Term structure modelling using the Nelson-Siegel framework

Rui Chen

Discipline of Finance, Business School

University of Sydney

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

2012.12

I would like to dedicate this thesis to my parents, my wife, Ivy, and my beloved families for their constant support and unconditional love.

Statement of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Rui Chen

Acknowledgements

First and foremost, I would like to thank my mother for her moral support and undying love because without it, I do not think I would have been able to accomplish anything. I also wish to thank my Heavenly Father for the talents and abilities he has given me.

My deepest gratitude and sincere thanks goes to my advisors, Jiri Svec and Maurice Peat. They have been supportive of the direction that I have taken and have allowed me to take control of my research. I could always rely on them for wise advice. This thesis would not have been possible without their support and encouragement.

In addition to my advisors I have been fortunate to have many other great researchers here at USYD. Quan Gan and Juan Yao have given me lots of good advice on various issues from my research to career development. Hamish Malloch has suggested some interesting extensions. Graham Partington, Andrew Grant, Tro Kortian and Maxwell Stevenson have always been stimulating.

A special acknowledgement goes to my best friend and co-author Ke Du at Business School, UTS. He has been a true friend and is an amazing person in too many ways. I also thank Adam Corbett, my other great mate who has been supportive in every way.

Multiple generations of PhD students at Business School USYD have made my time here both more enjoyable and more productive. Many thanks go to Min, Peter, Xiaogang, Susana, Yibai, Max, Nate, Qian, Justin and Reza, amongst many others.

It is a pleasure to thank the research student program managers Mandy, Maryann and Violeta for looking after us and making the research centre a second home to us. Financial support from the Business School is also gratefully acknowledged.

And finally, I am eternally obliged to my lovely wife, Ivy for her understanding and support through the duration of my studies. She has always been here for me when I have been frustrated throughout my PhD research. Her support and encouragement is in the end what makes this thesis possible.

Publications and Working Papers Relevant to the Thesis

1. **Chen, R.**, K. Du (2012). A generalized arbitrage-free Nelson-Siegel model: The impact of unspanned stochastic volatility. Forthcoming at *Finance Research Letters*

2. **Chen, R.** (2012). Pricing interest rate derivatives in a generalized arbitrage-free Nelson-Siegel model. *Working paper*

3. **Chen, R.**, K. Du and J. Svec (2012). On the joint calibration of LIBOR/Swap and interest rate derivatives under a latent factor model. *Working paper*

4. **Chen, R.**, J. Svec and M. Peat (2011). Modeling the government bond in Australian fixed-income market. *under review*

Abstract

Nelson and Siegel [1987] introduce a smooth exponential function to analyze the term structure of interest rates. This greatly facilitates the literature in the Nelson-Siegel model family. In this thesis, the Nelson-Siegel framework is studied extensively in terms of arbitrage-free restrictions, pricing of interest rate derivatives and empirical performance of both in-sample fitting and out-of-sample forecasting. Chapter 1 introduces the motivations and contributions of this thesis. In the following chapter, I review the literature of interest rate models that are used in this thesis.

In chapter 3, I study the performance of the dynamic Nelson-Siegel model in fitting and forecasting the government bond yields in Australia. The model is improved by utilizing a more powerful and robust state-space framework estimated with a Kalman filter. I show that this approach outperforms a random walk and a two-step estimation dynamic Nelson-Siegel model in forecasting the Australian government term structure across various forecasting horizons.

Using the results from chapter 3 and Diebold and Li [2006], the dynamic Nelson-Siegel model provides exceptional in-sample fitting and out-of-sample forecasting of interest rates. However, the lack of theoretical background is criticized by academics and practitioners, such

as the absence of arbitrage-free pricing. In this chapter, I develop a general arbitrage-free Nelson-Siegel model under the HJM framework, see Heath, Jarrow, and Morton [1992]. It maintains a Nelson-Siegel factor loading structure and features unspanned stochastic volatility factors. The corresponding market price of risk is derived based on Duffie, Pan, and Singleton [2000], Trolle and Schwartz [2009a], and Christensen, Diebold, and Rudebusch [2011]. The price of interest rate contingent claims, such as caps, swaptions and bond options, are derived from the generalized arbitrage-free Nelson-Siegel model based on Schrager and Pelser [2006] approximation procedure.

