# The arbitrage-free Nelson-Siegel model of the yield curve

Wanele Gcinumuzi Mamba (muzi@aims.ac.za)
African Institute for Mathematical Sciences (AIMS)

Supervised by: Dr. Raouf Ghomrasni
AIMS, South Africa

22 May 2014

*Submitted in partial fulfillment of a structured masters degree at AIMS South Africa*

# Abstract

The Nelson-Siegel term structure model is extremely popular among financial institutions. It owes its popularity to its goodness-of-fit, parsimony, and the ability to forecast well. However, despite empirical superiority, the model has been criticised by Filipović (2009) and Björk and Christensen (1999) for lacking the ability to eliminate the possibility riskless arbitrage opportunities.

The purpose of this research project is to present the arbitrage free Nelson-Siegel (AFNS) term structure model. We elaborate how Christensen et al. (2011) incorporated the dynamic Nelson-Siegel (DNS) factor loadings into the affine term structure proposed by Duffie and Kan (1996), in continuous time, to produce the AFNS. Furthermore, we show the derivation of the AFNS model in discrete time after the work of Niu and Zeng (2012). Finally, we present the AFNS derivative models that account for time variation in volatility.

## Declaration

I, the undersigned, hereby declare that the work contained in this research project is my original work, and that any work done by others or by myself previously has been acknowledged and referenced accordingly.

---

Wanele Gcinumuzi Mamba, 22 May 2014

# Contents

# 1. Introduction

The scramble to accurately predict and forecast the yield curve has led to a diversification of both parametric and non-parametric models. Parametric models derive yield curves through a single-piece function, with the parameters typically estimated through the least-squares regression technique. The most popular parametric models are the Nelson and Siegel (1987) model and its extension by Svensson (1994). Their popularity is due to their relatively easiness to estimate. Also, they give a possibility of forecasting an accurate yield curve. However, those models have limited ability to fit irregular yield curve shapes, a tendency to take extreme values at the short end, and a relatively strong co-dependence of estimates in different segments of the yield curve.

Under non-parametric models, the yield curve is deduced from piecewise polynomials, where the individual segments are joined together continuously at specific points in time (called knot points). Such methods involve selecting a set of knot points, extracting the corresponding set of spot rates, and finally interpolating, in order to obtain spot rates for a continuum of maturities . These models are flexible and they can fit ideally to most curves. However, the yield curve, resulting from the non-parametric models, is rarely capable of exactly pricing back all inputs.

In the money market, yield curves have several uses, such as:

1. financial risk management,
2. allocating portfolios,
3. determining spot prices and
4. financial assets and derivative pricing.

Many models for investigating the dynamics of the yield curves have been brought forth, but most of them are either exhaustive on paper but inconsistent with reality or vice versa. So the major challenge is to device a model that is consistent both theoretical and empirical. In estimating the yield-curve, various quotes on the interest rate market are used depending on the type of model being applied. These quotes include the spot rate, the zero-coupon rates or the forward rates.

## 1.1 Objective of this Research Project

Understanding the dynamic evolution of the yield curve is important for conducting monetary policy and pricing financial derivatives. The Nelson–Siegel model is extremely popular in practice, among both financial market practitioners and central banks. In this research project, we will present the arbitrage-free extended Nelson-Siegel term structure model (AFENS) proposed by Christensen et al. (2011). Then we will also derive the model in discrete-time following the work of Niu and Zeng (2012). Considering that models for risk management must accurately capture variability of interest rates, we present AFNS models with stochastic volatility.

## 1.2 Fundamentals of Interest Rate Modelling

**1.2.1 Bonds.** These are debt investments in which an investor lends money to an entity (corporations and governments) that borrows the funds for a particular period of time at an interest rate (coupon). Bonds are classified under fixed-income securities.

**1.2.2 Zero Coupon Bonds.** A zero coupon bond is a bond that does not have coupon payment, that is, it does not disperse regular interest payment. Instead a potential investor buys the bond at a discounted price lower than the face value. A formal definition of a zero coupon bond with a maturity time $T$ is as follows.

**Definition 1.2.3.** A *$T$-maturity zero coupon bond* is a contract that guaranties its holder the payment of one unit of currency at time $T$, with no intermediate payments. The contract value, i.e. the price of a zero coupon bond at time $t < T$ is denoted by $P(t,T)$ with $P(T,T) = 1$ for all maturities $T$ and is equal to present value of the nominal amount which can be written as

$$P(t,T) = \frac{1}{1 + \frac{d}{360} r} \tag{1.2.1}$$

where $r$ is a deterministic interest rates and $d$ is the number of days left before the maturity of the zero coupon bond.

**1.2.4 Time to maturity $(T - t)$.** It is the amount of time (in years) from the present time $t$ to the maturity time $T > t$.

**1.2.5 Fixed Coupon Bonds.** These are bonds that pays the same amount of interest for the duration of their term. Formally they can be defined as

**Definition 1.2.6.** For fixed points in time $T_0, T_1, \ldots, T_n$, where $T_0$ is the issue date of the bond, the owner the fixed coupon bond receives the deterministic coupon $C_i$ for $i := \{1, 2, \ldots, n-1\}$. At the time $T_n$ the owner receives the face value.

**1.2.7 Continuously-compounded spot interest rate.** This is the constant rate or yield, prevailing at time $t$ for the maturity $T$ denoted by $y(t,T)$, at which a zero coupon bond of $P(t,T)$ units of currency at time $t$ accrues continuously to yield a unit amount of currency at maturity. It is given by the formula

$$y(t,T) = -\frac{\log P(t,T)}{\tau(t,T)}, \tag{1.2.2}$$

where $\tau(t,T)$ is usually the time to maturity $(T - t)$.

**1.2.8 Forward Rates.** Given three fixed time points $t < S < T$, a contract at time $t$ which allows an investment of a unit amount of currency at time $S$, and gives a riskless deterministic rate of interest over the future interval $[S,T]$ is called the *The forward rate*. The yield curve can also be expressed in terms of forward rates rather than yields. A forward rate is the yield that an investor would agree to today to make an investment over a specified period in the future.

**1.2.9 Term Structure.** The term structure of interest rates is defined as the relationship between the yields of default-free pure discount (zero-coupon) bonds and their time to maturity. The term structure is not always directly observable because, with the exception of short-term treasury-bills, most of the substitutes for default-free bonds (government bonds) are not pure discount bonds. Therefore, an estimation methodology is required to derive the zero coupon yield curves from observable data. If we deal with approximations of empirical data to create yield curves it is necessary to choose suitable mathematical functions.

**1.2.10 Affine Term Structure.** Affine term structure captures the risk for long maturity bonds in an arbitrage free financial market. Affine term structure models are arbitrage free models in which bond prices are affine in nature, i.e. they can be expressed as a constant plus a linear term.

## 1.3 Admissible Affine Term Structure Models

Dai and Singleton (2000) mentioned that to eliminate arbitrage opportunities, the price $P(t,T)$, at time $t$ of a $T$ maturity zero coupon bond is given by

$$P(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s ds\right)\right], \tag{1.3.1}$$

with $\mathbb{E}^{\mathbb{Q}}$ being the expectation under risk neutral measure. They further stated that given a vector of state variables $X_t = (X_t^1, X_t^2, \ldots, X_t^n)$, the instantaneous risk free rate is an affine function of the form

$$r_t = \rho_0(t) + \rho_1'(t)X_t \tag{1.3.2}$$

where $\rho_0 : [0,T] \mapsto \mathbb{R}$ and $\rho_1 : [0,T] \mapsto \mathbb{R}^n$. The state variable $X_t$ is assumed to be a Markov process, follows the affine diffusion process of Duffie and Kan (1996) under a filtered probability space $(\Omega, \mathcal{F}, \mathcal{F}_{t\geq 0}, \mathbb{Q})$ and evolves according to the stochastic differential equation

$$dX_t = K^{\mathbb{Q}}(t)[\theta^{\mathbb{Q}}(t) - X_t]dt + \Sigma(t)\begin{pmatrix} \sqrt{\alpha_1(t) + \beta^1(t)X_t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\alpha_n(t) + \beta^n(t)X_t} \end{pmatrix} dW_t^{\mathbb{Q}}(t) \tag{1.3.3}$$

where $W_t^{\mathbb{Q}}$ is a standard Brownian motion under $\mathbb{Q}$, $K^{\mathbb{Q}}(t)$ and $\Sigma(t)$ are $n \times n$ matrices, and $\theta^{\mathbb{Q}}(t) : [0,T] \mapsto \mathbb{R}^n$ is the drift term. Also $\alpha(t)$ and $\beta(t)$ are defined as

$$\alpha(t) = \begin{pmatrix} \alpha_1(t) \\ \vdots \\ \alpha_n(t) \end{pmatrix}, \qquad \beta(t) = \begin{pmatrix} \beta_1^1(t) & \cdots & \beta_n^1(t) \\ \vdots & \ddots & \vdots \\ \beta_1^n(t) & \cdots & \beta_n^n(t) \end{pmatrix}. \tag{1.3.4}$$

The prices of zero-coupon bonds, under the Duffie and Kan (1996) framework, were shown to be exponential affine functions of $X_t$,

$$P(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s ds\right)\right] = \exp\left(A(t,T) + B(t,T)'X_t\right) \tag{1.3.5}$$

where $B(t,T)$, with dimension $n \times 1$, and $A(t,T)$ solve the ordinary differential equations,

$$\frac{dB(t,T)}{dt} = (K^{\mathbb{Q}})'B(t,T) - \frac{1}{2}\sum_{i=1}^{n}(\Sigma' B(t,T)B(t,T)'\Sigma)_{i,i}(\beta^i)' + \rho_1, \tag{1.3.6}$$

$$\frac{dA(t,T)}{dt} = -B(t,T)'K^{\mathbb{Q}}\theta^{\mathbb{Q}} - \frac{1}{2}\sum_{i=1}^{n}(\Sigma' B(t,T)B(t,T)'\Sigma)_{i,i}(\alpha_i)' + \rho_0, \tag{1.3.7}$$

with the boundary conditions, $B(T,T) = A(T,T) = 0$. From equation (1.3.5) we have that the yield curve, $y(t,T)$, is given by

$$y(t,T) = -\frac{\log P(t,T)}{T-t} = -\frac{A(t,T)}{T-t} - \frac{B(t,T)'}{T-t}X_t. \tag{1.3.8}$$

In the next chapter, we show how Christensen et al. (2011) perfomed reverse engineering to embed the factor loadings of the Nelson-Siegel yield curve into an affine term structure framework to produce the AFNS.

# 2. Nelson-Siegel Model

In this chapter, we give a brief review of the original Nelsons-Siegel model before looking at the Dynamic Nelsons-Siegel that was suggested by Diebold and Li (2006). Then, we present the affine arbitrage-free Nelson-Siegel(AFNS) term structure model that was propose by (Christensen et al., 2011). Finally, we discuss some extension models of the Nelson-Siegel yield curve.

## 2.1 Yield Curve factors

The variations in returns of all fixed-income securities can be clearly explained in terms of three factors of the yield curve namely: level, steepness and curvature. Taking into account the effects, these factors have on portfolio, investors can derive better hedge positions compared to what they can achieve from holding a zero-duration portfolio. The three factors illustrate most of the return variability across the term structure (Litterman and Scheinkman, 1991).

