Level Dependence in Volatility in Linear-Rational Term Structure Models

Kalind Ramnarayan

A dissertation submitted to the Faculty of Commerce, University of Cape Town, in partial fulfilment of the requirements for the degree of Master of Philosophy.

July 29, 2019

MPhil in Mathematical Finance,
University of Cape Town.

Declaration

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Philosophy to the University of Cape Town. It has not before been submitted for any degree or examination.

Signed by candidate

Kalind Ramnarayan

July 29, 2019

Abstract

The degree of level dependence in interest rate volatility is analysed in the linear-rational term structure model. The linear-rational square-root (LRSQ) model, where level dependence is set a priori, is compared to a specification where the factor process follows CEV-type dynamics which allows a more flexible degree of level dependence. Parameters are estimated using an unscented Kalman filter in conjunction with quasi-maximum likelihood. An extended specification for the state price density process is required to ensure reliable parameter estimates. The empirical analysis indicates that the LRSQ model generally overestimates level dependence. Although the CEV specification captures the degree of level dependence in volatility more accurately, it has a trade-off with analytical tractability. The optimal specification, therefore, depends on the type of model implementation and general economic conditions.

Acknowledgements

I would like to thank Dr Alex Backwell for his excellent supervision. The insightful guidance and assistance he provided is greatly appreciated. I would also like to thank Associate Professor David Taylor, Associate Professor Peter Ouwehand, Associate Professor Thomas McWalter, Obeid Mahomed and AIFMRM for providing a simulating and enjoyable environment for learning and development. Finally, I would like to thank my family for their continuous support and encouragement throughout my academic career.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Interest rate volatility is time varying in nature. One of the reasons this is the case is due to level dependence in volatility — a tendency for interest rate volatility to change when the level of the interest rate itself changes. It is a stylised fact that interest rate volatility is positively related to the level of the rate (Piazzesi, 2010). Furthermore, volatility is more sensitive to changes in the rate when interest rates are low, with the largest degree of level dependence exhibited when rates are near zero (Filipović, Larsson and Trolle, 2017). Capturing these features in a model are important as they have a fundamental role in the valuation of contingent claims and hedging interest rate risk (Chan, Karolyi, Longstaff and Sanders, 1992).

The linear-rational term structure model, developed by Filipović *et al.* (2017), form a class of interest rate models with many interesting and advantageous features. The model is tractable as it has the ability to ensure non-negativity of interest rates by specifying a lower bound as well as to accommodate unspanned factors. Other models have the ability to accommodate one of the above features but rarely both. In addition, the linear rational term structure model can admit a semi-analytical expression to price swaptions, whereas affine models only give rise to approximate solutions.

This dissertation analyses how accurately the linear-rational term structure model specified with a square-root (SQRT) factor process, known as the linear-rational square-root (LRSQ) model, captures the degree of the level dependence in volatility observed in the market. Filipović *et al.* (2017) conduct their empirical work with a SQRT specification, but the model can be given a more general CEV-type specification. This specification is implemented and compared to the baseline LRSQ model.

Parameter estimation is performed via quasi-maximum likelihood in conjunction with an unscented Kalman filter using a British-pound data set. A three-factor model is used in the estimation procedure. The estimated parameters indicate that the degree of level dependence in volatility is significantly different in each specification. The empirical analysis finds that the LRSQ model overestimates level dependence in general. This conclusion requires consideration of the various factors, which each influence the short rate. The degree of level dependence in volatility is found to be roughly the same in the SQRT and CEV specification when the movement in the short rate is due to the first or second factor. However, the SQRT specification significantly overestimates level dependence when the short rate changes due to movement in the third factor. This holds considerable weight as the movement in the first two factors do not heavily impact the level of the short rate relative to movement in the third factor. Finally, although the CEV specification captures level dependence more accurately it has a trade-off with analytical tractability in option pricing.

Chapter 2 gives an overview of the linear-rational framework. This includes a description of the framework's extended specification which is a technique Filipović *et al.* (2017) use to model the market price of risk. Chapter 3 considers the concept of level dependence in volatility in the context of a general multi-factor model as well as the linear-rational model. The CEV specification for the factor process is then specified. Chapter 4 begins by explaining the parameter estimation methodology. A simulation exercise is then performed to illustrate the effectiveness of the estimation technique and finally, the empirical analysis is given. Chapter 5 concludes the dissertation.

Chapter 2

The Linear-Rational Term Structure Model

2.1 The Linear-Rational Term Structure Framework

The term structure of a financial variable refers to how it is related to its maturity (Filipović, 2009). A central example is the term structure of interest rates. The four most well-known approaches to modelling the interest rate market are short rate models (e.g., the Vasicek (1977) model), whole yield curve models (e.g., the Heath, Jarrow and Morton (1992) models), market models (e.g., the Brace, Gatarek and Musiela (1997) approach) and state price density models. The linear-rational term structure model follows the relatively unconventional state price density approach, originally proposed by Constantinides (1992), to modelling the term structure.

The linear-rational term structure model is defined by specifying a state price density process, a positive adapted process, also known as a pricing kernel. The state price density process, denoted $\{\zeta_t\}$, by definition results in the process $\{X_t\zeta_t\}$ being a martingale under the real-world measure $\mathbb{P}$, where $\{X_t\}$ is the price process of a tradable asset. Hence, if an asset has a value of X_T at time-T, then its value at an earlier time-t is given by[1]

$$X_t = \frac{1}{\zeta_t}\mathbb{E}_t^{\mathbb{P}}[X_T\zeta_T]. \tag{2.1}$$

Comparing Equation (2.1) to a more typical risk-neutral formula (discounted risk-neutral expectation) shows how the state price density process involves a measure change from the real-world measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$ as well as a discounting property. Hence, specifying a state price density process is equivalent to jointly specifying the interest rate and change of measure dynamics (this is discussed further in Appendix A.1). The model that is attained by the specification

[1] The following notation represents a $\mathcal{F}_t$ conditional expectation under the real world measure $\mathbb{P}$ assuming a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$.

of the state price density process is arbitrage free, as the existence of a state price density process is equivalent to the existence of a risk-neutral measure.[2]

A zero-coupon bond is a financial instrument that pays the holder a unit-nominal at maturity. Zero-coupon bonds are usually considered to be fundamental financial instruments as many other interest rate derivatives, such as swaps, can be expressed in term of zero-coupon bonds. Assuming a lack of default risk, zero-coupon bond prices are obtained by setting $X_T = 1$ in Equation (2.1). This shows that state price density models, in general, result in zero-coupon bond prices being rational functions, i.e., ratios of expressions involving the state price density process:

$$P(t,T) = \frac{\mathbb{E}_t^{\mathbb{P}}[\zeta_T]}{\zeta_t}, \tag{2.2}$$

where $P(t,T)$ denotes the time-t price of a unit-nominal zero-coupon bond maturing at time-T. In the linear-rational framework, the state price density process is specified by defining a multi-factor process with state space $E \subseteq \mathbb{R}^m$ as follows:

$$dZ_t = K(\theta - Z_t)dt + dM_t, \tag{2.3}$$

where $K \in \mathbb{R}^{m\times m}$, $\theta \in \mathbb{R}^m$ and $\{M_t\}$ is a m-dimensional martingale under $\mathbb{P}$. The state price density process is then defined as a linear function of the factors:

$$\zeta_t = e^{-\beta t}\left(\phi + \psi^\top Z_t\right), \tag{2.4}$$

where $\phi + \psi^\top Z_t > 0$; $\phi, \beta \in \mathbb{R}$ and $\psi \in \mathbb{R}^m$.

The expression for zero-coupon bond prices can now be obtained by using an integrating factor method to solve the stochastic differential equation (SDE) in Equation (2.3). Thus, resulting in

$$\mathbb{E}_t^{\mathbb{P}}[Z_T] = \theta + e^{-K(T-t)}(Z_t - \theta), \tag{2.5}$$

which holds with a matrix exponential for $m > 1$. The linear drift of the factor process in conjunction with the linearity of the state price density results in zero-coupon bond prices being expressed as linear rational functions of the factors. From Equations (2.2), (2.4) and (2.5), it follows that

$$P(t,T) = e^{-\beta(T-t)}\frac{\phi + \psi^\top\theta + \psi^\top e^{-K(T-t)}(Z_t - \theta)}{\phi + \psi^\top Z_t}. \tag{2.6}$$

Equation (2.6) is important as it represents the model price for a zero-coupon bond, the fundamental instrument of interest. Equation (2.6) indicates that if the value of

[2] This follows from the fundamental theorems of asset pricing in continuous time originally outlined by Harrison and Pliska (1981).

Z_t is known then the whole term structure can be determined at time-t. Furthermore, the term structure evolves as the process $\{Z_t\}$ evolves stochastically over time. This expression is critical in the parameter estimation procedure as it is the link between the factor process, $\{Z_t\}$, and market observable prices.

The expression for the short rate in the model can be determined by evaluating the identity $r_t = -\frac{\partial}{\partial T} \log P(t,T) \mid_{T=t}$ which Filipović *et al.* (2017) show to be given by

$$r_t = \beta - \frac{\psi^\top K(\theta - Z_t)}{\phi + \psi^\top Z_t}. \tag{2.7}$$

Equation (2.7) illustrates that the model can ensure non-negativity of interest rates by letting

$$\beta = \sup_{Z \in E} \frac{\psi^\top K(\theta - z)}{\phi + \psi^\top z}. \tag{2.8}$$

Although the linear-rational framework holds for any martingale $\{M_t\}$ in Equation (2.3), the linear-rational framework has mostly been considered when the factor process follows diffusion dynamics:

$$dZ_t = K(\theta - Z_t)dt + \sigma(t, Z_t)dB_t,$$

where $\{B_t\}$ is a m-dimensional standard Brownian motion under $\mathbb{P}$. In this Brownian setting the market price of risk (the Girsanov kernel involved in the change of measure between the real-world $\mathbb{P}$ and the risk-neutral $\mathbb{Q}$) is given as

$$\lambda_t = -\frac{\sigma(t, Z_t)^\top \psi}{\phi + \psi^\top Z_t}. \tag{2.9}$$

Appendix A.1 gives further details, including the derivation of the market price of risk, which follows from the specification of the state price density process. It is important to note that the specification of the state price density controls the degree of risk-aversion in the model.

2.1.1 The Linear-Rational Square-Root Model

The factor process in the linear-rational square-root (LRSQ) model is specified to follow square-root (SQRT) dynamics,

$$dZ_t = K(\theta - Z_t)dt + \text{diag}\left(\sigma_1\sqrt{Z_{1,t}}, ..., \sigma_m\sqrt{Z_{m,t}}\right)dB_t, \tag{2.10}$$

where, as before, $K \in \mathbb{R}^{m\times m}$, $\theta \in \mathbb{R}^m$ and where $\sigma_1, ..., \sigma_m$ are constants. The factor process satisfying Equation (2.10) has a mean reversion property. Specifically, θ and K can be interpreted as the real-world level of mean reversion or long run mean of the factor process and rate of mean reversion respectively.

The LRSQ model, sets $\phi = 1$ and $\psi = \mathbf{1}$ such that the state price density process is given by $\zeta_t = e^{-\beta t}\left(1 + \mathbf{1}^\top Z_t\right)$. Filipović *et al.* (2017) show that, in the above specification, negative rates can be avoided by setting $\beta = \max\{\mathbf{1}^\top K\theta, -\mathbf{1}^\top K_1, ..., -\mathbf{1}^\top K_m\}$ where K_i denotes the ith column of vector K.

2.2 The Extended State Price Density Framework

The market price of risk expressed in Equation (2.9) is an endogenous variable, in the sense that it depends on parameters that play other roles in the model. It cannot be controlled independently of the models parameters and behaviours. For instance, in the LRSQ model defined above, the market price is always nonpositive. This endogeneity can be restrictive. Filipović *et al.* (2017) show that the standard specification is too restrictive to capture observed bond risk premium dynamics. A flexible market price of risk is important to reduce mathematical restrictions in the model as well as allow the model to match real-world facts more accurately. The form of a model's market price of risk has been studied by, for instance, Duffee (2002) and Duarte (2003), who extend the market price of risk specification of the completely affine terms structure (first generation) models, originally characterised by Duffie and Kan (1996), to ensure that the model captured a certain set of stylised facts.

Filipović *et al.* (2017) propose an extension to the above framework that results in a more flexible market price of risk. The extension of the linear-rational term structure framework involves defining the above framework in terms of an auxiliary measure $\mathbb{A}$ equivalent to the real-world measure $\mathbb{P}$. Specifically, the framework initially specifies the state price density process under $\mathbb{A}$, denoted $\{\zeta_t^{\mathbb{A}}\}$, such that $\{X_t \zeta_t^{\mathbb{A}}\}$ is an $\mathbb{A}$-martingale. The valuation formula is, therefore, given by:

$$X_t = \frac{1}{\zeta_t^{\mathbb{A}}} \mathbb{E}_t^{\mathbb{A}}\left[X_T \zeta_T^{\mathbb{A}}\right], \tag{2.11}$$

where

$$\zeta_t^{\mathbb{A}} = e^{-\beta t}\left(\phi + \psi^\top Z_t\right),$$

and

$$dZ_t = K(\theta - Z_t)dt + dM_t^{\mathbb{A}},$$

where $\{M_t^{\mathbb{A}}\}$ is a martingale under $\mathbb{A}$. The expectation in valuation Equation (2.11) is now taken under $\mathbb{A}$ and the endogenous market price of risk given in Equation (2.9) is now understood with respect to $\mathbb{A}$. The zero-coupon bond price given by Equation (2.6) is still valid, but the parameters in the expression are now appearing in the $\mathbb{A}$-dynamics of the factor process.

In order to create a link to the real-world measure $\mathbb{P}$ a Girsanov kernel, denoted $\{\delta_t\}$, must be defined in order to specify the change of measure. The real-world measure is then formally specified by the Radon-Nikodým process of $\mathbb{P}$ with respect to $\mathbb{A}$:[3]

$$\mathbb{E}_t^{\mathbb{A}}\left[\frac{d\mathbb{P}}{d\mathbb{A}}\right] = e^{\int_0^t \delta_s dB_s^{\mathbb{A}} - \frac{1}{2}\int_0^t ||\delta_s||^2 ds}.$$

It follows that [4]

$$\lambda_t = -\frac{\sigma(t, Z_t)^\top \psi}{\phi + \psi^\top Z_t} + \delta_t. \tag{2.12}$$

The expression for the market price of risk under $\mathbb{P}$ is as expected as the market price of risk is shifted by the Girsanov kernel in a change of measure. Equation (2.12) illustrates that the extended specification results in a more parametrised market price of risk under $\mathbb{P}$. There is now exogenous control as the parameters controlling δ_t now give a way, independent of the other parameters, to govern the market price of risk. Therefore, allowing more flexibility in the specification of risk-aversion.

