefore fundamentally different from actual forward Libor and swap rates as traded in the market. The lognormal HJM model was also well known to exhibit unbounded behaviour (producing infinite values) in contrast to the use of lognormal Libor distribution in Blacks formula for caplets. The construction of a mathematically consistent theory of a term structure with discrete Libor rates being lognormal was achieved by Miltersen, Sandmann and Sondermann, and developed by Brace, Gatarek and Musiela (BGM), see [1]. Jamshidian developed an equivalent market model based on lognormal swap rates.

Since we will not refer to market models later on, we are not going into further details.

Chapter 2

Consistent Forward Rate Curves

2.1 Motivation

A standard procedure when dealing with concrete interest rate models on a high frequency basis can be described as follows:

1. At time $t = 0$, use market data to fit (calibrate) the model to the observed bond prices.

2. Use the calibrated model to compute prices of various interest rate derivatives.

3. The following day ($t = 1$), repeat the procedure in 1. above in order to recalibrate the model, etc..

To carry out the calibration in step 1. above, the analyst has to produce a forward rate curve $\{f^*(0,T); T \geq 0\}$ from observed data. Since only a finite number of bonds trade in the market, the data consist of a discrete set of points, and a need to fit a curve to these points arises. This curve fitting procedure can use splines or a number of parameterized families of smooth forward curves. Once the curve $\{f^*(0,T); T \geq 0\}$ has been obtained, the parameters of the interest rate model may be calibrated to this.

The problem to be studied concerns the **consistency** between, on the one hand, the dynamics of a given interest rate model, and, on the other hand, the forward curve family employed.

We say that the pair $(\mathcal{M}, \mathcal{G})$ is **consistent** if all forward curves which may be produced by the interest rate model $\mathcal{M}$ are contained within the family $\mathcal{G}$, provided that the initial forward curve is in $\mathcal{G}$. Otherwise, the pair $(\mathcal{M}, \mathcal{G})$ is inconsistent.

We now have a number of natural problems to study.

I. Given an interest rate model $\mathcal{M}$ and a family of forward curves $\mathcal{G}$, what are necessary and sufficient conditions for consistency?

II. Take as given a specific family $\mathcal{G}$ of forward curves (e.g. the Nelson–Siegel family, see 2.5.1). Does there exist any interest rate model $\mathcal{M}$ which is consistent with $\mathcal{G}$?

III. Take as given a specific interest rate model $\mathcal{M}$ (e.g. the Hull–White model). Does there exist any finitely parameterized family of forward curves $\mathcal{G}$ which is consistent with $\mathcal{M}$?

2.2 Formal Problem Statement

Let us consider ourselves under the HJMM model and assume as given a parameterized family of forward rate curves

$$G : \mathcal{Z} \to C[0, \infty), \tag{2.1}$$

with $\mathcal{Z} \subseteq R^d$ the parameter space. By slight abuse of notation we can write the curve as

$$x \longmapsto G(x; z), \tag{2.2}$$

where $x = T - t$. The main problem is to determine under which conditions an interest rate model is **consistent** with a parameterized family of forward curves.

Remark. Note that we take the volatility structure $\sigma(t, T)$ as given. This structure is then kept fixed, so the calibration procedure only concerns the choice of initial forward rate curve. Put into measure theoretic terms this means that, by fixing the volatility structure, we have fixed a class of equivalent martingale measures, and the calibration step will then pin down a particular member of this class.

Definition 2.1 (Forward curve manifold). Given a smooth mapping $G : \mathcal{Z} \to C[0, \infty)$, we define the *forward curve manifold* $\mathcal{G}$ by

$$\mathcal{G} = ImG, \tag{2.3}$$

i.e.

$$\mathcal{G} = \{G(\cdot; z) \in C[0, \infty) : z \in \mathcal{Z}\}. \tag{2.4}$$

Definition 2.2 (Invariant manifold). Let a forward curve manifold $\mathcal{G}$ and a forward rate process $r(t, x)$ be given. We say that $\mathcal{G}$ is **invariant under the action of** r if, for every fixed initial time s, the condition $r(s, \cdot) \in \mathcal{G}$ implies that $r(t, \cdot) \in \mathcal{G}, \forall t \geq s$, Q–a.s.

The main problem, following the results in [4], becomes:

Suppose that we are given
(i) a forward rate model $\mathcal{M}$, specifying a process r;
and
(ii) a forward curve manifold $\mathcal{G}$.
Is $\mathcal{G}$ then invariant under the action of r?

Thus, the pair $(\mathcal{M}, \mathcal{G})$ is consistent if and only if the manifold $\mathcal{G}$ is invariant under the action of r, and the question we pursue is when does this happen.

