# Modeling a Two-currency Affine Arbitrage-free Nelson-Siegel Term Structure Model

by Yi Yu

M.Litt. in Finance

Department of Economics, Finance &Accounting

National University of Ireland, Maynooth

May, 2012

Supervisor: Professor Gregory Connor

**Abstract:** This thesis contains two papers. In the first paper, we provide a general overview of the most popular term structure of interest rate models. In order to understand different features of each model, we classify by means of general characteristics, from single-factor to multi-factor and forward rate based models. Each of these existing term structure models has its own advantages and disadvantages. We also highlight the recently advocated models in the literature: the Nelson-Siegel model, the affine and the quadratic arbitrage-free model. In the second paper we extend the affine arbitrage-free Nelson-Siegel model to a two-currency (3+1) factor structure model that incorporates the properties of interest rate term structure and foreign exchange rates simultaneously within one arbitrage-free framework by decomposing the pricing kernel into two independent portions: one portion contains three factors that model the affine Nelson-Siegel term structure of interest rate, the other portion contains one factor that captures the effect of the currency movement, which is independent of the term structure.

**Contents**

**Term Structure Model of Interest Rates-- A Literature Review**


## Abstract

Term structure modeling has enjoyed rapid growth during the last two decades. Given a large number of existing term structure models and a vast array of issues in the field, we attempt to provide a general overview of the most popular term structure of interest rate models. In order to understand different features of each model we classify by means of general characteristics from single-factor to multi-factor and forward rate based models. Each of these existing term structure models has its own advantages and disadvantages. We also highlight the recent advocated models in the literature: the Nelson-Siegel model, the affine and the quadratic arbitrage-free model.

## Introduction

The term structure of interest rates, also known as the yield curve, plays a central role both theoretically and practically in the economy. It gives the relationship between the yield on an investment and the term maturity of the investment. Yield curve modeling literature has at the origin the need of explaining interest rate behavior. Both short term and long term interest rates have an important role in financial markets, for different reasons and purposes. They are used for the price of borrowing or lending money; they are needed to price bonds and to price derivatives on bonds and other fixed income instruments. Thus understanding and modeling the term structure of interest rates has been one of the most challenging topics of financial research.

The characteristic features of interest rate models can be generally categorized into eight types: continuous or discrete models, single or multi-factor models, fitted (to the initial term structure) or non-fitted models and arbitrage-free or equilibrium model.

In order to keep the scope manageable, the aim of this paper is to provide an analysis of the most popular term structure models of interest rates that are applicable to the default-free zero-coupon bonds. We propose these models in a common framework and explain their merits and drawbacks from an overview perspective.

This paper is organized as follows: section 1 introduces the definitions and notations will be used throughout the paper; section 2 reviews simple factor interest rate models in the literature with both time-invariant and time varying parameters; section 3 considers the extension of the single factor model to multi-factor model. Section 4 reviews the forward rate based Heath, Jarrow and Morton (1992) model which models the entire term structure and provides a richer volatility pattern for predicting and controlling future volatilities. We highlight the most recent three advocated term structure models for bond yields in section 5 and conclude in section 6.

## 1. Definitions and notations

The models will generally be set up in a filtered probability space $(\Omega, F_t, \mathrm{P})$, where $\Omega$ is the sample space, $F_t$ is the sigma-field generated by a standard Brownian motion

W(t). P indicates the historical (physical) probability measure on the sample space Ω.

We denote $P(t,T)$ as the price at time t of a default-free zero-coupon bond with principal one dollar maturing at time T. It follows that $P(T,T)=1$. At time t, the yield to maturity $y(t,T)$ of the zero-coupon bond is the continuously compounded rate of return that causes the bond price to rise to one at time T. Yields are solved by $y(t,T) = -\frac{\ln P(t,T)}{T-t}$.

For a fixed time t, the shape of the yield $y(t,T)$ as T increases determines the term structure of interest rates. Since we only work with the zero-coupon bonds, the yield curve is the same as the term structure of interest rates.

The instantaneous risk-free rate also called short rate/short term rate is denoted as $r(t)$ and $r(t) = \lim_{t \to T} y(t,T)$.

Define $f(t,T_1,T_2)$ as the forward rate at time t for the period between time $T_1$ and $T_2$. The relationship between forward rate and zero-coupon bond is given by:

$$f(t,T_1,T_2) = \frac{\ln P(t,T_1) - \ln P(t,T_2)}{T_2 - T_1} \tag{1}$$

The instantaneous forward rate is the rate that one contracts at time t for a loan starting at time T for an instantaneous period of time $f(t,T) \equiv f(t,T,T)$.

The bond price can be defined in terms of forward rate as

$$P(t,T) = e^{-\int_t^T f(t,s)du} \tag{2}$$

The relationship between short rate and forward rate is given by $r(t) = f(t,t)$, so the short rate is a specific forward rate.

## 2. Single factor models

Factor models assume that the term structure of interest rates is driven by a set of state variables or factors. A principle component analysis[1] can be used to decompose the motion of the interest rate term structure into three independent factors: shift, twist and butterfly of the term structure (Wilson, 1994). As the first principle component (shift) explains a large fraction of the yield curve movement, it is tempting to reduce the problem to a single factor model.

Single factor models assume that all information about the term structure at any point in time can be summarized by one single factor – the short rate $r(t)$. As a consequence, only the short rate and time to maturity will affect the price of the zero-coupon bonds. There are two basic methodologies for pricing interest rate contingent claims in a single factor framework, the partial differential equation and the martingale approach. The former creates an instantaneous risk-free portfolio to obtain a second order partial differential equation that interest rate contingent claim must satisfy. The latter proposed by Harrison and Kreps (1979) and extended by Heath, Jarrow and Morton (1992) uses the result that in a complete market, in the absence of arbitrage, there exists an equivalent martingale measure under which asset prices can be computed as an expectation. These two approaches are equivalent by the theorem of Feynman-Kac.

There are three particular versions of the single factor models: the affine class models, the Gaussian models and the lognormal models.

The affine models named by Duffie and Kan (1996) have the following exponential affine form of the zero-coupon bond price

$$P(t,T) = \exp\left[a(t,T)r(t) + b(t,T)\right] \tag{3}$$

where $a(t,T)$ and $b(t,T)$ are deterministic functions that can be calculated via Riccati ordinary differential equation (obtained from the partial differential equation of the bond pricing). The term structure of interest rates is an affine function of the short rate:

$$y(t,T) = \frac{-a(t,T)}{T-t} r(t) + \frac{b(t,T)}{T-t} \tag{4}$$

[1] A principal component analysis is a statistical technique that identifies the best factors from historical yields, where the term best is in the sense of the two conditions: 1) the factors ought to explain a very large proportion of the variation of the yields of bonds at various horizons; 2) the factors should be independent of each other.

If under the risk-neutral probability, the mean and volatility are affine in $r(t)$, then we say that the model has an affine version.

A short term interest rate model is said to be Gaussian if it can be written as the following linear stochastic differential equation (SDE):

$$dr(t) = \mu_r(t, r(t))dt + \sigma_r(t, r(t))dW(t) \tag{5}$$

where $\mu_r$ and $\sigma_r$ are the drift and the volatility/standard deviation of the short rate, respectively. Gaussian model is a particular class of affine models, and $r(t)$ is normally distributed.

A short term interest rate model is said to be lognormal if and only if $\ln r(t)$ is Gaussian. The advantage of lognormal models over Gaussian is that by definition, lognormal rate models cannot generate negative interest rates. However, they generally lack analytical tractability.

So far we have discussed the versions of the single factor models, now we turn to some examples of the single factor models in the literature.

### 2.1 One-factor time invariant/ equilibrium models

**2.1.1 Merton (1973)** was the first to propose a general stochastic process as a model for the short rate. Under the historical probability measure P, the short rate has the following SDE:

$$dr(t) = \mu dt + \sigma dW(t) \tag{6}$$

where $\mu$ and $\sigma$ are constant.

The short rate is solved by $r(t) = r(s) + \mu t + \sigma \int_s^t dW(s)$ for any $t \geq s$.

Given the set of information at time s, the short term rate $r(t)$ is normally distributed with mean $r(s) + \mu(t - s)$ and variance $(t - s)\sigma^2$. The unboundedness of the first and second moment of the distribution allows the rate to become negative or infinite.

**2.1.2 Vasicek (1977)** proposes to model the short term interest rate as a Gaussian Ornstein-Uhlenbeck (mean-reverting) process:

$$dr(t) = \kappa\left(\theta - r(t)\right)dt + \sigma dW(t) \tag{7}$$

where $\kappa, \theta$ and $\sigma$ are constants. This model incorporates mean reversion. The short rate is pulled to a level $\theta$ at the mean reversion rate $\kappa$.

The explicit solution to the SDE in (7) gives us the short term rate $r(t) = \theta + \left(r(s) - \theta\right)e^{-\kappa(t-s)} + \sigma\int_s^t e^{-\kappa(t-s)} dW(u)$ for any $t \geq s$.

Given the set of information at time s, the short term rate $r(t)$ is normally distributed with mean $\theta + \left(r(s) - \theta\right)e^{-\kappa(t-s)}$ and variance $\frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right)$. Again, the short term interest rate can become negative.

**2.1.3 Cox, Ingersoll and Ross (1985)** have developed an alternative model where rates are always non-negative. The short term rate satisfies the following SDE:

$$dr(t) = \kappa\left(\theta - r(t)\right)dt + \sigma\sqrt{r(t)}dW(t) \tag{8}$$

where $\kappa, \theta$ and $\sigma$ are constants. This model has the same mean-reverting drift as the Vasicek model, but the variance of the change in the short rate in a short period of time is proportional to the short rate $r$ rather than constant. This means that, as the short term interest rate increases, its standard deviation increases.

The positive short term rate is obtained by solving the SDE in (8), $r(t) = \theta + \left(r(s) - \theta\right)e^{-\kappa(t-s)} + \sigma e^{-\kappa(t-s)}\int_s^t e^{\kappa(u-s)}\sqrt{r(u)}dW(u)$ for any $t \geq s$.

Given the set of information at time s, the short term rate $r(t)$ is distributed as a non-central chi-squared (Feller, 1951).

The one-factor time-invariant/equilibrium models we have reviewed above have the disadvantage that they do not automatically fit today's term structure of interest rates. The drift of the short rate as shown above is not usually a function of time. This leads us to the one-factor time-varying/arbitrage-free models. The essential difference between these two types of models is that in a time-varying/arbitrage-free model, the drift is, in general, dependent on time. This is because the shape of the initial zero curve governs the average path taken by the short rate in the future in an arbitrage-free model. In turns out that some equilibrium models can be converted to arbitrage-free models by including a function of time in the drift of the short rate.

## 2.2 One-factor time-varying/arbitrage-free models

**2.2.1 Ho and Lee** proposed the first arbitrage-free model of the term structure in a paper in 1986. They presented the model in the form of binomial tree of bond prices with two parameters: the short rate standard deviation and the market price of risk of the short rate. It has been shown by Jamshidian (1991a) that the continuous-time limit of the short rate is driven under the risk-neutral probability by the SDE:

$$dr(t) = \theta(t)dt + \sigma dW(t) \tag{9}$$

where the volatility term $\sigma$ is a constant and the drift term $\theta(t)$ is a function of time chosen to ensure the model fits the initial term structure.

**2.2.2 Hull and White one-factor model (1990)** explored an extension of the Vasicek(1977) model that provide an exact fit to the initial term structure.

$$dr(t) = \kappa\left(\frac{\theta(t)}{\kappa} - r(t)\right)dt + \sigma dW(t) \tag{10}$$

where $\kappa$ and $\sigma$ are constants. It can be characterized as the Ho-Lee model with mean reversion at rate $\kappa$ and the Vasicek model with a time-dependent reversion level. At time t, the short rate reverts to $\frac{\theta(t)}{\kappa}$ at rate $\kappa$.

**2.2.3 Hull-White (1993)** form a general specification of the short rate

$$dr(t) = \left(\theta(t) - \kappa r(t)\right)dt + \sigma r^{\alpha}(t)dW(t) \tag{11}$$

where $\kappa$ and $\sigma$ are constants, $\theta(t)$ is time-varying. There are a few existing models that can be extended in this framework. These models are such as the extended Vasicek model with $\alpha = 0$, the extended Cox, Ingersoll and Ross model with $\alpha = 0.5$ and the extended Brennan and Schwartz (1977) and Courtadon (1982) model with $\alpha = 1$.

The models we have seen so far (except Cox, Ingersoll and Ross model which has a square root process) are modeled as Gaussian processes; the popularity of using the Gaussian process is due to its analytical tractability. However, this process implies that there is a positive probability of negative rates. This leads to our next subsection of the one-factor model with lognormal rates to avoid the negative rates.

### 2.3 One-factor time-varying lognormal models

**2.3.1 Black, Derman and Toy (1987)** propose a one factor binomial model whose continuous time version has the form of

$$d \ln r(t) = \left(\theta(t) - \kappa \ln r(t)\right) dt + \sigma_r dW(t) \tag{12}$$

It assumes a lognormal process for the short rate, which precludes negative value.

They extended the model to allow for time dependent volatility in the 1990 paper.

$$d \ln r(t) = \left(\theta(t) - \kappa \ln r(t)\right) dt + \sigma_r(t) dW(t) \tag{13}$$

The model does not have as much analytical tractability as the Gaussian process model. It is not possible to produce formulas for valuing bonds in terms of the short rate using the model.