[3] The specification of the change of measure is done in more detail in Appendix A.3.

[4] the market price of risk under $\mathbb{P}$ is explicitly derived in Appendix A.4.

Chapter 3

Level Dependence in Volatility

Level dependence in interest rate volatility refers to the tendency for interest rate volatility to change when the interest rate itself changes. It is a stylised fact that the volatility of interest rates are positively correlated with the level of the rates themselves (Piazzesi, 2010). The positive relation between interest rate levels and volatility is also documented by Trolle and Schwartz (2014), Chan *et al.* (1992), Ait-Sahalia (1996), Conley, Hansen, Luttmer and Scheinkman (1997) and Stanton (1997). Kim and Singleton (2012) conducted an empirical analysis in a low interest rate environment and found a degree of level dependence larger than found in previous studies. Filipović *et al.* (2017) then showed, through conditional regression, that volatility has stronger level dependence when interest rates are low than when they are high (i.e., that volatility is more sensitive to a change in the interest rate when rates are low).

Chan *et al.* (1992) indicate that level dependence is an important feature in term structure modelling as it has fundamental significance in valuing contingent claims and hedging interest rate risk. The concept of level dependence in interest rate volatility can be illustrated by considering two classical single-factor short rate models. The short rate process $\{r_t\}$ proposed by Vasicek (1977) satisfies

$$dr_t = \kappa(\theta - r_t)dt + \sigma dB_t,$$

where κ, θ and σ are constants. The volatility function in the above SDE is a constant, therefore, the model does not incorporate level dependence as volatility in the model is insensitive to changes in the short rate. A more plausible short rate model with regards to level dependence was proposed by Cox, Ingersoll and Ross (1985), where $\{r_t\}$ satisfies

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t.$$

The volatility term in the above SDE is an increasing concave function with respect to the short rate. The gradient of the volatility function is positive implying a positive degree of level dependence. Furthermore, the gradient is steep for low values of r_t and then gradually flattens as r_t increases. The CIR model therefore captures both stylised facts with regards to level dependence in interest rate volatility.

The only source of noise/volatility in the linear-rational term-structure originates from the factor process. Hence, level dependence in volatility with regards to the interest rate is captured through the local volatility function, $\sigma(t, Z_t)$, of the factor process (refer to the Brownian linear-rational model's dynamics on page 5). This motivates adding additional flexibility into the local volatility function in order to analyse how accurately the SQRT specification captures level dependence in interest rate volatility.

3.1 The CEV Specification

The constant elasticity of variance (CEV) model is a diffusion model introduced by Cox (1975). A mean-reverting factor process with CEV-type dynamics, similar to a specification originally considered by Andersen and Piterbarg (2007), is characterised by the following SDE:

$$dZ_t = K(\theta - Z_t)dt + \text{diag}\left(\sigma_1 Z_{1,t}^{\alpha_1}, ..., \sigma_m Z_{m,t}^{\alpha_m}\right) dB_t, \tag{3.1}$$

where $0 < \alpha_i < 1$ for $i = 1, ..., m$. The local volatility function in Equation (3.1) comprises of constant volatility components, σ_i, and level dependent components, $Z_{i,t}^{\alpha_i}$, where α_i is interpreted as the degree of level dependence. The interpretation follows from the fact that α_i impacts the gradient of the local volatility function, i.e., the sensitivity of volatility to changes in the level of the factors. A smaller α_i results in a higher level dependence for very low factor values and lower level dependence for the other factor values. The magnitude of the local volatility function is also influenced by α_i; however, this is not the primary role of α_i. The general magnitude volatility is captured though the σ_i parameters.

The incorporation of level dependence in interest rate volatility can be slightly more complicated in multi-factor models. Firstly, notice that we are specifying factor dynamics rather than short rate dynamics explicitly. Therefore, the correlation between the short rate and the factors may not be obvious. Secondly, the volatility of the short rate is now a function of the factors. Thus, complicating the interpretation of the relation between the volatility of the short rate and its current level. In affine term structure models, the short rate is usually defined to be a linear function of the factors. This results in a monotone correlation between the short rate and each factor as well as the variance of the short rate being a linear function of the factors. This allows for a more straight-forward incorporation of level dependence.

In the linear-rational term-structure model, the partial derivatives of the short rate with respect to the factors are not necessarily monotone functions (illustrated by Appendix A.5). This complicates how level dependence in volatility of the factors are translated to level dependence with regards to the interest rate. However, Filipović *et al.* (2017) find a robust estimation feature in the linear-rational term structure model, where the link between the factors and the short rate becomes more straightforward. A detailed explanation is given in the empirical analysis of this dissertation.

The CEV dynamics generalise those in the LRSQ model, which are recovered when $\alpha_i = \frac{1}{2}$ for $i = 1, ..., m$. This illustrates that the LRSQ model sets the degree of level dependence a priori whereas the CEV specification allows for a more flexible degree of level dependence. Following Filipović *et al.* (2017), we set the mean reversion matrix K in Equation (3.1) as lower bi-diagonal. Equation (3.1) can then be written in scalar form as follows:

$$dZ_{i,t} = \begin{cases} K_{11}(\theta_1 - Z_{1,t})dt + \sigma_1 Z_{1,t}^{\alpha_1} dB_{1,t}, & \text{for } i = 1; \\ [K_{i(i-1)}(\theta_{(i-1)} - Z_{(i-1),t}) + K_{ii}(\theta_i - Z_{i,t})]dt + \sigma_i Z_{i,t}^{\alpha_i} dB_{i,t}, & \text{for } i = 2, ..., m. \end{cases} \tag{3.2}$$

The above specification can be generalised to incorporate the extended state price density specification described in Section 2.2. The specification in Equation (3.2) becomes the $\mathbb{A}$-dynamics of the factor process. The real-world dynamics of the factor process is now given by

$$dZ_{i,t} = \begin{cases} [K_{11}(\theta_1 - Z_{1,t}) + \sigma_1 Z_{1,t}^{\alpha_1}\delta_{1,t}]dt + \sigma_1 Z_{1,t}^{\alpha_1} dB_{1,t}^{\mathbb{P}}, & \text{for } i = 1; \\ [K_{i(i-1)}(\theta_{(i-1)} - Z_{(i-1),t}) + K_{ii}(\theta_i - Z_{i,t}) + \sigma_i Z_{i,t}^{\alpha_i}\delta_{i,t}]dt & \\ +\sigma_i Z_{i,t}^{\alpha_i} dB_{i,t}^{\mathbb{P}}, & \text{for } i = 2, ..., m. \end{cases}$$

In the LRSQ model, Filipović *et al.* (2017) specifies (refer to the Radon-Nikodým process equation in Section 2.2)

$$\delta_t = \left(\delta_1 \sqrt{Z_{1,t}}, \ldots, \delta_m \sqrt{Z_{m,t}}\right)^\top,$$

to ensure that $\mathbb{P}$-dynamics for the factor process remains a square-root process. Generalising this approach to the CEV specification, define

$$\delta_t = \left(\delta_1 Z_{1,t}^{-\alpha_1+1}, \ldots, \delta_m Z_{m,t}^{-\alpha_m+1}\right)^\top,$$

to ensure that the $\mathbb{P}$-dynamics and $\mathbb{A}$-dynamics both follow mean-reverting CEV-type dynamics. This specification results in the $\mathbb{P}$-dynamics given as:

$$dZ_{i,t} = \begin{cases} \left[\bar{K}_{11}(\bar{\theta}_1 - Z_{1,t})\right] dt + \sigma_1 Z_{1,t}^{\alpha_1} dB_{1,t}^{\mathbb{P}} & \text{for } i = 1; \\ \left[K_{i(i-1)}(\theta_{(i-1)} - Z_{(i-1),t}) + \bar{K}_{ii}(\bar{\theta}_i - Z_{i,t})\right] dt + \sigma_i Z_{i,t}^{\alpha_i} dB_{i,t}^{\mathbb{P}}, & \text{for } i = 2, ..., m, \end{cases}$$

where

$$\bar{K}_{ii} = K_{ii} - \sigma_i \delta_i$$

and

$$\bar{\theta}_i = \frac{\theta_i K_{ii}}{K_{ii} - \sigma_i \delta_i}.$$

The above can be expressed in matrix form as

$$dZ_t = \hat{K}(\hat{\theta} - Z_t)dt + \text{diag}\left(\sigma_1 Z_{1,t}^{\alpha_1}, ..., \sigma_m Z_{m,t}^{\alpha_m}\right) dB_t^{\mathbb{P}}, \tag{3.3}$$

where $\hat{K}_{ii} = \bar{K}_{ii}$, $\hat{K}_{ij} = K_{ij}$ for $i \neq j$ and $\hat{\theta} = \hat{K}^{-1}K\theta$. The $\mathbb{P}$-dynamics of a two-factor model[1] is explicitly expressed as

$$\begin{bmatrix} dZ_{1,t} \\ dZ_{2,t} \end{bmatrix} = \begin{bmatrix} \bar{K}_{11} & 0 \\ K_{21} & \bar{K}_{22} \end{bmatrix} \left[\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \end{bmatrix} - \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} \right] dt + \begin{bmatrix} \sigma_1 Z_{1,t}^{\alpha_1} & 0 \\ 0 & \sigma_2 Z_{2,t}^{\alpha_2} \end{bmatrix} \begin{bmatrix} dB_{1,t}^{\mathbb{P}} \\ dB_{2,t}^{\mathbb{P}} \end{bmatrix},$$

where

$$\hat{\theta}_1 = \bar{\theta}_1$$

and

$$\hat{\theta}_2 = \bar{\theta}_2 - \frac{\sigma_1 \delta_1 K_{21} \theta_1}{\bar{K}_{11} \bar{K}_{22}}.$$

An additional reason to use the extended specification is that the parameter estimation results are skewed under the initial framework. Specifically, the long-run mean parameters θ_i given in Equation (3.1) are falsely estimated. For example consider a 1-dimensional factor process such that $Z_t < 1$, $\forall t$, but the parameter estimate is $\theta = 1.5$, this is a contradiction given the interpretation of θ. This is due to θ having two roles under the initial specification. Firstly, as mentioned, it is the real-world long-run mean but it is also a pricing parameter explicitly involved in the zero-coupon bond price expression given by Equation (2.6). This results in tension between the cross-sectional data and time series data in the estimation procedure. The cross-sectional data ends up dominating the time series data. The extended specification assists in distinguishing the long-run mean and pricing parameters (the parameters that control the prices in any particular cross-section). Thus ensuring a true estimate of the real-world long-run mean, which ensures a more reliable analysis of the level dependence in volatility in this model.

[1] An explicit expression for a three-factor model is given in Appendix A.6.

The real-world long-run mean in the extended specification is given by $\hat{\theta}_i$ in Equation (3.3). Its explicit expression, given above, clarifies how the extended specification distinguishes the long-run mean from pricing parameters. In the estimation process θ_i will be given a value that optimally fits the cross-sectional data as it is a parameter in the zero-coupon bond price expression given by Equation (2.6). Although, in the extended specification θ_i is no longer the long-run mean. The role of long-run mean is now given by $\hat{\theta}_i$ which is dependent on variables, δ_i, not involved in pricing. Hence, the cross-sectional data will not dominate the time series data in the estimation of $\hat{\theta}_i$ as the time-series information will be captured through the δ_i parameters.

Chapter 4

Parameter Estimation

The market data that is used for the parameter estimation are fair swap rates. Ignoring the effects of default risk, a fixed-for-floating forward starting swap with tenor structure $\mathcal{T} = \{T_0 < T_1 < ... < T_N\}$ has the following fair swap rate at time $t < T_0$:

$$S(t, T_N) = \frac{P(t, T_0) - P(t, T_N)}{\sum_{i=1}^{N} P(t, T_i)\tau_i}, \tag{4.1}$$

where $\tau_i = T_i - T_{i-1}$. Equation (4.1) therefore links the state of the factor process to the market data as fair swap rates are expressed in terms of zero coupon bond prices. This follows from zero-coupon bond prices, given by Equation (2.6), being linear rational functions of the factors.

4.1 The Unscented Kalman Filter

We follow Filipović *et al.* (2017) in using the unscented Kalman filter, developed by Julier and Uhlmann (1996), to estimate the parameters of the model. The Kalman filter is an algorithm that is used to improve the estimation of a dynamics system's propagation, over time, by taking into account observations (in our case, swap rates) that are dependent on the dynamic system's state (in our case, the state variable values Z_t). The Kalman filter requires a prediction equation (which estimates the system's next state by using the value of its current state), a measurement equation (which relates the value of current observations to the current state) and a time series of observation values. The Euler-Maruyama approximation of the SDE of the factor process, given by Equation (3.3), gives the prediction equation for the system and the formula for a fair swap rate, given by Equation (4.1), gives the measurement equation.

The classical Kalman filter is used when the prediction and measurement equation are linear functions of the current state. In the context of this implementation the measurement equation is not a linear function of the current state (resulting form Equations (4.1) and (2.6)). Therefore, an unscented Kalman filter is used to overcome the issue of calculating mean vectors and covariance matrices. The unscented Kalman filter uses an unscented transform to approximate the mean vector and covariance matrix using 'sigma points' and corresponding weights.

In the following two subsections, we give detailed descriptions of the unscented transform and filtering algorithm, respectively. This includes an explanation of how the filtering method can be used to calculate a quasi-likelihood function, which can be maximised in order to determine the model parameters.

4.1.1 The Unscented Transform

The unscented Kalman filter algorithm is similar to the standard Kalman filter procedure. The major difference is that an unscented transform is used to approximate the mean and variance of a non-linearly transformed random variable (the standard Kalman filter only handles the special case of linear transforms, where means and variances can be determined analytically).