2.3 The Finite–Dimensional Case

In this section, in order to gain some geometrical intuition, we will analyze the finite-dimensional version of the problem.

2.3.1 The Finite–Dimensional Deterministic Case

We take as given:

• A deterministic $n-$dimensional "process" $Y : R_+ \to R^n$,

$$Y(t) = \begin{bmatrix} Y_1(t) \\ Y_2(t) \\ . \\ . \\ . \\ Y_n(t) \end{bmatrix}, \tag{2.5}$$

with differential given by

$$\frac{dY}{dt} = \mu(t, Y(t)), \tag{2.6}$$

where $\mu : R_+ \times R^n \to R^n$ is some smooth vector field.

• A smooth mapping $G : \mathcal{Z} \to R^n$,

$$G(z) = \begin{bmatrix} G_1(z) \\ G_2(z) \\ . \\ . \\ . \\ G_n(z) \end{bmatrix}. \tag{2.7}$$

The forward rate process r corresponds to the process Y, each entry in $Y(t)$ containing a certain specified x-coordinate of the infinite-dimensional vector $r(t,x)$. This is the interest rate model $\mathcal{M}$.

Now, we define the manifold $\mathcal{G}$ as

$$\mathcal{G} = \{G(z) : z \in \mathcal{Z}\}. \tag{2.8}$$

and we assume $Y(s) \in \mathcal{G}$, i.e. for some $z_0 \in \mathcal{Z}$, the initial term structure is given by $G(z_0)$, and we wish to determine when the following relation holds $Q-$a.s.

$$Y(t) \in \mathcal{G}, \forall t \geq s. \tag{2.9}$$

Geometrically, (2.9) is equivalent to

$$\frac{dY}{dt} \in T_{Y(t)}(\mathcal{G}), \forall t \geq s. \tag{2.10}$$

or to

$$\mu(t, G(z)) \in Im[DG(z)]; \tag{2.11}$$

here, $T_{Y(t)}(\mathcal{G})$ represents the tangent space to the manifold $\mathcal{G}$ at the point $Y(t)$. $DG(z)$ defines the Fréchet derivative (Jacobian) of G at z, i.e., the columns of the matrix representation of DG are the tangent vectors $\frac{\partial G(z)}{\partial z_i}$,$i = 1, ..., d$, so the tangent space $T_{Y(t)}(\mathcal{G})$ to $\mathcal{G}$ at $y = G(z)$ coincides with the image $Im[DG(z)]$.

We can conclude with the following result for the finite dimensional deterministic case:

Proposition 2.1 (Simple invariance). *The manifold $\mathcal{G}$ is invariant under the action of Y if and only if*

$$\mu(t, G(z)) \in Im[DG(z)], \tag{2.12}$$

for each $z \in \mathcal{Z}$ and each $t \geq 0$. Thus, the pair $(\mathcal{M}, \mathcal{G})$ is consistent under the indicated condition.

2.3.2 The Finite–Dimensional Controlled Case

Let us go further on to the controlled case:

$$\frac{dY}{dt} = \mu(t, Y(t)) + \sigma(t, Y(t))u(t), \tag{2.13}$$

where the deterministic control u acts as input to the system, with $u(t) \in R^m$ and σ an $n \times m$–matrix, and we are allowed to choose the control freely among some (large) class of R^m–valued functions $u(\cdot)$. This is, in fact, the standard “reachability problem“ from control theory.

Proposition 2.2 (Controlled invariance). *The manifold $\mathcal{G}$ is invariant under the action of Y if and only if*

$$\mu(t, G(z)) + \sigma(t, G(z))u \in Im[DG(z)], \forall u \in R^m, \tag{2.14}$$

for each $z \in \mathcal{Z}$ and each $t \geq 0$.

It is easily seen that (2.14) is equivalent to the conditions

$$\mu(t, G(z)) \in Im[DG(z)], \tag{2.15}$$

$$\sigma(t, G(z)) \in Im[DG(z)], \tag{2.16}$$

or to the existence of $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d \times m}$ such that

$$\mu(t, G(z)) = DG(z)\gamma(t, z) \tag{2.17}$$

$$\sigma(t, G(z)) = DG(z)\psi(t, z) \tag{2.18}$$

for every $(t, z) \in R_+ \times Z$.

2.3.3 The Finite–Dimensional Stochastic Case

We are now in the case when $Y(t)$ possesses a stochastic differential of the form

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t))dW(t) \tag{2.19}$$

For a more geometrical intuition we can write equation (2.19) as

$$\frac{dY}{dt} = \mu(t, Y(t)) + \sigma(t, Y(t))\dot{W}(t), \tag{2.20}$$

where $\dot{W}(t)$ is the “white noise”.