**2.3.2 Black and Karasinski (1991)** propose an extension of the Black, Derman and Toy (1987) model with a time-varying mean reversion rate $\kappa(t)$. The model has the form of

$$d \ln r(t) = \left(\theta(t) - \kappa(t) \ln r(t)\right) dt + \sigma_r(t) dW(t) \tag{14}$$

Again, the model lacks analytical properties.

Modeling lognormally distributed rates is the simplest way to avoid negative rates, but no closed form solution to the zero-coupon bonds can be found for these models.

To close this subsection, we list some other one-factor models that are not as popular as the ones we have discussed. These models are Dothan (1978), Brennan and Schwartz

(1977, 1980), Courtadon (1982), Rendleman and Bartter (1980) and Cox, Ingersoll and Ross (1980).

## 3. Multi-factor models

So far we have reviewed the single factor models where the short rate is the only explanatory variable. Most of these models are characterized by their analytical tractability. However, these models often fail to match observed prices. For an economic point of view, it seems unreasonable to assume that the entire term structure is governed only by the short rate. So using more than one explanatory factor to model the interest rate is quite useful. Most multi-factor models are in fact based on two factors. These models are such as Cox, Ingersoll and Ross (1985b), Longstaff and Schwartz (1991), Fong and Vasicek (1991), Chen (1994) and Duffie and Kan (1996). In this subsection, we exam some popular multi-factor models in more detail.

**3.1 Cox, Ingersoll and Ross (1985b)** presents a model in which the term structure of interest rates is determined by two factors: the real short rate $q(t)$ and the expected instantaneous inflation rate $\pi(t)$. Both factors are assumed to follow independent diffusion process:

$$\begin{aligned} dq(t) &= \mu_q(t)dt + \sigma_q(t)dW_q(t) \\ d\pi(t) &= \mu_\pi(t)dt + \sigma_\pi(t)dW_\pi(t) \end{aligned} \qquad (15)$$

where $W_q$ and $W_q$ are two independent Brownian motions. They obtain a complicated, but analytical solution for the zero-coupon bond price. Similar framework is proposed by Brennan and Schwartz (1982), in which the term structure of interest rates depends on both the short term rate $r(t)$ and the long term rate $l(t)$.

**3.2 Longstaff and Schwartz (1992)** developed an equilibrium model of the economy and derived from there a two-factor term structure model. The two factors are the short term rate $r(t)$ and the variance of changes in the short term rate $v(t)$.

In their framework, the representative investor has a logarithmic utility and has the choice between investing and consuming the only good available in the economy, whose price $P(t)$ follows the SDE:

$$d\frac{P(t)}{P(t)} = \left(\mu X(t) + \theta Y(t)\right)dt + \sigma\sqrt{Y}dW_1(t) \qquad (16)$$

where $X(t)$ and $Y(t)$ are two specific economic factors. $X(t)$ is the expected return part that is unrelated to the Brownian motion $W_1(t)$; $Y(t)$ is the factor correlated with $dP(t)$. The dynamics of the two factors are given by

$$\begin{aligned} dX(t) &= \left(a - bX(t)\right)dt + c\sqrt{X(t)}dW_2(t) \\ dY(t) &= \left(d - eY(t)\right)dt + f\sqrt{Y(t)}dW_3(t) \end{aligned} \qquad (17)$$

where $W_2(t)$ and $W_3(t)$ are uncorrelated Brownian motions and $a,b,c,d,e,f > 0$.

Longstaff and Schwartz do not provide any intuitive interpretation for these two factors, but they show that $X(t)$ and $Y(t)$ can be related to observable quantities, as the equilibrium instantaneous interest rate $r(t)$ and the variance of its changes $v(t)$ are given by a weighted sum of these two factors.

$$\begin{aligned} r(t) &= \mu c^2 X(t) + \left(\theta - \sigma^2\right) f^2 Y(t) \\ v(t) &= \mu^2 c^4 X(t) + \left(\theta - \sigma^2\right)^2 f^4 Y(t) \end{aligned} \qquad (18)$$

so that $r(t)$ and $v(t)$ are non-negative for all feasible values of state variables.

This model can be seen as an affine two factor model, in which one is the short rate and the other one is its volatility.

In a series of papers, Fong and Vasicek (1991,1992a,1992b) have derived a two-factor model using the same factors as this model. In their framework, the short rate and its variance evolve under the risk-neutral probability

$$\begin{aligned} dr(t) &= \beta\left(\bar{r} - r(t)\right)dt + \sqrt{v(t)}dW_1(t) \\ dv(t) &= \varsigma\left(\bar{v} - v(t)\right)dt + \xi\sqrt{v(t)}dW_2(t) \end{aligned} \qquad (19)$$

where $\bar{r}$ is the long term mean of the short rate and $v(t)$ is its instantaneous volatility, and $\bar{v}$ is the long term mean of the volatility. The two Brownian motions are correlated. This model does not preclude the positive probability of negative short term rates.

**3.3 Chen (1996)** proposed a three factor model of the term structure. In his model, the short rate dynamics depends on the current short rate, the stochastic mean of the short rate, and the stochastic volatility of the short rate.

$$\begin{aligned} dr(t) &= \kappa\left(\theta(t)-r(t)\right)dt+\sqrt{\sigma(t)}\sqrt{r(t)}dW_2(t) \\ d\theta &= v\left(\bar{\theta}-\theta(t)\right)dt+\xi\sqrt{\theta}dW_1(t) \\ d\sigma(t) &= \mu\left(\bar{\sigma}-\sigma(t)\right)dt+\eta\sqrt{\sigma}dW_3(t) \end{aligned} \tag{20}$$

**3.4 Duffie and Kan (1996)** introduced an N-factor affine term structure model which is obtained under the assumptions that the instantaneous short rate $r(t)$ is an affine function of a vector of unobserved state variables $\mathrm{X}(t)$.

$$r(t)=\rho_0+\sum_{i=1}^{n}\rho_i\,\mathrm{X}_i(t)\equiv\rho_0+\rho_X\,'\mathrm{X}(t) \tag{21}$$

The state variable vector follows an affine diffusion[2] :

$$d\mathrm{X}(t)=\kappa\left(\theta-\mathrm{X}(t)\right)dt+\Sigma\sqrt{S(t)}dW(t) \tag{22}$$

where W(t) is an N-dimensional independent standard Brownian motion under the risk-neutral probability, $\kappa$ and $\Sigma$ are N∗N matrices, which may be non-diagonal and asymmetric, and S(t) is a diagonal matrix with the $i^{th}$ diagonal element given by

$$\left[S(t)\right]_{ii}=\alpha_i+\beta_i\mathrm{X}(t) \tag{23}$$

Both the drifts in equation (22) and the conditional variances in equation (23) of the state variables are affine in X(t).

---

[2] Gaussian process and square-root process are the best known examples of affine diffusions. Gaussian process has a constant volatility, while the square-root processes introduce conditional heteroskedasticity by allowing the volatility function to depend on the state variables.

## 4. The Heath-Jarrow-Morton (1992) Model

Even with a multi-factor model, the term structure of interest rates has a rather limited number of degrees of freedom. An alternative approach to single and multi-factor interest rate modeling is to specify the entire term structure/yield curve of interest rates. Rather than using a finite number of state variables, some authors use one state variable of infinite dimension, namely, the term structure itself. The first contribution to this approach was made by the Ho and Lee (1986) binomial model in a discrete time. It was the first to model movements in the entire term structure. Heath, Jarrow and Morton (1992) have significantly extended the Ho and Lee (1986) model by considering forward rates rather than bond prices as their building block; it also extended it from one factor model to a multi-factor model.

Heath, Jarrow and Morton framework firstly developed a class of models that are derived by directly modeling the dynamics of instantaneous forward rates. It models the entire term structure as a state variable, providing conditions in a general framework that incorporates all the principles of arbitrage-free pricing and zero-coupon bond dynamics. The Heath, Jarrow and Morton model shows that there is a link between the drift and standard deviation of an instantaneous forward rate. The drift of the forward rates under the risk-neutral probability is entirely determined by its volatility, which is the major contribution of this model.

$$df(t,T) = \sum_{i=1}^{n} \sigma_T(t,T) \int_t^T \sigma(t,\tau) d\tau + \sum_{i=1}^{n} \sigma_T(t,T) dW(t) \quad (24)$$

Since the short term rate is a specific forward rate, the short rate in the Heath, Jarrow and Morton model can be written in an integral form as

$$dr(t) = f(0,t) + \int_0^t \sigma_f(s,t) \int_0^t \sigma_f(s,u) du ds + \int_0^t \sigma_f(s,t) dW(s) \quad (25)$$

Note that the difficulty of estimating the Heath, Jarrow and Morton model will arise because of the non-Markovin term in equation (25), which depends on the history of the process from time 0 to time t.

## 5. Final remarks

In this section, we highlight the most recent popular term structure models of bond yields advocated in the financial literature.

The three main classes of term structure models are the affine term structure model; the Nelson-Siegel model and the quadratic term structure models: the affine term structure model, originally introduced by Duffie and Kan (1996), classified by Dai and Singleton [3](2000) and extended to the essentially affine specification by Duffee (2002); the dynamic Nelson-Siegel model introduced by Diebold and Li (2006), which build on Nelson and Siegel (1987); and the class of quadratic term structure models classified by Ahn *et al.*(2002) and Leippold and Wu (2002).

The Nelson- Siegel model provides an intuitive description of the yield curve at each point in time. The dynamic Nelson-Siegel model (Diebold and Li, 2006) is easy to estimate and fits yield curve data well in-sample and produces good out-of sample forecasts. In contrast to arbitrage-free term structure models, this model class does not preclude arbitrage opportunities. However, an extension of the Nelson-Siegel model that is arbitrage-free does exist and this is done by Christensen, Diebold and Rudebusch (2007, 2008).

The affine and quadratic term structure models both have the arbitrage-free property; they derive the dynamic yield curve under a risk-neutral probability measure. The existence of risk-neutral probability measure implies that bond prices are arbitrage-free and the observed yield curve evolution is a result of the yield behavior under a historical probability measure. The transition from the risk-neutral to the historical measure is established via the market price of risk. Dai and Singleton (2000) provide the admissibility conditions and suggest a specification for completely affine term structure model. Duffee (2002) points out the restriction of the completely affine specification,

[3]. Within the family of Duffie and Kan affine term structure model, there is a trade-off between flexibility in modeling the conditional correlations and volatilities of the risk factors. This trade-off is formalized by their classification of N-factor affine family into N + 1 non-nested subfamilies of models. Vasicek (1977), Chen (1996), and Cox, Ingersoll and Ross(1985) models are classified into distinct subfamilies of the affine models.

and presents a broader class of essentially affine models, in which the market price of risk specification is more flexibly formulated.

Ahn *et al.* (2002) describe the classification and canonical representation of the quadratic term structure models analogously to the classification of affine models in Dai and Singleton (2000). They show that the quadratic model specification can capture the conditional volatility of yields better than the affine class.

Nyholm and Vidova-Koleva (2011) conduct an extensive out-of-sample forecasting experiment among quadratic, affine and dynamic Nelson-Siegel models using US yields curve monthly data from 1970 to 2000. They found that the quadratic three factor models provide the best in-sample-fit; the families of affine three-factor models and dynamic Nelson-Siegel models perform equally well in the out-of sample forecasting experiment, and that these two models produce better forecast than the quadratic model.

## 6. Conclusion

In this paper, we have reviewed a number of specifications of diffusion based term structure of interest rates models. (A summary of these models is provided in the Appendix). We have presented an overview of the most popular models by means of some general characteristics. From single factor to multi-factor models, forward rate based models and the most recent empirically advocated models; each of these models has its own advantages as well as disadvantages. On the whole, an ideal interest rate model should be theoretically consistent, flexible, well-specified and realistic; it should also provide good in-sample fit (to the data) and out-of-sample forecasting.