To overcome a computationally inefficient pricing scheme in chapter 4, chapter 5 derives a consistent pricing framework of interest rate caps and swaptions based on Carr and Madan [1999] and Duffie, Pan, and Singleton [2000]. This exploits the potential to jointly model the interest rates and their derivatives. I calibrate the model on an extensive panel data including Libor/Swap rate, At-The-Money(ATM) caps and swaptions. By casting the entire model into a state-space form, the extended Kalman filter is employed to calibrate the model. The results show that the model prices both interest rates and their derivatives accurately.

Finally, Chapter 6 concludes the thesis and discusses potential extensions.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Why interest rate models?

Interest rate models are one of the most sophisticated financial models applied to daily business. Consequently, a vast amount of research has been devoted to interest rate modelling. From my perspective, the reasons are three fold; analyzing the dynamics of interest rates, understanding the transmission mechanism of monetary policy and the pricing and hedging of fixed-income derivatives. Firstly, the ability to model the dynamics of interest rates is of crucial importance in many areas of finance. Equally important are accurate forecasts of future interest rates. Monetary policy is a second reason for modelling interest rates. The short-term rate is controlled by the central banks in most countries. The long-term yield, however, is the main factor required to understand and predict the future economy. For instance, mortgage decisions are often influenced by long-term mortgage rates instead of the short rate controlled by the central banks. Interest rate models are therefore designed to understand how the transmission mechanism works and how the central bank conducts monetary policy. The third reason is the pricing and hedging of fixed-income derivatives, the future payoffs of which are contingent on future interest rates. Standard products of interest rate derivatives, including interest rate caps and swaptions, are among the most liquid derivatives traded in financial markets. The exotic interest rate derivatives, which are designed to meet the particular needs of clients, can also be priced through specialized interest rate models. In addition, interest rate models provide a consistent framework to jointly model interest rates and their derivatives. For example, casting the interest rates and their derivatives into a consistent model provides additional information from fixed-income derivative markets for studying the dynamics of interest rates.

1.2 Interest rate models

There are various types of interest rate models, each with their own focus. This thesis employs three types of interest rate models: the Nelson-Siegel family, affine term structure models and Heath, Jarrow and Morton (HJM) models. This section introduces the literature on these models.

1.2.1 Nelson-Siegel framework

Nelson and Siegel [1987] introduce a parsimonious parametric model[1] to analyse the term structure of interest rates. Instead of modelling the yield of all maturities, the Nelson-Siegel model reduces the number of dimensions to be modelled

[1] Fit is a more appropriate description of the Nelson-Siegel contribution. However, to be consistent with literature the thesis refers to the Nelson-Siegel model.

to just three factors. The model consistently empirically fits the interest rate curves because it can generate most types of yield curve shapes. According to Bank for International Settlements [2005], nine out of thirteen central banks estimate their yield curves using the Nelson-Siegel model or its variants. Since its inception, the Nelson-Siegel model has been studied extensively by changing the numbers of parameters or using different specifications. Examples are the two-factor model (Diebold, Piazzesi, and Rudebusch [2005]), the three-factor model (Bliss [1996]) and the four-factor model (Svensson [1994], Björk and Christensen [1999]). The development of the Nelson-Siegel model has been covered in Diebold and Rudebusch [2013].

The dynamic Nelson-Siegel model estimated via a two-step estimation procedure is proposed by Diebold and Li [2006]. Diebold and Li [2006] show that the three-factor Nelson-Siegel model not only provides good in-sample fittings, but can also forecast the future interest rate by incorporating time-varying factors, especially for longer forecast horizons. Since then many studies have confirmed these results in different markets, for example government zero-coupon bond rates in Japan, Germany, and the UK, (Diebold, Li, and Yue [2008]), corporate bond rates in the US market, (Yu and Salyards [2009]), government bond rates in Brazil, (Vicente and Tabak [2008]) and government bond rates in China, (Luo, Han, and Zhang [2012]).