## 2.2 The Dynamic Nelson-Siegel Model

The Nelson-Siegel term structure model, pioneered by Nelson and Siegel (1987), is one of the most renowned parametric model. Given a set of yields $y(\tau)$, where $\tau = T - t$ is the yield to maturity, Nelson and Siegel (1987) were able to show the cross-section of yields at any point in time are described by the curve, which is expresses as,

$$y(\tau) = \beta_0 + \beta_1 \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_2 \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right), \tag{2.2.1}$$

where $\beta_0, \beta_1, \beta_2$ are model parameters that are determined from initial conditions and $\lambda$, the decay parameter, is constant. A dynamic version of the Nelson-Siegel curve (2.2.1) was suggested by Diebold and Li (2006).

$$y(\tau) = L_t + S_t \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left( \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \tag{2.2.2}$$

In their work, they were able to show that the time varying constants $L_t, S_t$ and $C_t$ respectively correspond to level, slope and curvature with regard to the Nelson-Siegel factor loadings. The Dynamic Nelson-Siegel provides a good empirical performance but it does not, like many other parametric models, impose the desirable theoretical restriction to absence of arbitrage. Christensen et al. (2011) came up with a remedy, to the problem, by deriving the arbitrage free Nelson-Siegel model.

## 2.3 Arbitrage-Free Nelson-Siegel (AFNS) Term Structure Model

The starting point of the Arbitrage-Free Nelson-Siegel (AFNS) derivation, by Christensen et al. (2011), was the standard continuous-time affine arbitrage free structure shown in section 1.3. Considering the three factor model ($n = 3$) with constant volatility matrix $\Sigma$, $\theta^{\mathbb{Q}}(t)$, $\alpha(t)$ and $\beta(t)$ taken to be the same as in the $\mathbb{A}_0(3)$ model proposed by (Dai and Singleton, 2000), they were able to prove the following proposition.

**Proposition 2.3.1.** Suppose that the instantaneous risk-free rate is defined by

$$r_t = L_t + S_t \tag{2.3.1}$$

and suppose that the state variables $X_t = (L_t, S_t, C_t)$ evolve according to the system of stochastic differential equation (SDE), below, under the risk-neutral $\mathbb{Q}$-measure

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{Q}} \\ \theta_S^{\mathbb{Q}} \\ \theta_C^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \Sigma \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0 \tag{2.3.2}$$

where $\theta^{\mathbb{Q}}$ is the mean of the latent factors and $\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}$ is the volatility matrix. We have the price of the zero-coupon given by

$$\begin{aligned} P(t,T) &= \mathbb{E}_t^{\mathbb{Q}} \left[ \exp\left( -\int_t^T r_s ds \right) \right] \\ &= \exp(B_1(t,T)L_t + B_2(t,T)S_t + B_3(t,T)C_t + A(t,T)) \end{aligned} \tag{2.3.3}$$

where $B_1(t,T)$, $B_2(t,T)$, $B_3(t,T)$ and $A(t,T)$ are solutions to the following system of ordinary differential equations (ODEs)

$$\frac{d}{dt} \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \\ B_3(t,T) \end{pmatrix} = \rho_1 + (K^{\mathbb{Q}})' B(t,T) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \\ B_3(t,T) \end{pmatrix} \tag{2.3.4}$$

and

$$\frac{\mathrm{d}A(t,T)}{dt} = -B(t,T)' K^{\mathbb{Q}} \theta^{\mathbb{Q}} - \frac{1}{2} \sum_{i=1}^{3} (\Sigma' B(t,T) B(t,T)' \Sigma)_{i,i}, \tag{2.3.5}$$

with boundary conditions given by $B_1(T,T) = B_2(T,T) = B_3(T,T) = A(T,T) = 0$. The solutions of the system ODEs (2.3.4) are

$$B_1(t,T) = -(T-t) \tag{2.3.6}$$

$$B_2(t,T) = -\frac{1 - e^{-\lambda(T-t)}}{\lambda} \tag{2.3.7}$$

$$B_3(t,T) = (T-t)e^{-\lambda(T-t)} - \frac{1 - e^{-\lambda(T-t)}}{\lambda} \tag{2.3.8}$$

with $B_1(t,T)$, $B_2(t,T)$, $B_3(t,T)$ known as factor loadings. The solution of (2.3.5) is given by

$$A(t,T) = (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_2 \int_t^T B_2(s,T) ds + (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_3 \int_t^T B_3(s,T) ds + \frac{1}{2} \sum_{i=1}^{3} \int_t^T (\Sigma' B(s,T) B(s,T)' \Sigma)_{i,i} ds \tag{2.3.9}$$

where $K^{\mathbb{Q}}$ is the mean reversion matrix. Lastly, we have the zero-coupon bond yields given by

$$y(t,T) = L_t + \left( \frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)} \right) S_t + \left( \frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)} \right) C_t - \frac{A(t,T)}{T-t}. \tag{2.3.10}$$

*Proof.* We want to show that $B(t,T)$ correspond to the factor loadings of the Nelson-Siegel yield curve model. Considering the system of ODEs for $B(t,T)$ in (1.3.6) with $\theta^{\mathbb{Q}} = 0$ and $\beta^i(t) = 0$, for $i := \{1,2,3\}$, we have

$$\frac{dB(t,T)}{dt} = \rho_1 + (K^{\mathbb{Q}})' B(t,T), \qquad B(T,T) = 0. \tag{2.3.11}$$

Rearranging equation (2.3.11) we get

$$\frac{dB(t,T)}{dt} - (K^{\mathbb{Q}})' B(t,T) = \rho_1, \qquad B(T,T) = 0. \tag{2.3.12}$$

Multiplying both the RHS and LHS by the integrating factor, $e^{(K^{\mathbb{Q}})'(T-t)}$,

$$e^{(K^{\mathbb{Q}})'(T-t)} \frac{dB(t,T)}{dt} - (K^{\mathbb{Q}})' e^{(K^{\mathbb{Q}})'(T-t)} B(t,T) = e^{(K^{\mathbb{Q}})'(T-t)} \rho_1 \tag{2.3.13}$$

$$\implies \frac{d}{dt}\left(e^{(K^{\mathbb{Q}})'(T-t)} B(t,T)\right) = e^{(K^{\mathbb{Q}})'(T-t)} \rho_1. \tag{2.3.14}$$

Integrating both sides from time $t$ to maturity time $T > t$,

$$\int_t^T \frac{d}{ds}\left\{e^{(K^{\mathbb{Q}})'(T-s)} B(s,T)\right\} ds = \int_t^T e^{(K^{\mathbb{Q}})'(T-s)} \rho_1 ds, \tag{2.3.15}$$

Evaluating the integrals and using the boundary conditions we get

$$B(t,T) = -e^{-(K^{\mathbb{Q}})'(T-t)} \int_t^T e^{(K^{\mathbb{Q}})'(T-s)} \rho_1 ds. \tag{2.3.16}$$

For $(K^{\mathbb{Q}})'$ and $\rho_1$, Christensen et al. (2011) imposed the structures below,

$$(K^{\mathbb{Q}})' = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \quad \text{and} \quad \rho_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}. \tag{2.3.17}$$

Then, we have that

$$e^{(K^{\mathbb{Q}})'(T-t)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\lambda(T-t)} & 0 \\ 0 & -\lambda(T-t)e^{\lambda(T-t)} & e^{\lambda(T-t)} \end{pmatrix} \tag{2.3.18}$$

and hence

$$e^{-(K^{\mathbb{Q}})'(T-t)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda(T-t)} & 0 \\ 0 & \lambda(T-t)e^{-\lambda(T-t)} & e^{-\lambda(T-t)} \end{pmatrix}. \tag{2.3.19}$$

Now substituting (2.3.18) and (2.3.19) into (2.3.15) we get:

$$\begin{aligned} B(t,T) &= -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda(T-t)} & 0 \\ 0 & \lambda(T-t)e^{-\lambda(T-t)} & e^{-\lambda(T-t)} \end{pmatrix} \int_t^T \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\lambda(T-s)} & 0 \\ 0 & -\lambda(T-s)e^{\lambda(T-s)} & e^{\lambda(T-s)} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} ds \\ &= -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda(T-t)} & 0 \\ 0 & \lambda(T-t)e^{-\lambda(T-t)} & e^{-\lambda(T-t)} \end{pmatrix} \int_t^T \begin{pmatrix} 1 \\ e^{\lambda(T-s)} \\ -\lambda(T-s)e^{\lambda(T-s)} \end{pmatrix} ds \\ &= -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda(T-t)} & 0 \\ 0 & \lambda(T-t)e^{-\lambda(T-t)} & e^{-\lambda(T-t)} \end{pmatrix} \begin{pmatrix} T-t \\ -\frac{1-e^{\lambda(T-t)}}{\lambda} \\ -(T-t)e^{\lambda(T-t)} - \frac{1-e^{\lambda(T-t)}}{\lambda} \end{pmatrix}. \end{aligned} \tag{2.3.20}$$

Multiplying out the matrices in (2.3.20) we get

$$B(t,T) = \begin{pmatrix} -(T-t) \\ -\frac{1-e^{\lambda(T-t)}}{\lambda} \\ (T-t)e^{\lambda(T-t)} - \frac{1-e^{\lambda(T-t)}}{\lambda} \end{pmatrix}. \tag{2.3.21}$$

Thus

$$\begin{pmatrix} B_1(t,T) \\ B_2(t,T) \\ B_3(t,T) \end{pmatrix} = \begin{pmatrix} -(T-t) \\ -\frac{1-e^{\lambda(T-t)}}{\lambda} \\ (T-t)e^{\lambda(T-t)} - \frac{1-e^{\lambda(T-t)}}{\lambda} \end{pmatrix} \tag{2.3.22}$$

as proposed. Therefore

$$y(t,T) = -\frac{\log P(t,T)}{T-t} \tag{2.3.23}$$

$$= -\frac{B(t,T)'}{T-t}X_t - \frac{A(t,T)}{T-t}$$

$$= -\frac{1}{T-t}\begin{pmatrix} -(T-t) & -\frac{1-e^{\lambda(T-t)}}{\lambda} & (T-t)e^{\lambda(T-t)} - \frac{1-e^{\lambda(T-t)}}{\lambda} \end{pmatrix}\begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} - \frac{A(t,T)}{T-t} \tag{2.3.24}$$

which simplifies to (2.3.10). □

Proposition 2.3.1 has the following implications.

- The state variables, $X_t = (L_t, S_t, C_t)$, are Gauss-Markov processes. Also, comparing the Dynamic Nelson-Siegel model (2.2.1) with (2.3.10), we note that the instantaneous interest rate, $r_t$, is function of the level and slope of the yield curve with curvature factor serving as a stochastic time-varying mean of the slope factor. Unfortunately, having $L_t$ and $S_t$ following a Gaussian process means $r_t$ can have negative values.
- The parameter $\lambda$, curvature and slope factors mean-reversion rate, should be considered constant.
- The AFNS dynamics only has a notable structure under the risk-neutral probability measure $\mathbb{Q}$, but not under a $\mathbb{P}$-measure. Hence there are an infinite number of possibilities that can be considered.
- The yields have a normal distribution, hence the volatility is constant, which is not a desirable property for model used in risk management.