## Appendix

### Table: A partial list of the term structure models

| Model | Model Description |
|---|---|
| Merton (1970) | $dr(t) = \mu dt + \sigma dW(t)$ |
| Vasicek (1977) | $dr(t) = \kappa\left(\theta - r(t)\right)dt + \sigma dW(t)$ |
| Dothan (1978) | $dr(t) = \sigma r(t) dW(t)$ |
| Brennan and Schwartz (1979) | $dr(t) = \theta_1\left(r,\lambda,t\right)dt + \sigma_1\left(r,\lambda,t\right)dW_1(t)$ <br> $d\lambda(t) = \theta_2\left(r,\lambda,t\right)dt + \sigma_2\left(r,\lambda,t\right)dW_2(t)$ |
| Courtadon (1982) | $dr(t) = \kappa\left(\theta - r(t)\right)dt + \sigma r(t) dW(t)$ |
| Cox, Ingersoll and Ross (1985) | $dr(t) = \kappa\left(\theta - r(t)\right)dt + \sigma\sqrt{r(t)}dW(t)$ |
| Cox, Ingersoll and Ross (1985b) | $dq(t) = \mu_q(t)dt + \sigma_q(t)dW_q(t)$ <br> $d\pi(t) = \mu_\pi(t)dt + \sigma_\pi(t)dW_\pi(t)$ |
| Ho and Lee (1986) | $dr(t) = \theta(t)dt + \sigma dW(t)$ |
| Nelson and Siegel (1986) | $f(\tau) = \beta_1 + \beta_2 e^{-\lambda\tau} + \beta_3 \frac{1}{\lambda\tau} e^{-\lambda\tau}$ |
| Hull and White (1990) | $dr(t) = \kappa\left(\frac{\theta(t)}{\kappa} - r(t)\right)dt + \sigma dW(t)$ |
| Black and Karasinski (1992) | $d\ln r(t) = \left(\theta(t) - \kappa(t)\ln r(t)\right)dt + \sigma_r(t)dW(t)$ |
| Hull and White (1993) | $dr(t) = \left(\theta(t) - \kappa r(t)\right)dt + \sigma r^{\alpha}(t)dW(t)$ |
| Black, Derman and Toy (1990) | $d\ln r(t) = \left(\theta(t) - \kappa\ln r(t)\right)dt + \sigma_r dW(t)$ |
| Heath, Jarrow and Morton (1992) | $df(t,T) = \sum_{i=1}^{n}\sigma_T(t,T)\int_t^T \sigma(t,\tau)d\tau + \sum_{i=1}^{n}\sigma_T(t,T)dW(t)$ |
| Longstaff and Schwartz (1992) and Chen and Scott (1992) | $dX(t) = \left(a - bX(t)\right)dt + c\sqrt{X(t)}dW_1(t)$ <br> $dY(t) = \left(d - eY(t)\right)dt + f\sqrt{Y(t)}dW_2(t)$ |
| Fong and Vasicek (1991, 1992a, 1992b) | $dr(t) = \beta\left(\bar{r} - r(t)\right)dt + \sqrt{v(t)}dW_1(t)$ <br> $dv(t) = \varsigma\left(\bar{v} - v(t)\right)dt + \xi\sqrt{v(t)}dW_2(t)$ |
| Chen (1996) | $dr(t) = \kappa\left(\theta(t) - r(t)\right)dt + \sqrt{\sigma(t)}\sqrt{r(t)}dW_2(t)$ <br> $d\theta = v\left(\bar{\theta} - \theta(t)\right)dt + \xi\sqrt{\theta}dW_1(t)$ <br> $d\sigma(t) = \mu\left(\bar{\sigma} - \sigma(t)\right)dt + \eta\sqrt{\sigma}dW_3(t)$ |
| Duffie and Kan (1996) | $d\mathrm{X}(t) = \kappa(\theta - \mathrm{X}(t)dt + \Sigma\sqrt{S(t)}dW(t)$ <br> $r(t) = \rho_0 + \rho_X\,'\mathrm{X}(t)$ |

| | |
|---|---|
| Ahn et al. (2002) | $dX(t) = \kappa(\theta - X(t)dt + \Sigma dW(t)$<br>$r(t) = \alpha + \beta' X(t) + X(t)'\psi X(t)$ |
| Christensen et al. (2007) | $dX(t) = \kappa(\theta - X(t)dt + \Sigma dW(t)$<br>$r(t) = L(t) + S(t)$ |

## References

[1] Ahn D-H, Dittmar RF and Gallant RA. 2002,' Quadratic term structure models: theory and evidence', Review of Financial Studies 15(1): 243–288.

[2] Artzner P., Delbean F. 1989, 'Term structure of interest rates: the martingale approach', Advanced in Applied Mathematics, vol.10, 99.95-129.

[3] Backus, D., S. Foresi, and C. Telmer, 2001, 'Affine term structure models and the forward premium anomaly', Journal of Finance, 56, 279-304.

[4] Bank for International Settlements, Monetary and Economic Department 2005, 'Zero-coupon yield curves: Technical documentation', BIS Papers, No.25.

[5] Bjork, T. and B.J. Christensen, 1999. 'Interest rate dynamics and consistent forward rate curves', Mathematical Finance, 9(4), 323-348.

[6] Bliss, R., 1997a. 'Movements in the term structure of interest rates', Economic Review, Federal Reserve Bank of Atlanta, 82, 16-33.

[7] Black F., Derman E. and Toy W. 1990, 'A one factor model of interest rates and its application to treasury bond options', Financial Analysts Journal, pp.33-39.

[8] Black F. and Karasinski P. 1991, 'Bond and option pricing when short rates are lognormal', Financial Analysts Journal, pp.52-59.

[9] Brennan M.J. and Schwartz E. S. 1977, 'Saving bonds, retractable bonds and callable bonds', Journal of Financial Economics, vol. 5, no. 1, pp. 67-88.

[10] Brennan M.J. and Schwartz E. S. 1979, 'A continuous time approach to the pricing of bonds', Journal of Banking and Finance, vol. 3, pp. 135-155.

[11] Brennan M.J. and Schwartz E. S. 1982, 'An equilibrium model of bond pricing and a test of market efficiency', Journal of Financial and Quantitative Analysis, vol. 15, pp. 907-929.

[12] Brown R. and Dybvig P.H. 1986, 'The empirical implications of the Cox, Ingersoll, Ross theory of the term structure of interest rates', Journal of Finance, vol. XLI, pp. 617-630.

[13] Carverhill A. 1995, 'A note on the models of Hull and White for pricing options on the term structure', Journal of Fixed Income, pp. 89-96.

[14] Chen L. 1996,'stochastic mean and stochastic volatility: a three factor model of the term structure of interest rates and its application to the pricing of interest rate derivatives', Blackwell publishers.
[15] Chen R.R. and Scott L. 1992, 'Pricing interest rate options in a two factor Cox, Ingersoll and Ross model of the term structure', Review of Financial Studies, vol. 5, pp.613-636.
[16] Christensen, J. H. E., F.X. Diebold and G..D. Rudebusch, 2007, 'The affine arbitrage-free class of Nelson-Siegel term structure models', American Economic Review, 95(2), 415-420.
[17] Christensen, J. H. E., F.X. Diebold and G..D. Rudebusch, 2008, 'An arbitrage-free generalized Nelson-Siegel term structure model', Econometrics Journal, 01, 1-31.
[18] Constantinides, G.1992, 'A Theory of the Nominal Structure of Interest Rates,' Review of Financial Studies, 5, 531-552.
[19] Courtadon G. 1982, ' A theory of the nominal structure of interest rates', Review of Financial Studies, vol. 5 ,pp. 531-552.
[20] Cox, J., J. Ingersoll, and S. Ross, 1985a, 'An Intertemporal General Equilibrium Model of Asset Prices,' Econometrica, 53, 363-384.
[21] Cox, J., J. Ingersoll, and S. Ross, 1985b, 'A Theory of the Term Structure of Interest Rates,' Econometrica, 53, 385-406.
[22] Dai, Q., and K. Singleton, 2000, 'Specification Analysis of Affine Term Structure Models,' Journal of Finance, 55, 1943-1978.
[23] Dai, Q. and K. Singleton, 2001, 'Expectations Puzzles, Time-Varying Risk Premia, and Dynamic Models of the Term Structure,' working paper, Stanford University.
[24] Diebold, F. X. and C. Li, 2006,'Forecasting the term structure of government bond yields', Journal of Econometrics, 130(2), 337-364.
[25] Diebold, F. X., Ji, L and Yue, V. 2007,'Global yield curve dynamics and interactions: A generalized Nelson-Siegel approach', Working paper, University of Pennsylvania.

[26] Diebold, F. X., Rudebusch, G.D. and Aruoba, S.B, 2006,'The macroeconomy and the yield curve: A dynamic latent factor approach`, Journal of Economics, 131, 309-338.
[27] Dothan Uri L. 1978, 'On the term structure of interest rates', Journal of Financial Economics, vol. 6, pp. 59-69.
[28] Duffee, G. 2000, 'Term Premia and Interest Rate Forecasts in Affine Models,' working paper, University of California, Berkeley.
[29] Duffie, D. and R. Kan, 1996, 'A Yield-Factor Model of Interest Rates,' Mathematical Finance, 6, 379-406.
[30] Duffie, D. and K. Singleton, 1997, "An Econometric Model of the Term Structure of Interest Rates," Journal of Finance, 52, 1287-1321.
[31] Feller 1951, 'Two singular diffusion problems', Ann. Math., vol. 54, pp. 173-181.
[32] Filipovic, Damir, 1999, ' A note on the Nelson-Siegel family', Mathematical Finance, 9(4).
[33] Fong H.G and Vasicek O.A. 1991,'Fixed income volatility management', Journal of Portfolio Management, Summer, pp. 41-46.
[34] Fong H.G and Vasicek O.A. 1992a, 'Omission imporssible', Risk, vol. 5, no.2, pp.62-65.
[35] Fong H.G and Vasicek O.A. 1992b, 'Interest rate volatility as a stochastic factor', Gifford Fong Associates working paper.
[36] Harrison, M., and D. Kreps, 1979, "Martingales and Arbitrage in Multiperiod Security Markets," Journal of Economic Theory, 20, 381-408.
[37] Harrison, J. and Pliska, S. 1981,' Martingales and stochastic integrals in the theory of continuous trading, stochastic processes and their applications', 11, 215–60.
[38] Heath, D., Jarrow, R. and Morton, A. 1990a, 'Bond pricing and the term structure of interest rates: a discrete time approximation', Journal of Financial and Quantitative Analysis, 25, 419–40.
[39] Heath, D., Jarrow, R. and Morton, A. 1990b,' Contingent claim valuation with a random evolution of interest rates, Review of Futures Markets', 9, 54–76.

[40] Heath, D., Jarrow, R. and Morton, A. 1992,' Bond pricing and the term structure of interest rates: a new Methodology', Econometrica, 60, 77–105.
[41] Ho, T.S.Y. and Lee, S.B. 1986,' Term structure movements and the pricing of interest rate contingent claims',
Journal of Finance, 41, 1011–29.
[42] Hull, J. and White, A. 1990,' Pricing interest rate derivative securities', Review of Financial Studies, 3, 573–92.
[43] Hull, J. and White, A. 1993,' One-factor interest rate models and the valuation of interest rate derivative securities', Journal of Financial and Quantitative Analysis, 28, 235–54.
[44] Hull J. and White A, 1994a, 'Numerical procedures for implementing term structure models I: single-factor models', Journal of Derivatives, Autumn, 7-16.
[45] Hull J. and White A, 1994b, 'Numerical procedures for implementing term structure models I: two-factor models', Journal of Derivatives, Winter, 37-49
[46] Jamshidian F, 1991, 'Forward Induction and construction of yield curve diffusion models', Journal of Fixed Income, 1, 62-74
[47] Kim, D.H and Jonathan H. Wright, 2005, 'An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates', Finance and Economics Discussion Series Working paper, 2005-33.
[48] Leippold M, Wu L. 2002,'Asset pricing under the quadratic class', Journal of Financial and Quantitative Analysis 37: 271–295.
[49] Leippold M, Wu L. 2003,' Design and estimation of quadratic term structure models', European Finance Review 7:47–73.
[50] Leippold M, Wu L. 2007, 'Design and estimation of multi-currency quadratic models', Review of Finance 11:167–207.
[51] Longstaff, F. 1989, 'A Nonlinear General Equilibrium Model of the Term Structure of Interest Rates', Journal of Finance, 23, 1259-1282.
[52] Longstaff, F. and E. Schwartz, 1992, "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibrium Model," Journal of Finance, 47, 1259-1282.

[53] Merton R.C. 1973, ‘Theory of rational option pricing’, Bell Journal of Economics and Management Science, vol. 4, pp. 141-183.
[54] Miltersen, K., Sandmann, K. and Sondermann, D. 1997,’ Closed form solutions for term structure derivatives with lognormal interest rates’, Journal of Finance, 52, 409–30.
[55] Nelson, C.R, and A.F. Siegel, 1987, ’Parsimonious modeling of yield curves’, The Journal of Business, 60(4), 473-489.
[56] Rendleman R.and Batter B. 1980, ‘The pricing of options on debt securities’, Journal of Financial and Quantitative Analysis, vol. 15, pp. 11-24.
[57] Richard S. 1978, ‘An arbitrage model of the term structure of interest rates’, Journal of Financial Economics, vpl.6, pp. 33-57.
[58] Vasicek, O. 1977, ‘An equilibrium characterization of the term structure’, Journal of Financial Economics, 5, 177-88.

**A Joint Affine Arbitrage-free Nelson-Siegel (3+1) Factor Model of the Term Structure and Exchange Rates**


## Abstract

This paper extends the three-factor affine arbitrage-free Nelson-Siegel term structure model to a two-currency environment by assuming the level factor be the currency-specific factor for each country and common slope and curvature factors for both countries. The pricing kernel is the building block of theories of fixed-income securities, in the sense that a description of the kernel is sufficient to characterize prices of risk-free bonds. Therefore, we decompose the pricing kernel into two independent portions to join the interest rate term structure and the currency movement in one framework. One portion contains three factors that model the affine arbitrage-free Nelson-Siegel term structure of interest rate. The other portion contains one factor that captures the effect of the currency movement, which is independent of the term structure. This paper derives the joint two-currency affine arbitrage-free Nelson-Siegel (3+1) factor structure model that incorporates the properties of interest rate term structure and foreign exchange rates simultaneously within one arbitrage-free framework.



**Key words:** affine arbitrage-free Nelson-Siegel model, risk premium, pricing kernel, exchange rate

## Introduction

The term structure of interest rates measures the relationship among the yields on default-free securities that differ only in their term to maturity. Understanding and modeling the term structure of interest rates represents one of the most challenging topics of financial research.