Although the dynamic Nelson-Siegel model forecasts interest rates well, it has a theoretical drawback that is not arbitrage free, see Filipovic [1999], Filipovic [2000], Diebold, Piazzesi, and Rudebusch [2005] and Krippner [2006]. Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011] propose a new class of affine arbitrage-free models by adding an additional arbitrage-free term. The proposed affine arbitrage-free Nelson-Siegel model combines the advantages of affine term structure models and the dynamic Nelson-Siegel models.

1.2.2 Affine term structure model

Affine term structure models, which are originally proposed by Vasicek [1977] and Cox, Ingersoll, and Ross [1985], are extensively analyzed in Duffie and Kan [1996] and Dai and Singleton [2000]. The term, affine, refers to the zero-coupon rates and their stochastic differential equations under both real and risk-neutral measures described by an affine function of a group of state variables. The closed-form solutions of interest rate derivatives greatly facilitates research on the affine term structure models. Affine models have numerous advantages, such as the correlated and mean reverting factors, stochastic volatility and closed-form zero-coupon bond price. Due to its affine structure of zero-coupon bond rate, a vast amount of research has focussed on estimating various forms of affine interest rate models, (e.g., Chan, Karolyi, Longstaff, and Sanders [1992], Chen and Scott [1992], Chen and Scott [1993], Duan and Simonato [1993], Longstaff and Schwartz [1993], Brown and Schaefer [1994], Duffie and Singleton [1997], Duan and Simonato [1999], Dai and Singleton [2000], Bali [2003], Duffee and Stanton [2004]).

Moreover, the conditional characteristic function of an affine process is known in closed-form, see Liu, Pan, and Pedersen [1999], Duffie, Pan, and Singleton [2000], and Singleton [2001], which allows the pricing of exotic interest rate derivatives. The affine model is particularly attractive for pricing and hedging interest rate derivatives. For example, Chen [1996] prices bond options and more sophisticated interest rate contingent claims by adding a stochastic mean reverting factor and a volatility factor. Munk [1999] shows that the price of a European coupon-bearing bond option can be approximated accurately through a group of European zero-coupon bond options by introducing the stochastic duration method. Singleton and Umantsev [2002] provide an semi closed-form solution for European options on coupon bonds using a numerically accurate and computationally fast method identical to the Fourier inversion method used for pricing European zero-coupon bond options in Duffie, Pan, and Singleton [2000]. Collin-Dufresne and Goldstein [2002] propose an Edgeworth expansion of the coupon bond density to price the swaptions approximately, whose moments are calculated by the joint moments of the associated zero-coupon bonds. Schrager and Pelser [2006] extend the literature on European coupon bond options pricing in both speed and accuracy based on the affine structure of approximate swap rate dynamics. The approximate method used in Schrager and Pelser [2006] is similar to Munk [1999].

1.2.3 HJM model

In contrast to the affine term structure model by modelling the dynamics of the short rate, Heath, Jarrow, and Morton [1992] provide a framework to directly model the dynamics of the instantaneous forward rate. The paper presents the no-arbitrage conditions that must be satisfied by the term structure models. The original idea comes from Richard [1978]. Heath, Jarrow, and Morton [1992] provide major breakthrough in the pricing of interest rate contingent claims by building a framework that must be satisfied by all of the models. By modelling the forward rate directly, the yield curve fitting is contained naturally in the HJM model. Extensive studies have been undertaken assuming Gaussian instantaneous forward rates, for example, DeMunnik [1992], Frachot, Janci, and Lacoste [1992], El Karoui and Lacoste [1992], Frachot [1995] and Miltersen [1994].

The applications of the HJM model, such as interest rate contingent claims pricing and computational methods, are extensively studied by Brace and Musiela [1995], Carverhill and Pang [1995], Jeffrey [1995], Rutkowski [1996], Miltersen and Persson [1999], Chiarella and Kwon [2000], Trolle and Schwartz [2009a] and Trolle and Schwartz [2009b].

Heath, Jarrow, and Morton [1992] derive the no-arbitrage condition which is guaranteed by the particular relationship between the drift term and the volatility term in the stochastic differential equation of the instantaneous forward rate. However, HJM models are non-Markovian in general. It is well known that interest rate derivatives cannot be priced in non-Markovian models using a PDE-based approach. The Markovian transformation method has been analysed by Björk and Svensson [2001], Chiarella and Kwon [2001] and Chiarella and Kwon [2003].