**2.3.2 Yield Adjustment Term.** Worth noting is that the factor loadings of the yield functions (2.2.1) and (2.3.10) are identical. The major difference between the two functions is that arbitrage-free model contains an additional maturity dependant term $-\frac{A(t,T)}{T-t}$, known as the *yield adjustment term*. It is of the form

$$-\frac{A(t,T)}{T-t} = -\frac{1}{2}\frac{1}{T-t}\sum_{i=1}^{3}\int_t^T (\Sigma' B(s,T)B(s,T)'\Sigma)_{i,i}ds, \tag{2.3.25}$$

and the explicit solution can be obtained only when the mean parameters of the state variables, $X_t$ under the equivalent martingale measure $\mathbb{Q}$ is zero i.e. $\theta^{\mathbb{Q}} = 0$. Thus, the *yield adjustment term* can be derived as shown below.

$$\begin{aligned}
\frac{A(t,T)}{T-t} &= \frac{1}{2}\frac{1}{T-t}\int_t^T \sum_{i=1}^{3}(\Sigma' B(s,T)B(s,T)'\Sigma)_{i,i}ds \\
&= \frac{1}{2}\frac{1}{T-t}\int_t^T \sum_{i=1}^{3}\left[\begin{pmatrix}\sigma_{11} & \sigma_{21} & \sigma_{31}\\ \sigma_{12} & \sigma_{22} & \sigma_{32}\\ \sigma_{13} & \sigma_{23} & \sigma_{33}\end{pmatrix}\begin{pmatrix}B_1(s,T)\\ B_2(s,T)\\ B_3(s,T)\end{pmatrix}\right. \\
&\qquad \left.\times \begin{pmatrix}B_1(s,T) & B_2(s,T) & B_3(s,T)\end{pmatrix}\begin{pmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\ \sigma_{21} & \sigma_{22} & \sigma_{23}\\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{pmatrix}\right]_{i,i} ds \\
&= \frac{\bar{A}}{2}\frac{1}{T-t}\int_t^T B_1(s,T)^2 ds + \frac{\bar{B}}{2}\frac{1}{T-t}\int_t^T B_2(s,T)^2 ds + \frac{\bar{C}}{2}\frac{1}{T-t}\int_t^T B_3(s,T)^2 ds \\
&\quad + \bar{D}\frac{1}{T-t}\int_t^T B_1(s,T)B_2(s,T)ds + \bar{E}\frac{1}{T-t}\int_t^T B_1(s,T)B_3(s,T)ds \\
&\quad + \bar{F}\frac{1}{T-t}\int_t^T B_2(s,T)B_3(s,T)ds
\end{aligned}$$

Evaluating the integrals, we get the simplified representation for $\frac{A(t,T)}{T-t}$, which is

$$\begin{aligned}
\frac{A(t,T)}{T-t} = &\bar{A}\left(\frac{(T-t)^2}{6}\right) \\
&+ \bar{B}\left(\frac{1}{2\lambda^2} - \frac{1}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{1}{4\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right) \\
&+ \bar{C}\left(\frac{1}{2\lambda^2} + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{4\lambda}(T-t)e^{-2\lambda(T-t)} - \frac{3}{4\lambda^2}e^{-2\lambda(T-t)} - \frac{2}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t}\right. \\
&\qquad \left. + \frac{5}{8\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right) \\
&+ \bar{D}\left(\frac{1}{2\lambda}(T-t) + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t}\right) \\
&+ \bar{E}\left(\frac{3}{\lambda^2}e^{-\lambda(T-t)} + \frac{1}{2\lambda}(T-t) + \frac{1}{\lambda}(T-t)e^{-\lambda(T-t)} - \frac{3}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t}\right) \\
&+ \bar{F}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda^2}e^{-\lambda(T-t)} - \frac{1}{2\lambda^2}e^{-2\lambda(T-t)} - \frac{3}{\lambda^3}\frac{1-e^{-\lambda(T-t)}}{T-t} + \frac{3}{4\lambda^3}\frac{1-e^{-2\lambda(T-t)}}{T-t}\right),
\end{aligned}$$

where

- $\bar{A} = \sigma_{11}^2 + \sigma_{12}^2 + \sigma_{13}^2$,
- $\bar{B} = \sigma_{21}^2 + \sigma_{22}^2 + \sigma_{23}^2$,
- $\bar{C} = \sigma_{31}^2 + \sigma_{32}^2 + \sigma_{33}^2$,
- $\bar{D} = \sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22} + \sigma_{13}\sigma_{23}$,

- $\bar{E} = \sigma_{11}\sigma_{31} + \sigma_{12}\sigma_{32} + \sigma_{13}\sigma_{33}$,
- $\bar{F} = \sigma_{21}\sigma_{31} + \sigma_{22}\sigma_{32} + \sigma_{23}\sigma_{33}$.

The above derivation implies two vital results. First, the yields in the AFNS class of models are given by an analytical formula and that greatly facilitates their empirical implementation. Second, the maximally flexible underlying volatility parameters in the AFNS specification that can be identified is a triangular volatility matrix (Diebold and Rudebusch, 2012),

$$\Sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}. \tag{2.3.26}$$

The relationship between real-world dynamics under the $\mathbb{P}$-measure and the risk neutral dynamics under the $\mathbb{Q}$-measure of the AFNS models in continuous time, is given by measure change,

$$dW_t^{\mathbb{Q}} = dW_t^{\mathbb{P}} + \Gamma_t dt \tag{2.3.27}$$

where $\Gamma_t$ is the risk premium. To keep the affine dynamics under $\mathbb{P}$-measure, as suggested by Duffie (2002), we set the affine risk premium specification as

$$\Gamma_t = \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}, \tag{2.3.28}$$

which gives us the allowance of choosing any mean-reversion matrix $K^{\mathbb{P}}$ and the mean vector $\theta^{\mathbb{P}}$, under the $\mathbb{P}$-measure, and still retain the affine structure observed under the $\mathbb{Q}$-measure. Hence we can express the SDE (2.3.2), under the $\mathbb{P}$-measure as

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = K^{\mathbb{P}} \left[ \begin{pmatrix} \theta_1^{\mathbb{P}} \\ \theta_2^{\mathbb{P}} \\ \theta_3^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \Sigma \begin{pmatrix} dW_t^{1,\mathbb{P}} \\ dW_t^{2,\mathbb{P}} \\ dW_t^{3,\mathbb{P}} \end{pmatrix}. \tag{2.3.29}$$

From the above dynamic representation, we formulate two AFNS models, the independent-factor and correlated-factor models, by imposing mean-reversion matrix $K^{\mathbb{P}}$ and the state volatility matrix $\Sigma$. For the Independent factor Arbitrage Free Nelson Siegel model, the mean reversion and state volatility matrix are

$$K^{\mathbb{P}} = \begin{pmatrix} \kappa_{11} & 0 & 0 \\ 0 & \kappa_{22} & 0 \\ 0 & 0 & \kappa_{33} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \tag{2.3.30}$$

and in the correlated-factor Arbitrage Free Nelson Siegel matrices take these forms

$$K^{\mathbb{P}} = \begin{pmatrix} \kappa_{11} & \kappa_{21} & \kappa_{13} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}. \tag{2.3.31}$$

The measurement equation resulting from both AFNS models is

$$y(\tau_i) = L_t + \left( \frac{1 - e^{-\lambda \tau_i}}{\lambda \tau_i} \right) S_t + \left( \frac{1 - e^{-\lambda \tau_i}}{\lambda \tau_i} - e^{-\lambda \tau_i} \right) C_t - \frac{A(\tau_i)}{\tau_i} + \epsilon_t(\tau_i), \quad i \in \{1, 2, \ldots, N\} \tag{2.3.32}$$

where $\epsilon_t(\tau_i)$, the measurement errors, are i.i.d. noise.

**Corollary 2.3.3.** The forward rate curve corresponding to the arbitrage free Nelson-Siegel yield curve is given by

$$f_t(\tau) = L_t + e^{-\lambda\tau}S_t + \lambda\tau e^{-\lambda\tau}C_t - \frac{\partial A(\tau)}{\partial\tau} \tag{2.3.33}$$

with

$$\frac{\partial A(\tau)}{\partial\tau} = \sigma_{11}^2\frac{\tau^2}{2} + \sigma_{22}^2\frac{(1-e^{-\lambda\tau})^2}{2\lambda^2} + \sigma_{33}^2\frac{(1-(1+\lambda\tau)e^{-\lambda\tau})^2}{2\lambda^2} \tag{2.3.34}$$

where $\tau$ is the time to maturity and the three latent factors are taken to be independent.

*Proof.* The relationship between the forward rate curve and yield curve, according to Steeley (2014), is given by

$$f_t(\tau) = y(\tau) + \tau\frac{dy(\tau)}{d\tau} \tag{2.3.35}$$

Consider

$$\begin{aligned}
\frac{dy(\tau)}{d\tau} &= S_t\left(\frac{\lambda^2\tau e^{-\lambda\tau} - \lambda(1-e^{-\lambda\tau})}{\lambda^2\tau^2}\right) + C_t\left(\frac{\lambda^2\tau e^{-\lambda\tau} - \lambda(1-e^{-\lambda\tau})}{\lambda^2\tau^2} + \lambda e^{-\lambda\tau}\right) \\
&\quad + \frac{A(\tau)}{\tau^2} - \frac{1}{\tau}\frac{\partial A(\tau)}{\partial\tau} \\
&= \frac{S_t}{\tau}\left(e^{-\lambda\tau} - \frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \frac{C_t}{\tau}\left(e^{-\lambda\tau} - \frac{1-e^{-\lambda\tau}}{\lambda\tau} + \lambda\tau e^{-\lambda\tau}\right) + \frac{A(\tau)}{\tau^2} - \frac{1}{\tau}\frac{\partial A(\tau)}{\partial\tau} \\
&= \frac{1}{\tau}\left[S_t e^{-\lambda\tau} - S_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + C_t\lambda\tau e^{-\lambda\tau} - C_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \frac{A(\tau)}{\tau} - \frac{\partial A(\tau)}{\partial\tau}\right] \\
&= \frac{1}{\tau}\left[-S_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) - C_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \frac{A(\tau)}{\tau} + S_t e^{-\lambda\tau} + C_t\lambda\tau e^{-\lambda\tau} - \frac{\partial A(\tau)}{\partial\tau}\right]
\end{aligned}$$

which implies that

$$\frac{dy(\tau)}{d\tau} = \frac{1}{\tau}\left[(-y(\tau) + L_t) + S_t e^{-\lambda\tau} + C_t\lambda\tau e^{-\lambda\tau} - \frac{\partial A(\tau)}{\partial\tau}\right] \tag{2.3.36}$$

Substituting (2.3.36) into (2.3.35), we get

$$\begin{aligned}
f_t(\tau) &= y(\tau) + \tau\frac{1}{\tau}\left[(-y(\tau) + L_t) + S_t e^{-\lambda\tau} + C_t\lambda\tau e^{-\lambda\tau} - \frac{\partial A(\tau)}{\partial\tau}\right] \\
&= L_t + e^{-\lambda\tau}S_t + \lambda\tau e^{-\lambda\tau}C_t - \frac{\partial A(\tau)}{\partial\tau}.
\end{aligned} \tag{2.3.37}$$