The pricing of the bonds is made using the risk-neutral probability measure. Under the risk–neutral probability measure, the no-arbitrage condition is satisfied. This is a fundamental element, which assures that the model is consistent. The no-arbitrage condition provides the foundation for a large literature on arbitrage-free models that starts with the models of Vasicek (1977) and Cox, Ingersoll and Ross (1985). These frameworks specify the risk-neutral evolution of the underlying yield curve factors and provide an explicit definition of the term structure risk premium which reflects the relative price of different maturities across time.

The affine specification of the arbitrage-free term structure models, originally introduced by Duffie and Kan (1996), classified by Dai and Singleton (2000) and extended to the essentially affine specification by Duffee (2002), stand out as the most popular class due to its analytical tractability. Yields are linear functions of underlying latent state variables (or latent factors) with factor loadings that can be calculated via (Riccati) ordinary differential equations. However, the canonical affine arbitrage-free model exhibits very poor empirical time-series performance; especially when forecasting future yields (Duffee, 2002). The main reason that causes these empirical problems is the over-parameterization of the underlying model.

In contrast to the arbitrage-free term structure models, the dynamic Nelson-Siegel model introduced by Diebold and Li (2006), which builds on Nelson and Siegel (1987), provides a parsimonious framework that corresponds to a modern three-factor model of time-varying level, slope and curvature. Empirically, the dynamic Nelson-Siegel model fits both cross-section and time series of yields remarkably well and performs better in out-of-sample forecasting than the random walk and a large set of time series models applied directly to yields, as well as slope regression models (Diebold and Li, 2006). Unfortunately, this model class theoretically does not preclude arbitrage opportunities.

The affine arbitrage-free Nelson-Siegel model developed by Christensen, Diebold and Rudebusch (2007) combines the best of both the arbitrage-free and the dynamic Nelson-Siegel yield curve traditions. It maintains the affine arbitrage-free modeling tradition; the Nelson-Siegel structure helps to identify the latent level, slope and curvature factors, so the affine arbitrage-free Nelson-Siegel model can be easily and robustly estimated via the Kalman filter. Moreover, the affine arbitrage-free Nelson-Siegel model exhibits superior empirical forecasting performance. The application of affine arbitrage-free Nelson-Siegel model in the literature is quite limited, especially in an international context. Thus, the objective of this paper is to extend the affine arbitrage-free Nelson-Siegel model to a foreign currency market.

The challenge of international term structure modeling is to incorporate the properties of interest rate term structure and foreign exchange rates within an arbitrage-free framework. A sizeable research literature has analyzed interest rates under the international context. The early version of modeling exchange rates movements as diffusion processes are based on geometric Brownian motion with constant exchange rate volatility, along with constant interest rates. (Biger and Hull, 1983) (Garman and Kohlhagen, 1983). The recent literatures often models the interest rate and the exchange rate movement with the same set of state variables, the exchange rate movement between any two countries is assumed to be controlled by the term structure of interest rates in the two countries. These studies are such as Saá-Requejo (1993), Dewacher and Maes (2001), Backus, Foresi and Telmer (2001) and Ahn (2004). In contrast to these studies, a few studies such as Brandt and Santa-Clara (2001) and Han and Hammond (2003) allow independent exchange rates movements. Unfortunately, the model either introduces an internal inconsistency (Brandt and Santa-Clara 2001) or parametric constraints have to be imposed on the exchange rate dynamics to preclude arbitrage opportunities (Han and Hammond, 2003).

In this paper, we present the affine arbitrage-free Nelson-Siegel term structure model in a two-currency environment by allowing an independent currency movement that guarantees internal consistency (fundamental asset pricing relation holds) without imposing any constraints on the exchange rate dynamics. The independent currency movement factor is captured by a martingale component in the pricing kernel that is orthogonal to the pricing of the term structure of interest rate. Thus, the innovations of the interest rate and currency movement are uncorrelated with each other. The exchange rate risk premium in turn comprises of two independent components, one is driven from the interest rate risk and the other one is driven by the currency risk factor. Therefore, this joint (3+1) factor structure model releases the tension between the exchange rate movement and the term structure of interest rates

The remainder of the paper is organized as follows: section 1 briefly reviews the original affine arbitrage-free Nelson-Siegel framework; section 2 elaborates the relation of the pricing kernel to interest rates and exchange rates; section 3 is the assumptions of the (3+1) factor structure to join interest rates and exchange rates; we derive a two-currency affine arbitrage-free Nelson-Siegel (3+1) factor structure model in section 4 and the paper concludes in section 5.

## 1. The affine arbitrage-free Nelson-Siegel model

Within the literature on arbitrage-free term structure models, the affine class expounded by Duffie and Kan (1996) has become very popular. For models of this family, bond yields are affine functions of the driving state variables. This property follows from a linear state process and an adequately chosen pricing kernel.

In this section, we briefly review the dynamic Nelson-Siegel and the affine arbitrage-free Nelson-Siegel framework that preserves the Nelson-Siegel factor loadings structure.

### 1.1 The dynamic Nelson-Siegel framework

The original Nelson-Siegel model provides a parsimonious parameterization of the instantaneous forward rate curve given as follows:

$$f(\tau) = \beta_1 + \beta_2 e^{-\lambda\tau} + \beta_3 \frac{1}{\lambda\tau} e^{-\lambda\tau} \tag{1}$$

where $f(\tau)$ is the instantaneous forward rate, τ indicates the time to maturity.

The model has four parameters: three $\beta$s are the state variables and $\lambda > 0$ is the speed of convergence of the term structure toward the long-horizon rate.

The zero-coupon yields consistent with the forward rates given by equation (1) and solved using the integral form of the instantaneous forward rates:

$$y(\tau) = \frac{\int_t^T f(s)ds}{\tau} = \beta_1 + \beta_2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \tag{2}$$

The original Nelson-Siegel model is a static representation which is commonly used to fit the yield curve at a point in time. However, to understand the evolution of the bond market over time, a dynamic version is required. Diebold and Li (2006) suggest allowing the $\beta$ parameters to vary over time. In addition, these three parameters are interpreted as time-varying level, slope and curvature factors given their corresponding Nelson-Siegel loadings. Rewriting the static Nelson-Siegel framework, equation (2), gives us the dynamic Nelson-Siegel model:

$$y_t(\tau) = L_t + S_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + C_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \tag{3}$$

where $L_t, S_t, C_t$ represent the time-varying level, slope and curvature factor respectively. The level factor with its loading of 1 has the same impact on the whole yield curve. The term $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right)$, which is the loading on the slope factor, captures short-term movements that mainly affect yield on the short end of the curve. The curvature factor is a medium term factor and its factor loading, $\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right)$, captures how curvy the yield curve is at the medium maturities.

Since the dynamic Nelson-Siegel model does not preclude arbitrage opportunities, it is directly formulated under the empirical physical (also called historical or data-generating) probability measure P. The extension of the dynamic Nelson-Siegel model that is arbitrage free does exist and that is the affine arbitrage-free Nelson-Siegel model which will be elaborated in the next subsection.

## 1.2 The affine arbitrage-free Nelson-Siegel framework

Christensen, Diebold and Rudebusch (CDR, 2007) develop a class of arbitrage-free affine dynamic term structure models that approximate the widely-used Nelson-Siegel yield-curve specification. The theoretical analysis relates this model to the standard stochastic continuous time affine arbitrage-free (AF, hereafter) framework, originally introduced by Duffie and Kan (1996). This affine diffusion process is represented in a filtered probability space $(\Omega, F_t, Q)$, where $\Omega$ is the sample space, $F_t$ is the sigma-field generated by the set of all possible events up to time t. $Q$ indicates the risk-neutral probability measure on the sample space $\Omega$.

Bonds are usually priced with the help of a so-called ‘risk-neutral probability measure $Q$. Under the risk-neutral probability measure, expected excess returns on bonds are zero. Put differently, the expected rate of return on a long bond equals the risk-free rate. The big advantage of pricing bonds in continuous time is Ito’s lemma. Ito’s lemma allows us to turn the problem of solving the conditional expectation in the bond pricing equation into the problem of solving a partial differential equation for the bond price. The model for the yield curve is defined through two elements: the change of measure from risk-neutral to physical; and the short rate dynamics.

Here we summarize the structure of the affine term structure with three state variables and Gaussian process under the risk-neutral probability measure $Q$:

A1. The instantaneous risk-free rate is assumed to be an affine function of the state variables $X_t$:

$$r_t = \rho_0 + \rho_1 ' X_t \tag{4}$$

where $\rho_0 \in R, \rho_1 \in R^3$ are the parameters, $X_t$ is a vector of state variables.

A2. The state variables $X_t$ follow Gaussian Ornstein-Uhlenbeck processes (also called mean-reverting processes) with a constant variance-covariance matrix $\Sigma$, they can be described by the following system of stochastic differential equations (SDEs):

$$dX_t = \kappa\ (\theta - X_t)dt + \Sigma\ dW_t \tag{5}$$

where $\kappa \in R^{3*3}$ is the mean reversion matrix, $\theta \in R^3$ is a vector of the long-run mean of the interest rates, $W_t$ is a vector of Brownian motions for the state variables. $\Sigma \in R^{3*3}$ is a constant variance-covariance matrix.

A3. The zero-coupon bond prices in the affine AF framework are exponential affine functions of the state variables. This result is proved by Duffie and Kan (1996).

$$P(t,T) = E_t^Q\left[\exp\left(-\int_t^T r_u du\right)\right] = \exp\left[A(t,T) + B(t,T)' X_t\right] \tag{6}$$

where A(t,T) and B(t,T) are the solutions to the ordinary differential equations (ODEs)

$$\begin{aligned} \frac{dA(\tau)}{d\tau} &= -\frac{dA(t,T)}{dt} = -\rho_0 + B(\tau)'\kappa\theta + \frac{1}{2}\sum_{i=1}^{3}\left(\Sigma' B_i(\tau) B_i(\tau)'\Sigma\right) \\ \frac{dB(\tau)}{d\tau} &= -\frac{dB(t,T)}{dt} = -\rho_1' - \kappa' B(\tau) \end{aligned} \tag{7}$$

where $\tau = T - t$ is the time to maturity.

A4. The zero-coupon yields are then obtained by

$$y_t = -\frac{\ln P(\tau)}{\tau} = -\frac{B(\tau)}{\tau} X_t - \frac{A\ (\tau)}{\tau} \tag{8}$$

Comparing this affine AF yield to the dynamic Nelson-Siegel yield in equation (3), we observe that the key difference between these two equations is an extra and unavoidable yield-adjustment term, $-\frac{A\ (\tau)}{\tau}$, which depends only on the maturity in equation (8). Thus, it is possible to extend the dynamic Nelson-Siegel to be AF. The target is to find the affine AF model with factor loadings $\frac{B(\tau)}{\tau}$ in equation (8) that exactly matches the Nelson-Siegel ones in equation (3).

CDR proposed an assumption that the instantaneous risk-free rate is the sum of the level and slope factor to extract the Nelson-Siegel factor loadings from the affine AF framework and develop the so-called affine arbitrage-free Nelson-Siegel (AFNS, hereafter) model.

The framework of the affine AFNS model is described below where B1- B4 are corresponding to A1-A4 in the affine AF framework under the risk-neutral probability measure and B5-B8 are more specific to the AFNS model.

B1. The three state variables in the AFNS model are defined as level, slope and curvature; they can be denoted in a vector form as $X_t = (L_t, S_t, C_t)$ . The instantaneous risk-free rate is assumed to be the sum of level and slope factors:

$$r_t = L_t + S_t \tag{10}$$

where the parameters $\rho_0$ and $\rho_1$ in the affine AF model (A1) are assumed to be $\rho_0 = 0, \rho_1 = (1,1,0)'$ respectively; and $X_t = (L_t, S_t)$.

B2. The state variables follow Gaussian Ornstein-Uhlenbeck processes, they are assumed to be described by the following system of SDEs:

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left( \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} \right) dt + \Sigma \begin{pmatrix} dW_t^L \\ dW_t^S \\ dW_t^C \end{pmatrix} \tag{11}$$

$$dX_t = \kappa\ (\theta - X_t)dt + \Sigma\ dW_t \tag{11'}$$

Equation (11) can be written in short as equation (11'), which is the same form as equation (5) in A2, where $\kappa = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix}$ is the mean reversion matrix that is assumed to contain only zero and parameter $\lambda$ . Parameter $\lambda$ is a positive constant, namely, the mean reversion rate of the curvature and slope factors as well as the scale by which a deviation of the curvature factor from its mean affects the mean of the slope factor. $\Sigma$ is the variance-covariance matrix, it will be specified in B5.