In order to be able to carry on with geometrical arguments we will write the analysis in terms of Stratonovich integrals instead of the Ito integrals, see Appendix A.

Proposition 2.3 (Stochastic invariance) *Assume that the process Y possesses a Stratonovich differential given by*

$$dY(t) = \mu(t, Y(t))dt + \sigma(t, Y(t)) \circ dW(t). \tag{2.21}$$

Then the manifold $\mathcal{G}$ is invariant for Y if and only if there exist mappings $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d\times m}$ such that

$$\mu(t, G(z)) = DG(z)\gamma(t, z) \tag{2.22}$$

$$\sigma(t, G(z)) = DG(z)\psi(t, z), \tag{2.23}$$

for each $t \geq 0$ and $z \in \mathcal{G}$.

2.4 Invariant Forward Rate Models

We are now back to the infinite-dimensional stochastic case of section 2.2. The Itô dynamics for the forward rate was given by:

$$\begin{aligned} dr(t,x) = \Big\{\frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du\Big\} dt + \\ + \sigma_0(t,x) dW(t) \end{aligned} \tag{2.24}$$

This gives us the following Stratonovich dynamic:

$$\begin{aligned} dr(t,x) = \Big\{\frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du\Big\} dt \\ -\frac{1}{2} d < \sigma_0(\cdot, x), W > (t) + \sigma_0(t,x) \circ dW(t) \end{aligned} \tag{2.25}$$

The quadratic variation correction can be written:

$$-\frac{1}{2} d < \sigma_0(\cdot, x), W > (t) = \varphi(t,x) dt$$

Definition 2.3 ($r-$invariance). Consider a given interest rate model $\mathcal{M}$, specifying a forward rate process $r(t,x)$, as well as a forward curve manifold $\mathcal{G}$. We say that $\mathcal{G}$ is **r−invariant** under the action of the forward rate process $r(t,x)$ if there exists a stochastic process Z with state space $\mathcal{Z}$ and possessing a Stratonovich differential of the form

$$dZ(t) = \gamma(t, Z(t))dt + \psi(t, Z(t)) \circ dW(t), \tag{2.26}$$

such that, for every fixed choice of initial time s, whenever $y(s,\cdot) \in \mathcal{G}$, the stochastic process defined by

$$y(t,x) = G(x; Z(t)), \forall t \geq s, x \geq 0, \tag{2.27}$$

solves the SDE (2.24) with initial condition $r(s,\cdot) = y(s,\cdot)$.

Theorem 2.1 (Main Theorem). *The forward curve manifold $\mathcal{G}$ is $r-$invariant for the forward rate process $r(t,x)$ in $\mathcal{M}$ if and only if*

$$G_x(\cdot; z) + \sigma_0(t,\cdot) \int_0^{\cdot} \sigma_0(t,u)' du + \varphi(t,\cdot) \in Im[G_z(\cdot; z)] \tag{2.28}$$

$$\sigma_0(t,\cdot) \in Im[G_z(\cdot; z)] \tag{2.29}$$

for all $(t,z) \in R_+ \times \mathcal{Z}$.

(2.28) is *the consistent drift condition*

(2.29) is *the consistent volatility condition.*

Here, G_z and G_x denote the Frechet derivatives of G with respect to z and x, respectively, which are assumed to exist.

Proof.(from [4]) To prove necessity we assume **r**–invariance. Then we may take the differential in equation (2.27), producing

$$dy(t,x) = G_z(x; Z(t))\gamma(t, Z(t))dt + G_z(x; Z(t))\psi(t, Z(t)) \circ dW(t).$$

Comparing this with equation (2.25) and equating coefficients yields our conclusion.

To prove sufficiency, assume that (2.28) and (2.29) hold. Then we may select $\gamma : R_+ \times \mathcal{Z} \to R^d$ and $\psi : R_+ \times \mathcal{Z} \to R^{d\times m}$ such that

$$\begin{aligned} G_x(\cdot; z) + \sigma_0(t,\cdot) \int_0^{\cdot} \sigma_0(t,s)' ds + \varphi(t,\cdot) = G_z(\cdot; z)\gamma(t,z), \\ \sigma_0(t,\cdot) = G_z(\cdot; z)\psi(t,z), \end{aligned} \tag{2.30}$$

for all $(t,z) \in R_+ \times \mathcal{Z}$. For $y(s,\cdot) \in \mathcal{G}$, we define Z as the solution to (2.26) with the initial condition $Z(s) = z_0$ and the infinite–dimensional process $y(t,x)$ by $y(t,s) = G(x; Z(t))$.