Consider an n-dimensional random variable with mean vector μ and covariance matrix Σ. The unscented transform involves defining a set of sample points called sigma points as follows:

$$\mathcal{X}_1 = \mu,$$

$$\mathcal{X}_i = \mu + \left(\sqrt{(n+\lambda)\Sigma}\right)_{i-1},$$

$$\mathcal{X}_{i+n} = \mu - \left(\sqrt{(n+\lambda)\Sigma}\right)_{i-1},$$

for $i = 2, ..., n+1$, where $\left(\sqrt{(n+\lambda)\Sigma}\right)_i$ is defined to be the ith column of the matrix square root of $(n+\lambda)\Sigma$ and λ is a scaling parameter. An example of a matrix square root is the Cholesky decomposition, it is commonly used in implementation as it has numerical stability. These sigma points can be thought of as samples that capture the mean and covariance information of the random variable. Weights corresponding to the sigma points are defined. These can be thought of as probabilities, forming a discrete distribution together with the sigma points. Two sets of weights $\{W_i^m\}$ and $\{W_i^c\}$ are in fact defined, such that

$$\mu = \sum_{i=1}^{2n+1} W_i^m \mathcal{X}_i,$$

$$\Sigma = \sum_{i=1}^{2n+1} W_i^c (\mathcal{X}_i - \mu)(\mathcal{X}_i - \mu)^\top,$$

i.e., in such a way that the mean and variance of the original random variable are also exhibited by the sigma point distribution. The above set of equations does not have a unique solution hence the weights are expressed in terms of some free parameters. The expressions for the corresponding weights are given as follows:

$$W_1^m = \frac{\lambda}{n+\lambda},$$

$$W_1^c = \frac{\lambda}{n+\lambda} + \left(1 - \alpha^2 + \beta\right),$$

$$W_i^m = W_i^c = \frac{1}{2(n+\lambda)},$$

for $i = 2, ..., 2n+1$, where $\lambda = \alpha^2(n+k) - n$. The free parameters α, k influence the distance of the sigma points from its mean vector and β can be used to capture further distributional properties of the random variable if additional information is known about the current distribution.

The mean vector and covariance matrix of a non-linear transform $f : \mathbb{R}^n \to \mathbb{R}^m$ of the current random variable are approximated as follows:

$$\mu_f \approx \sum_{i=1}^{2n+1} W_i^m f(\mathcal{X}_i),$$

$$\Sigma_f \approx \sum_{i=1}^{2n+1} W_i^c (f(\mathcal{X}_i) - \mu_f)(f(\mathcal{X}_i) - \mu_f)^\top.$$

The intuition is simply that the transform is applied to the discrete sigma points, from which the mean and variance are calculated directly, and are used as approximations to the truly transformed random variable.

4.1.2 The Filtering Steps

Bolder (2001) gives a description of how the standard Kalman filter is implemented on affine term structure models.[1] This approach is generalised to implement the unscented Kalman filter on the CEV specification of the linear-rational term structure model from Section 3.1.

1. **Initialisation of the factor process**
   The algorithm requires an initial input for, the time $t = 0$, values of the state variables and a measure of certainty of these inputs. Bolder (2001) proposes

[1] Bolder (2001) acknowledges that the use of the Kalman filter on affine term-structure models was popularised by Duan and Simonato (1999), Lund (1997), Geyer and Pichler (1999), De Jong (2000) and Babbs and Nowman (1999).

the use of the long run mean vector and covariance matrix. The Kalman filter is usually robust to the initialisation (i.e, the initialisation does not heavily influence the path of the filter). The initial mean and covariance matrix is denoted μ_0 and Σ_0 respectively. Steps $2-4$ will be repeated for $t = 1, ..., N$.

2. **Initial factor process forecast**

   An initial forecast of the factor process at time t is obtained by a function $F : \mathbb{R}^n \to \mathbb{R}^n$, which denotes the conditional expectation of the prediction equation given information at time $t-1$. The initial forecast vector and covariance matrix of the factor process are calculated — using the time $t-1$ filtered forecast vector μ_{t-1} and covariance matrix Σ_{t-1} obtained in step 4 — as follows:

$$\mu_t^* = F_0 + F_1\mu_{t-1},$$
$$\Sigma_t^* = F_1\Sigma_{t-1}F_1^\top + Q_{t-1}.$$

   In the context of this dissertation's implementation, following from Equation (3.3), we have

$$F_0 = \hat{K}\hat{\theta}\Delta t,$$
$$F_1 = I - \hat{\theta}\Delta t,$$
$$Q_{t-1} = \text{diag}\left(\sigma_i^2 \mu_{i,t-1}^{2\alpha_i}\right)_{i=1,\ldots,n}.$$

   The basic intuition is that μ_t^* is the best guess for the state at time t, given our filtered estimate μ_{t-1} at time $t-1$, and Σ_t^* is the variance associated with that guess (based on the incoming variance Σ_{t-1} and the new uncertainty in the movement of the factor process reflected by Q_{t-1}).

3. **Observation forecast**

   A matrix of sigma points are formed from the initial forecast (predicted mean) vector and covariance matrix above and are given as

$$\hat{\mathcal{X}}_t = \left[\mu_t^*, \mu_t^* + \sqrt{(n+\lambda)\Sigma_t^*}, \mu_t^* - \sqrt{(n+\lambda)\Sigma_t^*}\right],$$

   where $\hat{\mathcal{X}}_t$ is a $n \times (2n+1)$-dimensional matrix and the ith sigma point is given by, $\hat{\mathcal{X}}_{t,i}$, the ith column of $\hat{\mathcal{X}}_t$. The above sigma points are passed through a function $H : \mathbb{R}^n \to \mathbb{R}^m$, where $m \geq n$[2], which denotes the measurement equation. The advanced sigma points (sigma points passed through the measurement equation) are denoted by

$$\hat{\mathcal{Y}}_{t,i} = H(\hat{\mathcal{X}}_{t,i}),$$

[2] Assuming an m-dimensional observation vector, the dimension of the observation vector needs to be at least the same dimension as the factor process to ensure the system is not under-determined.

for $i = 1, ..., 2n + 1$. The observation forecast is then calculated by the unscented transform and denoted by

$$m_t = \sum_{i=1}^{2n+1} W_i^m \hat{\mathcal{Y}}_{t,i}.$$

Intuitively the observation forecast is what we expect the observations to be, given the current estimate of the state itself, because it is the mean of the transformed/measured state estimate. The associated covariance matrix for the forecast as well as the covariance matrix between the factor process and observations are then, respectively, calculated as follows:

$$V_t = \sum_{i=1}^{2n+1} W_i^c (\hat{\mathcal{Y}}_{t,i} - m_t)(\hat{\mathcal{Y}}_{t,i} - m_t)^\top + R_t,$$

$$C_t = \sum_{i=1}^{2n+1} W_i^c (\hat{\mathcal{X}}_{t,i} - \mu_t^*)(\hat{\mathcal{Y}}_{t,i} - m_t)^\top,$$

where R_t is a diagonal matrix of identical elements, denoted by σ^2_{rates}. This term results from the assumption that the observed swap rates are given by model-implied ones plus some independent noise of the data, example bid-ask spreads, data entry errors and non-simultaneous observations (Bolder, 2001). Furthermore, the addition of observation noise allows the use of more cross-sectional data without resulting in an overdetermined system (Piazzesi, 2010).[3]

4. **Updated factor process forecast**
The true observation value is compared to the observation forecast calculated in step 3 to determine the prediction error. The prediction error is then used to update the initial estimate of the factor process. The weight placed on the new information gained from the actual observation values are represented by the Kalman gain matrix,

$$K_t = C_t V_t^{-1}.$$

The time-t filtered forecast (mean vector) and covariance matrix are, respectively, given by

$$\mu_t = \mu_t^* + K_t[\mathcal{Y}_t - m_t],$$

$$\Sigma_t = \Sigma_t^* - K_t V_t K_t^\top,$$

where $\mathcal{Y}_t$ is the time-t actual observation.

[3] Piazzesi (2010) refers to this as a stochastic singularity problem.

5. **The construction of the likelihood function**
   Steps 2-4 results in the construction of a time series of approximate mean vectors and covariance matrices for the observation values. Assuming that the observations are normally distributed a log-likelihood value can be calculated as follows[4]

$$l(\rho) = -\frac{1}{2}\sum_{t=1}^{N}[m\log(2\pi) + \log(|V_t|) + (\mathcal{Y}_t - m_t)^\top V_t^{-1}(\mathcal{Y}_t - m_t)],$$

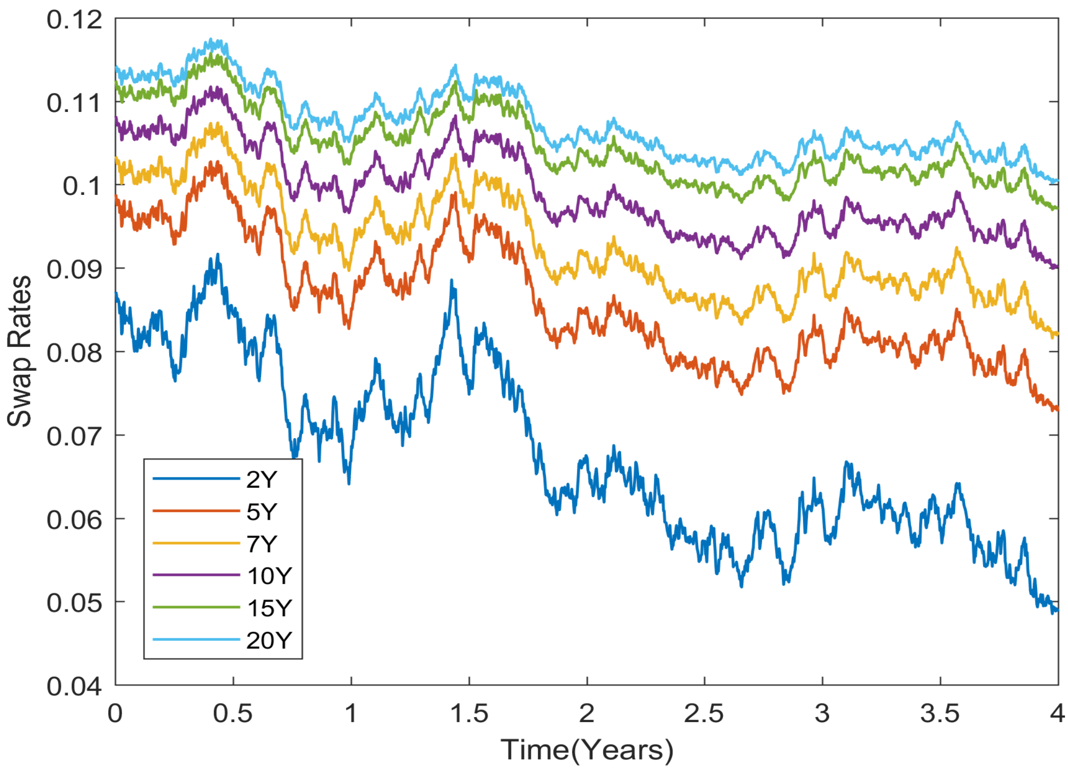
   for some parameter set ρ. It is important to note that the above steps require a parameter set, but this can be varied until the likelihood is maximised. This is done numerically through an optimisation scheme.

[4] $|V_t|$ denotes the determinant of matrix V_t.

4.2 Parameter Recovery

The effectiveness of the unscented Kalman filter technique with regards to parameter estimation can be seen through a simulation exercise. A CEV factor process is simulated, based on a known parameter set, using an Euler-Maruyama approximation of the SDE in Equation (3.3). A truncation scheme is used to prevent the factor process from becoming negative. Observation data is then constructed over a 4-year time horizon recording daily observations, i.e, $\Delta t = \frac{1}{365}$. The observations simulated are fair swap rates with quarterly payments and maturities of 2, 5, 7, 10, 15 and 20 years. The simulation of the observations is done using Equation (4.1), and then observational noise is added to the fair swap rates. Non-negativity of interest rates is ensured by setting β according to the expression in Section 2.1.1. The simulated data using a CEV factor process is displayed in Figure 4.1.

Fig. 4.1: Simulated swap rates using a CEV factor process



The unscented Kalman filter is then used to recover the known parameters given the simulated data. It is important to note that numerically maximising the likelihood function is a multi-dimension optimisation problem, therefore the estimation procedure is broken into two steps to improve its robustness. The estimation is first done using a SQRT factor process, a special case of the CEV factor process, where the degree of level dependence of volatility is set a priori.[5] Estimation is then performed using a CEV factor process, using the estimates of θ, K and σ_{rates} from the SQRT process as an initial guess. The estimation results for a two-factor model are presented in Table 4.1.

Tab. 4.1: Two-factor model parameter recovery

Parameters	True Value	CEV Estimate	SQRT Estimate
α_1	0.4000	0.3932	0.5000$*$
α_2	0.6000	0.5661	0.5000$*$
σ_1	0.3500	0.3487	0.4734
σ_2	0.2500	0.2344	0.1987
θ_1	0.6000	0.5986	0.5946
θ_2	0.4000	0.3942	0.3982
K_{11}	0.0900	0.0909	0.0911
K_{22}	0.3750	0.3769	0.3772
K_{21}	-0.1000	-0.0986	-0.1017
δ_1	-1.6000	-1.5494	-1.1216
δ_2	-2.5000	-1.1966	-1.4182
$\sigma_{\text{rates}} \times 10^4$	1.0000	0.9927	0.9934

The estimated parameters are accurate, except for σ in the SQRT specification and δ estimates. However, there are logical explanations for these estimates. In this simulation it is the case that $0 < Z_{i,t} < 1$ (illustrated in Figure 4.4), therefore the local volatility function of the SQRT process underestimates the magnitude of volatility attributed to the level dependent component for the first factor, $Z_{1,t}^{\alpha_1}$, and overestimates the attribution to $Z_{2,t}^{\alpha_2}$. The constant volatility component, σ_1, of the first factor is therefore given a higher estimate to compensate for the underestimation of volatility. Similarly σ_2 is given a lower estimate to compensate for overestimation of volatility. Regarding δ, recall that it features in the form for the market price of risk (refer to Equation (2.12)). It is important to see that one should not expect to accurately estimate parameters involved solely in the market price of risk (that is, such parameter estimates have large standard errors). This is documented in detail by Duffee and Stanton (2012) who show that while parameters involved in the risk-neutral drift can be accurately estimated, parameters involved only in the real-world drift (or, equivalently, only in the market price of risk, which links the risk-neutral and real-world drifts) cannot. Recalling Equation (2.12), some of the

[5] In Table 4.1 $*$ denotes parameters fixed a priori and not estimated.

parameters involved in the market price of risk play other roles in the model (the σ parameters control the volatility of the factor process), while delta has no other role, and is therefore difficult to estimate accurately. Indeed, Filipović *et al.* (2017) report very large standard errors for delta. Intuitively, accurate estimation of δ is difficult as the likelihood function is insensitive to changes in δ. Furthermore, the likelihood function captures cross-sectional information more accurately than time series information.