B3. The zero-coupon bond prices in the affine AFNS model are exponential affine functions of the state variables.

$$P(t,T) = \exp\left[A(t,T) + B_1(t,T)L_t + B_2(t,T)S_t + B_3(t,T)C_t\right] \tag{12}$$

where $A(t,T)$ and $B(t,T)$ are the solutions to the following ODEs:

$$\begin{aligned} \frac{dA(\tau)}{d\tau} &= -\frac{dA(t,T)}{dt} = B(\tau)'\kappa\theta + \frac{1}{2}\sum_{i=1}^{3}\left(\Sigma' B_i(\tau) B_i(\tau)'\Sigma\right) \\ \frac{dB(\tau)}{d\tau} &= -\frac{dB(t,T)}{dt} = -\rho_1' - \kappa' B(\tau) \end{aligned} \tag{13}$$

The only parameters in the system of ODEs for $B(t,T)$ functions are $\rho_1$ and $\kappa$, the coefficient of $r_t$ and the mean reversion structure for the state variables under the risk-neutral probability measure $Q$. This explains the crucial assumptions of B1 and B2 for deriving the AFNS model. With the boundary condition that $A(T,T) = B_1(T,T) = B_2(T,T) = B_3(T,T) = 0$, the unique solution to equation (13) is given by

$$\begin{aligned} B_1(\tau) &= -\tau \\ B_2(\tau) &= -\frac{1-e^{-\lambda\tau}}{\lambda} \\ B_3(\tau) &= -\frac{1-e^{-\lambda\tau}}{\lambda} + \tau e^{-\lambda\tau} \end{aligned} \tag{14}$$

$$A(\tau) = B(\tau)'\kappa\theta + \frac{1}{2}\sum_{i=1}^{3}\int_t^T \left(\Sigma' B(u) B(u)'\Sigma\right)_{i,i} du \tag{15}$$

B4. The AFNS zero-coupon yields are then obtained by substituting equation (14) into equation (8),

$$y_t = L_t + S_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + C_t\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) - \frac{A\ (\tau)}{\tau} \tag{16}$$

The factor loadings of each factor are exactly the Nelson-Sigel factor loadings in equation (3). The proof of equation (16) is provided in Appendix A.

B5. Define the value of the yield-adjustment term in the AFNS model.

From equation (15), we observe that the value of the yield-adjustment term depends on the choice of the variance-covariance matrix $\Sigma$ given the value of factor loadings (Bs) obtained in equation (14). CDR identify the AFNS model by fixing $\theta = 0$ under the risk-neutral $Q$-measure. They claim that the maximally flexible AFNS specification can be identified as an upper or lower triangular volatility matrix. With the lower triangular form of the volatility matrix, this gives the correlated-factor specification of the AFNS model. In this paper, we study another specific case of the volatility matrix, a parsimonious diagonal matrix, where

$$\Sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \tag{17}$$

This gives the independent-factor specification of the AFNS model; the yield-adjustment term is given by

$$\begin{aligned} \frac{A(\tau)}{\tau} &= \sigma_{11}^2 \frac{\tau^2}{6} + \sigma_{22}^2 \left( \frac{1}{2\lambda^2} - \frac{1-e^{-\lambda\tau}}{\lambda^3\tau} + \frac{1-e^{-2\lambda\tau}}{4\lambda^3\tau} \right) \\ &+ \sigma_{33}^2 \left( \frac{1}{2\lambda^2} - \frac{e^{-\lambda\tau}}{\lambda^2} - \frac{\tau e^{-2\lambda\tau}}{4\lambda} - \frac{3e^{-2\lambda\tau}}{4\lambda^2} - \frac{2(1-e^{-\lambda\tau})}{\lambda^3\tau} - \frac{5(1-e^{-2\lambda\tau})}{8\lambda^3\tau} \right) \end{aligned} \tag{18}$$

So far the dynamics of the AFNS model has been defined under the risk-neutral $Q$-measure. The existence of the risk-neutral probability measure implies that bond prices are arbitrage–free and the observed yield curve evolution is a result of the yield behavior under a physical P-measure. A transition from the risk-neutral to the physical measure is established via the market price of risk (also called the Sharpe ratio, which is the excess return per standard deviation).

B6. The AFNS model under the physical probability measure P is given by the following measure change (the Girsanoc theorem)[4]:

$$dW_t = dW_t^P + \Gamma_t dt \tag{19}$$

where $W_t$ and $W_t^P$ indicate the Brownian motion under the risk-neutral $Q$-measure and physical P-measure, respectively; and $\Gamma_t$ represents the market price of risk. From equation (19) we observe that the change of measure from the risk-neutral to physical only affects the drift term of the process by $\Gamma_t$, but not the volatility term.

B7. Define the specification of the market price of risk.

CDR apply the essentially affine risk premium specification which is introduced by Duffee (2002) to preserve the affine dynamics under the physical probability measure. Given this specification, the market price of risk follows the essentially affine form as

[4] The Girsanov theorem is explained in Appendix B.

$$\Gamma_t = \gamma_i^0 + \gamma_{ij}^1 X_t \tag{20}$$

where $\gamma_i^0 \in R^3, \gamma_{ij}^1 \in R^{3*3}$.

B8. The dynamics of the AFNS under the P-measure is then has the following general form

$$dX_t = \kappa^P(\theta^P - X_t)dt + \Sigma dW_t^P \tag{21}$$

where $\kappa^P = \kappa - \Sigma\gamma_{ij}^1, \theta^P = (\kappa - \Sigma\gamma_{ij}^1)^{-1}\Sigma\gamma_i^0, \kappa^P\theta^P = \Sigma\gamma_i^0$. We can rewrite the market price of risk in equation (20) as

$$\Gamma_t = \Sigma^{-1}\left(\kappa^P\theta^P + \left(\kappa - \kappa^P\right)X_t\right) \tag{22}$$

For the independent-factor specification, the three state variables are independent with the diagonal volatility matrix. Equation (21) can be written in a matrix form as

$$\begin{pmatrix} dL_t \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ 0 & \kappa_{22}^P & 0 \\ 0 & 0 & \kappa_{33}^P \end{pmatrix}\left(\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix}\right)dt + \begin{pmatrix} \sigma_{11}^P & 0 & 0 \\ 0 & \sigma_{22}^P & 0 \\ 0 & 0 & \sigma_{33}^P \end{pmatrix}\begin{pmatrix} dW_t^{L,P} \\ dW_t^{S,P} \\ dW_t^{C,P} \end{pmatrix} \tag{23}$$

## 2. The pricing kernel

The pricing kernel plays an important role in both interest rate and exchange rate dynamics; therefore we elaborate the relation of the pricing kernel to interest rates and exchange rates, respectively under the no-arbitrage condition in this section.

### 2.1 The pricing kernel and interest rates

The pricing kernel approach is originally derived by Constantinides (1992). The pricing kernel (also called state density deflator, stochastic discount factor), $M_t$, is defined as the unconditional change of measure discounted at the risk-free rate. Mathematically, it can be read as $M_t = \frac{m_t}{\beta_t}$, where $m_t$ (also called the Radon-Nikodym derivative $m_t = \frac{dQ}{dP}$) indicates the unconditional change of measure of the risk-neutral probability measure with respective to the physical probability measure, and $\beta_t$ is the compounded bank account return which has the value of $\exp\left(\int_0^t r_s ds\right)$ and it evolves according to $d\beta_t = r_t \beta_t dt$.

The dynamics of the pricing kernel can be obtained by Ito's lemma[5]:

$$\frac{dM_t}{M_t} = -r_t dt - \Gamma_t dW_t^P \tag{24}$$

Proof: Let $M_t = f(\beta_t, m_t) = \frac{m_t}{\beta_t}$ , by Ito's formula,

$$df(\beta_t, m_t) = \frac{\partial f(\beta_t, m_t)}{\partial \beta_t} d\beta_t + \frac{\partial f(\beta_t, m_t)}{\partial m_t} dm_t$$

$$+\frac{1}{2}\left[\frac{\partial^2 f(\beta_t, m_t)}{\partial \beta_t^2} Var_t(d\beta_t) + 2\frac{\partial^2 f(\beta_t, m_t)}{\partial \beta_t \partial m_t} Cov_t(d\beta_t, dm_t) + \frac{\partial^2 f(\beta_t, m_t)}{\partial m_t^2} Var_t(dm_t)\right]$$

$$= -\frac{m_t}{\beta_t^2} d\beta_t + \frac{1}{\beta_t} dm_t = -\frac{m_t}{\beta_t} r_t d_t + \frac{dm_t}{\beta_t}$$

so $\frac{dM_t}{M_t} = \frac{d(m_t / \beta_t)}{(m_t / \beta_t)} = -r_t dt + \frac{dm_t}{m_t} = -r_t dt - \Gamma_t dW_t^P$

From equation (24), we observe that the drift of the pricing kernel is the risk-free interest rate and the volatility of the pricing kernel is the market price of risk. A no-arbitrage condition restricts the pricing kernel drift to be negative of the instantaneous risk-free interest rate and the diffusion coefficient to be negative of the market price of risk.

In complete markets and the absence of arbitrage opportunities, there exists a unique positive pricing kernel such that the price at time t of a zero-coupon bond maturing at time T and with unit face value under the physical probability measure is:

$$P(t,T) = E_t^P\left[\frac{M_T}{M_t}\right] \tag{25}$$

Equation (24) and equation (25) imply that modeling the term structure of interest rates in an arbitrage-free way can be reduced to modeling the dynamics of the pricing kernel, which in turn can be reduced to specifying the dynamics of the risk-free rate and

[5] Ito's formula is the function of a stochastic process under continuous-time setting. The formula is $dX_t = \mu_t dt + \sigma_t dW_t$; $df(X_t) = f'(X_t)dX_t + \frac{1}{2} f''(X_t)(dX_t)^2$; $Var(dX_t) = (dX_t)^2$

the market price of risk. Since we are applying the AFNS framework in this paper, the risk-free rate is the sum of level and slope factor (see equation (10)); the market price of risk has an essentially affine function of the level, slope and curvature factors (equation (20)).

## 2.2 The pricing kernel and exchange rates

We assume that it is a two-currency based world, the assets can be denominated in either domestic currency or foreign currency, we further assume that there is a separate pricing kernel for each currency, which denotes as $M_t^d$ and $M_t^f$ , respectively. A no-arbitrage condition enforces a consistency of pricing for any security over time. According to equation (25), we now have a dynamic international bond pricing model as:

$$P_t^d(\tau) = E_t^P\left(\frac{M_{t+\tau}^d}{M_t^d}\right),\ P_t^f(\tau) = E_t^P\left(\frac{M_{t+\tau}^f}{M_t^f}\frac{S_t}{S_{t+\tau}}\right) \tag{26}$$

The dynamics of the pricing kernels (equation (24)) with denominating domestic and foreign currency have the following SDEs:

$$\frac{dM_t^d}{M_t^d} = -r_t^d dt - \left(\Gamma_t{}^d\right)' dW_t^{dP},\ \frac{dM_t^f}{M_t^f} = -r_t^f dt - \left(\Gamma_t{}^f\right)' dW_t^{fP} \tag{27}$$

The spot exchange rate, $S_t$ , is defined as the number of units of domestic currency per unit of foreign currency. As noted by Backus, Foresi and Telmer (2001), in a complete market, the exchange rate is uniquely determined by the ratio of the two pricing kernels.

$$S_t = \frac{M_t^f}{M_t^d} \tag{28}$$

By Ito's formula, the dynamic evolution of $S_t$ is given by

$$\frac{dS_t}{S_t} = \left[\left(r_t^d - r_t^f\right) + \Gamma_t^d{}'\left(\Gamma_t^d - \Gamma_t^f\right)\right]dt + \left(\Gamma_t^d - \Gamma_t^f\right)' dW_t^P \tag{29}$$

Proof: By Ito's formula,

$$dS_t = d\left(\frac{M_t^f}{M_t^d}\right) = \frac{\partial f\left(M_t^d, M_t^f\right)}{\partial M_t^f} dM_t^f + \frac{\partial f\left(M_t^d, M_t^f\right)}{\partial M_t^d} dM_t^d$$

$$+\frac{1}{2}\left[\frac{\partial^2 f\left(M_t^d, M_t^f\right)}{\partial\left(M_t^f\right)^2} Var_t(dM_t^f) + 2\frac{\partial^2 f\left(M_t^d, M_t^f\right)}{\partial M_t^f \partial M_t^d}\operatorname{cov}\left(dM_t^f, dM_t^d\right) + \frac{\partial^2 f\left(M_t^d, M_t^f\right)}{\partial (M_t^d)^2} Var_t\left(dM_t^d\right)\right]$$

$$= \frac{1}{M_t^d} dM_t^f - \frac{M_t^f}{\left(M_t^d\right)^2} dM_t^d - \frac{1}{\left(M_t^d\right)^2}\left(\left(\Gamma_t^d\right)'\Gamma_t^f\right) M_t^d M_t^f dt + \frac{M_t^f}{2\left(M_t^d\right)^3}\left(\left(\Gamma_t^d\right)'\Gamma_t^d\right) \quad (30)$$

Substitute (27) into (30), we have

$$dS_t = \frac{1}{M_t^d}\left(-r_t^f dt - \left(\Gamma_t^f\right)' dW_t^{fP}\right) M_t^f - \frac{M_t^f}{\left(M_t^d\right)^2}\left(-r_t^d dt - \left(\Gamma_t^d\right)' dW_t^{dP}\right) M_t^d$$

$$-\frac{1}{\left(M_t^d\right)^2}\left(\left(\Gamma_t^d\right)'\Gamma_t^f\right) M_t^d M_t^f dt + \frac{M_t^f}{\left(M_t^d\right)^3}\left(\left(\Gamma_t^d\right)'\Gamma_t^d\right)\left(M_t^d\right)^2 dt$$

$$= \left(r_t^d - r_t^f + \Gamma_t^d{}'\left(\Gamma_t^d - \Gamma_t^f\right)\right) S_t dt - \left(\Gamma_t^d\right)' S_t dW_t^{dP} - \left(\Gamma_t^f\right)' S_t dW_t^{fP}$$

$$\therefore \frac{dS_t}{S_t} = \left[\left(r_t^d - r_t^f\right) + \Gamma_t^d{}'\left(\Gamma_t^d - \Gamma_t^f\right)\right] dt + \left(\Gamma_t^d - \Gamma_t^f\right)' dW_t^P$$