1.3 Contributions

This thesis extends the literature on the Nelson-Siegel model in several ways. They are classified into empirical extension and theoretical extension and are introduced in detail in the following sections.

1.3.1 Empirical extensions

The dynamic Nelson-Siegel model provides a better in-sample fitting and out-of-sample forecast, relative to competing models, including the affine term structure model (see Vicente and Tabak [2008]) and the Random Walk (see Diebold and Li [2006] and Koopman, Mallee, and Van der Wel [2010]). According to Litterman and Scheinkman [1991], most of the variation in the term structure can be explained by a three factor interest rate model and most interest rate models have focused on three-factor specifications, see Dai and Singleton [2002], Duffee and Stanton [2004] and Trolle and Schwartz [2009a]. These empirical findings have been confirmed across different markets. However, to best of my knowledge, the model has not been tested on the Australian government bond rates. A straightforward extension is therefore to apply the dynamic three-factor Nelson-Siegel model to the Australian government bond rates. The second extension can be employed in the estimation procedure. Although the two-step estimation procedure proposed by Diebold and Li [2006] is easy to implement, Diebold, Rudebusch, and Aruoba [2006] argue that the parameter estimation and signal extraction uncertainty associated with the first step are not considered in the second step. De Pooter [2007] also points out that the time series of the three factors can potentially be biased with outliers estimated in the first step. The second extension is therefore to test the performance of the one-step and two-step estimation procedures on the Australian government bond rates.

Chapter 3 studies the performance of the dynamic Nelson-Siegel model in fitting and forecasting the government bond yields in Australia. Using Australian Treasury notes from January 1994 to August 2010. The results are similar to the results obtained by Diebold and Li [2006] using US government bond yields from January 1985 to December 2000. In order to test the performance of two types of estimation procedures, the state-space Nelson-Siegel model is estimated via the Kalman filter to forecast the government bond yields in Australia. The results show that the one-step approach outperforms a random walk model and two-step approach in forecasting Australian government bond rates, which confirm the suggestions in Diebold, Rudebusch, and Aruoba [2006] and De Pooter [2007]. The dynamic Nelson-Siegel model via a one-step estimation procedure is therefore proved to be an appropriate framework to model and forecast the Australian government interest rates, which provides a basis for making policy decisions and investing in Australian fixed-income market.

1.3.2 Theoretical extensions

Although the dynamic Nelson-Siegel model performs well both in-sample and out-of-sample, see Diebold and Li [2006] and Koopman, Mallee, and Van der Wel [2010], Filipovic [1999] and Krippner [2006] show that it is not arbitrage-free. Without the arbitrage-free pricing, it is impossible to apply the Nelson-Siegel model to pricing interest rate derivatives. Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011] therefore present an arbitrage-free specification of the Nelson-Siegel model, which basically incorporates the Nelson-Siegel structure in the affine term structure framework by adding an additional arbitrage-free term. Interest rate volatility is assumed to be constant in Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011]. However, the interest rate is clearly stochastic and stochastic volatility contains important unspanned parts, see Li and Zhao [2006] and Trolle and Schwartz [2009a]. The first theoretical extension is therefore to derive an arbitrage-free Nelson-Siegel model that relaxes the constant volatility assumptions in the previous literature as well as incorporating the unspanned stochastic parts. The second theoretical extension is to price interest rate derivatives under the generalised arbitrage-free Nelson-Siegel model. According to Chacko and Das [2002], since the payoff of interest rate caps and swaptions can be manipulated into particular bond options with a specialized strike price, pricing the interest rate caps and swaptions is equivalent to pricing the bond options, including zero-coupon bonds and coupon-bearing bonds.