This concept is illustrated by Figure 4.2. The estimated CEV parameters, given in Table 4.1, are used in the calculation of the likelihood value. The parameter ranges in Figure 4.2 were selected for illustrative ease. The scale of the vertical axis is consistent among all graphs to allow comparison. It can be seen that the likelihood function is relatively insensitive to σ and δ parameters compared to K and θ parameters. This further highlights the significant weight placed on the cross-sectional data as K and θ are parameters involved in pricing. In the context of this implementation σ is not a cross-sectional parameter (as σ is absent in the zero-coupon bond price given by Equation (2.6)). However, if options were considered then σ would be involved in pricing and hence be a cross-sectional parameter. Estimation of σ, α and δ parameters are, therefore, more difficult as their estimate depends on the time-series data captured by the likelihood function. Relatively accurate σ and α estimates are obtained by using good estimates for K and θ as an initial guess thus allowing the optimiser to focus on estimating σ and α. This motivated breaking up the estimation procedure into two steps.

Figure 4.3 essentially magnifies a component from Figure 4.2. The vertical dotted lines indicate the true parameter values. Notice that the σ values where the likelihood obtains its maximum is very close to the true σ values whereas the δ values where the likelihood obtains its maximum is significantly different from the true δ values. This clarifies that obtaining accurate δ estimates though optimisation given the current data is a hopeless endeavour. More accurate δ estimates could be obtained by using more data, specifically a longer time horizon, with the intention of capturing more time-series information in the likelihood function.

Finally, the effectiveness of the Kalman filter with regards to state identification is illustrated by Figure 4.4. The estimated parameters displayed in Table 4.1 were used in the Kalman filter algorithm. It can be seen that the filtered factor process matches the simulated factor process very closely. Furthermore, this occurs with imprecise δ inputs. This indicates that the model fit is not heavily influenced by δ parameters. The results for a three-factor parameter recovery simulation are given in Appendix B.1.

Fig. 4.2: Log-likelihood value for various parameters

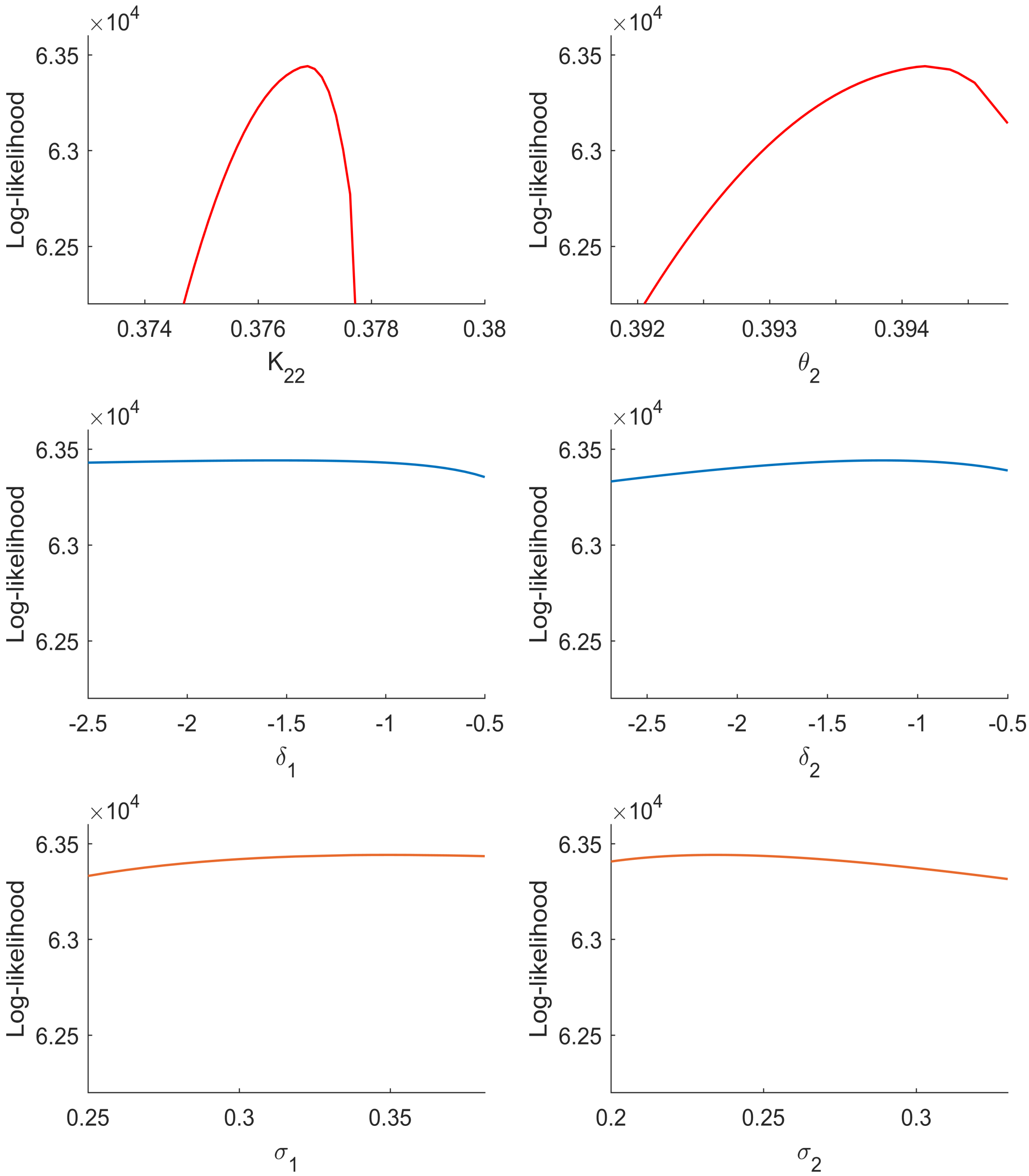


Fig. 4.3: Time-series data captured by log-likelihood function

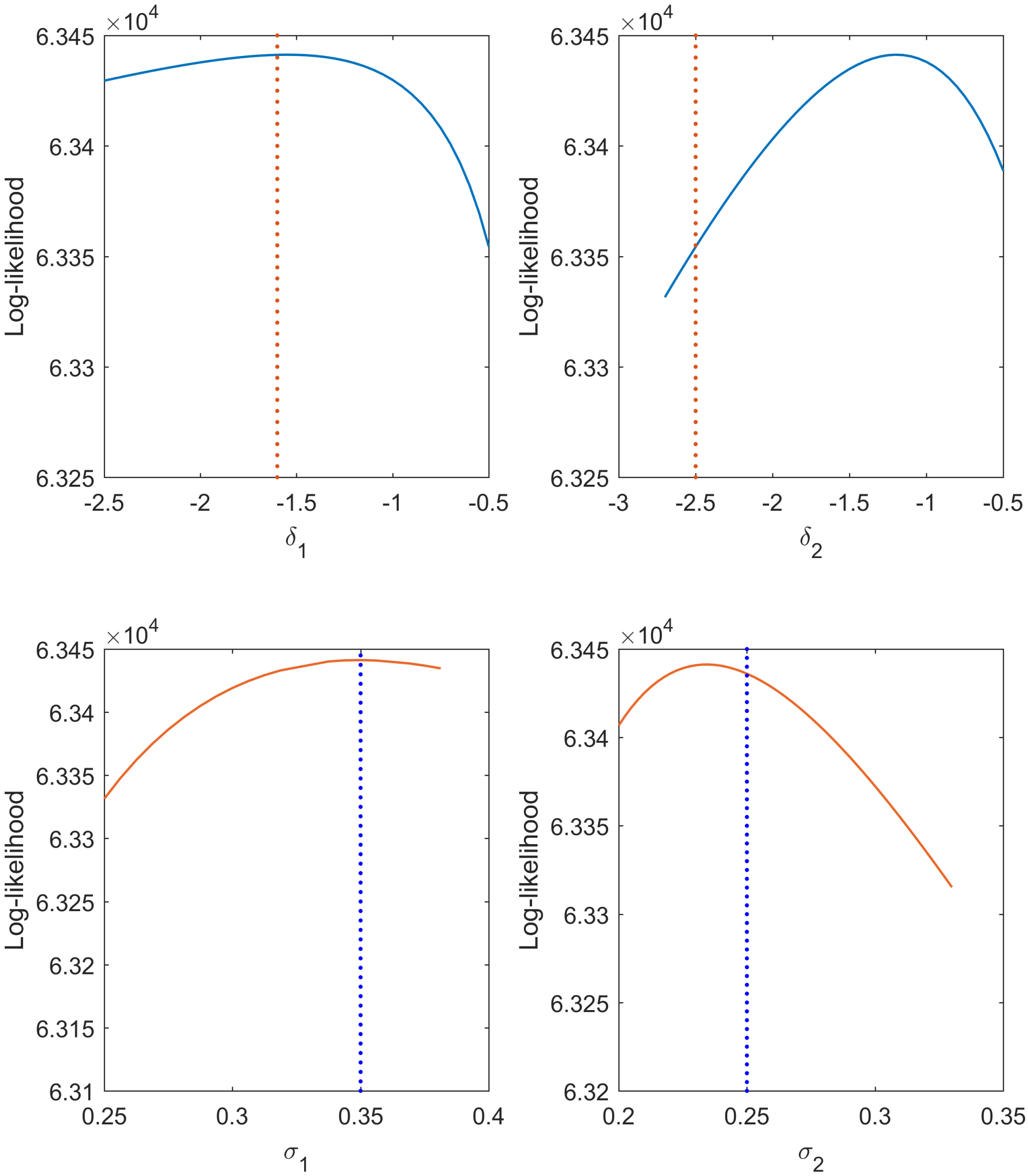
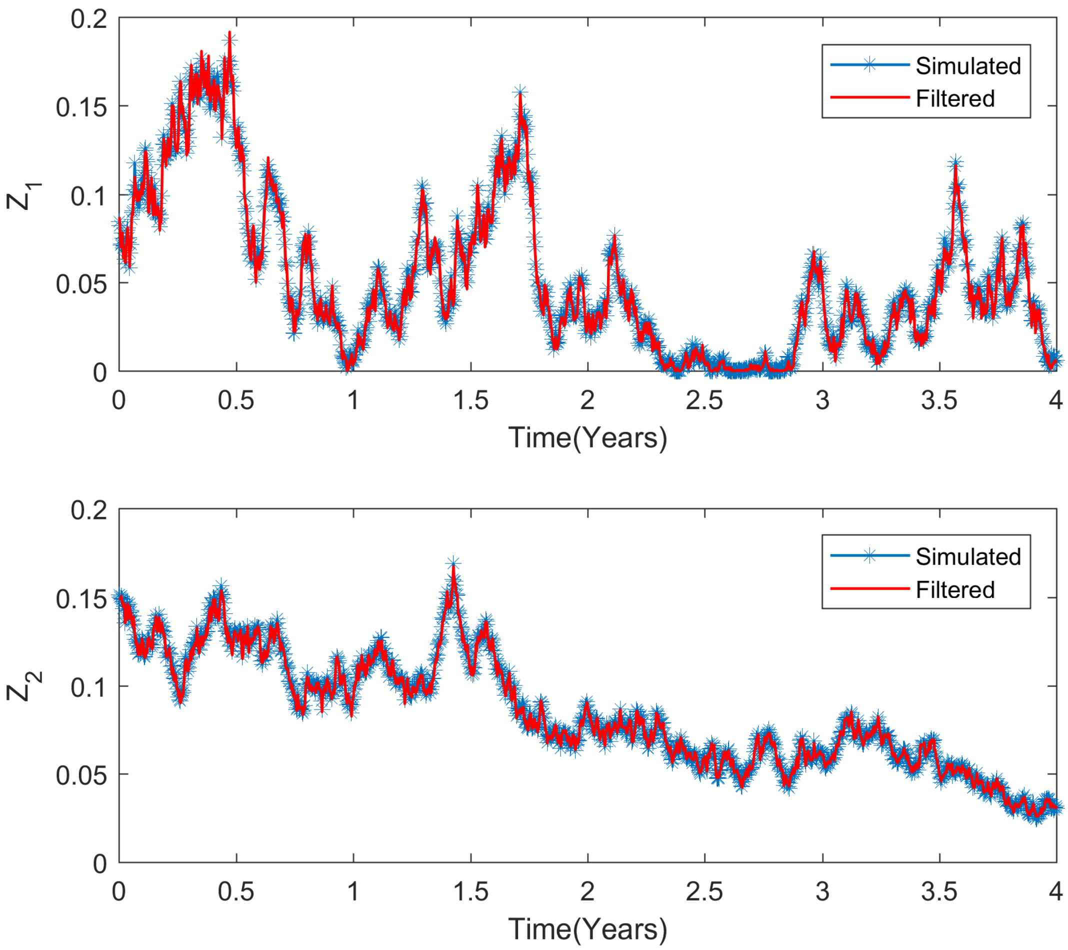


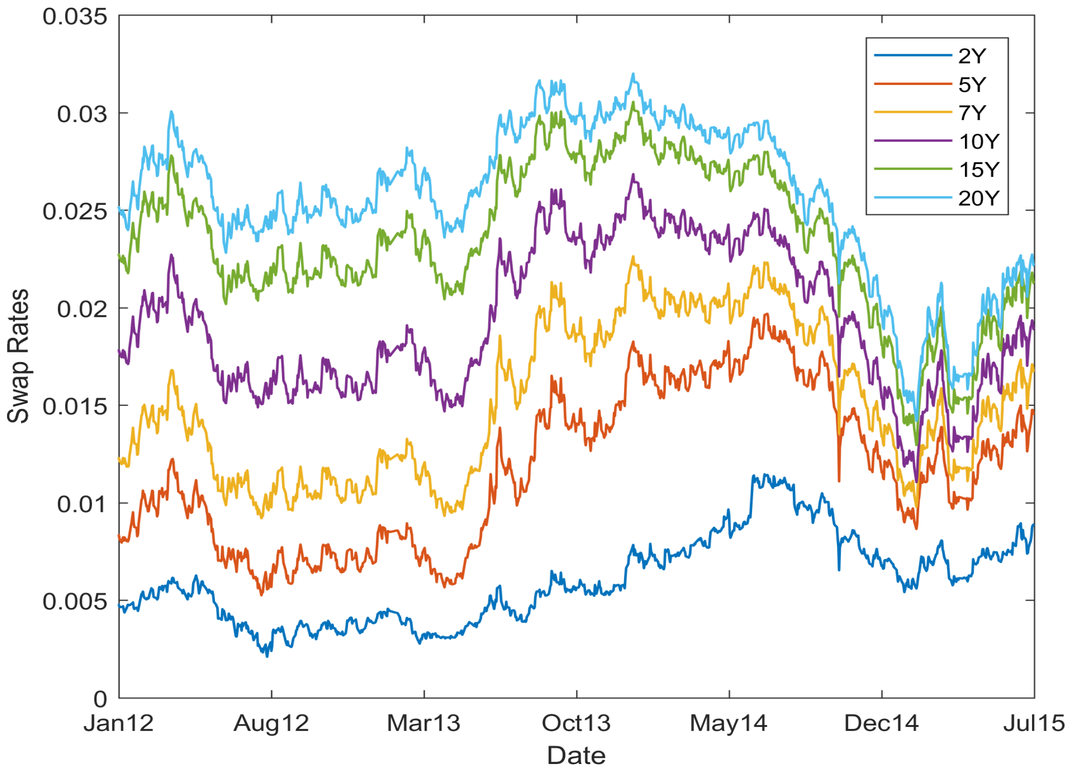
Fig. 4.4: Simulated and filtered CEV factor process, using the estimated parameters



4.3 Empirical Analysis

Empirical analysis is based on data from the British-pound swap market.[6] The time series consists of 897 trading days over the period 4 January 2012 to 29 July 2015. The observations considered are fair swap rates on spot-starting contracts with annual payments and maturities of 2, 5, 7, 10, 15 and 20 years. A benefit of calibrating directly to swap rates is that no bootstrapping of a yield curve is required. The data is displayed in Figure 4.5. Estimation is done using three factors.