Equation (29) in a logarithm form is given by

$$ds_t = \left\{\left(r_t^d - r_t^f\right) + \frac{1}{2}\left(\left(\Gamma_t^d\right)'\Gamma_t^d - \left(\Gamma_t^f\right)'\Gamma_t^f\right)\right\} dt + \left(\Gamma_t^d - \Gamma_t^f\right)' dW_t^P \qquad (31)$$

Proof:

$$\because d\log S_t = d\log M_t^f - d\log M_t^d = ds_t = \left(\mu_t - \frac{1}{2}\sigma_t^2\right) dt + \sigma_t dW_t^p,$$

and $\frac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t^p$

By equation (29), $\mu_t = \left(r_t^d - r_t^f\right) + \Gamma_t^d{}'\left(\Gamma_t^d - \Gamma_t^f\right), \sigma_t = \left(\Gamma_t^d - \Gamma_t^f\right)'$

$$\therefore ds_t = \left\{\left(r_t^d - r_t^f\right) + \Gamma_t^d{}'\left(\Gamma_t^d - \Gamma_t^f\right) - \frac{1}{2}\left(\left(\Gamma_t^d - \Gamma_t^f\right)'\left(\Gamma_t^d - \Gamma_t^f\right)\right)\right\} dt + \left(\Gamma_t^d - \Gamma_t^f\right)' dW_t^P$$

$$= \left\{\left(r_t^d - r_t^f\right) + \frac{1}{2}\left(\left(\Gamma_t^d\right)'\Gamma_t^d - \left(\Gamma_t^f\right)'\Gamma_t^f\right)\right\} dt + \left(\Gamma_t^d - \Gamma_t^f\right)' dW_t^P$$

From equation (31) we observe that the drift of the exchange rate dynamics contains two elements: the interest rate differential and a quadratic differential of the risk premia $\frac{1}{2}\left(\left(\Gamma_t^d\right)'\Gamma_t^d - \left(\Gamma_t^f\right)'\Gamma_t^f\right)$; the volatility of the exchange rate is the difference between the two market prices of risk.

So far we have shown how the pricing kernel is involved in both interest rates and exchange rates; also we have reviewed the AFNS term structure model. We are then going to join the AFNS model with the exchange rate through the decomposition of the pricing kernel in the next sections.

### 3. Assumptions for the joint model

***Assumption 1.*** Assume the vector of state variables has a form of (3+1), denote as $Z_t = \left[X_t, U_t\right]$, where $X_t = \left[L_t, S_t, C_t\right]$ is a state variable vector of interest rates in the AFNS model, $U_t$ denotes the currency movement factor, which is a state factor that captures the currency movement. $\left\langle X_t, U_t \right\rangle = 0$, in other terms, these two sets of state variables are independent of each other. Also the state variables are assumed to follow Gaussian Ornstein-Uhlenbeck processes which are controlled by the following SDEs:

$$\begin{aligned} dX_t &= \kappa^X \left(\theta^X - X_t\right)dt + \Sigma^X dW_t^X \\ dU_t &= \kappa^U \left(\theta^X - U_t\right)dt + \sigma^U dW_t^U \end{aligned} \tag{32}$$

where superscript X and U indicate the interest rate factors and the currency movement factor, respectively. Furthermore, the innovations of interest rates and the currency movement are independent of each other. i.e $< dW_t^X, dW_t^U >= 0$.

***Assumption 2.*** Orthogonal decomposition of the pricing kernel

The pricing kernel is assumed to be expressed as follows:

$$M_t = \psi(X_t)\phi(U_t) \tag{33}$$

where $\psi(X_t)$ and $\phi(U_t)$ are some deterministic, continuous function in the real space $R$ such that $\left\langle \psi(X_t), \phi(U_t) \right\rangle = 0$, in other words, these two components are orthogonal to each other. We simplify the notation by letting $\psi(X_t) = \psi_t$ and $\phi(U_t) = \phi_t$ hereafter.

***Assumption 3.*** The independent component of the pricing kernel $\phi_t$ is a martingale (driftless) under the physical probability measure, it has the following dynamics:

$$\frac{d\phi_t}{\phi_t} = -\Gamma_\phi(U_t)dW_t^U = -\Gamma_{\phi t}dW_t^U \tag{34}$$

where $\Gamma_{\phi t}$ is the market price of risk of the currency movement, and this is the diffusion (or shock) of the innovation for the currency factor.

***Assumption 4.*** The currency market price of risk $\Gamma_{\phi t}$ is assumed to have the following essentially affine specification in line with the AFNS market price of risk.

$$\Gamma_{\phi t} = \varphi^0 + \varphi^1 U_t \tag{35}$$

Under Assumption 3, the martingale component $\phi_t$ does not enter the pricing of zero-coupon bonds,

$$P(Z_t, \tau) = E_t^P\left[\frac{\psi_T}{\psi_t}\right]E_t^P\left[\frac{\phi_T}{\phi_t}\right] = E_t^P\left[\frac{\psi_T}{\psi_t}\right] = P(X_t, \tau) \tag{36}$$

The bond price at time t is a function of the state variable vector $X_t$ only; therefore, $\psi_t$ can be labeled as the term structure pricing kernel. The dynamics of this term structure pricing kernel is given by:

$$\frac{d\psi_t}{\psi_t} = -r_t dt - \Gamma_{\psi t} dW_t^P \tag{37}$$

Although $\phi_t$ does not enter the pricing of the term structure, it will determine the exchange rate dynamics between two currencies.

By Assumption 2, we have the (3+1) factor pricing kernel structure as:

$$\frac{dM_t}{M_t} = \frac{d\psi_t}{\psi_t} + \frac{d\phi_t}{\phi_t} \tag{38}$$

Proof:

By Assumption 2, $M_t = \psi(X_t)\phi(U_t)$

$$\begin{aligned}
dM_t &= d\psi(X_t)d\phi(U_t) = d\left(\psi_t \phi_t\right) \\
&= \psi_t d\phi_t + \phi_t d\psi_t + \Gamma_{\psi t}\psi_t \Gamma_{\phi t}\phi_t dW_t^X dW_t^U \\
&= \psi_t\left(-\Gamma_{\phi t}\phi_t dW_t^U\right) + \phi_t\left(-r_t\psi_t dt - \Gamma_{\psi t}\psi_t dW_t^X\right) + 0 \\
&= -\Gamma_{\phi t} M_t dW_t^U + -r_t M_t dt - \Gamma_{\psi t} M_t dW_t^X
\end{aligned}$$

$$\begin{aligned}
\therefore \frac{dM_t}{M_t} &= -r_t dt - \Gamma_{\psi t} dW_t^X - \Gamma_{\phi t} dW_t^U \\
&= \frac{d\psi_t}{\psi_t} + \frac{d\phi_t}{\phi_t}
\end{aligned}$$

## 4. The joint two-currency AFNS (3+1) factor model

### 4.1 A two-currency AFNS model

In section 2, we reviewed the AFNS model and this version actually can be viewed as a single (domestic) currency yield curve. In this subsection we add a foreign yield curve on to the AFNS model and construct a two-currency AFNS term structure model that is used to join the independent currency movement in section 4.2.

The structure of the two-currency AFNS framework is elucidated in line with the AFNS framework (B1-B8).

C1. For each yield curve, there are three latent state variables—level, slope and curvature. The instantaneous risk-free rates are modeled in nominal terms and are assumed to be the sum of a currency-specific (local) level factor for each country and a common slope factor for both countries under the risk-neutral probability measure with that associated currency. The curvature factor is needed to preserve the Nelson-Siegel loadings, so it is assumed to be a common factor as well. $\alpha$ is a scale to measure the relative magnitude of the common slope and curvature factors embedded in the foreign interest rate. Therefore, the domestic and foreign risk-free interest rates have the following form:

$$\begin{aligned} r_t^d &= L_t^d + S_t \\ r_t^f &= L_t^f + \alpha S_t \end{aligned} \tag{39}$$

C2. The state variables follow Gaussian Ornstein-Uhlenbeck processes with a constant volatility matrix Σ. The three factors in each yield curve are assumed to be orthogonal to each other. Therefore, the variance-covariance matrix is a parsimonious diagonal matrix. The state variables of the two-currency AFNS model have the following SDEs under the corresponding currency denominated risk-neutral $Q$-measure:

$$\begin{aligned} dX_t^d &= \kappa^d \left(\theta^d - X_t^d\right)dt + \Sigma^d dW_t^d \\ dX_t^f &= \kappa^f \left(\theta^f - X_t^f\right)dt + \Sigma^f dW_t^f \end{aligned} \tag{40}$$

where d and f indicate domestic and foreign currency, respectively. Equation (40) can be written in an equivalent matrix form as

$$
\begin{pmatrix} dL_t^d \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left( \begin{pmatrix} \theta_1^d \\ \theta_2 \\ \theta_3 \end{pmatrix} - \begin{pmatrix} L_t^d \\ S_t \\ C_t \end{pmatrix} \right) dt + \begin{pmatrix} \sigma_{11}^d & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{L^d} \\ dW_t^S \\ dW_t^C \end{pmatrix}
$$

$$
\begin{pmatrix} dL_t^f \\ dS_t \\ dC_t \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \alpha\lambda & -\alpha\lambda \\ 0 & 0 & \alpha\lambda \end{pmatrix} \left( \begin{pmatrix} \theta_1^f \\ \theta_2 \\ \theta_3 \end{pmatrix} - \begin{pmatrix} L_t^f \\ S_t \\ C_t \end{pmatrix} \right) dt + \begin{pmatrix} \sigma_{11}^f & 0 & 0 \\ 0 & \alpha\sigma_{22} & 0 \\ 0 & 0 & \alpha\sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{L^f} \\ dW_t^S \\ dW_t^C \end{pmatrix} \quad (41)
$$

C3.The zero-coupon bond prices are exponential affine functions of the state variables for each yield which follow the same procedure in B3 (equations (12)-(15)).

C4. The zero-coupon bond yields for each curve are then given by:

$$
\begin{aligned}
y_t^d &= L_t^d + S_t \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A^d(\tau)}{\tau} \\
y_t^f &= L_t^f + \alpha S_t \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \alpha C_t \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) - \frac{A^f(\tau)}{\tau}
\end{aligned} \quad (42)
$$

C5. Since we assume the factors in each yield are orthogonal to each other with the diagonal variance-covariance matrix, the independent-factor specification of the AFNS model applies, the value of the yield-adjustment term for each yield is

$$
\begin{aligned}
\frac{A^d(\tau)}{\tau} &= \left(\sigma_{11}^d\right)^2 \frac{\tau^2}{6} + \left(\sigma_{22}\right)^2 \left( \frac{1}{2\lambda^2} - \frac{1-e^{-\lambda\tau}}{\lambda^3\tau} + \frac{1-e^{-2\lambda\tau}}{4\lambda^3\tau} \right) \\
&\quad + \sigma_{33} \left( \frac{1}{2\lambda^2} - \frac{e^{-\lambda\tau}}{\lambda^2} - \frac{\tau e^{-2\lambda\tau}}{4\lambda} - \frac{3e^{-2\lambda\tau}}{4\lambda^2} - \frac{2(1-e^{-\lambda\tau})}{\lambda^3\tau} - \frac{5(1-e^{-2\lambda\tau})}{8\lambda^3\tau} \right) \\
\frac{A^f(\tau)}{\tau} &= \left(\sigma_{11}^f\right)^2 \frac{\tau^2}{6} + \left(\alpha\sigma_{22}\right)^2 \left( \frac{1}{2\lambda^2} - \frac{1-e^{-\lambda\tau}}{\lambda^3\tau} + \frac{1-e^{-2\lambda\tau}}{4\lambda^3\tau} \right) \\
&\quad + \alpha\sigma_{33} \left( \frac{1}{2\lambda^2} - \frac{e^{-\lambda\tau}}{\lambda^2} - \frac{\tau e^{-2\lambda\tau}}{4\lambda} - \frac{3e^{-2\lambda\tau}}{4\lambda^2} - \frac{2(1-e^{-\lambda\tau})}{\lambda^3\tau} - \frac{5(1-e^{-2\lambda\tau})}{8\lambda^3\tau} \right)
\end{aligned} \quad (43)
$$

C6. The two-currency AFNS model under the physical probability measure is given by the measure change, which has the same process as in B6, for each yield denominated in domestic and foreign currency, respectively.

$$
\begin{aligned}
dW_t^d &= dW_t^{dP} + \Gamma_{\psi t}^d dt \\
dW_t^f &= dW_t^{fP} + \Gamma_{\psi t}^f dt
\end{aligned} \quad (44)
$$

C7. The market price of risk has an essentially affine specification. For each yield curve, we have the market price of risk as

$$\Gamma_{\psi t}^{d} = \gamma_i^{0d} + \gamma_{ij}^{1d} X_t^d$$
$$\Gamma_{\psi t}^{f} = \gamma_i^{0f} + \gamma_{ij}^{1f} X_t^f \qquad (45)$$

where $\gamma_i^{0d} \in R^3, \gamma_{ij}^{1d} \in R^{3*3}, \gamma_i^{0f} \in R^3, \gamma_{ij}^{1f} \in R^{3*3}$