Chapter 4 tries to fill the gap in the literature by introducing a generalised arbitrage-free Nelson-Siegel model, which is derived from a more generalised HJM framework. According to Trolle and Schwartz [2009a], it is widely believed that the unspanned stochastic volatility can be modeled much easier in a HJM framework. The generalised arbitrage-free Nelson-Siegel model relaxes the constant volatility assumptions in Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011]. A Markovian transformation, see Chiarella and Kwon [2001], is introduced to model the additional arbitrage-free term, which is assumed to be locally riskless in my model. The generalised arbitrage-free Nelson-Siegel model maintains the Nelson-Siegel factor loading structure and features unspanned stochastic volatility factors. By assuming a constant volatility, the model in Christensen, Diebold, and Rudebusch [2011] can also be derived based on my model framework as a special case. The corresponding market price of risk is also derived based on Duffie, Pan, and Singleton [2000], Trolle and Schwartz [2009a] and Christensen, Diebold, and Rudebusch [2011].

The dynamic Nelson-Siegel model is incorporated into the HJM framework by featuring the unspanned stochastic volatility and thus represents an original contribution. Swaption prices are derived under my framework based on Schrager and Pelser [2006] approximation scheme. However, their approach is based on the affine term structure models, my swaption prices are derived under the HJM framework and thus represents another original contribution.

However, the method of Schrager and Pelser [2006] is computationally inefficient in the calibration procedure. Traders demand the models not only price accurately, but also price derivatives real-time. I therefore introduce a second pricing method according to a Fourier transformation in Carr and Madan [1999] and the transform method in Duffie, Pan, and Singleton [2000] in chapter 5. This method provides a consistent and efficient way to price the interest rate caps and swaptions, which also exploits the potential to jointly model interest rates and their derivatives. The joint framework is then calibrated on an extensive panel data including Libor/Swap rate, ATM caps and swaptions. By casting the entire model into a state-space form, I employ a extended Kalman filter to calibrate the model. The results show that the consistent model derived in chapter 5 price interest rates and their derivatives accurately. More importantly, it extends the Schrager and Pelser [2006] method on speed. This consistent framework can be applied to pricing and hedging interest rate derivatives. In addition, the semi-closed form solutions provide an accurate way to calculate the Greeks for interest rate derivatives, which are very important sensitivity measures for interest rate risk management.

The remainder of this thesis is organized as follows: In the next chapter, I introduce the Nelson-Siegel model framework and its extensions. In Chapter 3, I study the performance of the dynamic Nelson-Siegel model in modelling and forecasting the government bond yields in Australia. The model is further casted into a state-space model, which is then calibrated via Kalman filter algorithm. Chapter 4 presents a general arbitrage-free Nelson-Siegel model via the HJM framework and the price of interest rate contingent claims under the general arbitrage-free Nelson-Siegel model. In chapter 5, the price of interest rate derivatives is derived in a consistent and computationally efficient framework, and then the model is calibrated on an extensive panel data via an extended Kalman filter. Chapter 6 concludes the thesis and discusses several future extensions.

Chapter 2

Literature reviews

This chapter discusses the literature on the Nelson-Siegel framework. I first introduce the original Nelson-Siegel model. Followed by a extensive explanation of its variants, the Nelson-Siegel model family is summarised into a general state-space model form. The literature on the calibration of the Nelson-Siegel model family is also discussed.

2.1 Nelson-Siegel family of models

2.1.1 Nelson-Siegel model

Nelson and Siegel [1987] suggest an exponential function to analyse the dynamics of the forward rate,

$$f(t,T) = \beta_1 + \beta_2 e^{-\lambda\tau} + \beta_3 \lambda\tau e^{-\lambda\tau}$$

where $f(t,T)$ denotes the forward rate on time t with time to maturity $\tau = T - t$. The parameters $\beta_1, \beta_2, \beta_3$ can be interpreted as the level, slope and curvature, respectively. The decay parameter λ controls the exponential decay rate as well as where β_3 achieves its maximum curvature loading. Large values of λ produce fast decay and provide better fitting of the short-term rate, while small values generate slow decay and fit the yield curve better over longer-term horizons. The model is flexible enough to generate most empirical interest rate shapes and consistently fits various interest rate curves, see Diebold and Li [2006], Vicente and Tabak [2008] and Yu and Salyards [2009].