$$\begin{aligned} dy(t,x) &= G_z(x; Z(t))\gamma(t, Z(t))dt + G_z(x; Z(t))\psi(t, Z(t)) \circ dW(t) \\ &= \{G_x(x; Z(t)) + \sigma_0(t,x) \int_0^x \sigma_0(t,s)' ds + \varphi(t,s)\}dt + \sigma_0(t,x) \circ dW(t) \\ &= \Big\{\frac{\partial}{\partial x} y(t,x) + \sigma_0 \int_0^x \sigma_0(t,s)' ds + \varphi(t,s)\Big\}dt + \sigma_0(t,s) \circ dW(t). \end{aligned}$$

Thus, y solves the SDE (2.24).□

Definition 2.4 (Consistency). We say that the interest rate model $\mathcal{M}$ is **consistent** with the forward rate manifold $\mathcal{G}$ if the consistent drift and volatility conditions (2.28)–(2.29) hold.

2.5 Applications

According to a document from the Bank for International Settlements [2] the estimation methods used by some of the most important central banks are as follows:

Central bank	**Curve fitting procedure**
Belgium	Nelson–Siegel, Svenson
Canada	Svenson
Finland	Nelson–Siegel
France	Nelson–Siegel, Svenson
Germany	Svenson
Italy	Nelson–Siegel
Japan	Smoothing splines
Norway	Svenson
Spain	Nelson–Siegel (<1995), Svenson
Sweden	Svenson
UK	Svenson
USA	Smoothing splines

The shapes for the two prominent examples are:

- *The Nelson–Siegel family* (see [12]):$G_{NS}(x; z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x}$
- *The Svensson family* (see [13]):$G_S(x; z) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-z_6 x}$

In order to illustrate the previous ideas we will investigate the consistency between a concrete forward curve family and some interest rate models.

2.5.1 The Nelson–Siegel Family

The NS forward curve manifold $\mathcal{G}$ is parameterized by $z \in \mathcal{Z} = R^4$, the curve shape G being given by

$$G(x; z) = z_1 + z_2 e^{-z_4 x} + z_3 x e^{-z_4 x}. \tag{2.31}$$

For $z_4 \neq 0$ the Frechet derivatives are:

$$G_z(x; z) = [1, e^{-z_4 x}, x e^{-z_4 x}, -(z_2 + z_3 x) x e^{-z_4 x}], \tag{2.32}$$

$$G_x(x; z) = (z_3 - z_2 z_4 - z_3 z_4 x) e^{-z_4 x}. \tag{2.33}$$

For the degenerate case $z_4 = 0$:

$$G(x; z) = z_1 + z_2 + z_3 x \Longrightarrow G(x; z) = z_1 + z_3 x \tag{2.34}$$

$$G_z(x; z) = [1, x], \tag{2.35}$$

$$G_x(x; z) = z_3. \tag{2.36}$$

- *The Ho–Lee Model* (1986)

We will check if the HL model is consistent with the NS manifold.

$$dr(t) = \Theta(t)dt + \sigma dW(t), \tag{2.37}$$

$r(t) = f(t, t)$.

The HJM formulation of HL is

$$df(t, T) = \alpha(t, T)dt + \sigma dW(t). \tag{2.38}$$

By Theorem 2.1, consistency of the HL model $\mathcal{M}$ and any given manifold $\mathcal{G}$ requires that the consistent drift and volatility conditions be met.

$$G_x(x; z) + \sigma^2 x \in Im[G_z(x; z)], \tag{2.39}$$

$$\sigma \in Im[G_z(\cdot; z)]. \tag{2.40}$$

Proposition 2.4 (Nelson–Siegel and Ho–Lee).

(a) *The full Nelson–Siegel family is inconsistent with the Ho–Lee interest rate model.*

(b) *The degenerate Nelson–Siegel family $G(x; z) = z_1 + z_3 x$ is in fact consistent with Ho–Lee.*

- *The Hull–White Model* (1990)

If our the model, $\mathcal{M}$, is the Hull–White Model then

$$dr(t) = \{\Theta(t) - as(t)\}dt + \sigma dW(t), \tag{2.41}$$

where $a, \sigma > 0$.