Fig. 4.5: Time-series of British-pound swap rates



The model fit is illustrated by Figure 4.6 and Figure 4.7. The fitted swap rates closely match the pattern of the observed rates, indicating a good model fit. The root-mean-squared error (RMSE) was calculated at each observation date for the swap rates across their maturities. RMSE is defined as follows

$$\text{RMSE}_t = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_{i,t} - y_{i,t})^2},$$

[6] A developed economy with relatively low interest-rates.

Fig. 4.6: Time-series of model fitted swap rates under the CEV specification

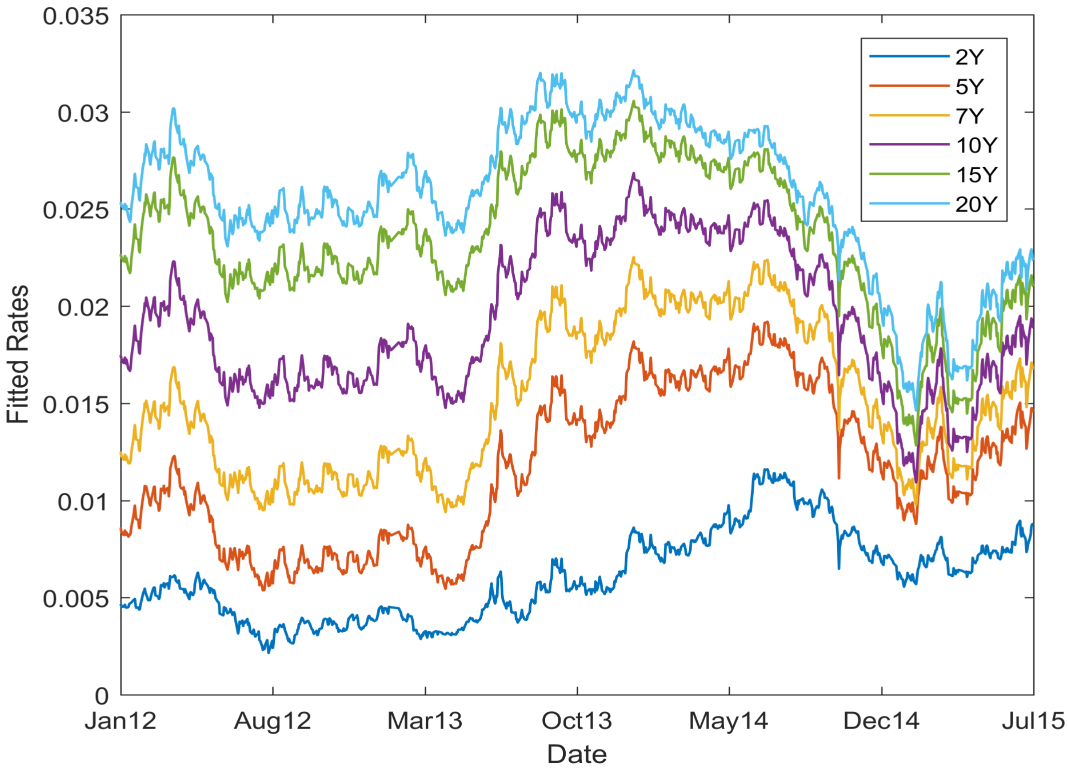
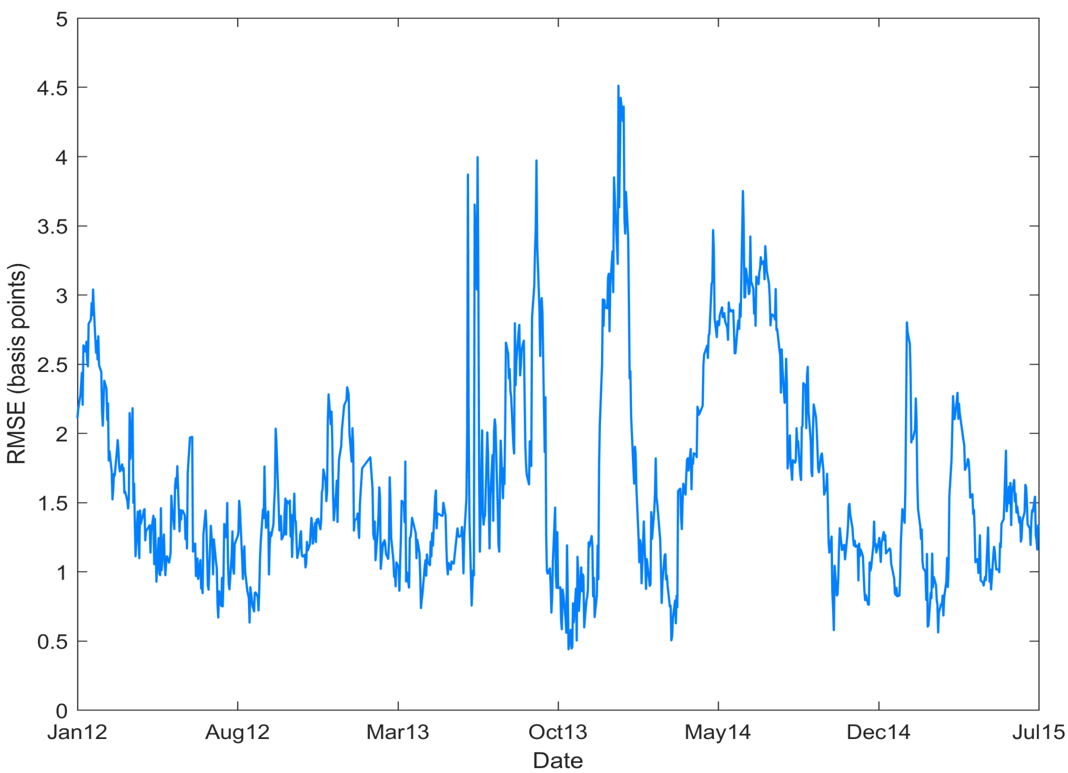


Fig. 4.7: Time-series of RMSE under the CEV specification



where $\hat{y}_{i,t}$ denotes time-t fitted values and $y_{i,t}$ denotes time-t observed values. RMSE is a measure of cross-sectional fit. The estimated parameters are presented in Table 4.2 and the filtered factor process is displayed in Figure 4.8.

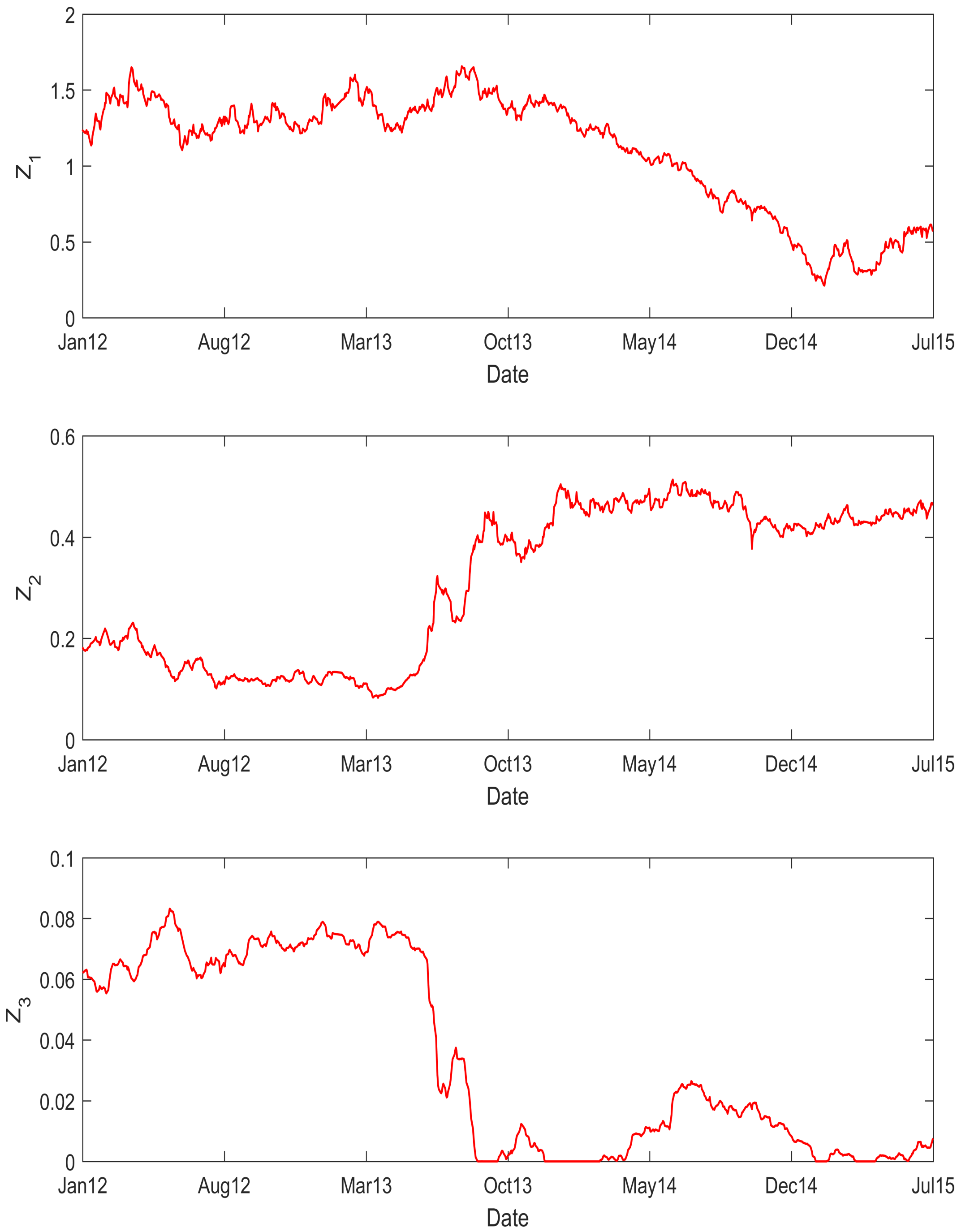
Tab. 4.2: Three-factor model parameter estimates

Parameters	CEV Estimate	SQRT Estimate
α_1	0.1911	0.5000
α_2	0.3064	0.5000
α_3	0.1051	0.5000
σ_1	0.4252	0.4556
σ_2	0.1666	0.2178
σ_3	0.0343	0.1436
θ_1	0.7729	1.2476
θ_2	0.5002	0.7407
θ_3	0.1960	0.2334
K_{11}	0.0709	0.0701
K_{22}	0.3522	0.3240
K_{33}	0.3519	0.3873
K_{21}	-0.1012	-0.1004
K_{32}	-0.3826	-0.3543
δ_1	0.0612	0.0497
δ_2	-4.8374	-2.9138
δ_3	33.0893	8.2568
$\sigma_{\text{rates}} \times 10^4$	1.9851	2.0168
Log-likelihood	36 213	36 162
Mean RMSE $\times 10^4$	1.6258	1.6200

The log-likelihood is greater in the CEV specification than in the SQRT specification, as expected, because the CEV specification is more parametrised. The higher likelihood is particularly due to a better time series fit as the CEV specification has about the same (slightly worse) cross-sectional fit. This follows from the mean RMSE reported on Table 4.2.