C8. The two-currency AFNS term structure model with independent factor specification has the dynamics under the physical P-measure in a matrix form as

$$\begin{bmatrix} dL_t^d \\ dS_t \\ dC_t \end{bmatrix} = \begin{bmatrix} \kappa_{11}^{dP} & 0 & 0 \\ 0 & \kappa_{22}^{P} & 0 \\ 0 & 0 & \kappa_{33}^{P} \end{bmatrix} \left[ \begin{pmatrix} \theta_1^{dP} \\ \theta_2^{P} \\ \theta_3^{P} \end{pmatrix} - \begin{pmatrix} L_t^d \\ S_t^d \\ C_t \end{pmatrix} \right] dt + \begin{bmatrix} \sigma_{11}^{d} & 0 & 0 \\ 0 & \sigma_{22}^{d} & 0 \\ 0 & 0 & \sigma_{33} \end{bmatrix} \begin{bmatrix} dW_t^{P,L^d} \\ dW_t^{P,S} \\ dW_t^{P,C} \end{bmatrix}$$

$$\begin{bmatrix} dL_t^f \\ dS_t \\ dC_t \end{bmatrix} = \begin{bmatrix} \kappa_{11}^{fP} & 0 & 0 \\ 0 & \alpha\kappa_{22}^{P} & 0 \\ 0 & 0 & \alpha\kappa_{33}^{P} \end{bmatrix} \left[ \begin{pmatrix} \theta_1^{fP} \\ \theta_2^{P} \\ \theta_3^{P} \end{pmatrix} - \begin{pmatrix} L_t^f \\ S_t^f \\ C_t \end{pmatrix} \right] dt + \begin{bmatrix} \sigma_{11}^{f} & 0 & 0 \\ 0 & \alpha\sigma_{22}^{f} & 0 \\ 0 & 0 & \alpha\sigma_{33} \end{bmatrix} \begin{bmatrix} dW_t^{P,L^f} \\ dW_t^{P,S} \\ dW_t^{P,C} \end{bmatrix} \qquad (46)$$

or equivalently, $$dX_t^{dP} = \kappa^{dP}(\theta^{dP} - X_t^d)dt + \Sigma^{dP} dW_t^{dP}$$
$$dX_t^{fP} = \kappa^{fP}(\theta^{fP} - X_t^f)dt + \Sigma^{fP} dW_t^{fP} \qquad (47)$$

where $\kappa^{dP}\theta^{dP} = \Sigma^d \gamma_i^{0d}, \kappa^{fP}\theta^{fP} = \Sigma^f \gamma_i^{0f}$

The two-currency AFNS model has been constructed as two individual yield curves with local level factor move along with the common slope and curvatures; the next step is to jointly model interest rates and currency movement within one arbitrage-free framework.

### 4.2 Joint two-currency AFNS (3+1) factor model

In the two-currency AFNS model, there is a separate pricing kernel for each currency, $M_t^d$ and $M_t^f$, respectively.

Given the orthogonal decomposition in the pricing kernel and equation (38), the dynamics of the pricing kernel of domestic and foreign currency are in the following (3+1) structure:

$$\frac{dM_t^d}{M_t^d} = \frac{d\psi_t^d}{\psi_t^d} + \frac{d\phi_t^f}{\phi_t^f} = -r_t^d dt - \Gamma_t^d(Z_t)' dW_t$$
$$= -r_t^d - \left(\Gamma_{\psi t}^d\right)' dW_t^{X^d} - \left(\Gamma_{\phi t}^f\right)' dW_t^U$$
$$\frac{dM_t^f}{M_t^f} = \frac{d\psi_t^f}{\psi_t^f} + \frac{d\phi_t^d}{\phi_t^d} = -r_t^f dt - \Gamma_t^f(Z_t)'$$
$$= -r_t^f - \left(\Gamma_{\psi t}^f\right)' dW_t^{X^f} - \left(\Gamma_{\phi t}^d\right)' dW_t^U \qquad (48)$$

where $\Gamma_{\phi t}^f$ indicates the foreign currency shock to the domestic pricing kernel, while $\Gamma_{\phi t}^d$ is the domestic currency shock to the foreign pricing kernel.

The domestic local (level) factor affects only the asset bonds in the domestic market, while the foreign local (level) factor influences only the bond values in the foreign market. Any uncertainty associated with the domestic (foreign) local factor does not change the values of foreign (domestic) bonds. However, the domestic (foreign) local factor appears only in the diffusion terms as $\left(\Gamma_{\phi t}^d\right)$ and $\left(\Gamma_{\phi t}^f\right)$ of the SDEs expressed in foreign (domestic) currency in equation (48), but not the drift term. This is because of that the domestic (foreign) local shocks are martingale (Assumption 3).

### 4.3 The exchange rate dynamics in the joint AFNS (3+1) factor structure framework

In section 2.2, we have shown the dynamic evolution of the exchange rate. We apply the Ito's lemma to the joint (3+1) framework and obtain the following new exchange rate dynamics

$$\frac{dS_t}{S_t} = \left[\left(r_t^d - r_t^f\right) + \left(\Gamma_{\psi t}^d\right)'\left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) + \left(\Gamma_{\phi t}^d\right)'\left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)\right] dt$$
$$+ \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) dW_t^X + \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right) dW_t^U \qquad (49)$$

Proof:

Substitute (48) into (30), we have

$$dS_t = \frac{1}{M_t^d}\left(-r_t^f dt - \left(\Gamma_{\psi t}^f\right)' dW_t^{X^f} - \left(\Gamma_{\phi t}^f\right)' dW_t^U\right) M_t^f - \frac{M_t^f}{\left(M_t^d\right)^2}\left(-r_t^d dt - \left(\Gamma_{\psi t}^d\right)' dW_t^{X^d} - \left(\Gamma_{\phi t}^d\right)' dW_t^U\right) M_t^d$$

$$-\frac{1}{\left(M_t^d\right)^2}\left[\left(\Gamma_{\psi t}^d\right)'\Gamma_{\psi t}^f + \left(\Gamma_{\phi t}^d\right)'\Gamma_{\phi t}^f\right] M_t^d M_t^f dt + \frac{M_t^f}{\left(M_t^d\right)^3}\left[\left(\Gamma_{\psi t}^d\right)^2 + \left(\Gamma_{\phi t}^d\right)^2\right]\left(M_t^d\right)^2 dt$$

$$= \left(r_t^d - r_t^f + \Gamma_{\psi t}^d{}'\left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) + \Gamma_{\phi t}^d{}'\left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)\right) S_t dt - \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right)' S_t dW_t^X - \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)' S_t dW_t^U$$

$$\frac{dS_t}{S_t} = \left[\left(r_t^d - r_t^f\right) + \left(\Gamma_{\psi t}^d\right)'\left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) + \left(\Gamma_{\phi t}^d\right)'\left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)\right] dt + \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) dW_t^X + \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right) dW_t^U$$

where $r_t^d - r_t^f = L_t^d - L_t^f + (1-\alpha) S_t$

Equation (49) in a more familiar logarithmic form is:

$$\begin{aligned} ds_t = &\left\{\left(r_t^d - r_t^f\right) + \frac{1}{2}\left(\left(\Gamma_{\psi t}^d\right)'\Gamma_{\psi t}^d - \left(\Gamma_{\psi t}^f\right)'\Gamma_{\psi t}^f\right) + \frac{1}{2}\left(\left(\Gamma_{\phi t}^d\right)'\Gamma_{\phi t}^d - \left(\Gamma_{\phi t}^f\right)'\Gamma_{\phi t}^f\right)\right\} dt \\ &+ \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right)' dW_t^X + \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)' dW_t^U \end{aligned} \tag{50}$$

Proof:

$$\because d\log S_t = d\log M_t^f - d\log M_t^d = ds_t = \left(\mu_t - \frac{1}{2}\sigma_t^2\right) dt + \sigma_t dW_t^p,$$

and $\dfrac{dS_t}{S_t} = \mu_t dt + \sigma_t dW_t^p$

By equation (49), we have

$$\mu_t = \left(r_t^d - r_t^f\right) + \left(\Gamma_{\psi t}^d\right)'\left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) + \left(\Gamma_{\phi t}^d\right)'\left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right),\ \sigma_t = \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right) + \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)$$

$$\begin{aligned} \therefore ds_t = &\left\{\left(r_t^d - r_t^f\right) + \frac{1}{2}\left(\left(\Gamma_{\psi t}^d\right)'\Gamma_{\psi t}^d - \left(\Gamma_{\psi t}^f\right)'\Gamma_{\psi t}^f\right) + \frac{1}{2}\left(\left(\Gamma_{\phi t}^d\right)'\Gamma_{\phi t}^d - \left(\Gamma_{\phi t}^f\right)'\Gamma_{\phi t}^f\right)\right\} dt \\ &+ \left(\Gamma_{\psi t}^d - \Gamma_{\psi t}^f\right)' dW_t^X + \left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)' dW_t^U \end{aligned}$$

The new exchange rate dynamics in my joint AFNS (3+1) factor structure framework are characterized by a drift term that is the sum of the interest rate differential and an exchange rate risk premium. Comparing equation (50) to equation (31), the exchange rate risk premium has an extra term $\frac{1}{2}\left(\left(\Gamma_{\phi t}^d\right)'\Gamma_{\phi t}^d - \left(\Gamma_{\phi t}^f\right)'\Gamma_{\phi t}^f\right)$ which is captured by the independent currency factor. The diffusion term also contains an extra term $\left(\Gamma_{\phi t}^d - \Gamma_{\phi t}^f\right)$ that is the differential of the market price of currency risks. In addition, the independent currency factor $U_t$ does not influence the pricing of interest rate, but it does enter the pricing of the exchange rate between the two countries.

When the absolute value of the domestic market price of risk of a factor risk is higher in the domestic market than in the foreign market, the exchange rate is expected to appreciate and vice versa. This implies that the exchange rate is determined so as to equalize the market prices of factor risks in the two markets. The additional currency factor plays a crucial role in releasing the tension between interest rate and exchange rate movements.

**5. Conclusion**

In this article, we present a two-currency AFNS model with a (3+1) pricing kernel factor structure framework. This joint model can simultaneously model the term structure of interest rates as well as the exchange rates between them. The AFNS form enables us to integrate Gaussian state variables, essentially affine market price of risk and exchange rate dynamics into one arbitrage-free framework. While maintaining internal consistency, the (3+1) factor structure explicitly accounts for the fact that a predominant portion of the currency movement is independent of the movements in the term structure of interest rates in either country.

## Appendix

### A: Proof of the AFNS model

$$P(t,T)=E_t^Q\left[\exp\left(-\int_t^T r_u du\right)\right]=\exp\left[A(t,T)+B(t,T)'X_t\right]$$

The partial differential equation for P(t,T) is given by

$$\frac{\partial P}{\partial t}-\kappa X_t\frac{\partial P}{\partial X_t}+\frac{1}{2}\Sigma'\Sigma\frac{\partial^2 P}{\partial X_t^{\ 2}}-rP=0$$

This gives

$$\frac{dA(t,T)}{dt}+\frac{dB(t,T)}{dt}X_t+\kappa^T X_t B(\tau)+\frac{1}{2}\left(\Sigma' B(\tau)B(\tau)'\Sigma\right)-\rho_1' X_t=0$$

By the matching principle, we obtain the following ODEs:

$$\frac{dB(\tau)}{d\tau}=-\frac{dB(t,T)}{dt}=-\rho_1'-\kappa' B(\tau)$$

$$\frac{dA(\tau)}{d\tau}=-\frac{dA(t,T)}{dt}=B(\tau)'\kappa\theta+\frac{1}{2}\sum_{i=1}^{3}\left(\Sigma' B_i(\tau)B_i(\tau)'\Sigma\right)$$

To solve $B(\tau)$, the system of ODE for $B(\tau)$ is given by $\frac{dB(\tau)}{d\tau}=-\rho_1'-\kappa' B(\tau)$ with boundary condition $B(T,T)=0$

Since $\frac{d}{dt}[e^{\kappa'\tau}B(\tau)]=e^{\kappa'\tau}\frac{dB(\tau)}{dt}+\kappa' e^{\kappa'\tau}B(\tau)=[\rho_1'+\kappa' B(u)]e^{\kappa' u}-\kappa' e^{\kappa' u}B(u)$

integrating both sides gives $\int_0^\tau \frac{d}{du}\left(e^{\kappa' u}B(u)\right)=\int_0^\tau \rho_1' e^{\kappa' u}du$

so we get $B(\tau)=-e^{-\kappa'\tau}\int_0^\tau \rho_1' e^{\kappa' u}du$

Imposing the structure of $\kappa'$ and $\rho_1 : \kappa'=\begin{bmatrix}0&0&0\\0&\lambda&0\\0&-\lambda&\lambda\end{bmatrix}, \rho_1=\begin{bmatrix}1\\1\\0\end{bmatrix}$

we have $e^{\kappa'\tau}=\exp\begin{pmatrix}0&0&0\\0&\lambda\tau&0\\0&-\lambda\tau&\lambda\tau\end{pmatrix}=\exp\begin{pmatrix}0&0&0\\0&\lambda\tau&0\\0&0&\lambda\tau\end{pmatrix}\exp\begin{pmatrix}0&0&0\\0&0&0\\0&-\lambda\tau&0\end{pmatrix}=\exp(a)+\exp(b)$