The HJM forward rate formulation becomes

$$df(t,T) = \alpha(t,T)dt + \sigma e^{-a(T-t)}dW(t). \tag{2.42}$$

and we have the following result:

Proposition 2.5 (Nelson–Siegel and Hull–White). *The Hull–White model is inconsistent with the Nelson–Siegel family.*

We can define an augmented NS manifold $\mathcal{G}_a$ by the curve shape

$$G(x;z) = z_1 + z_2 e^{-ax} + z_3 x e^{-ax} + z_4 e^{-2ax}. \tag{2.43}$$

Proposition 2.6 (Augmented Nelson–Siegel and Hull–White). *The augmented Nelson–Siegel family* (2.43) *is consistent with the Hull–White model.*

Remark. The minimal consistent family is in fact given by

$$G(x;z) = z_1 e^{-ax} + z_2 e^{-2ax}. \tag{2.44}$$

2.5.2 The Svensson Family

The Svensson forward curve manifold $\mathcal{G}$ is parameterized by $z \in \mathcal{Z} = R^6$, the curve shape G being given by

$$G(x;z) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x e^{-z_6 x}. \tag{2.45}$$

For $z_6 \neq 0$ the Frechet derivatives are:

$$\begin{aligned} G_z(x;z) &= [1, e^{-z_5 x}, x e^{-z_5 x}, x e^{-z_5 x}, -(z_2 + z_3 x) x e^{-z_5 x}, -z_4 x^2 e^{-z_6 x}], \\ G_x(x;z) &= (z_3 - z_2 z_5 - z_3 z_5 x) e^{-z_5 x} + (z_4 - z_4 z_6 x) e^{-z_6 x}. \end{aligned} \tag{2.46}$$

Proposition 2.7 (Svensson and Ho–Lee).

(a) *The full Svensson family is inconsistent with the Ho–Lee interest rate model.*

(b) *The degenerate Svensson family* $G(x;z) = z_1 + (z_2 + z_3 x)e^{-z_5 x} + z_4 x$ *is in fact consistent with Ho–Lee.*

Proof. In order for the full Svensson family to be consistent with the Ho–Lee model we need

$$G_x(x;z) + \sigma^2 x \in Im[G_z(x;z)], \tag{2.47}$$

$$\sigma \in Im[G_z(\cdot;z)]. \tag{2.48}$$

In order for (2.47) to hold we need $z_6 = 0$ and this gives us:

$$\begin{aligned} G_z(x;z) &= [1, e^{-z_5 x}, xe^{-z_5 x}, x, -(z_2 + z_3 x)xe^{-z_5 x}], \\ G_x(x;z) &= (z_3 - z_3 z_5)e^{-z_5 x} + z_4. \end{aligned} \tag{2.49}$$

We are looking for coefficients A, B, C, D and E such that

$$(z_3 - z_3 z_5)e^{-z_5 x} + z_4 + \sigma^2 x = A + Be^{-z_5 x} + Dx - Ex(z_2 + z_3 x)e^{-z_5 x}. \tag{2.50}$$

This is verified by $A = z_4$, $C = 0$, $D = \sigma^2$ and $E = 0$.$\square$

In a similar way we can prove the following proposition.

Proposition 2.8 (Svensson and Hull–White). *The Hull–White model is inconsistent with the Svensson family.*

2.5.3 The Exponential–Polynomial Family

In this section, following [4] again, we look at r-dynamics

$$\begin{aligned} dr(t,x) = \Big\{ \frac{\partial}{\partial x} r(t,x) + \sigma_0(t,x) \int_0^x \sigma_0(t,u)' du \Big\} dt + \\ + \sigma_0(t,x) dW(t) \end{aligned} \tag{2.51}$$

with deterministic volatility.

We fix an integer K and a vector $n = (n_1, ..., n_K)$ with non–negative integers as entries.

Definition 2.5 (Exponential–polynomial family). The forward curve manifold $EP(K,n)$ is defined as the set of all curves of the form

$$G(x) = \sum_{i=1}^{K} p_i(x) e^{-\alpha_i x}, \tag{2.52}$$

where $\alpha_i \in R$ for all i, and where p_i is any polynomial with $deg(p_i) \leq n_i$ for all i.

We can write the polynomial p_i as

$$p_i(x) = \sum_{i=1}^{n_i} z_{ij} x^j,$$

and we see that it can be determined by its (n_i+1)-dimensional vector of coefficients $z_i = (z_{i0}, ... z_{in_i})$. Then $EP(K,n)$ is defined as

$$G : R^{|n|+K} \times R^K \to C[0,\infty),$$

where $|n| = \sum n_i$, and $G(x; z, \alpha)$ is given by the right hand side of (2.52), with $\alpha = (\alpha_1, ..., \alpha_K)$, and $z = (z_1, ... z_K)$.