It is immediately evident, whilst comparing the estimated parameters between the two models, that the degree of level dependence with respect to the factors is significantly different. Because the estimated α_i parameters in the CEV case are all less than a half, the SQRT specification (where α_i is fixed) underestimates level dependence in volatility of the factor process for very low levels and overestimates it for all other levels. Although, as mentioned earlier, the implications on the level dependence in interest rate volatility is not yet clear. The interpretation of level

Fig. 4.8: Filtered factor process under the CEV specification



dependence with respect to the interest rate becomes more straightforward due to a robust estimation feature in the model observed by Filipović *et al.* (2017). The estimated drift parameters accurately approximate the following identity

$$\beta = \mathbf{1}^\top K\theta = -\mathbf{1}^\top K_1 = -\mathbf{1}^\top K_2 > -\mathbf{1}^\top K_3,$$

where K_i denotes the i^{th} column of vector of K. Appendix B.2 shows that this results in the short rate being expressed as

$$r_t = \frac{(\beta + K_{33})Z_{3,t}}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}}. \tag{4.2}$$

The $\mathbb{P}$-dynamics of the short rate, obtained through an application of Ito's lemma, is then given as follows:

$$dr_t = (\cdots)dt + \sigma_r(Z_t)dB_t^{\mathbb{P}},$$

where

$$\sigma_r(Z_t) = \left[\sigma_1 Z_{1,t}^{\alpha_1}\frac{\partial r_t}{\partial Z_{1,t}}, \sigma_2 Z_{2,t}^{\alpha_2}\frac{\partial r_t}{\partial Z_{2,t}}, \sigma_3 Z_{3,t}^{\alpha_3}\frac{\partial r_t}{\partial Z_{3,t}}\right],$$

$$\frac{\partial r_t}{\partial Z_{1,t}} = \frac{-(\beta + K_{33})Z_{3,t}}{(1 + Z_{1,t} + Z_{2,t} + Z_{3,t})^2},$$

$$\frac{\partial r_t}{\partial Z_{2,t}} = \frac{-(\beta + K_{33})Z_{3,t}}{(1 + Z_{1,t} + Z_{2,t} + Z_{3,t})^2},$$

$$\frac{\partial r_t}{\partial Z_{3,t}} = \frac{(\beta + K_{33})(1 + Z_{1,t} + Z_{2,t})}{(1 + Z_{1,t} + Z_{2,t} + Z_{3,t})^2}.$$

The partial derivatives above, based on the robust estimation approximations, of the short rate with respect to each factor are monotone functions. Specifically, the first two factors are negatively correlated with the short rate and the third factor is positively correlated with the short rate. The instantaneous volatility of the short rate,

$$||\sigma_r(Z_t)|| = \sqrt{\left(\sigma_1 Z_{1,t}^{\alpha_1}\frac{\partial r_t}{\partial Z_{1,t}}\right)^2 + \left(\sigma_2 Z_{2,t}^{\alpha_2}\frac{\partial r_t}{\partial Z_{2,t}}\right)^2 + \left(\sigma_3 Z_{3,t}^{\alpha_3}\frac{\partial r_t}{\partial Z_{3,t}}\right)^2},$$

can now be analysed to determine the degree of level dependence in each model. This is done holding each factor constant at their average filtered values, then the instantaneous volatility of the short rate is calculated by varying each factor between their minimum and maximum filtered values. Each factor is then transformed according to Equation (4.2), such that it becomes the short rate, in order to determine the influence of a particular factor on the short rate. Figures 4.9, 4.10 and 4.11 illustrate the relation between the instantaneous volatility of the short rate

Fig. 4.9: The relation between the volatility of the short rate and the first factor

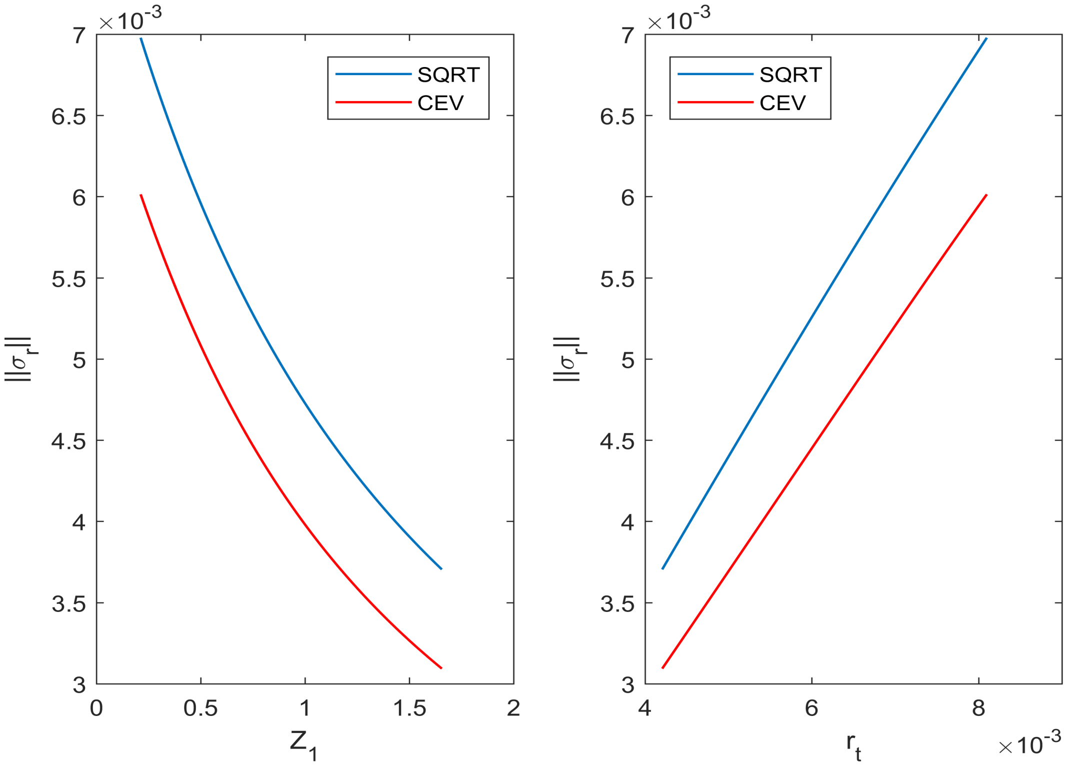


Fig. 4.10: The relation between the volatility of the short rate and the second factor

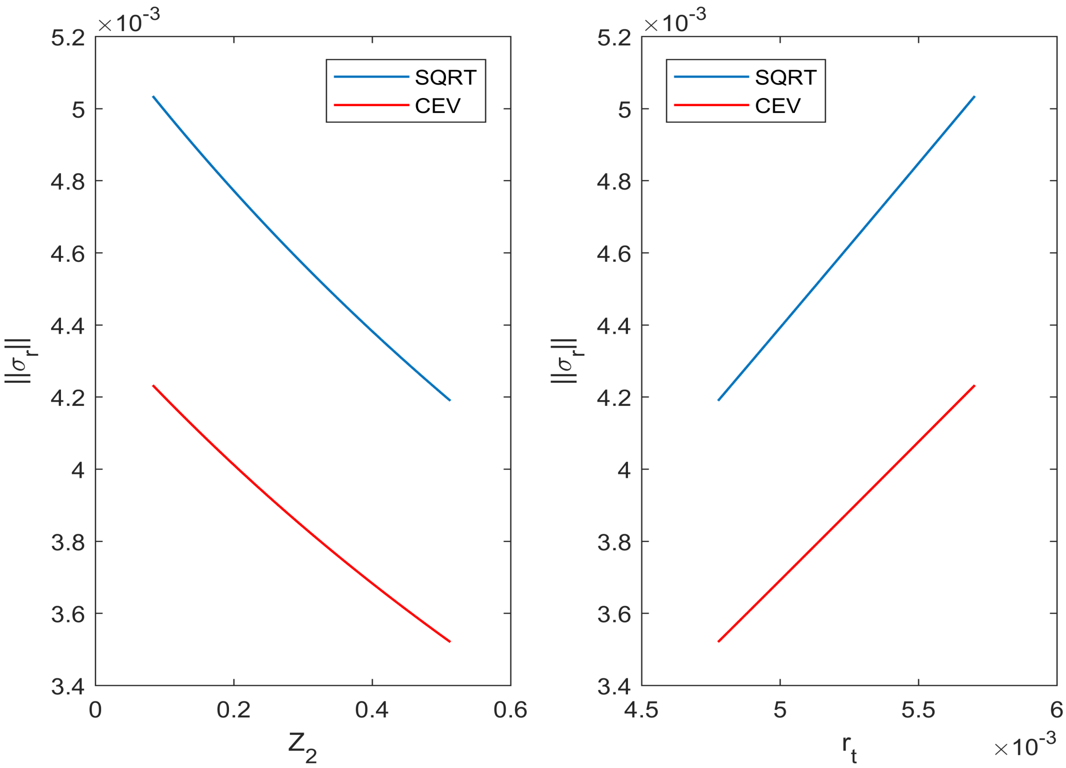
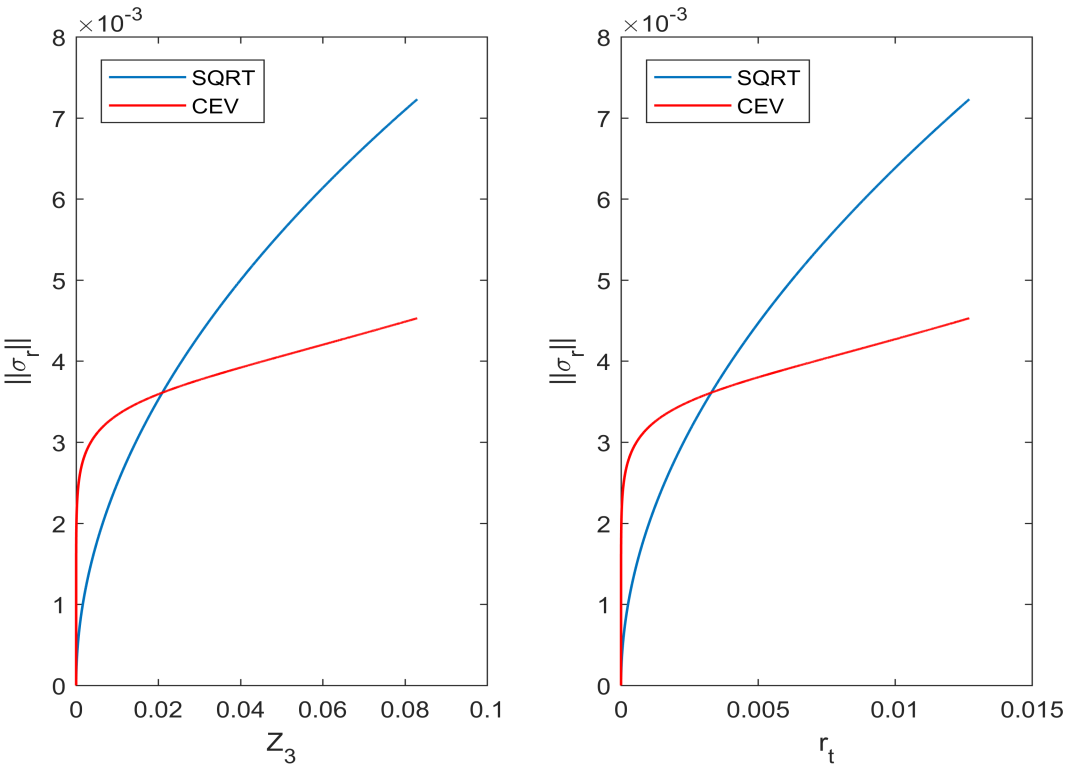


Fig. 4.11: The relation between the volatility of the short rate and the third factor



and each factor respectively. The stylised facts regarding interest rate volatility are captured by both specifications. This is illustrated by the above figures as the volatility of the short rate is an increasing function with respect to r_t regardless of which factor influenced the change in the short rate. Furthermore, the volatility of the short rate is a concave function with respect to r_t, when the change in r_t is attributed to the third factor, thus ensuring stronger level dependence when the interest rate is low. The volatility of the short rate is, however, a linear function of r_t when the movement in the short rate is due to changes in the first or second factor. This implies that there is a constant level dependence in volatility of the short rate, when the short rate changes due to movements in the first or second factor.

Movement in the first two factors do not heavily impact the level of the short rate relative to movement in the third factor. The above figures illustrate that a movement of Z_1 on the interval $[0.212, 1.656]$ and a movement of Z_2 on $[0.083, 0.513]$ corresponds to a movement of the short rate on $[0.0042, 0.0081]$ and $[0.0048, 0.0057]$ respectively, whereas a movement of Z_3 on $[0, 0.083]$ corresponds to a movement of the short rate on $[0, 0.0127]$. The minimal influence of the first two factors on the short rate can also be seen by considering the partial derivatives of the short rate with respect to each factor in conjunction with the filtered factor levels. The magnitude of the partial derivative with respect to the third factor is approximately 73 times larger than the partial derivatives with respect to the first two factors if the average filtered values are used in the calculation. This is consistent with Filipović *et al.* (2017) who document that when Z_3 is low the short rate is constrained near zero allowing Z_1 and Z_2 flexibility to affect longer-term interest rates without significantly impacting the short rate.

The degree of level dependence in volatility is roughly the same in the SQRT and CEV specification when the movement in the short rate is due to the first or second factor. This is illustrated by Figures 4.9 and 4.10 as the gradient of the local volatility function is essentially the same in both specifications. The first two factor do not heavily influence the short rate, therefore the above finding does not hold as much weight as results referring to the difference in level dependence when the short rate changes due to a movement in the third factor.

Level dependence in volatility is significantly different when the short rate changes due to movement in the third factor. Specifically, the SQRT specification overestimates level dependence. The difference in the average gradient of the volatility of the short rate between the SQRT and CEV specification given in Figure 4.11 is relatively large. This finding holds further weight given that interest rates were significantly low in the data set considered. Specifically, given that the SQRT specification significantly overestimates level dependence when interest rates are low, it will further overestimate level dependence when interest rates are higher.

It should be kept in mind that although the CEV specification captures the degree of level dependence in volatility more accurately, it has a trade-off with analytical tractability. Unlike the SQRT specification, the CEV specification does not allow affine approximations to the characteristic function. The most appropriate specification is therefore dependent on the type of model implementation as well as the general economic conditions.

Chapter 5

Conclusion

This dissertation analysed how accurately the linear-rational term structure model, specified with a factor process following SQRT dynamics, captures level dependence in volatility. This was done by comparing the SQRT specification, where the degree of level dependence is set a priori, to the CEV specification which allows a more flexible degree of level dependence. An extended specification for the state price density process was required to ensure reliable parameter estimates. Parameter estimation was performed on a three-factor model using a British pound data set. The empirical analysis suggests that the SQRT specification overestimates the degree of level dependence in volatility. This finding holds further significance as estimation was done using data where the general level of interest rates was low. It can, therefore, be inferred — taking into account the stylised fact that there is stronger level dependence when rates are low — that the SQRT model will further overestimate level dependence when rates are higher. Although the CEV specification captures level dependence more accurately , it trades off analytical tractability. Pricing options will be more challenging, because affine approximations to the characteristic function are no longer available. Therefore, the optimal specification will depend on the type of model implementation as well as the general economic conditions.

Bibliography

Ait-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate, *The Review of Financial Studies* **9**(2): 385–426.

Andersen, L. B. and Piterbarg, V. V. (2007). Moment explosions in stochastic volatility models, *Finance and Stochastics* **11**(1): 29–50.

Babbs, S. H. and Nowman, K. B. (1999). Kalman filtering of generalized Vasicek term structure models, *Journal of Financial and Quantitative Analysis* **34**(1): 115–130.

Bolder, D. (2001). Affine term-structure models: Theory and implementation, *SSRN Working Paper* .

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

Chan, K. C., Karolyi, G. A., Longstaff, F. A. and Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate, *The Journal of Finance* **47**(3): 1209–1227.

Conley, T. G., Hansen, L. P., Luttmer, E. G. and Scheinkman, J. A. (1997). Short-term interest rates as subordinated diffusions, *The Review of Financial Studies* **10**(3): 525–577.

Constantinides, G. M. (1992). A theory of the nominal term structure of interest rates, *The Review of Financial Studies* **5**(4): 531–552.