Now, we consider the independent factor AFNS model. The constants for the yield adjustment term will be $\bar{A} = \sigma_{11}^2, \bar{B} = \sigma_{22}^2, \bar{C} = \sigma_{33}^2, \bar{D} = 0, \bar{E} = 0, \bar{F} = 0$, then

$$\begin{aligned}
A(\tau) &= \sigma_{11}^2\left(\frac{\tau^3}{6}\right) + \sigma_{22}^2\left(\frac{\tau}{2\lambda^2} - \frac{1}{\lambda^3}(1-e^{-\lambda\tau}) + \frac{1}{4\lambda^3}(1-e^{-2\lambda\tau})\right) \\
&\quad + \sigma_{33}^2\left(\frac{\tau}{2\lambda^2} + \frac{\tau}{\lambda^2}e^{-\lambda\tau} - \frac{1}{4\lambda}\tau^2 e^{-2\lambda\tau} - \frac{3}{4\lambda^2}e^{-2\lambda\tau} - \frac{2}{\lambda^3}(1-e^{-\lambda\tau}) + \frac{5}{8\lambda^3}(1-e^{-2\lambda\tau})\right).
\end{aligned}$$

Thus

$$\begin{aligned}\frac{\partial A(\tau)}{\partial \tau} &= \sigma_{11}^2\left(\frac{\tau^2}{2}\right)+\sigma_{22}^2\left(\frac{1}{2\lambda^2}-\frac{1}{\lambda^3}\lambda e^{-\lambda\tau}+\frac{1}{2\lambda^3}\lambda e^{-2\lambda\tau}\right)\\ &+\sigma_{33}^2\left(\frac{1}{2\lambda^2}-\frac{\tau}{\lambda}e^{-\lambda\tau}+\frac{1}{\lambda^2}e^{-\lambda\tau}-\frac{\tau}{2\lambda}e^{-2\lambda\tau}+\frac{\tau^2}{2}e^{-2\lambda\tau}-\frac{2}{4\lambda^2}e^{-2\lambda\tau}\right.\\ &\left.\quad+\frac{3\tau}{2\lambda}e^{-2\lambda\tau}-\frac{2}{\lambda^2}e^{-\lambda\tau}+\frac{5}{4\lambda^2}e^{-2\lambda\tau}\right)\\ &= \sigma_{11}^2\left(\frac{\tau^2}{2}\right)+\sigma_{22}^2\left(\frac{1-2e^{-\lambda\tau}+(e^{-2\lambda\tau})^2}{2\lambda^2}\right)\\ &+\sigma_{33}^2\left(\frac{1-2(1+\lambda\tau)e^{-\lambda\tau}+((1+\lambda\tau e^{-2\lambda\tau})^2}{2\lambda^2}\right),\end{aligned}$$

which simplifies to

$$\frac{\partial A(\tau)}{\partial \tau}=\sigma_{11}^2\frac{\tau^2}{2}+\sigma_{22}^2\frac{(1-e^{-\lambda\tau})^2}{2\lambda^2}+\sigma_{33}^2\frac{(1-(1+\lambda\tau)e^{-\lambda\tau})^2}{2\lambda^2}. \tag{2.3.38}$$

□

### 2.3.4 . A shape-based Decomposition of the Yield-Adjustment Term

Christensen et al. (2011) demonstrated that the yield adjustment term adds considerable flexibility to the shape of the yield curve. In plotting it, they found that it is monotonically downward concave in the case when the three factors are independent, whereas for correlated factors the yield adjustment term still has a downward slope but has an upward change in curvature in the range 10 to 20 years to maturity.

Steeley (2014) was able to shed some light on how the yield adjustment term behaves in that way. In his work, he first provided an alternative representation of the yield adjustment term that enables one to clearly recognize how it affects the flexibility of the yield curve,

$$\frac{A(t,T)}{T-t}=ANS(t,T)+CDR5(t,T)+FNS(t,T)+QF(t,T) \tag{2.3.39}$$

where

$$\begin{aligned}ANS(t,T) &= \left[\frac{\bar{B}+\bar{C}+2\bar{F}}{2\lambda^2}\right]-\left[\frac{\bar{B}+\bar{C}+2\bar{F}}{\lambda^2}\right]\left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}\right)\\ &-\left[\frac{\bar{C}+\bar{D}+3\bar{E}+\bar{F}}{\lambda^2}\right]\left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}-e^{-\lambda(T-t)}\right)\end{aligned} \tag{2.3.40}$$

$$\begin{aligned}CDR5(t,T) &= \left[\frac{2\bar{B}+2\bar{C}+4\bar{F}}{\lambda_2^2}\right]\left(\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)}\right)\\ &-\left[\frac{3\bar{C}+2\bar{F}}{\lambda_2^2}\right]\left(\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)}-e^{-\lambda_2(T-t)}\right)\end{aligned} \tag{2.3.41}$$

$$FNS(t,T) = \bar{E}\left(\frac{(T-t)e^{-\lambda(T-t)}}{\lambda}\right)-\left[\frac{\bar{C}}{2}\right]\left(\frac{(T-t)e^{-\lambda_2(T-t)}}{\lambda_2}\right) \tag{2.3.42}$$

$$QF(t,T) = \left[\frac{\bar{D}+\bar{E}}{2\lambda}\right](T-t)+\left[\frac{\bar{A}}{6}\right](T-t)^2 \tag{2.3.43}$$

and $\lambda_2 = 2\lambda$. The $ANS(t,T)$ component consists of three functions, with each imposing adjustment on the level, slope and curvature loadings in the AFNS term structure model. From (2.3.40), it is evident that the level factor of the yield curve is being adjusted by $\left[\frac{\bar{B}+\bar{C}+2\bar{F}}{2\lambda^2}\right]$ and the slope and the curvature by $\left[\frac{\bar{B}+\bar{C}+2\bar{F}}{\lambda^2}\right]$ and $\left[\frac{\bar{C}+\bar{D}+3\bar{E}+\bar{F}}{\lambda^2}\right]$ respectively. The $CDR5(t,T)$ component consists of additional slope and curvature terms which are $\left[\frac{2\bar{B}+2\bar{C}+4\bar{F}}{\lambda_2^2}\right]\left(\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)}\right)$ and $\left[\frac{3\bar{C}+2\bar{F}}{\lambda_2^2}\right]\left(\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)} - e^{-\lambda_2(T-t)}\right)$.

The $FNS(t,T)$ component consists of curvature terms relative to the Nelson-Siegel forward rate curve. Given that the forward rate curve is related to the rate at which the spot rate changes, the $FNS(t,T)$ terms can be thought of as representing change in curvature along the yield curve. Lastly, the $QF(t,T)$ component consists of a quadratic component function of maturity.

## 2.4 Nelson-Siegel Model Extensions

It has been observed that for sensible choice of $\lambda$, in the three factors Nelson-Siegel model, the slope and curvature factor loadings decay rapidly to zero as maturity approaches. This leaves only the level factor to fit over yields with maturities of ten plus years. Some researchers have made improvements on the Nelson-Siegel model to address the problem for in-sample fitting of yields with longer maturities. For example, Svensson (1994) presented his modified version, called the Nelson-Siegel Svensson model, with an additional curvature variable,

$$y(\tau) = \beta_1 + \beta_2\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\right) + \beta_3\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right) + \beta_4\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right) \tag{2.4.1}$$

which dethrones the NS in fitting yields with longer maturities. Mimicking the procedure by Diebold and Li (2006), Christensen et al. (2009) replaced the coefficients $(\beta_1, \beta_2, \beta_3, \beta_4)$ with time-varying coefficients $(L_t, S_t, C_t^1, C_t^2)$ to produce the following dynamic Nelson-Siegel Svensson (NSS) model

$$y(\tau) = L_t + S_t\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\right) + C_t^1\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right) + C_t^2\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right). \tag{2.4.2}$$

However making the NSS dynamic, did not render it arbitrage-free. Efforts to directly apply the procedure in Proposition 2.3.1 to make the DNSS arbitrage-free proved to be pointless since the mechanics require pairing of curvature and slope factors of the same-rate mean reversion. Hence, Christensen et al. (2009) suggested that a slope factor be added in equation (2.4.2) to match the curvature factor. This led to a five factor model, extension of the NS model, called The dynamic generalised Nelson-Siegel Model (DGNS):

$$\begin{aligned} y(\tau) = L_t &+ S_t^1\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau}\right) + S_t^2\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau}\right) + C_t^1\left(\frac{1-e^{-\lambda_1\tau}}{\lambda_1\tau} - e^{-\lambda_1\tau}\right) \\ &+ C_t^2\left(\frac{1-e^{-\lambda_2\tau}}{\lambda_2\tau} - e^{-\lambda_2\tau}\right) \end{aligned} \tag{2.4.3}$$

where $L_t$ is the level factor, $S_t^1$ and $S_t^2$ are slope factors and $C_t^1$ and $C_t^2$ are curvature factors. Without loss of generality, the following condition are imposed on the decay parameters, $0 < \lambda_2 < \lambda_1$. From the DGNS, Christensen et al. (2009) proposed the arbitrage-free generalised Nelson-Siegel model (AFGNS) by extending Proposition 2.3.1.

**Proposition 2.4.1.** Suppose that the instantaneous risk-free rate is defined by

$$r_t = L_t + S_t^1 + S_t^2, \tag{2.4.4}$$

and also suppose that the state variables $X_t = (L_t, S_t^1, S_t^2, C_t^1, C_t^2)$ evolve according to following SDEs, under the risk-neutral $\mathbb{Q}$-measure

$$d\begin{pmatrix} L_t \\ S_t^1 \\ S_t^2 \\ C_t^1 \\ C_t^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & -\lambda_1 & 0 \\ 0 & 0 & \lambda_2 & 0 & -\lambda_2 \\ 0 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{Q}} \\ \theta_{S^1}^{\mathbb{Q}} \\ \theta_{S^2}^{\mathbb{Q}} \\ \theta_{C^1}^{\mathbb{Q}} \\ \theta_{C^2}^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t^1 \\ S_t^2 \\ C_t^1 \\ C_t^2 \end{pmatrix} \right] dt + \Sigma d \begin{pmatrix} W_t^{L,\mathbb{Q}} \\ W_t^{S^1,\mathbb{Q}} \\ W_t^{S^2,\mathbb{Q}} \\ W_t^{C^1,\mathbb{Q}} \\ W_t^{C^2,\mathbb{Q}} \end{pmatrix}, \quad 0 < \lambda_2 < \lambda_1, \tag{2.4.5}$$

where $\theta^{\mathbb{Q}}$ is the mean of the latent factors and $\Sigma$ is a constant volatility matrix. The price of the zero-coupon is given by

$$\begin{aligned} P(t,T) &= \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left(-\int_t^T r_s ds\right)\right] \\ &= \exp(B_1(t,T)L_t + B_2(t,T)S_t^1 + B_3(t,T)S_t^2 + B_4(t,T)C_t^1 + B_5(t,T)C_t^2 + A(t,T)) \end{aligned}$$

where $B_1(t,T)$, $B_2(t,T)$, $B_3(t,T)$ , $B_4(t,T)$, $B_5(t,T)$ and $A(t,T)$ are solutions to the following system of ODEs