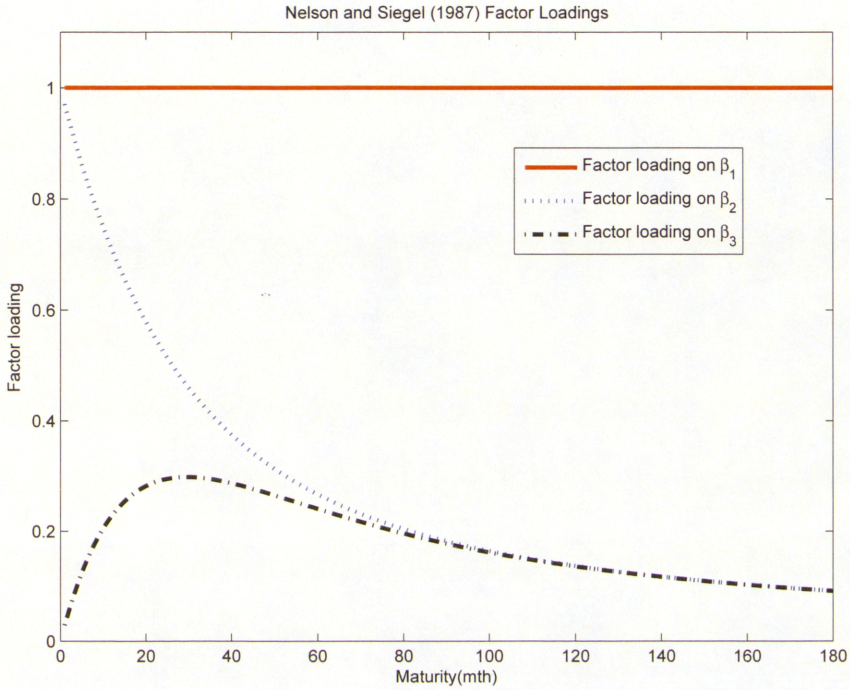
The corresponding zero-coupon rate $y(t,T)$ in the Nelson-Siegel model is given by,

$$y(t,T) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

where $y(t,T)$ denotes the zero-coupon rate on time t with a maturity at $\tau = T - t$. The three factor loadings, 1, $\frac{1-e^{-\lambda\tau}}{\lambda\tau}$, $\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}$ with a fixed λ ($\lambda = 0.0609$) are illustrated in Figure 2.1. The loading of β_1 is one; does not decay to zero in the limit which suggests that it is a long-term factor. The loading of β_2 decays monotonically and quickly from one to zero. Therefore, the loading suggests a short-term factor. Finally, the loading of β_3 increases from zero to the maximum value and then decays relatively slowly to zero; hence, it may be viewed as a mid-term factor.

Although the Nelson and Siegel [1987] model lacks a theoretical background, Diebold, Rudebusch, and Aruoba [2006] find that the level factor is a good proxy for inflation and the slope factor is highly correlated with real economic activity.

Figure 2.1: The three factor loadings of the Nelson and Siegel model ($\lambda = 0.0609$)



The model is known for its parsimonious yet good approximation of the yield curve of government bonds with over 99% explanation of the variation in yields. According to Bank for International Settlements [2005], the model is widely used in industry and nine out of thirteen central banks estimate their yield curves using the Nelson-Siegel model or one of its extensions.

2.1.2 Extensions of the Nelson-Siegel model

The Nelson-Siegel model has been widely extended in the literature. The variants of the Nelson-Siegel model are introduced in this section, such as Svensson [1994], Bliss [1996], Björk and Christensen [1999] and Diebold, Piazzesi, and Rudebusch [2005]. The corresponding factors and factor loadings are discussed in detail.