Cox, J. C. (1975). Notes on option pricing i: Constant elasticity of variance diffusions, *Stanford University, Graduate School of Business* .

Cox, J. C., Ingersoll, J. E. and Ross, S. A. (1985). A theory of the term structure of interest rates, *Econometrica* **53**: 385–407.

De Jong, F. (2000). Time series and cross-section information in affine term-structure models, *Journal of Business & Economic Statistics* **18**(3): 300–314.

Duan, J.-C. and Simonato, J.-G. (1999). Estimating and testing exponential-affine term structure models by Kalman filter, *Review of Quantitative Finance and Accounting* **13**(2): 111–135.

Duarte, J. (2003). Evaluating an alternative risk preference in affine term structure models, *The Review of Financial Studies* **17**(2): 379–404.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models, *The Journal of Finance* **57**(1): 405–443.

Duffee, G. R. and Stanton, R. H. (2012). Estimation of dynamic term structure models, *The Quarterly Journal of Finance* **2**(02): 111–135.

Duffie, D. and Kan, R. (1996). A yield-factor model of interest rates, *Mathematical Finance* **6**(4): 379–406.

Filipović, D. (2009). *Term-Structure Models. A Graduate Course.*, Springer.

Filipović, D., Larsson, M. and Trolle, A. B. (2017). Linear-rational term structure models, *The Journal of Finance* **72**(2): 655–704.

Geyer, A. L. and Pichler, S. (1999). A state-space approach to estimate and test multifactor Cox-Ingersoll-Ross models of the term structure, *Journal of Financial Research* **22**(1): 107–130.

Harrison, J. M. and Pliska, S. R. (1981). Martingales and stochastic integrals in the theory of continuous trading, *Stochastic Processes and their Applications* **11**(3): 215–260.

Heath, D., Jarrow, R. and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica: Journal of the Econometric Society* pp. 77–105.

Julier, S. J. and Uhlmann, J. K. (1996). A general method for approximating nonlinear transformations of probability distributions, *Technical report*, Robotics Research Group, Department of Engineering Science, University of Oxford.

Kim, D. H. and Singleton, K. J. (2012). Term structure models and the zero bound: an empirical investigation of Japanese yields, *Journal of Econometrics* **170**(1): 32–49.

Lund, J. (1997). *Econometric analysis of continuous-time arbitrage-free models of the term structure of interest rates*, Department of Finance, Aarhus School of Business.

Piazzesi, M. (2010). Affine term structure models, *Handbook of Financial Econometrics* **1**: 691–766.

Stanton, R. (1997). A nonparametric model of term structure dynamics and the market price of interest rate risk, *The Journal of Finance* **52**(5): 1973–2002.

Trolle, A. B. and Schwartz, E. S. (2014). The swaption cube, *The Review of Financial Studies* **27**(8): 2307–2353.

Vasicek, O. (1977). An equilibrium characterization of the term structure, *Journal of Financial Economics* **5**(2): 177–188.

Appendix A

Appendix for Chapter 2

A.1 The dynamics of the state price density and its link to the risk-neutral measure

The dynamics of the state price density process, $\{\zeta_t\}$, can be obtained by applying Ito's lemma to, $\zeta_t = f(t, Z_t)$, the expression given by Equation (2.4). Now $f_t = -\beta\zeta_t$, $f_x = e^{-\beta t}\psi^\top$ and $f_{xx} = 0$ hence,

$$\begin{aligned}
d\zeta_t &= -\beta\zeta_t dt + e^{-\beta t}\psi^\top dZ_t \\
&= -\beta\zeta_t dt + e^{-\beta t}\psi^\top (K(\theta - Z_t)dt + \sigma(t, Z_t)dB_t) \\
&= -\left[\beta\zeta_t - e^{-\beta t}\psi^\top K(\theta - Z_t)\frac{\zeta_t}{\zeta_t}\right] dt + \psi^\top \sigma(t, Z_t)e^{-\beta t}\frac{\zeta_t}{\zeta_t}dB_t \\
&= -\left[\beta\zeta_t - \frac{\psi^\top K(\theta - Z_t)\zeta_t}{\phi + \psi^\top Z_t}\right] dt + \frac{\psi^\top \sigma(t, Z_t)\zeta_t}{\phi + \psi^\top Z_t}dB_t.
\end{aligned}$$

Therefore the dynamics are expressed, by definition of a state price density, as follows:

$$\frac{1}{\zeta_t}d\zeta_t = -r_t dt - \lambda_t^\top dB_t, \tag{A.1}$$

where the market price of risk is given as

$$\lambda_t = -\frac{\sigma(t, Z_t)^\top \psi}{\phi + \psi^\top Z_t}.$$

The dynamics of the state price density given in Equation (A.1) helps to see why the state price density involves both a change of measure and discounting property. The drift term involves the discounting whereas the martingale term involves the change of measure. The Radon-Nikodým process involved in the change of measure can be obtained by accumulating the state price density process.

Let $\bar{\zeta}_t = f(t, \zeta_t) = e^{\int_0^t r_t dt}\zeta_t$ then applying Ito's lemma the follow dynamics for the accumulated state price density process is obtained and is given by:

$$d\bar{\zeta}_t = -\lambda_t^\top \bar{\zeta}_t dB_t. \tag{A.2}$$

The Girsanov kernel in Equation (A.2) is the negative market price of risk. Hence, supporting that the change of measure is to the risk neutral measure $\mathbb{Q}$.

Notice that both the short rate and market price of risk are functions of the factor process, Z_t, indicating that the specification of state price density jointly specifies the interest rate and change of measure. Alternatively, the change of measure and discounting property of the state price density can be seen by simultaneously considering Equation (2.1) and the risk-neutral valuation formula. The following relation is obtained:

$$\mathbb{E}_t^{\mathbb{Q}}\left[X_T \frac{A_t}{A_T}\right] = \mathbb{E}_t^{\mathbb{P}}\left[X_T \frac{\zeta_T}{\zeta_t}\right], \tag{A.3}$$

where $A_t = e^{\int_0^t r_t dt}$. The expectation on the left-hand-side (LHS) of Equation (A.3) is under $\mathbb{Q}$ whereas it's under $\mathbb{P}$ on right-hand-side (RHS), furthermore on the LHS the expectation is of the discounted time-T value of an asset or attainable claim. Therefore, it can be seen that the state price density process has both a discounting and change of measure property.

Now applying Girsanov's theorem, the $\mathbb{Q}$-dynamics of the factor process is expressed as:

$$\begin{aligned} dZ_t &= [K(\theta - Z_t) - \sigma(t, Z_t)\lambda_t]dt + \sigma(t, Z_t)dB_t^{\mathbb{Q}} \\ &= \left[K(\theta - Z_t) + \frac{\sigma(t, Z_t)\sigma(t, Z_t)^\top \psi}{\phi + \psi^\top Z_t}\right] dt + \sigma(t, Z_t)dB_t^{\mathbb{Q}}. \end{aligned}$$

A.2 The relation between the $\mathbb{P}$ and $\mathbb{A}$ state price density process in the extended specification

The state price density under $\mathbb{P}$, denoted $\{\zeta_t^{\mathbb{P}}\}$, is not explicitly required to understand the extended specification at a high level. However, it plays an important role in the derivation of the market price of risk under $\mathbb{P}$ as well as in the technical intricacies of pricing kernel models. The state price density under $\mathbb{P}$ is defined as follows:

$$\zeta_t^{\mathbb{P}} = \mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} e^{-\int_0^t r_s ds}\right].$$

Therefore, the valuation formula is given by:

$$X_t = \frac{1}{\zeta_t^{\mathbb{P}}}\mathbb{E}_t^{\mathbb{P}}\left[X_T \frac{d\mathbb{Q}}{d\mathbb{P}} e^{-\int_0^T r_s ds}\right]. \tag{A.4}$$

Rearranging Equation (A.4) we get the following:

$$\zeta_t^{\mathbb{P}} = \frac{1}{X_t}\mathbb{E}_t^{\mathbb{P}}\left[X_T \frac{d\mathbb{Q}}{d\mathbb{P}} e^{-\int_0^T r_s ds}\right],$$

then substituting X_t from Equation (2.11) results in

$$\zeta_t^{\mathbb{P}} = \frac{\zeta_t^{\mathbb{A}}\mathbb{E}_t^{\mathbb{P}}\left[X_T \frac{d\mathbb{Q}}{d\mathbb{P}} e^{-\int_0^T r_s ds}\right]}{\mathbb{E}_t^{\mathbb{A}}\left[X_T \frac{d\mathbb{Q}}{d\mathbb{A}} e^{-\int_0^T r_s ds}\right]}.$$

Baye's theorem implies:

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbb{A}}{d\mathbb{P}}\right]\mathbb{E}_t^{\mathbb{A}}[Z] = \mathbb{E}_t^{\mathbb{P}}\left[Z\frac{d\mathbb{A}}{d\mathbb{P}}\right],$$

for some measurable random variable Z. Now if we let $Z = X_T\frac{d\mathbb{Q}}{d\mathbb{A}}e^{-\int_0^T r_s ds}$ then $Z\frac{d\mathbb{A}}{d\mathbb{P}} = X_T\frac{d\mathbb{Q}}{d\mathbb{P}}e^{-\int_0^T r_s ds}$ which results in the final expression:

$$\zeta_t^{\mathbb{P}} = \zeta_t^{\mathbb{A}}\mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbb{A}}{d\mathbb{P}}\right].$$

A.3 The specification of the change of measure from $\mathbb{A}$ to $\mathbb{P}$ in the extended framework

The extended specification initially defines the framework under an auxiliary measure $\mathbb{A}$. In order to create a link to the real-world measure $\mathbb{P}$ a change of measure from $\mathbb{A}$ to $\mathbb{P}$ must be specified. In this Brownian setting a change of measure is performed via the well known Girsanov's theorem. The theorem requires defining a predictable process[1], denoted $\{\delta_t\}$, called the Girsanov kernel. The Radon-Nikodým derivative of $\mathbb{P}$ with respect to $\mathbb{A}$ that specifies the measure change is then defined as follows:[2]

$$\frac{d\mathbb{P}}{d\mathbb{A}} = \mathcal{E}(\delta \bullet B)_* = e^{\int_0^* \delta_t dB_t^{\mathbb{A}} - \frac{1}{2}\int_0^* ||\delta_t||^2 dt},$$

where $\mathcal{E}(\cdot)$ denotes the Doléan exponential of a stochastic process and $(\delta \bullet B)_* = \int_0^* \delta_t dB_t^{\mathbb{A}}$ such that Novikov's condition holds:

$$\mathbb{E}^{\mathbb{A}}\left[e^{\frac{1}{2}\int_0^* ||\delta_t||^2 dt}\right] < \infty.$$

The Radon-Nikodým derivative defined above is a martingale. This follows from $(\delta \bullet B)_*$ being a martingale in conjunction with the Doléan exponential of a martingale being a martingale. Hence,

$$\mathbb{E}_t^{\mathbb{A}}\left[\frac{d\mathbb{P}}{d\mathbb{A}}\right] = \mathcal{E}(\delta \bullet B)_t.$$

The expression relating $\mathbb{A}$ and $\mathbb{P}$ Brownian motion is then given by:

$$dB_t^{\mathbb{P}} = dB_t^{\mathbb{A}} - \delta_t dt.$$

Therefore, the $\mathbb{P}$-diffusion dynamics of the factor process is given by:

$$dZ_t = [K(\theta - Z_t) + \sigma(t, Z_t)\delta_t]dt + \sigma(t, Z_t)dB_t^{\mathbb{P}}.$$

[1] A predictable process can intuitively be thought of as a left continuous adapted process.

[2] The "$*$" in the following equations acts as a place holder as we do not get into the differences between finite and infinite time models.

A.4 The $\mathbb{P}$ dynamics of the state price density in the extended framework

Appendix A.2 showed that the real-world state price density is expressed as

$$\zeta_t^{\mathbb{P}} = \zeta_t^{\mathbb{A}} \mathbb{E}_t^{\mathbb{P}} \left[\frac{d\mathbb{A}}{d\mathbb{P}} \right].$$

Appendix A.3 then specified the change of measure from $\mathbb{A}$ to $\mathbb{P}$ through $\frac{d\mathbb{P}}{d\mathbb{A}}$. However, the expression for $\mathbb{E}_t^{\mathbb{P}} \left[\frac{d\mathbb{A}}{d\mathbb{P}} \right]$ is not yet obvious and is derived as follows. Applying the link between $\mathbb{A}$ and $\mathbb{P}$ Brownian motion we obtain $d(\delta \bullet B)_t = \delta_t (dB_t^{\mathbb{P}} + \delta_t dt)$. Therefore, $(\delta \bullet B)_* = \int_0^* \delta_t dB_t^{\mathbb{P}} + \int_0^* ||\delta_t||^2 dt$ allowing $\frac{d\mathbb{P}}{d\mathbb{A}}$ to be expressed in terms of $\mathbb{P}$-Brownian Motion,

$$\frac{d\mathbb{P}}{d\mathbb{A}} = e^{\int_0^* \delta_t dB_t^{\mathbb{P}} + \frac{1}{2}\int_0^* ||\delta_t||^2 dt}.$$

Now since $\frac{d\mathbb{P}}{d\mathbb{A}}$ exists we know that $\frac{d\mathbb{A}}{d\mathbb{P}}$ exists and is given by the reciprocal of the above expression:

$$\frac{d\mathbb{A}}{d\mathbb{P}} = e^{\int_0^* -\delta_t dB_t^{\mathbb{P}} - \frac{1}{2}\int_0^* ||\delta_t||^2 dt} = \mathcal{E}(-\delta \bullet B)_*.$$

Therefore, by the martingale property we get $\mathbb{E}_t^{\mathbb{P}} \left[\frac{d\mathbb{A}}{d\mathbb{P}} \right] = \mathcal{E}(-\delta \bullet B)_t$.