$$\frac{d}{dt}\begin{pmatrix} B_1(t,T) \\ B_2(t,T) \\ B_3(t,T) \\ B_4(t,T) \\ B_5(t,T) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 0 & 0 \\ 0 & -\lambda_1 & 0 & \lambda_1 & 0 \\ 0 & 0 & -\lambda_2 & 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} B_1(t,T) \\ B_2(t,T) \\ B_3(t,T) \\ B_4(t,T) \\ B_5(t,T) \end{pmatrix} \tag{2.4.6}$$

and

$$\frac{\mathrm{d}A(t,T)}{dt} = -B(t,T)'K^{\mathbb{Q}}\theta^{\mathbb{Q}} - \frac{1}{2}\sum_{i=1}^{5}(\Sigma' B(t,T)B(t,T)'\Sigma)_{i,i}, \tag{2.4.7}$$

with boundary conditions given by $B_1(T,T) = B_2(T,T) = B_3(T,T) = B_4(T,T) = B_5(T,T) = A(T,T) = 0$. The solution of the system of ODEs (2.4.6) is

$$B_1(t,T) = -(T-t), \tag{2.4.8}$$

$$B_2(t,T) = -\frac{1-e^{-\lambda_1(T-t)}}{\lambda_1}, \tag{2.4.9}$$

$$B_3(t,T) = -\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2}, \tag{2.4.10}$$

$$B_4(t,T) = (T-t)e^{-\lambda_1(T-t)} - \frac{1-e^{-\lambda_1(T-t)}}{\lambda_1}, \tag{2.4.11}$$

$$B_5(t,T) = (T-t)e^{-\lambda_2(T-t)} - \frac{1-e^{-\lambda_2(T-t)}}{\lambda_2}. \tag{2.4.12}$$

$B_1(t,T)$, $B_2(t,T)$, $B_3(t,T)$ , $B_4(t,T)$, $B_5(t,T)$ are factor loadings. The solution of (2.4.7) is given by

$$\begin{aligned} A(t,T) = (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_2 \int_t^T B_2(s,T)ds + (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_3 \int_t^T B_3(s,T)ds + (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_5 \int_t^T B_4(s,T)ds \\ + (K^{\mathbb{Q}}\theta^{\mathbb{Q}})_5 \int_t^T B_5(s,T)ds + \frac{1}{2}\sum_{i=1}^{5}\int_t^T (\Sigma' B(s,T)B(s,T)'\Sigma)_{i,i}ds \end{aligned} \tag{2.4.13}$$

where $K^{\mathbb{Q}}$ is the mean reversion matrix. Lastly, we have the yields for zero-coupon given by

$$y(t,T) = L_t + \frac{1-e^{-\lambda_1(T-t)}}{\lambda_1(T-t)}S_t^1 + \frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)}S_t^2 + \left[\frac{1-e^{-\lambda_1(T-t)}}{\lambda_1(T-t)} - e^{-\lambda_1(T-t)}\right]C_t^1 + \left[\frac{1-e^{-\lambda_2(T-t)}}{\lambda_2(T-t)} - e^{-\lambda_2(T-t)}\right]C_t^2 - \frac{A(t,T)}{T-t} \tag{2.4.14}$$

The proof is also a straightforward extension of the proof for Proposition 2.3.1. As in the AFNS, the yield-adjustment term take the form

$$\frac{A(t,T)}{T-t} = \frac{1}{2}\frac{1}{T-t}\sum_{i=1}^{5}\int_t^T (\Sigma' B(s,T)B(s,T)'\Sigma)_{i,i}ds \tag{2.4.15}$$

and the maximally flexible AFGNS specification triangular volatility matrix takes the form

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} \\ 0 & \sigma_{22} & \sigma_{23} & \sigma_{24} & \sigma_{25} \\ 0 & 0 & \sigma_{33} & \sigma_{34} & \sigma_{35} \\ 0 & 0 & 0 & \sigma_{44} & \sigma_{45} \\ 0 & 0 & 0 & 0 & \sigma_{55} \end{pmatrix}. \tag{2.4.16}$$

# 3. Affine Arbitrage-free Nelson-Siegel model in Discrete Time

Discrete-time models are an integral part of micro-finance analysis since theories and methods in time-series analysis are usually in discrete-time. In the preceding chapter, we presented the arbitrage-free Nelson-Siegel term structure model in continuous time. Following the work Niu and Zeng (2012) and Ang and Piazzesi (2003) we demonstrate the essential affine arbitrage free model in discrete time.

## 3.1 General Set-up for Term Structure Model with Micro-Factors

- The short rate is assumed to be an affine function of all state variables $X_t$, that is,

$$r_t = \rho_0 + \rho_1' X_t, \tag{3.1.1}$$

  where $\rho_0 \in \mathbb{R}$ and $\rho_1 \in \mathbb{R}^K$.

- The state dynamics $X_t$, under physical measure $\mathbb{P}$, follow a first order Gaussian VAR(1) process

$$X_t = \mu + \Phi X_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \Sigma\Sigma'). \tag{3.1.2}$$

- The state dynamics under risk-neutral measure $\mathbb{Q}$ follow the first order Gaussian VAR(1) process

$$X_t = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1} + \epsilon_t^{\mathbb{Q}}. \tag{3.1.3}$$

- Under constant volatility, the affine model assumes a risk price, $\lambda_t$, which takes an affine form of the states variables, that is,

$$\lambda_t = \lambda_0 + \lambda_1 X_t \tag{3.1.4}$$

  where $\lambda_0 \in \mathbb{R}^K$ and $\lambda_1 \in \mathbb{R}^{K\times K}$

- Given that $P_t^n$ is the price of $n$-period zero coupon bond, under no-arbitrage conditions the relationship between bond prices of maturities $n+1$ and $n$ is given by

$$P_t^{n+1} = \mathbb{E}_t \left[ m_{t+1} P_{t+1}^n \right], \tag{3.1.5}$$

  where $m_{t+1}$ is the pricing kernel and according to Ang and Piazzesi (2003) is defined by

$$m_{t+1} = \exp\left( -r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1} \right). \tag{3.1.6}$$

The no-arbitrage recursive relations can be deduced such that the bonds are exponential affine functions of $X_t$ and the yields are affine on the states. Thus the zero coupon bond prices are given by

$$P_t^n = \exp(B_n' X_t + A_n). \tag{3.1.7}$$

where $B_n$ and $A_n$ follow the difference equations below

$$\begin{aligned} B_{n+1}' &= B_n'(\Phi - \Sigma\lambda_1) - \rho_1 \\ &= -\rho_1 \sum_{k=0}^{n} (\Phi - \Sigma\lambda_1)^k \end{aligned} \tag{3.1.8}$$

$$A_{n+1} = A_n + B_n'(\mu - \Sigma\lambda_0) + \frac{1}{2} B_n' \Omega B_n - \rho_0 \tag{3.1.9}$$

with $\Omega = \Sigma\Sigma'$, $B_1 = -\rho_1$ and $A_1 = -\rho_0$. Then the continuously compounded affine function of yields, $y_t^n$, on an $n$-period zero coupon bond is given by

$$y_t^n = -\frac{\log P_t^n}{n} = -\frac{1}{n}(B_n' X_t + A_n) \tag{3.1.10}$$

Under the equivalent martingale measure $\mathbb{Q}$,

$$P_t^{n+1} = \mathbb{E}_t^{\mathbb{Q}}\left[\exp(-r_t)P_{t+1}^n\right], \tag{3.1.11}$$

where $\mathbb{E}_t^{\mathbb{Q}}$ means the expectation under $\mathbb{Q}$. The dynamics of $X_t$ are characterized by $\mu^{\mathbb{Q}}$ and autoregressive matrix $\Phi^{\mathbb{Q}}$. Between physical measure and risk neutral measure we have the following relationships

$$\Phi^{\mathbb{Q}} = \Phi - \Sigma\lambda_1, \tag{3.1.12}$$
$$\mu^{\mathbb{Q}} = \mu - \Sigma\lambda_0. \tag{3.1.13}$$

Hence, under the above risk neutral parameters equations (3.1.8) and (3.1.9) become

$$\begin{aligned} B'_{n+1} &= B'_n\Phi^{\mathbb{Q}} + B'_1 \\ &= B'_1\sum_{k=0}^{n}(\Phi^{\mathbb{Q}})^k, \end{aligned} \tag{3.1.14}$$

$$A_{n+1} = A_n + B'_n\mu^{\mathbb{Q}} + \frac{1}{2}B'_n\Omega B_n + A_1. \tag{3.1.15}$$

The dynamic Nelson-Siegel yield function of an $n$-period zero coupon bond is given by

$$y_t^n = L_t + \left(\frac{1-e^{-\lambda n}}{\lambda n}\right)S_t + \left(\frac{1-e^{-\lambda n}}{\lambda n} - e^{-\lambda n}\right)C_t. \tag{3.1.16}$$

## 3.2 Making the DNS Arbitrage-Free in Discrete Time

Given the zero bond pricing function, for the three factors affine model with state variables $X_t = (L_t, S_t, C_t)$, a yield function close to the Nelson-Siegel yield function is of the form

$$y_t^n = L_t + \left(\frac{1-e^{-\lambda n}}{\lambda n}\right)S_t + \left(\frac{1-e^{-\lambda n}}{\lambda n} - e^{-\lambda n}\right)C_t - \frac{A_n}{n}, \tag{3.2.1}$$

with the system of difference equations for $B_n$ having the solution

$$B_n = \begin{pmatrix} -n \\ -\frac{1-e^{-\lambda n}}{\lambda} \\ ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}. \end{pmatrix} \tag{3.2.2}$$

It is obvious that the factors loading exactly match DNS, then there is the yield-adjustment term, $-\frac{A_n}{n}$ which satisfies the following difference equation for $n > 1$

$$A_n = A_{n-1} + B'_{n-1}\mu^{\mathbb{Q}} + \frac{1}{2}B'_{n-1}\Omega B_{n-1} + A_1 \tag{3.2.3}$$

Next, following the framework of Niu and Zeng (2012) we establish the sufficient and necessary conditions for the DNS model to be arbitrage free.

**Proposition 3.2.1 (sufficient conditions).** Given that the short rate is

$$r_t = \rho_0 + \rho_1' X_t, \tag{3.2.4}$$

where $\rho_1 = \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{pmatrix}$ and assume the state variables $X_t = (L_t, S_t, C_t)$ follow first order Gaussian VAR process under the equivalent martingale measure $\mathbb{Q}$

$$X_t = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}} X_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \Omega), \tag{3.2.5}$$

where $\Phi^{\mathbb{Q}}$ takes the specific form

$$\Phi^{\mathbb{Q}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{pmatrix}. \tag{3.2.6}$$

The zero coupon bond price is given by

$$P_t^n = \exp(B_n' X_t + A_n), \tag{3.2.7}$$

and

$$B_n = \begin{pmatrix} -n \\ -\frac{1-e^{-\lambda n}}{\lambda} \\ n e^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda} \end{pmatrix} \tag{3.2.8}$$

is the Nelson-Siegel factor loadings. For $n > 1$, $A_n$ is the solution to the difference equation

$$A_n = A_{n-1} + B_{n-1}' \mu^{\mathbb{Q}} + \frac{1}{2} B_{n-1}' \Omega B_{n-1} + A_1 \tag{3.2.9}$$

with $B_1 = -\rho_1$.