The partial derivatives of our mapping G are

$$\frac{\partial G}{\partial z_{ij}} = x^j e^{-\alpha_i x}, \tag{2.53}$$

$$\frac{\partial G}{\partial \alpha_i} = -x p_i(x) e^{-\alpha_i x}, \tag{2.54}$$

$$\frac{\partial G}{\partial x} = \sum_{i=1}^{K} (p_i'(x) - \alpha_i p_i(x)) e^{-\alpha_i x}. \tag{2.55}$$

Lemma 2.1 *Fix (z,α) as above. Then we have*

$$g(\cdot) \in Im[G_{z,\alpha}(\cdot; z, \alpha)]$$

if and only if g is of the form

$$g(x) = \sum_{i=1}^{K} q_i(x) e^{-\alpha_i x}$$

where q_i is an arbitrary polynomial with $deg(q_i) \leq n_i + 1$.

Using the previous lemma we can prove the following:

Proposition 2.9 (Inconsistency) *No non–trivial forward rate model of the form* (2.51) *is consistent with the forward curve family* $EP(K, n)$*.*

Proof. From the consistent volatility condition of Theorem 2.1 we have $\sigma_0(t, \cdot) \in Im[G_z(\cdot; z)]$ and using Lemma 2.1 we conclude that $\sigma_0(x) = \sum_{i=1}^{K} q_i(x)e^{-\alpha_i x}$, for fixed α (and z). This can be satisfied for all choices of α if $\sigma_0 \equiv 0$.

Corollary 2.1 No non–trivial forward rate model with deterministic volatility structure is consistent with Nelson–Siegel.

We will look now at a more restricted family, $REP(K, n; \beta)$, defined as follows:

Definition 2.6 (Restricted family). Consider a fixed choice of (K, n) and a vector $\beta = (\beta_1, ..., \beta_K) \in R^K$. The restricted forward curve manifold $REP(K, n, \beta)$ is defined as the set of all curves of the form

$$G(x) = \sum_{i=1}^{K} p_i(x)e^{-\beta_i x}, \tag{2.56}$$

where p_i is any polynomial with $deg(p_i) \leq n_i$ for all i.

The family $REP(K, n, \beta)$ is defined as the mapping:

$$G : R^{|n|+K} \to C[0, \infty), \tag{2.57}$$

The dimension of the manifold is $|n| + K$ whereas, for $EP(K, n)$ is $|n| + 2K$.

Lemma 2.2 *Given* z*, we have*

$$g(\cdot) \in Im[G_z(\cdot; z)]$$

if and only if g *is of the form*

$$g(x) = \sum_{i=1}^{K} q_i(x)e^{-\beta_i x}$$

where q_i *is an arbitrary polynomial with* $deg(q_i) \leq n_i$*. Furthermore, for every* z *we have the relation*

$$G_x(\cdot; z) \in Im[G_z(\cdot; z)].$$

Proposition 2.10 (Consistency without polynomial) *Consider a fixed manifold $REP(K, n, \beta)$ with no purely polynomial part, i.e. $\beta_i > 0$, for all i. The volatility function $\sigma_0(x)$ is consistent with this manifold if and only if the following conditions hold.*

1. *The volatility function must be of the form*

$$\sigma_0(x) = \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x}, \tag{2.58}$$

with $L \leq K$ and $deg(\hat{p}_i) \leq n_i$ for all $i = 1, ..., L$.

2. *For all $i, j \in \{1, ..., L\}$ there exists an index $k \in \{1, ..., K\}$ such that*

$$\begin{aligned} \beta_i + \beta_j &= \beta_k, \\ deg(\hat{p}_i) + deg(\hat{p}_j) &\leq n_k. \end{aligned} \tag{2.59}$$

Proof. We are looking for a volatility function σ_0 for a one–factor model consistent with the given family. Again, from Theorem 2.1 and Lemma 2.2 we conclude that the volatility has to be of the form (2.58). For (2.55) we have $G_x(\cdot; z) \in Im[G_z(\cdot; z)]$, so the only thing we have to check it is if

$$\sigma_0(t, \cdot) \int_0^{\cdot} \sigma_0(t, u)' du \in Im[G_z(\cdot; z)].$$

Since $\int_0^x \sigma_0(u) du = \sum_{i=1}^{L} p_i^*(x) e^{-\beta_i x}$ with $deg(p_i^*) = deg(\hat{p}_i)$ we derive

$$\sigma_0(x) \int_0^x \sigma_0(u)' du = \sum_{i,j=1}^{L} \hat{p}_i(x) p_j^*(x) e^{(-\beta_i + \beta_j)x},$$

and $deg(\hat{p}_i p_j^*) = deg(\hat{p}_i) + deg(\hat{p}_j)$. This gives us the conclusions of the theorem.□

Corollary 2.2 Under the assumptions of Proposition 2.10, there exists a consistent volatility function σ_0 if and only if there exist two indices $i, k \leq K$ such that

$$2\beta_i = \beta_k.$$

If this condition is met, then any σ_0 of the form $\sigma_0(x) = \hat{p}_i(x) e^{-\beta_i x}$, with $deg(\hat{p}_i) \leq n_i$ and $2 \cdot deg(\hat{p}_i) \leq n_k$ is consistent.