The next step is to derive the dynamics for the $\mathbb{P}$-state price density process. Note that $\zeta_t^{\mathbb{P}} = f(\zeta_t^{\mathbb{A}}, Y_t)$ where $Y_t = \mathbb{E}_t^{\mathbb{P}}[\frac{d\mathbb{A}}{d\mathbb{P}}]$, therefore we can apply Ito's lemma (product rule) to obtain:

$$\begin{aligned}
d\zeta_t^{\mathbb{P}} &= \zeta_t^{\mathbb{A}} dY_t + Y_t d\zeta_t^{\mathbb{A}} + dY_t d\zeta_t^{\mathbb{A}} \\
&= -\zeta_t^{\mathbb{A}} Y_t \delta_t^{\top} dB_t^{\mathbb{P}} + Y_t \zeta_t^{\mathbb{A}} \left(-r_t dt - (\lambda_t^{\mathbb{A}})^{\top} dB_t^{\mathbb{A}} \right) + \zeta_t^{\mathbb{P}} (\lambda_t^{\mathbb{A}})^{\top} \delta_t dt \\
&= -\zeta_t^{\mathbb{P}} \delta_t^{\top} dB_t^{\mathbb{P}} + \zeta_t^{\mathbb{P}} \left(-r_t dt - (\lambda_t^{\mathbb{A}})^{\top} (dB_t^{\mathbb{P}} + \delta_t dt) \right) + \zeta_t^{\mathbb{P}} (\lambda_t^{\mathbb{A}})^{\top} \delta_t dt \\
&= -\zeta_t^{\mathbb{P}} r_t dt - \zeta_t^{\mathbb{P}} \left((\lambda_t^{\mathbb{A}})^{\top} + \delta_t^{\top} \right) dB_t^{\mathbb{P}}.
\end{aligned}$$

Therefore, the dynamics are expressed as follows:

$$\frac{d\zeta_t^{\mathbb{P}}}{\zeta_t^{\mathbb{P}}} = -r_t dt - \left(\lambda_t^{\mathbb{P}} \right)^{\top} dB_t^{\mathbb{P}},$$

where

$$\begin{aligned}
\lambda_t^{\mathbb{P}} &= \lambda_t^{\mathbb{A}} + \delta_t \\
&= -\frac{\sigma(t, Z_t)^{\top} \psi}{\phi + \psi^{\top} Z_t} + \delta_t.
\end{aligned}$$

A.5 Partial derivatives of the short rate with respect to the factors

A two-factor model is considered to show that the partial derivative of the short rate with respect to the factors are not necessarily monotone functions. In a two-factor model the short rate given by Equation (2.7) can be expressed as

$$r_t = \beta - \frac{K_1^*(\theta_1 - Z_{1,t}) + K_2^*(\theta_2 - Z_{2,t})}{1 + Z_{1,t} + Z_{2,t}},$$

where $K_i^* = \mathbf{1}^\top K_i$ and K_i denotes the i^{th} column of vector K. The partial derivatives with respect each factor are then given as follows

$$\frac{\partial r_t}{\partial Z_{1,t}} = \frac{K_1^*(1+\theta_1) + K_2^*\theta_2 - (K_2^* - K_1^*)Z_{2,t}}{(1 + Z_{1,t} + Z_{2,t})^2},$$

$$\frac{\partial r_t}{\partial Z_{2,t}} = \frac{K_2^*(1+\theta_2) + K_1^*\theta_1 - (K_1^* - K_2^*)Z_{1,t}}{(1 + Z_{1,t} + Z_{2,t})^2}.$$

It can therefore be seen that

$$\frac{\partial r_t}{\partial Z_{1,t}} > 0 \Leftrightarrow K_1^*(1+\theta_1) + K_2^*\theta_2 > (K_2^* - K_1^*)Z_{2,t},$$

$$\frac{\partial r_t}{\partial Z_{2,t}} > 0 \Leftrightarrow K_2^*(1+\theta_2) + K_1^*\theta_1 > (K_1^* - K_2^*)Z_{1,t}.$$

The above demonstrates that the partial derivatives of the short rate with respect to each factor are not necessarily monotone functions as the sign of the partial derivatives depends on drift parameters of the factor process as well as the level of the factors.

A.6 Explicit expression of the $\mathbb{P}$-dynamics for a three-factor model

The explicit $\mathbb{P}$-dynamics for a three-factor model following the CEV specification is given as follows:

$$\begin{bmatrix} dZ_{1,t} \\ dZ_{2,t} \\ dZ_{3,t} \end{bmatrix} = \begin{bmatrix} \bar{K}_{11} & 0 & 0 \\ K_{21} & \bar{K}_{22} & 0 \\ 0 & K_{32} & \bar{K}_{33} \end{bmatrix} \left[\begin{bmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_3 \end{bmatrix} - \begin{bmatrix} Z_{1,t} \\ Z_{2,t} \\ Z_{3,t} \end{bmatrix} \right] dt + \begin{bmatrix} \sigma_1 Z_{1,t}^{\alpha_1} & 0 & 0 \\ 0 & \sigma_2 Z_{2,t}^{\alpha_2} & 0 \\ 0 & 0 & \sigma_3 Z_{3,t}^{\alpha_3} \end{bmatrix} \begin{bmatrix} dB_{1,t}^{\mathbb{P}} \\ dB_{2,t}^{\mathbb{P}} \\ dB_{3,t}^{\mathbb{P}} \end{bmatrix},$$

where

$$\hat{\theta}_1 = \bar{\theta}_1,$$

$$\hat{\theta}_2 = \bar{\theta}_2 - \frac{\sigma_1 \delta_1 K_{21} \theta_1}{\bar{K}_{11} \bar{K}_{22}},$$

and

$$\hat{\theta}_3 = \bar{\theta}_3 - \frac{\sigma_2 \delta_2 K_{32} \theta_2}{\bar{K}_{22} \bar{K}_{33}} + \frac{\sigma_1 \delta_1 K_{21} K_{32} \theta_1}{\bar{K}_{11} \bar{K}_{22} \bar{K}_{33}}.$$

Appendix B

Appendix for Chapter 3

B.1 Parameter recovery for a three-factor model

Fig. B.1: Simulated swap rates using a three-factor CEV process

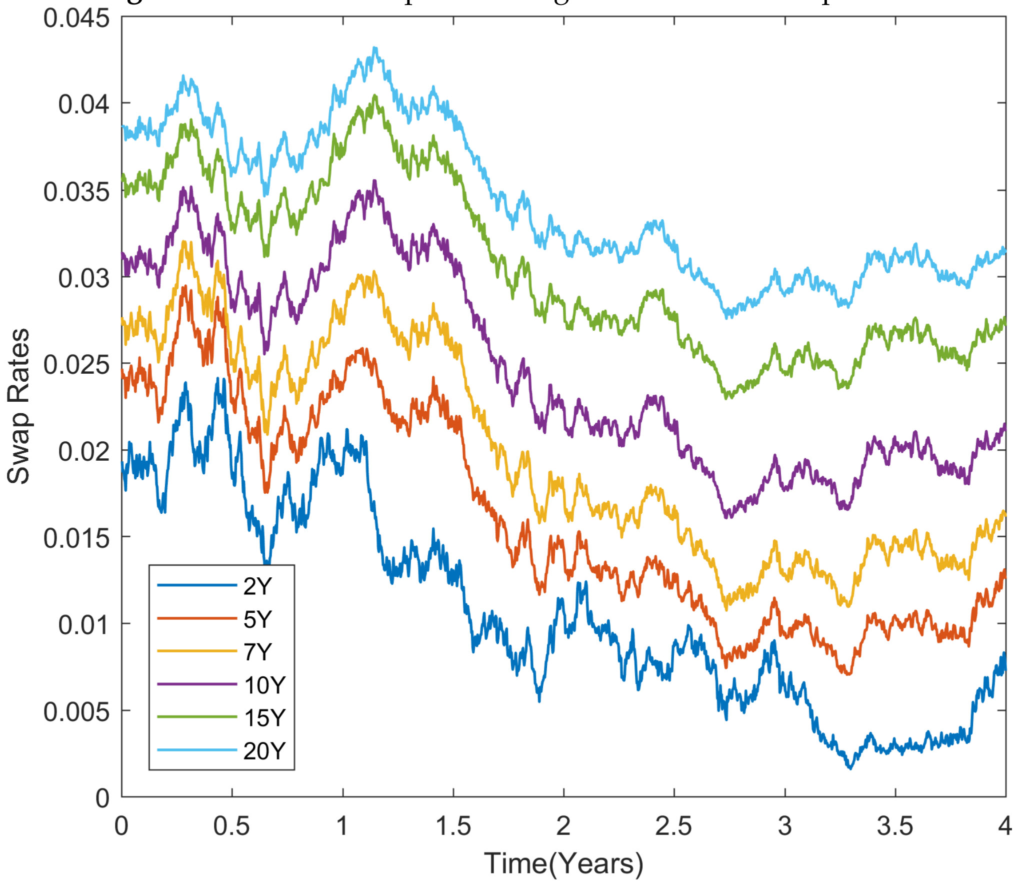
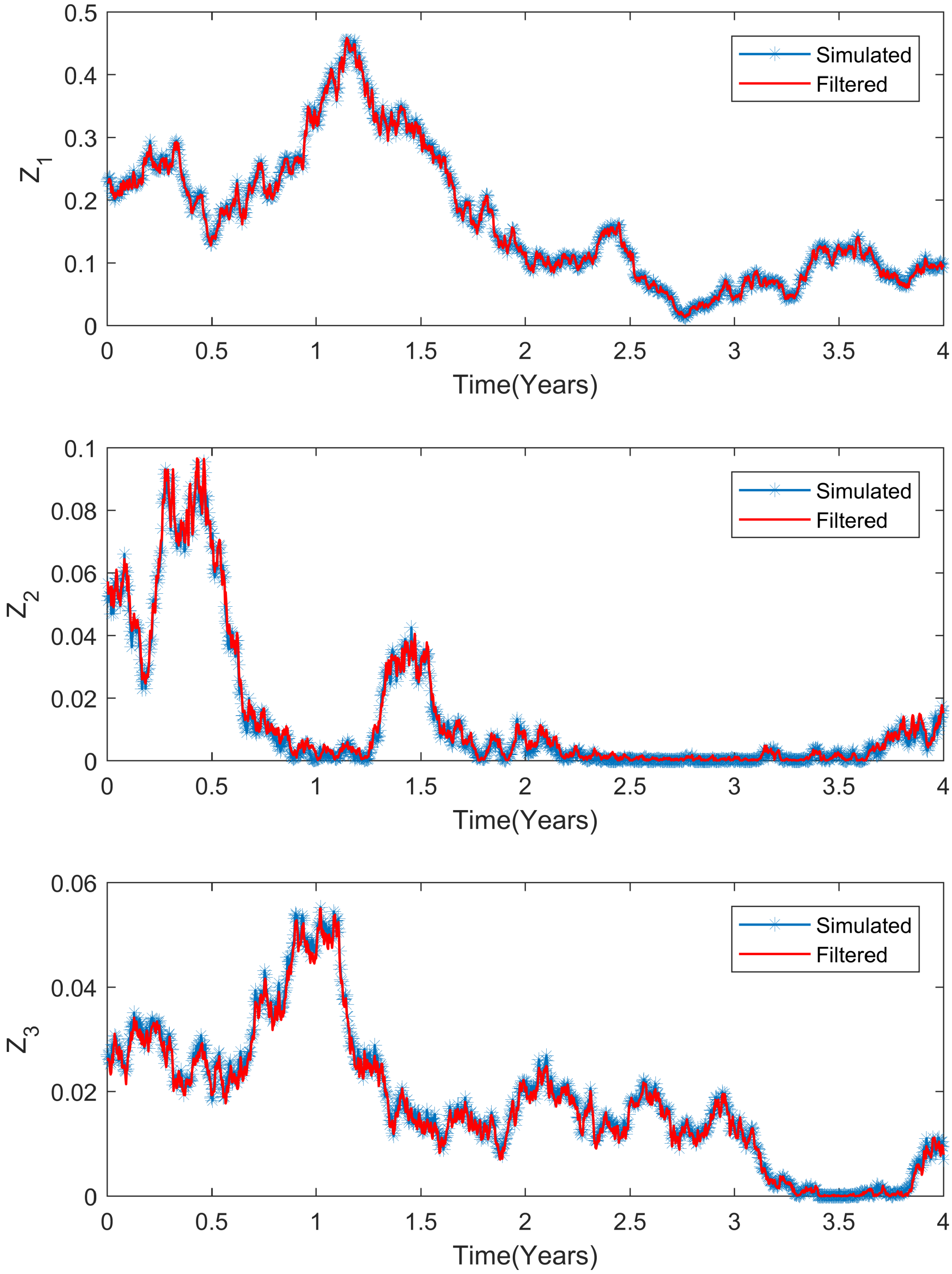


Fig. B.2: Simulated and filtered CEV factor process, using the estimated parameters



Tab. B.1: Three-factor model parameter recovery

Parameters	True Value	CEV Estimate	SQRT Estimate
α_1	0.6000	0.5349	0.5000
α_2	0.5000	0.4891	0.5000
α_3	0.4000	0.3805	0.5000
σ_1	0.3500	0.3078	0.2875
σ_2	0.3000	0.2756	0.2853
σ_3	0.1010	0.0956	0.1558
θ_1	0.6053	0.5963	0.5987
θ_2	0.2227	0.2269	0.2218
θ_3	0.1610	0.1562	0.1609
K_{11}	0.0978	0.0990	0.0988
K_{22}	0.4267	0.4156	0.4262
K_{33}	0.6724	0.6898	0.6692
K_{21}	-0.1570	-0.1581	-0.1579
K_{32}	-0.4859	-0.4747	-0.4854
δ_1	-0.4513	-0.5213	-0.5586
δ_2	-0.7609	-0.8087	-0.7643
δ_3	-3.0441	-3.3114	-2.0698
$\sigma_{\text{rates}} \times 10^4$	1.0000	1.0072	1.0089

B.2 The expression of the short rate given the robust estimation feature

The short rate, originally expressed in Equation (2.7), given the robust estimation feature

$$\beta = \mathbf{1}^\top K\theta = -\mathbf{1}^\top K_1 = -\mathbf{1}^\top K_2 > -\mathbf{1}^\top K_3,$$

can then be expressed as

$$\begin{aligned}
r_t &= \beta - \frac{\psi^\top K(\theta - Z_t)}{\phi + \psi^\top Z_t} \\
&= \beta - \frac{\mathbf{1}^\top K(\theta - Z_t)}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}} \\
&= \beta - \frac{\beta - \mathbf{1}^\top K Z_t}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}} \\
&= \beta - \frac{\beta - [\mathbf{1}^\top K_1, \mathbf{1}^\top K_2, \mathbf{1}^\top K_3] Z_t}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}} \\
&= \beta - \frac{\beta + \beta(Z_{1,t} + Z_{2,t}) - K_{33} Z_{3,t}}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}} \\
&= \frac{(\beta + K_{33}) Z_{3,t}}{1 + Z_{1,t} + Z_{2,t} + Z_{3,t}}.
\end{aligned}$$