*Proof.* Let $B_1 = \begin{pmatrix} -1 \\ -\frac{1-e^{-\lambda}}{\lambda} \\ e^{-\lambda} - \frac{1-e^{-\lambda}}{\lambda} \end{pmatrix}$ for $n > 1$, from the difference equations (3.1.14) and (3.1.15) we have

$$\begin{aligned} B_n' &= B_1' \sum_{k=0}^{n-1} (\Phi^{\mathbb{Q}})^k, & (3.2.10) \\ A_n &= A_{n-1} + B_{n-1}' \mu^{\mathbb{Q}} + \frac{1}{2} B_{n-1}' \Omega B_{n-1} + A_1. & (3.2.11) \end{aligned}$$

Then

$$\begin{aligned} (\Phi^{\mathbb{Q}})^k = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{pmatrix}^k = \left( e^{-\lambda} \begin{pmatrix} e^{\lambda} & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix} \right)^k &= e^{-k\lambda} \begin{pmatrix} e^{\lambda} & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{pmatrix}^k \\ &= e^{-k\lambda} \begin{pmatrix} e^{k\lambda} & 0 & 0 \\ 0 & 1 & k\lambda \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-k\lambda} & k\lambda e^{-k\lambda} \\ 0 & 0 & e^{-k\lambda} \end{pmatrix}, \end{aligned}$$

which implies that

$$\sum_{k=0}^{n-1}(\Phi^{\mathbb{Q}})^k = \sum_{k=0}^{n-1}\begin{pmatrix}1 & 0 & 0\\ 0 & e^{-k\lambda} & k\lambda e^{-k\lambda}\\ 0 & 0 & e^{-k\lambda}\end{pmatrix} = \begin{pmatrix}n & 0 & 0\\ 0 & \sum_{k=0}^{n-1} e^{-k\lambda} & \sum_{k=0}^{n-1} k\lambda e^{-k\lambda}\\ 0 & 0 & \sum_{k=0}^{n-1} e^{-k\lambda}\end{pmatrix}$$
$$= \begin{pmatrix}n & 0 & 0\\ 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} & \lambda\left[\frac{e^{-\lambda}-e^{-n\lambda}}{(1-e^{-\lambda})^2} - \frac{(n-1)e^{-n\lambda}}{1-e^{-\lambda}}\right]\\ 0 & 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}}\end{pmatrix}.$$

From (3.2.10) we have that

$$\begin{aligned}
B_n' &= B_1'\sum_{k=0}^{n-1}(\Phi^{\mathbb{Q}})^k\\
&= \begin{pmatrix}-1 & -\frac{1-e^{-\lambda}}{\lambda} & e^{-\lambda}-\frac{1-e^{-\lambda}}{\lambda}\end{pmatrix}\begin{pmatrix}n & 0 & 0\\ 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} & \lambda\left[\frac{e^{-\lambda}-e^{-n\lambda}}{(1-e^{-\lambda})^2} - \frac{(n-1)e^{-n\lambda}}{1-e^{-\lambda}}\right]\\ 0 & 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}}\end{pmatrix}\\
&= \begin{pmatrix}-n & -\frac{1-e^{-n\lambda}}{\lambda} & -\Lambda\end{pmatrix}
\end{aligned}$$

where

$$\begin{aligned}
\Lambda &= \frac{e^{-\lambda}-e^{-n\lambda}}{1-e^{-\lambda}} - (n-1)e^{-n\lambda} + \frac{1-e^{-n\lambda}}{\lambda} - \frac{e^{-\lambda}-e^{-(n+1)\lambda}}{1-e^{-\lambda}}\\
&= ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}.
\end{aligned}$$

Therefore

$$B_n = \begin{pmatrix}-n\\ -\frac{1-e^{-\lambda n}}{\lambda}\\ ne^{-\lambda n}-\frac{1-e^{-\lambda n}}{\lambda}\end{pmatrix}. \tag{3.2.12}$$

□

**Proposition 3.2.2 (necessary conditions).** Suppose the DNS with a constant yield adjustment term, $a_n$, fits the yield curve

$$y_t^n = a_n + L_t + \left(\frac{1-e^{-\lambda n}}{\lambda n}\right)S_t + \left(\frac{1-e^{-\lambda n}}{\lambda n} - ne^{-\lambda n}\right)C_t, \tag{3.2.13}$$

thoroughly and suppose the Nelson-Siegel latent factors follow a VAR(1) process

$$X_t = \mu^{\mathbb{Q}} + \Phi^{\mathbb{Q}}X_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0,\Omega), \tag{3.2.14}$$

then the risk-free rate satisfies the affine process

$$r_t = \rho_0 + \rho_1' X_t, \tag{3.2.15}$$

where $\rho_1 = \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{pmatrix}$, $\rho_0 = A_1$ and the risk neutral dynamic coefficient matrix $\Phi^{\mathbb{Q}}$ is given by

$$\Phi^{\mathbb{Q}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{pmatrix}. \tag{3.2.16}$$

*Proof.* For $n = 1$, equation (3.2.15) is fulfilled by (3.2.13). Then, for $n > 1$, from equation (3.1.10) the yield curve for affine term structure model is given by

$$y_t^n = -\frac{1}{n}(A_n + B_n' X_t), \tag{3.2.17}$$

where

$$B_{n+1}' = B_n'\Phi^{\mathbb{Q}} + B_1' \tag{3.2.18}$$

$$A_{n+1} = A_n + B_n'\mu^{\mathbb{Q}} + \frac{1}{2}B_n'\Omega B_n + A_1. \tag{3.2.19}$$

Comparing equations (3.2.17) and (3.2.13) we deduce that

$$B_n' = \begin{pmatrix} -n & -\frac{1-e^{-\lambda n}}{\lambda} & ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda} \end{pmatrix}, \tag{3.2.20}$$

$$A_n = -na_n. \tag{3.2.21}$$

Note that the difference equation (3.2.18) can be written as

$$B_{n+1} = (\Phi^{\mathbb{Q}})' B_n + B_1. \tag{3.2.22}$$

Now, let $\Phi^{\mathbb{Q}} = \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{pmatrix}$, we substitute (3.2.20) into (3.2.22) to get

$$-\begin{pmatrix} n+1 \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} - (n+1)e^{-\lambda(n+1)} \end{pmatrix} = -\begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{pmatrix}\begin{pmatrix} n \\ \frac{1-e^{-\lambda n}}{\lambda} \\ \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n} \end{pmatrix} - \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{pmatrix},$$

$$\begin{pmatrix} n+1 \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} - (n+1)e^{-\lambda(n+1)} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{pmatrix}\begin{pmatrix} n \\ \frac{1-e^{-\lambda n}}{\lambda} \\ \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n} \end{pmatrix} + \begin{pmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{pmatrix}.$$

Rearranging the above equation we get

$$\begin{pmatrix} n \\ \frac{e^{-\lambda n}-e^{-\lambda(n+1)}}{\lambda} \\ \frac{e^{-\lambda n}-e^{-\lambda(n+1)}}{\lambda} - (n+1)e^{-\lambda(n+1)} + e^{-\lambda} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{pmatrix}\begin{pmatrix} n \\ \frac{1-e^{-\lambda n}}{\lambda} \\ \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n} \end{pmatrix}. \tag{3.2.23}$$

Now we solve the following resulting equations;

$$\begin{aligned}
\phi_{11}n + \phi_{21}\left(\frac{1-e^{-\lambda n}}{\lambda}\right) + \phi_{31}\left(\frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n}\right) &= n, &(3.2.24)\\
\phi_{12}n + \phi_{22}\left(\frac{1-e^{-\lambda n}}{\lambda}\right) + \phi_{32}\left(\frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n}\right) &= \frac{e^{-\lambda n} - e^{-\lambda(n+1)}}{\lambda}, &(3.2.25)\\
\phi_{13}n + \phi_{23}\left(\frac{1-e^{-\lambda n}}{\lambda}\right) + \phi_{33}\left(\frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n}\right) &= \frac{e^{-\lambda n} - e^{-\lambda(n+1)}}{\lambda}\\
&\quad -(n+1)e^{-\lambda(n+1)} + \ e^{-\lambda}. &(3.2.26)
\end{aligned}$$

Equation (3.2.24) implies that

$$\phi_{11} = 1, \quad \phi_{21} = 0 \quad \text{and} \quad \phi_{31} = 0.$$

Equation (3.2.25) implies that

$$\begin{aligned}
\phi_{12} &= 0;\\
\phi_{22}\frac{1}{\lambda} + \phi_{32}\frac{1}{\lambda} - (\phi 22 + \phi_{32})\frac{e^{-\lambda n}}{\lambda} - \phi_{32}ne^{-\lambda n} &= \frac{e^{-\lambda}}{\lambda} - \frac{e^{-\lambda}e^{-\lambda n}}{\lambda}. &(3.2.27)
\end{aligned}$$

By inspection, equation (3.2.27) hold when:

$$\phi_{22} = e^{-\lambda} \quad \text{and} \quad \phi_{32} = 0$$

Thus, equation (3.2.26) has the implication

$$\begin{aligned}
\phi_{13} &= 0 \qquad \text{and}\\
\phi_{23}\frac{1}{\lambda} + \phi_{33}\frac{1}{\lambda} - (\phi_{23} + \phi_{33})\frac{e^{-\lambda n}}{\lambda} - \phi_{33}ne^{-\lambda n} &= \frac{e^{-\lambda}}{\lambda} + e^{-\lambda} - (e^{-\lambda} + \lambda e^{-\lambda})\frac{e^{-\lambda n}}{\lambda}\\
&\quad -e^{-\lambda}ne^{-\lambda n} &(3.2.28)
\end{aligned}$$

and equation (3.2.28) hold when

$$\phi_{23} = e^{-\lambda} \quad \text{and} \quad \phi_{33} = \lambda e^{-\lambda}.$$

Hence we have

$$\Phi^{\mathbb{Q}} = \begin{pmatrix} 1 & 0 & 0\\ 0 & e^{-\lambda} & \lambda e^{-\lambda}\\ 0 & 0 & e^{-\lambda} \end{pmatrix}. \qquad (3.2.29)$$

□

# 4. Affine Nelson-Siegel model with Stochastic Volatility

Several interesting facts about interest rate volatility have been unravelled, including that it is understandably stochastic and consist of unspanned components. Also, some literature reveals that there is a correlation between changes in interest rate volatility and the change in interest rates, themselves (Trolle and Schwartz, 2009).

In chapter 2 we presented the AFNS model under constant volatility. However, in forecasting it is recommendable to have yields normally distributed with time variation in the volatility. So this chapter focus on presenting the class of affine Nelson-Siegel models with stochastic volatility.

## 4.1 Models with One Stochastic Volatility

The first class of stochastic volatility models permits only one factor to exhibit interest rate volatility. Although we have three possibilities under this class, only two are admissible. The first one with level as the driving factor for stochastic volatility and the other with curvature.