Proposition 2.11 ("Zero–beta" consistency). *Consider a fixed family $REP(K, n, \beta)$ with a non–trivial purely polynomial part. The volatility function $\sigma_0(x)$ is consistent with this family if and only if the following conditions hold.*

1. *The volatility function must be of the form*

$$\sigma_0(x) = \hat{p}_0(x) + \sum_{i=1}^{L} \hat{p}_i(x) e^{-\beta_i x}, \tag{2.60}$$

where $deg(\hat{p}_i) \le n_i$ *for all* $i = 1, ..., L$.

2. *For all* $i, j \in \{1, ..., L\}$ *there exists an index* $k \in \{1, ..., K\}$ *such that*

$$\begin{aligned} \beta_i + \beta_j &= \beta_k, \\ deg(\hat{p}_i) + deg(\hat{p}_j) &\le n_k. \end{aligned} \tag{2.61}$$

If $\hat{p}_0(x) = 0$ *these are all the conditions. If* $\hat{p}_0(x) \ne 0$ *the following conditions must be added.*

3. *For every* $i \in \{1, ...L\}$ *it holds that*

$$deg(\hat{p}_i) + deg(\hat{p}_0) + 1 \le n_i.$$

4. *For* $i = 0$ *we have*

$$2 \cdot deg(\hat{p}_0) + 1 \le n_i.$$

Proof. 1. and 2. can be easily verified as in Proposition 2.10. As before, $deg(p_j^*) = deg(\hat{p}_j)$, $j \ne 0$ and also, $deg(p_0^*) = deg(\hat{p}_0) + 1$, $(\hat{p}_0 \ne 0)$. This gives us 3. and 4.

In this case we have the following simple test for the existence of a consistent volatility function σ_0.

Corollary 2.3 Under the assumptions of Proposition 2.11, there exists a consistent volatility function $\sigma_0(x)$ if and only if either

(i) there exist some $i, j \in \{1, ..., K\}$ such that

$$2 \cdot \beta_i = \beta_j \tag{2.62}$$

or

(ii) we have

$$n_0 \le 1. \tag{2.63}$$

Corollary 2.4 Ho–Lee is the only model with deterministic volatility consistent with affine forward rate curves.

Corollary 2.5 Hull–White is the only model with deterministic volatility consistent the augmented Nelson–Siegel forward curve family.

Until now we just looked at one–factor models. As in [4] we can combine the results for the Ho–Lee and Hull–White models into a two–factors model, the so called HJM model (not the HJM forward rate framework). Combining the HJM formulation for HL (2.38) and the one for HW (2.42) we have:

$$df(t,T) = \alpha(t,T)dt + \sigma_1 dW_1(t) + \sigma_2 e^{-a(T-t)} dW_2(t), \tag{2.64}$$

where W_1 and W_2 are independent Wiener processes.

A forward curve manifold consistent with the HJM model $\mathcal{M}$ must combine the characteristics of the degenerate and augmented NS manifolds $\mathcal{G}_0$ and $\mathcal{G}_a$. We define the linear–exponential manifold $\hat{\mathcal{G}}_a$ to be the set of curves

$$G(x;z) = z_1 + z_2 x + z_3 e^{-ax} + z_4 x e^{-ax} + z_5 e^{-2ax}. \tag{2.65}$$

The consistent volatility condition in this case requires the existence of A, B, C, D and E such that

$$\{\sigma_1, \sigma_2 e^{-ax}\} = A + Bx + Ce^{-ax} + Dxe^{-ax} Ee^{-2ax}, \tag{2.66}$$

$A, ..., E$ may be different for the two functions on the left hand side. The consistent drift condition requires that

$$(z_4 - z_3 a - z_4 ax)e^{-ax} - 2z_5 ae^{-2ax} + \sigma_1^2 x + \frac{\sigma_2^2}{a}(e^{-ax} - e^{-2ax}) \in Im[G_z(\cdot;z)],$$

i.e. the left hand side must be represented in the same form as the right hand side of (2.66), but possibly for different constants $A, ..., E$. Analyzing the conditions as before, we derive the following conclusion.