Since $\kappa'\tau$ has the same real eigenvalue 1 and $b^2=0$ which means b is a nilpotent matrix, exp(b) can be calculated directly from the definition of the exponential matrix

$$\exp(b) = I + b + \frac{b^2}{2!} + ... = \sum_{i=1}^{\infty} \frac{b^i}{i!}$$

$$\exp(b) = \exp\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\lambda\tau & 0 \end{pmatrix} = I + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -\lambda\tau & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\lambda\tau & 1 \end{pmatrix}$$

$$e^{\kappa'\tau} = \begin{pmatrix} e^0 & 0 & 0 \\ 0 & e^{\lambda\tau} & 0 \\ 0 & 0 & e^{\lambda\tau} \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\lambda\tau & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\lambda\tau} & 0 \\ 0 & -\lambda\tau e^{\lambda\tau} & e^{\lambda\tau} \end{pmatrix} .$$

and therefore, $e^{-\kappa^T\tau} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda\tau} & 0 \\ 0 & \lambda\tau e^{-\lambda\tau} & e^{-\lambda\tau} \end{pmatrix}$

Inserting these into the ODEs:

$$B(\tau) = -e^{-\kappa'\tau}\int_0^\tau \rho_1{}' e^{\kappa' u} du = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda\tau} & 0 \\ 0 & \lambda\tau e^{-\lambda\tau} & e^{-\lambda\tau} \end{pmatrix} \int_0^\tau \left[\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{\lambda u} & 0 \\ 0 & -\lambda u e^{\lambda u} & e^{\lambda u} \end{pmatrix}\right] du$$

$$= -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda\tau} & 0 \\ 0 & \lambda\tau e^{-\lambda\tau} & e^{-\lambda\tau} \end{pmatrix} \int_0^\tau \begin{pmatrix} 1 \\ e^{\lambda u} \\ -\lambda u e^{\lambda u} \end{pmatrix} du$$

where $\int_0^\tau 1 du = \tau$

$$\int_0^\tau e^{\lambda u} du = \frac{1}{\lambda} e^{\lambda u} \Big|_0^\tau = -\frac{1-e^{\lambda\tau}}{\lambda}$$

$$\int_0^\tau -\lambda u e^{\lambda u} du = -\tau e^{\lambda\tau} + \frac{1}{\lambda} e^{\lambda u} \Big|_0^\tau = -\tau e^{\lambda\tau} - \frac{1-e^{\lambda\tau}}{\lambda}$$

so $B(\tau) = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda\tau} & 0 \\ 0 & \lambda\tau e^{-\lambda\tau} & e^{-\lambda\tau} \end{pmatrix} \begin{pmatrix} \tau \\ -\frac{1-e^{\lambda\tau}}{\lambda} \\ -\tau e^{\lambda\tau} - \frac{1-e^{\lambda\tau}}{\lambda} \end{pmatrix} = \begin{pmatrix} -\tau \\ -\frac{1-e^{-\lambda\tau}}{\lambda} \\ -\frac{1-e^{\lambda\tau}}{\lambda} + \tau e^{-\lambda\tau} \end{pmatrix}$

The unique solution for this system of ODEs is therefore given by:

$$B^L(\tau) = -\tau$$

$$B^S(\tau) = -\frac{1-e^{-\lambda\tau}}{\lambda}$$

$$B^C(\tau) = -\frac{1-e^{-\lambda\tau}}{\lambda} + \tau e^{-\lambda\tau}$$

The zero-coupon bond yields are then have the form of:

$$y_t = -\frac{\ln P(\tau)}{\tau} = L_t + S_t(\frac{1-e^{-\lambda\tau}}{\lambda\tau}) + C_t(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}) - \frac{A\ (\tau)}{\tau}$$

For the independent- factor AFNS specification, with diagonal variance-covariance matrix, the yield-adjustment term is

$$\frac{A(\tau)}{\tau} = \frac{1}{2\tau}\sum_{i=1}^{3}\int_0^\tau (\Sigma^T B(u)B(u)^T\Sigma)_{j,j}\,du$$

$$= \frac{1}{2\tau}\left\{\sigma_{11}^2\int_0^\tau B^L(u)^2 du + \sigma_{22}^2\int_0^\tau B^S(u)^2 du + \sigma_{33}^2\int_0^\tau B^C(u)du\right\}$$

$$= \sigma_{11}^2\frac{\tau^2}{6} + \sigma_{22}^2\left(\frac{1}{2\lambda^2} - \frac{1-e^{-\lambda\tau}}{\lambda^3\tau} + \frac{1-e^{-2\lambda\tau}}{4\lambda^3\tau}\right)$$

$$+\sigma_{33}^2\left(\frac{1}{2\lambda^2} - \frac{e^{-\lambda\tau}}{\lambda^2} - \frac{\tau e^{-2\lambda\tau}}{4\lambda} - \frac{3e^{-2\lambda\tau}}{4\lambda^2} - \frac{2(1-e^{-\lambda\tau})}{\lambda^3\tau} - \frac{5(1-e^{-2\lambda\tau})}{8\lambda^3\tau}\right)$$

where $\int_0^\tau B^L(u)^2 du = \int_0^\tau s^2 ds = \frac{1}{3}\tau^3$

$$\int_0^\tau B^S(u)^2 du = \int_0^\tau \frac{1-2e^{-\lambda s}+e^{-2\lambda s}}{\lambda^2} ds = \frac{1}{\lambda^2}\left(s + 2\frac{e^{-\lambda s}}{\lambda} + \frac{e^{-2\lambda s}}{2\lambda}\right)\Bigg|_0^\tau$$

$$= \frac{\tau}{\lambda^2} - \frac{2(1-e^{-\lambda\tau})}{\lambda^3} + \frac{1-e^{-2\lambda\tau}}{2\lambda^3}$$

$$\int_0^\tau B^C(u)du = \int_0^\tau\left(s^2e^{-2\lambda s} - 2se^{-\lambda s}\frac{1-e^{-\lambda s}}{\lambda} + \frac{1-2e^{-\lambda s}+e^{-2\lambda s}}{\lambda^2}\right)ds$$

$$\int_0^\tau\left(s^2e^{-2\lambda s}\right)ds = \left(\frac{-\tau^2e^{-2\lambda\tau}}{2\lambda} - \frac{\tau e^{-2\lambda\tau}}{2\lambda} + \frac{1-e^{-2\lambda\tau}}{4\lambda^2}\right) - \frac{2}{\lambda}\int_0^\tau\left(se^{-\lambda s} - se^{-2\lambda s}\right)ds$$

$$= -\frac{2}{\lambda}\left(\frac{-\tau e^{-\lambda\tau}}{\lambda} + \frac{1-e^{-\lambda\tau}}{\lambda^2} + \frac{\tau e^{-2\lambda\tau}}{2\lambda} - \frac{1-e^{-2\lambda\tau}}{4\lambda^2}\right)$$

**B. The Girsanov theorem**

The Girsanov theorem describes the distribution of the stochastic process under a new probability measure. Define this new Ito process $W^Q$ as $dW_t^Q = \Gamma_t dt + dW_t^P$. The Girsanov theorem claims that

1) the only measure under which $dW_t^Q$ is a martingale is given by the Radon-Nikodym derivative of the risk-neutral measure $Q$ with respect to the physical measure P: $m_T = \frac{dQ}{dP}$ with $m_T = \exp\left(-\int_0^T \frac{1}{2}\left(\frac{\mu-r}{\sigma}\right)^2 dt - \int_0^T\left(\frac{\mu-r}{\sigma}\right)dW_t^P\right)$

2) $W_t^Q$ is a Brownian motion under risk-neutral measure $Q$, that is, $Var_t^Q(dW_t^Q) = dt$.

Note: we use $dW_t$ to indicate $dW_t^Q$ throughout this paper.

**References**


[1] Ahn. D. 2004, ‘Common factors and local factors: Implications for term structures and exchange rates’, Journal of Financial and Quantitative Analysis, 39, 69-102.

[2] Bank for International Settlements, Monetary and Economic Department 2005, ’Zero-coupon yield curves: Technical documentation’, BIS Papers, No.25.

[3] Bjork, T., and B.J. Christensen, 1999, ’Interest rate dynamics and consistent forward rate curves’, Mathematical Finance, 9(4), 323-348.

[4] Bliss, R. 1997a, ‘Movements in the term structure of interest rates’, Economic Review, Federal Reserve Bank of Atlanta, 82, 16-33.

[5] Bliss, R. 1997b,’Testing term structure estimation methods’, Advanced in futures and options research, 9, 97-231.

[6] Backus, D., A. Gregory, and C. Telmer, 1993, ‘Accounting for forward rates in markets for foreign currency’, Journal of Finance, 48, 1887-1908.

[7] Backus, D., S. Foresi, and C. Telmer, 2001, ‘Affine term structure models and the forward premium anomaly’, Journal of Finance, 56, 279-304.

[8] Biger N. and J. Hull, 1983, ‘The valuation of currency options’, Financial Management, 12, 24-28 (Spring).

[9] Brandt, M. and Santa-Clara, P. 2002,’ Simulated likelihood estimation of diffusions with an application to exchange rate dynamics in incomplete markets’, Journal of Financial Economics 63, 161–210.

[10] Christensen, J. H. E., F.X. Diebold and G.D. Rudebusch, 2007, ’The affine arbitrage-free class of Nelson-Siegel term structure models’, American Economic Review, 95(2), 415-420.

[11] Christensen, J. H. E., F.X. Diebold and G.D. Rudebusch, 2008, ’An arbitrage-free generalized Nelson-Siegel term structure model’, Econometrics Journal, 01, 1-31.

[12] Constantinides, G. M. 1992, ‘ A theory of the nominal term structure of interest rates’, Review of Financial Studies 5, 531–552.

[13] Coroneo, Laura, Ken Nyholm and Rositsa Vidova-Koleva, 2008, ‘How arbitrage-free is the Nelson-Siegel model?’, European Central Bank, No.874, February, 2008.

[14] Cox, John C., Jonathan E.Ingersoll and Stephen A.Ross, 1985,’A theory of the term structure of interest rates’, Econometrica, 53, 385-407.

[15] Dai, Q. and K. J. Singleton, 2000,’Specification Analysis of affine term structure models’, Journal of Finance, 55, 1943-1978.

[16] Dai, Q. and K. J. Singleton, 2002,’Expectation puzzles, time-varying risk premia, and affine models of the term structure’, Journal of Financial Economics, 63, 415-441.

[17] Dewachter, H. and K. Maes. 2001,’An admissible affine model for joint term structure dynamics of interest rates’, University of Leuven. Photocopy.

[18] Diebold, F. X., and C. Li, 2006,’Forecasting the term structure of government bond yields’, Journal of Econometrics, 130(2), 337-364.

[19] Diebold, F. X., Ji, L and Yue, V. 2007,’Global yield curve dynamics and interactions: A generalized Nelson-Siegel approach’, Working paper, University of Pennsylvania.

[20] Diebold, F. X., Rudebusch, G.D.and Aruoba, S.B, 2006,’The macroeconomy and the yield curve: A dynamic latent factor approach`, Journal of Economics, 131, 309-338.

[21] Duan, Jin-Chuan and Jean-Guy Simonato,1999, ’Estimating and testing exponential-affine term structure models by Kalman filter’, Review of Quantitative Finance and Accounting,13, 111-135.

[22] Duffee, G. R. 2002,’Term premia and interest rate forecasts in affine models’, Journal of Finance, 57(1), 405-443.

[23] Duffie, Darrell, and Rui Kan, 1996,’A yield-factor model of interest rates’, Mathematical Finance, 52, 1287-1321.

[24] Filipovic, Damir, 1999, ’ A note on the Nelson-Siegel family’, Mathematical Finance, 9(4).

[25] Han. B and P.Hammond, 2003, ‘Affine models of the joint dynamics of exchange rates and interest rates’, Working paper, University of Calgary and Stanford University.

[26] Kim, D.H and Jonathan H. Wright, 2005,’An arbitrage-free three-factor term structure model and the recent behavior of long-term yields and distant-horizon forward rates’, Finance and Economics Discussion Series Working paper, 2005-33.

[27] Krippner, Leo, 2006, ‘A theoretically consistent version of the Nelson and Siegel class of yield curve models’, Applied Mathematical Finance, 13(1), 39-59

[28] Nelson, C.R, and A.F. Siegel, 1987,’Parsimonious modeling of yield curves’, The Journal of Business, 60(4), 473-489.

[29] Nielsen, L.T., and J. Saá-Requejo, 1993, ‘Exchange rate and term structure dynamics and the pricing derivative securities’, Working paper, INSEAD.

[30] Piazzesi, M. 2003, ‘Affine Term Structure Models.’, Forthcoming in Handbook of Financial Econometrics, edited by Y. Ait-Sahalia and L. P. Hansen.

[31] Saá-Requejo, J. 1993, ‘The Dynamics and the Term Structure of Risk Premia in Foreign Exchange Markets’, Working paper, INSEAD

[32] Vasicek, O. 1977, ‘An equilibrium characterization of the term structure’, Journal of Financial Economics, 5, 177-88.