**4.1.1 Volatility driven by level factor ($AFNS_1 - L$).** The state variables $X_t = (L_t, S_t, C_t)$ under this model, are described by the system of stochastic differential equations (SDEs) below, under the equivalent martingale measure (EMM) $\mathbb{Q}$,

$$\begin{aligned}\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{\mathbb{Q}} & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{Q}} \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{L_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21} L_t} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31} L_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0\end{aligned}$$

where the level factor, $L_t$, drives stochastic volatility and it induces instantaneous volatility for slope $(S_t)$ and curvature $(C_t)$ factors through $\beta_{21}$ and $\beta_{31}$ volatility sensitivities, respectively.

The factor loadings for the level factor is a solution to the ODE

$$\begin{aligned}\frac{dB_1(t,T)}{dt} &= 1 + \kappa_{11}^{\mathbb{Q}} B_1(t,T) - \frac{1}{2}\sigma_{11}^2 B_1(t,T)^2 - \frac{1}{2}\sigma_{21}^2 B_2(t,T)^2 - \frac{1}{2}\sigma_{31}^2 B_3(t,T)^2 \\ &\quad - \sigma_{11}\sigma_{21} B_1(t,T) B_2(t,T) - \sigma_{11}\sigma_{31} B_1(t,T) B_3(t,T) - \sigma_{21}\sigma_{31} B_2(t,T) B_3(t,T) \\ &\quad - \frac{1}{2}\beta_{21}\left[\sigma_{22}^2 B_2(t,T)^2 + \sigma_{32}^2 B_3(t,T)^2 + 2\sigma_{22}\sigma_{32} B_2(t,T) B_3(t,T)\right] - \frac{1}{2}\beta_{31}\sigma_{33}^2 B_3(t,T)^2,\end{aligned}$$

while the factor loadings for slope $(B_2(t,T))$ and curvature $(B_3(t,T))$ are given by

$$B_2(t,T) = -\frac{1 - e^{-\lambda(T-t)}}{\lambda}, \tag{4.1.1}$$

$$B_3(t,T) = (T-t)e^{-\lambda(T-t)} - \frac{1 - e^{-\lambda(T-t)}}{\lambda}. \tag{4.1.2}$$

The yield adjustment term $A(t,T)$ is a solution to the ODE

$$\frac{dA(t,T)}{dt} = -\kappa_{11}^{\mathbb{Q}}\theta_L^{\mathbb{Q}}B_1(t,T) - \frac{1}{2}(\sigma_{32}^2+\sigma_{33}^2)B_3(t,T)^2 - \sigma_{22}\sigma_{32}B_2(t,T)B_3(t,T) - \frac{1}{2}\sigma_{22}^2B_2(t,T)^2.$$

A similar model was proposed by Ohnishi and Sim, but with $\beta_{21}=\beta_{31}=0$. They allowed only the level factor to follow the process by Cox et al. (1985). The solution for slope and curvature factor loadings remained as in equations (4.1.1) and (4.1.2), and are identical to those of the AFNS. However, for the level factor loading, they obtained the explicit solution

$$B^1(t,T) = -\frac{2-e^{-\eta(T-t)}}{\eta+\kappa_{11}^{\mathbb{Q}}+(\eta-\kappa_{11}^{\mathbb{Q}})e^{-\eta(T-t)}}, \tag{4.1.3}$$

where $\eta=\sqrt{(\kappa_{11}^{\mathbb{Q}})^2+2\sigma_{11}^2}$.Hence the resultant yield curve is given by

$$\begin{aligned} y(t,T) &= \left(\frac{2-e^{-\eta(T-t)}}{(T-t)[\eta+\kappa_{11}^{\mathbb{Q}}+(\eta-\kappa_{11}^{\mathbb{Q}})e^{-\eta(T-t)}]}\right)L_t + \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)}\right)S_t \\ &\quad + \left(\frac{1-e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}\right)C_t - \frac{A(t,T)}{T-t}. \end{aligned}$$

To finalise the model, we use the affine risk premium specification in (2.3.28) and under physical measure, $\mathbb{P}$ the dynamics are

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ \kappa_{21}^{\mathbb{P}} & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}} \\ \kappa_{31}^{\mathbb{P}} & \kappa_{32}^{\mathbb{P}} & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{P}} \\ \theta_S^{\mathbb{P}} \\ \theta_C^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{L_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}L_t} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}L_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}. \end{aligned}$$

To eliminate riskless arbitrage, the Feller conditions

$$\kappa_{11}^{\mathbb{P}}\theta_L^{\mathbb{P}} > \frac{1}{2}\sigma_{11}^2 \quad \text{and} \quad \kappa_{11}^{\mathbb{Q}}\theta_L^{\mathbb{Q}} > \frac{1}{2}\sigma_{11}^2$$

must be satisfied (Christensen et al., 2014; Cheridito et al., 2007).

**4.1.2 Volatility driven by curvature factor ($AFNS_1-C$).** In this model, the state variables are described, under the EMM $\mathbb{Q}$, by following system of SDEs

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} 0 \\ 0 \\ \theta_C^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{11} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{1+\beta_{13}C_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{23}C_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0, \end{aligned}$$

where the curvature factor, $C_t$, drives stochastic volatility and it induces instantaneous volatility for level $(L_t)$ and slope $(S_t)$ factors via $\beta_{13}$ and $\beta_{23}$ volatility sensitivities.

For level and slope the factor loadings are

$$\begin{aligned} B_1(t,T) &= -(T-t), \\ B_2(t,T) &= -\frac{1-e^{-\lambda(T-t)}}{\lambda}. \end{aligned}$$

and the factor loadings for the curvature factor is given by the solution to

$$\begin{aligned} \frac{dB_3(t,T)}{dt} &= \lambda B_3(t,T) - \lambda B_2(t,T) - \frac{1}{2}\beta_{31}\sigma_{11}^2 B_1(t,T)^2 \\ &\quad -\frac{1}{2}\beta_{32}\left[\sigma_{12}^2 B_1(t,T)^2 + \sigma_{22}^2 B_2(t,T)^2 + 2\sigma_{12}\sigma_{22}B_1(t,T)B_2(t,T)\right] \\ &\quad -\frac{1}{2}\left[\sigma_{13}B_1(t,T) + \sigma_{23}B_2(t,T) + \sigma_{33}B_3(t,T)\right]^2. \end{aligned}$$

The yield adjustment term is given by the solution to the ODE

$$\begin{aligned} \frac{dA(t,T)}{dt} &= \lambda\theta_C^{\mathbb{Q}}[B_2(t,T) - B_3(t,T)] - \frac{1}{2}\sigma_{11}^2 B_1(t,T)^2 \\ &\quad -\frac{1}{2}\left[\sigma_{12}^2 B_1(t,T)^2 + \sigma_{22}^2 B_2(t,T)^2 + 2\sigma_{12}\sigma_{22}B_1(t,T)B_2(t,T)\right]. \end{aligned}$$

Considering the affine risk premium specification, the $\mathbb{P}$-dynamics are given by

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & \kappa_{12}^{\mathbb{P}} & \kappa_{13}^{\mathbb{P}} \\ \kappa_{21}^{\mathbb{P}} & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}} \\ 0 & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{P}} \\ \theta_S^{\mathbb{P}} \\ \theta_C^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{11} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \sqrt{1+\beta_{13}C_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{23}C_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}, \end{aligned}$$

The parameters must obey the following Feller conditions

$$\lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \kappa_{33}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{33}^2$$

## 4.2 Models with Two Stochastic Volatility

Under the second class of stochastic volatility model, two factors are allowed to drive the stochastic volatility of interest rates. Christensen et al. (2014) explained that we can only have stochastic volatility driven by level and curvature factors or slope and curvature factors.

**4.2.1 Volatility driven by level and curvature factor ($AFNS_2 - LC$).** The factor loadings evolution of the $AFNS_2 - LC$ is governed, under the EMM $\mathbb{Q}$, by the system of SDEs

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{Q}} \\ 0 \\ \theta_C^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \sqrt{L_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}L_t+\beta_{23}C_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0, \end{aligned}$$

where $L_t$ and $C_t$ drive the interest rate volatility, $\beta_{13}$ and $\beta_{23}$ measure the sensitivities of the two factors towards the slope factor.

The factor loadings for level and curvature factors are solutions to the the ODEs

$$\begin{aligned} B_1(t,T) &= 1+\epsilon B_1(t,T)-\frac{1}{2}\sigma_{11}^2B_1(t,T)^2-\frac{1}{2}\sigma_{21}^2B_2(t,T)^2 \\ &\quad -\sigma_{11}\sigma_{21}B_1(t,T)B_2(t,T)-\frac{1}{2}\beta_{21}\sigma_{22}^2B_2(t,T)^2, \\ B_3(t,T) &= -\lambda B_2(t,T)+\lambda B_3(t,T)-\frac{1}{2}\beta_{23}\sigma_{22}^2B_2(t,T)^2 \\ &\quad -\frac{1}{2}[\sigma_{23}^2B_2(t,T)^2+\sigma_{33}^2B_3(t,T)^2+2\sigma_{23}\sigma_{33}B_2(t,T)B_3(t,T)] \end{aligned}$$

The factor loading for slope is given by

$$B_2(t,T)=-\frac{1-e^{-\lambda(T-t)}}{\lambda}.$$

The yield adjustment term solves the ODE

$$\frac{dA(t,T)}{dt}=-\epsilon\theta_L^{\mathbb{Q}}+\lambda\theta_C^{\mathbb{Q}}[B_2(t,T)-B_3(t,T)]-\frac{1}{2}\sigma_{22}B_2(t,T)^2.$$

Applying the affine risk premium specification, we have the $\mathbb{P}$-dynamics as follows

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & 0 & 0 \\ \kappa_{21}^{\mathbb{P}} & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}} \\ \kappa_{31}^{\mathbb{P}} & 0 & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{P}} \\ \theta_S^{\mathbb{P}} \\ \theta_C^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \sqrt{L_t} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}L_t+\beta_{23}C_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}. \end{aligned}$$

The Feller conditions are imposed on the curvature factor, that is,

$$\lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \kappa_{31}^{\mathbb{P}}\theta_L^{\mathbb{P}}+\kappa_{33}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{33}^2 \tag{4.2.1}$$

and the constraint $\kappa_{31}^{\mathbb{P}} \leq 0$ is imposed for viability of the model.

**4.2.2 Volatility driven by slope and curvature factor ($AFNS_2-SC$).** The factor loadings $AFNS_2-SC$ evolves, under the EMM $\mathbb{Q}$, according to following system of SDEs

$$\begin{aligned} \begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} 0 \\ \theta_S^{\mathbb{Q}} \\ \theta_C^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\quad \times \begin{pmatrix} \sqrt{1+\beta_{12}S_t+\beta_{13}C_t} & 0 & 0 \\ 0 & \sqrt{S_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0 \end{aligned}$$

where $S_t$ and $C_t$ drives the interest rate volatility and $\beta_{12}$ and $\beta_{13}$ measure the sensitivities of the two factors towards the level factor.