Proposition 2.12 (HJM and the linear–exponential family). *The HJM model is inconsistent with the Nelson–Siegel family (and with both the degenerate and augmented variations of it). However, the HJM model (as well as both HL and HW) is consistent with the linear–exponential family $\hat{\mathcal{G}}_a$ in* (2.65).

Proof. The proof follows easily from the consistent drift condition and the consistent volatility condition.

Remark. Similar results are obtained for the general HJMM framework of interest rates [5].

Chapter 3

Conclusions

The results gathered in this essay form a basis for further empirical research on the term structure of interest rates. We only answered the first problem stated at the beginning of Chapter 2. The second problem was studied only by examples and the third one was not approached at all.

Further results in the geometrical theory of interest rates are obtained by Tomas Björk [5], Damir Filipović [7] and Josef Teichmann [8]. Tomas Björk developed the theory in a general Hilbert space framework, using the finite dimensional realization concepts from control theory. He used the Hilbert space of real analytic functions as the space of forward rate curves. This particular space is very small and it does not even include the forward rate curves generated by Cox–Ingersoll–Ross model. The extension to a much larger space of forward rates was done by Filipović and Teichmann. In particular, they proved that any forward rate model admitting a finite dimensional realization must necessarily have an affine term structure.

Bibliography

[1] D. Gatarek A. Brace and M. Musiela. The market model of interest rate dynamics. *Math. Finance*, 7(2):127–155, 1997.

[2] BIS. Zero-coupon yield curves: Technical documentation. Technical report, Bank for International Settlements, Basle, March 1999.

[3] T. Björk. *Arbitrage Theory in Continuous Time.* Oxford University Press, 1998.

[4] T. Björk and B. J. Christensen. Interest rate dynamics and consistent forward rate curves. *Math. Finance*, 9:323–348, 1999.

[5] T. Björk and L. Svensson. On the existence of finite dimensional realizations for nonlinear forward rate models. *Working paper, Stockoholm School of Economics, sbumitted*, 1997.

[6] R. Jarrow D. Heath and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for continget claim valuation. *Econometrica*, 60:77–105, 1992.

[7] D. Filipović. *Consistency Problems for HJM Interest Rate Models.* PhD thesis, ETH Zurich, 2000.

[8] D. Filipović and J. Teichmann. Existence of finite dimensional realizations for stochastic equations. *Working paper, TU Wien*, 2000.

[9] J.Ingersoll J. Cox and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–408, 1985.

[10] H. Lee. Similarities and differences. *Interest Rate Risk*, 2000. Risk Waters Group Ltd.

[11] M. Musiela. Stochastic pdes and term structure models. *Journees Internationales de Finance, IGR-AFFI, La Baule*, 1993.

[12] C. Nelson and A. Siegel. Parsimonious modeling of yield curves. *Journal of Business*, 60:473–489, 1987.

[13] L. E. O. Svensson. Estimating and interpreting forward interest rates: Sweden 1992-1994. *IMF Working Paper No. 114*, September 1994.

[14] O. Vasicek. An equilibrium characterization of the term structure. *J. Finan. Econom.*, 5:177–188, 1977.

Appendix A

Stratonovich Integral

This appendix, based on [4], is meant to offer some information on Stratonovich integral. The main difference between Itô formula and the regular chain rule is the appearance of an extra term, also called Itô term, or drift term. A variation of the Itô integral can be obtained so that the classical chain rule is satisfied.

Definition A.1 (Stratonovich integral). For given semimartingales X and Y, the **Stratonovich integral** of X with respect to Y, is defined as

$$\int_0^t X(s) \circ dY(s) = \int_0^t X(s)dY(s) + \frac{1}{2} < X, Y >_t . \tag{A.1}$$

The first term in the right hand side is the Itô integral. In the present case, with only Wiener processes as driving noise, we can define the "quadratic variation process" $< X, Y >$ in (A.1) by

$$< X, Y >= \int_0^t dX(s)dY(s), \tag{A.2}$$

with the usual "multiplication rules":

$$dW \cdot dt = dt \cdot dt = 0, dW \cdot dW = dt.$$

The main result for the Stratonovich integral is the following:

Proposition A.1 (Chain rule). *Assume that the function $F(t, y)$ is smooth. Then we have*

$$dF(t, Y(t)) = \frac{\partial F}{\partial t}(t, Y(t))dt + \frac{\partial F}{\partial y} \circ dY(t). \tag{A.3}$$

Thus, in Stratonovich calculus, the Itô formula takes the form of the standard chain rule of ordinary calculus.

Even though the Itô integrals are martingales, thus important computational advantage, we prefer Stratonovich integrals. Stratonovich integrals behave nicely under transformations, so they are natural to use for examples in connection with the stochastic differential equations on manifolds.