The factor loadings on the level factor is given by

$$\frac{dB_1(t,T)}{dt} = -(T-t),$$

while the factor loadings for slope ($B_2(t,T)$) and curvature ($B_3(t,T)$) are solutions to the ODEs

$$\begin{aligned}
\frac{B_2(t,T)}{dt} &= 1+\lambda B_2(t,T) - \frac{1}{2}\sigma_{22}^2 B_2(t,T)^2 - \frac{1}{2}\sigma_{12}^2 B_1(t,T)^2 \\
&\quad -\sigma_{12}\sigma_{22}B_1(t,T)B_2(t,T) - \frac{1}{2}\beta_{12}\sigma_{11}^2 B_1(t,T)^2, \\
\frac{B_3(t,T)}{dt} &= -\lambda B_2(t,T) + \lambda B_3(t,T) - \frac{1}{2}\sigma_{13}^2 B_1(t,T)^2 - \frac{1}{2}\sigma_{33}^2 B_3(t,T)^2 \\
&\quad -\sigma_{13}\sigma_{33}B_1(t,T)B_3(t,T) - \frac{1}{2}\beta_{13}\sigma_{11}^2 B_1(t,T)^2.
\end{aligned}$$

The yield adjustment term is a solution of the ODE

$$\frac{dA(t,T)}{dt} = -\lambda[(\theta_S^{\mathbb{Q}} - \theta_C^{\mathbb{Q}})B_2(t,T) + \theta_C^{\mathbb{Q}} B_3(t,T)] - \frac{1}{2}\sigma_{11}^2 B_1(t,T)^2.$$

To finalise the model, we apply the extended affine risk premium specification. Then the $\mathbb{P}$-dynamics are given by

$$\begin{aligned}
\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{\mathbb{P}} & \kappa_{12}^{\mathbb{P}} & \kappa_{13}^{\mathbb{P}} \\ 0 & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}} \\ 0 & \kappa_{32}^{\mathbb{P}} & \kappa_{33}^{\mathbb{P}} \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{P}} \\ \theta_S^{\mathbb{P}} \\ \theta_C^{\mathbb{P}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \sqrt{1+\beta_{12}S_t+\beta_{13}C_t} & 0 & 0 \\ 0 & \sqrt{S_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{P}} \\ dW_t^{S,\mathbb{P}} \\ dW_t^{C,\mathbb{P}} \end{pmatrix}.
\end{aligned}$$

The Feller conditions

$$\kappa_{22}^{\mathbb{P}}\theta_S^{\mathbb{P}} + \kappa_{23}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{22}^2; \quad \lambda\theta_S^{\mathbb{Q}} - \lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{22}^2; \quad \kappa_{32}^{\mathbb{P}}\theta_S^{\mathbb{P}} + \kappa_{33}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{33}^2; \quad \text{and} \quad \lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{33}^2 \quad (4.2.2)$$

are imposed to eliminate arbitrage and for admissibility, we require that $\kappa_{32}^{\mathbb{P}} \le 0$ and $\kappa_{33}^{\mathbb{P}} \le 0$.

## 4.3 Models with Three Stochastic Volatility ($AFNS_3$)

All three factors are allowed to drive volatility in the $AFNS_3$ specification. The state variables evolve, under the risk neutral $\mathbb{Q}$-measure, relative the following system of SDEs

$$\begin{aligned}
\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} &= \begin{pmatrix} \epsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[ \begin{pmatrix} \theta_L^{\mathbb{Q}} \\ \theta_S^{\mathbb{Q}} \\ \theta_C^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \sqrt{L_t} & 0 & 0 \\ 0 & \sqrt{S_t} & 0 \\ 0 & 0 & \sqrt{C_t} \end{pmatrix} \begin{pmatrix} dW_t^{L,\mathbb{Q}} \\ dW_t^{S,\mathbb{Q}} \\ dW_t^{C,\mathbb{Q}} \end{pmatrix}, \quad \lambda > 0.
\end{aligned}$$

The factor loadings solves the ODEs

$$\begin{aligned}\frac{B_1(t,T)}{dt} &= 1+\epsilon B_1(t,T)-\frac{1}{2}\sigma_{11}^2 B_1(t,T)^2\\ \frac{B_2(t,T)}{dt} &= 1+\lambda B_2(t,T)-\frac{1}{2}\sigma_{22}^2 B_2(t,T)^2\\ \frac{B_3(t,T)}{dt} &= -\lambda B_2(t,T)+\lambda B_3(t,T)-\frac{1}{2}\sigma_{33}^2 B_3(t,T)^2.\end{aligned}$$

The yield adjustment term, $A(t,T)$, is a solution to

$$\frac{dA(t,T)}{dt} = -\epsilon\theta_L^{\mathbb{Q}} B_1(t,T) - \lambda[(\theta_S^{\mathbb{Q}}-\theta_C^{\mathbb{Q}})B_2(t,T)+\theta_C^{\mathbb{Q}}B_3(t,T)].$$

Using the extended affine risk premium specification, we have the $\mathbb{P}$-dynamics given by

$$\begin{aligned}\begin{pmatrix} dL_t\\ dS_t\\ dC_t\end{pmatrix} &= \begin{pmatrix}\kappa_{11}^{\mathbb{P}} & \kappa_{12}^{\mathbb{P}} & \kappa_{13}^{\mathbb{P}}\\ \kappa_{21}^{\mathbb{P}} & \kappa_{22}^{\mathbb{P}} & \kappa_{23}^{\mathbb{P}}\\ \kappa_{31}^{\mathbb{P}} & \kappa_{32}^{\mathbb{P}} & \kappa_{33}^{\mathbb{P}}\end{pmatrix}\left[\begin{pmatrix}\theta_L^{\mathbb{P}}\\ \theta_S^{\mathbb{P}}\\ \theta_C^{\mathbb{P}}\end{pmatrix} - \begin{pmatrix}L_t\\ S_t\\ C_t\end{pmatrix}\right]dt + \begin{pmatrix}\sigma_{11} & 0 & 0\\ 0 & \sigma_{22} & 0\\ 0 & 0 & \sigma_{33}\end{pmatrix}\\ &\quad\times\begin{pmatrix}\sqrt{L_t} & 0 & 0\\ 0 & \sqrt{S_t} & 0\\ 0 & 0 & \sqrt{C_t}\end{pmatrix}\begin{pmatrix}dW_t^{L,\mathbb{P}}\\ dW_t^{S,\mathbb{P}}\\ dW_t^{C,\mathbb{P}}\end{pmatrix}.\end{aligned}$$

To ensure that the model is free from arbitrage, we impose the Feller condition

$$\begin{aligned}\epsilon\theta_L^{\mathbb{Q}} > \frac{1}{2}\sigma_{11}^2 \quad &\text{and}\quad \kappa_{11}^{\mathbb{P}}\theta_L^{\mathbb{P}}+\kappa_{12}^{\mathbb{P}}\theta_S^{\mathbb{P}}+\kappa_{13}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{11}^2;\\ \lambda\theta_S^{\mathbb{Q}}-\lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{22}^2 \quad &\text{and}\quad \kappa_{21}^{\mathbb{P}}\theta_L^{\mathbb{P}}+\kappa_{22}^{\mathbb{P}}\theta_S^{\mathbb{P}}+\kappa_{23}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{22}^2;\\ \lambda\theta_C^{\mathbb{Q}} > \frac{1}{2}\sigma_{33}^2 \quad &\text{and}\quad \kappa_{31}^{\mathbb{P}}\theta_L^{\mathbb{P}}+\kappa_{32}^{\mathbb{P}}\theta_S^{\mathbb{P}}+\kappa_{33}^{\mathbb{P}}\theta_C^{\mathbb{P}} > \frac{1}{2}\sigma_{33}^2,\end{aligned}$$

and we further impose the constraints $\kappa_{12}^{\mathbb{P}} \le 0$, $\kappa_{13}^{\mathbb{P}} \le 0$, $\kappa_{21}^{\mathbb{P}} \le 0$, $\kappa_{23}^{\mathbb{P}} \le 0$, $\kappa_{31}^{\mathbb{P}} \le 0$ and $\kappa_{32}^{\mathbb{P}} \le 0$.

# 5. Conclusion

In the pricing of fixed income securities a superior term structure model guarantees better portfolio returns and improve risk management. The Nelson-Siegel model is known for empirical superiority even though it lacks theoretical rigour. In this research project, we have looked at the Nelson-Siegel that imposes the desirable condition of no arbitrage. We have shown how Christensen et al. (2011) derived the model, from the work of Duffie and Kan (1996), in continuous time. We also have derived the discrete time version of the AFNS after the work of Niu and Zeng (2012). One of the shortcomings of AFNS model is that it assumes constant volatility, which is not a desirable characteristic for a model to be used for risk management. Hence, we have further investigated the descendants of the AFNS model that impose stochastic volatility in one, two or three of the latent factors.

# Acknowledgements

First, I would like to extend my sincere gratitude to my supervisor Dr. Raouf Ghomrasni for his guidance throughout the process of producing this research project. His critiques and suggestions helped me a lot. I would also like to thank Adriaan Van Vuuren for all the help he gave me. Many thanks to the AIMS family for providing a conducive environment for me to carrying out this research project. Lastly, I would like to thank my friends and colleagues for their emotional support.

# References


A. Ang and M. Piazzesi. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, 50(4):745–787, 2003.

T. Björk and B. J. Christensen. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance*, 9(4):323–348, 1999.

P. Cheridito, D. Filipović, and R. L. Kimmel. Market price of risk specifications for affine models: Theory and evidence. *Journal of Financial Economics*, 83(1):123–170, 2007.

J. Christensen, J. Lopez, and G. Rudebusch. Can spanned term structure factors drive stochastic yield volatility? *Manuscript, Federal Reserve Bank of San Francisco*, 2014.

J. H. E. Christensen, F. X. Diebold, and G. D. Rudebusch. An arbitrage-free generalized Nelson-Siegel term structure model. *Journal of Econometrics*, 12:33–64, 2009.

J. H. E. Christensen, F. X. Diebold, and G. D. Rudebusch. The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics*, 164:4 – 20, 2011.

J. C. Cox, J. E. Ingersoll Jr, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, pages 385–407, 1985.

Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.

F. X. Diebold and C. Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130:337–364, 2006.

F. X. Diebold and G. D. Rudebusch. *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton University, 2012.

D. Duffie. Term premia and interest rates forecast in affine models. *Journal of Finance*, 57:405–443, 2002.

D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.

D. Filipović. *Term Structure Models*. Spingler-Verlag, 2009.

C. Li, L. Niu, and G. Zeng. A generalized arbitrage-free Nelson-Siegel term structure model with macroeconomic fundamentals. *Available at SSRN 1787262*, 2011.

R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1: 49–53, 1991.

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

L. Niu and G. Zeng. The discrete-time framework of the arbitrage-free Nelson-Siegel class of term structure models, 2012.

M. Ohnishi and D. Sim. A modified arbitrage-free Nelson Siegel model: An alternative affine term structure model of interest rates.

J. M. Steeley. A shape-based decomposition of the yield adjustment term in the arbitrage-free Nelson and Siegel (AFNS) model of the yield curve. *Applied Financial Economics*, (ahead-of-print):1–9, 2014.

L. Svensson. Estimating and interpreting forward interest rates: Sweden. *International Monetary Fund*, 4:114, 1994.

A. B. Trolle and E. S. Schwartz. A general stochastic volatility model for the pricing of interest rate derivatives. *Review of Financial Studies*, 22(5):2007–2057, 2009.