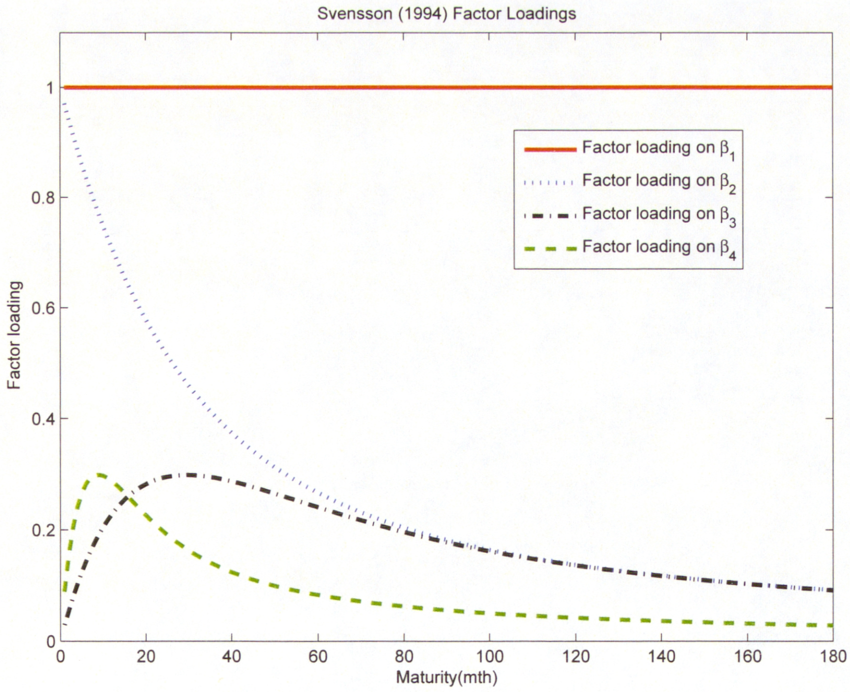
2.1.2.1 Svensson (1994)

Svensson [1994] proposes a four factor model by incorporating an additional curvature factor with a different λ value,

$$f(t,T) = \beta_1 + \beta_2 e^{-\lambda_1 \tau} + \beta_3 \lambda_1 \tau e^{-\lambda_1 \tau} + \beta_4 \lambda_2 \tau e^{-\lambda_2 \tau}$$

where $f(t,T)$ denotes the instantaneous forward rate on time t with time to maturity $\tau = T - t$. The parameters $\beta_1, \beta_2, \beta_3, \beta_4$ can be interpreted as the level, slope, first curvature and second curvature factor. The additional fourth factor increases the model flexibility and improves the in-sample fitting. Svensson [1994]

Figure 2.2: The four factor loadings of the Svensson model ($\lambda_1 = 0.0609, \lambda_2 = 0.177$)



demonstrates that the extension model easily fits the yield curve, particularly when the term structure of interest rates is abnormally shaped, by using the Sweden forward rate from 1992 to 1994. The corresponding spot rate $y(t, T)$ can be derived by integrating the instantaneous forward rate. Let $y(t, T)$ denote the spot rate maturity at time T, for a given date t. The spot rate is given by,

$$
\begin{aligned}
y(t,T) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right) \\
+ \beta_4 \left(\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \right)
\end{aligned}
$$

The exponential decay rate as well as the maximum value of curvature are controlled by two different λs, which turns out to provide more flexibility to the model. Instead of choosing to fit the short-term or the long-term rate better, the two different λs can provide better fitting simultaneously by choosing two appropriate values. The four factor loadings, with two fixed λs, are illustrated in Figure 2.2. Similar to Nelson and Siegel [1987], the loading of β_1 can be viewed as a long-term factor loading, and the loading of β_2 can be viewed as a short-term loading. However, the loading of β_3 and β_4 can be viewed together as a mid-term loading.

2.1.2.2 Bliss (1997)

Bliss [1996] extends the Nelson-Siegel model by relaxing the restriction that the slope and curvature loadings should be governed by the same λ. Therefore, Bliss [1996] suggests different values of λ for slope and curvature factor loadings in the

Nelson-Siegel model,

$$f(t,T) = \beta_1 + \beta_2 e^{-\lambda_1 \tau} + \beta_3 \lambda_2 \tau e^{-\lambda_2 \tau}$$

where $f(t,T)$ and $\beta_1, \beta_2, \beta_3$ represent the instantaneous forward rate, level, slope and curvature factors, respectively. λ_1 is the decay parameter controls exponential decay rate of slope, and λ_2 controls the maximum value of the curvature loading. The corresponding zero-coupon bond rate $y(t,T)$ can be derived by,

$$y(t,T) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \right)$$

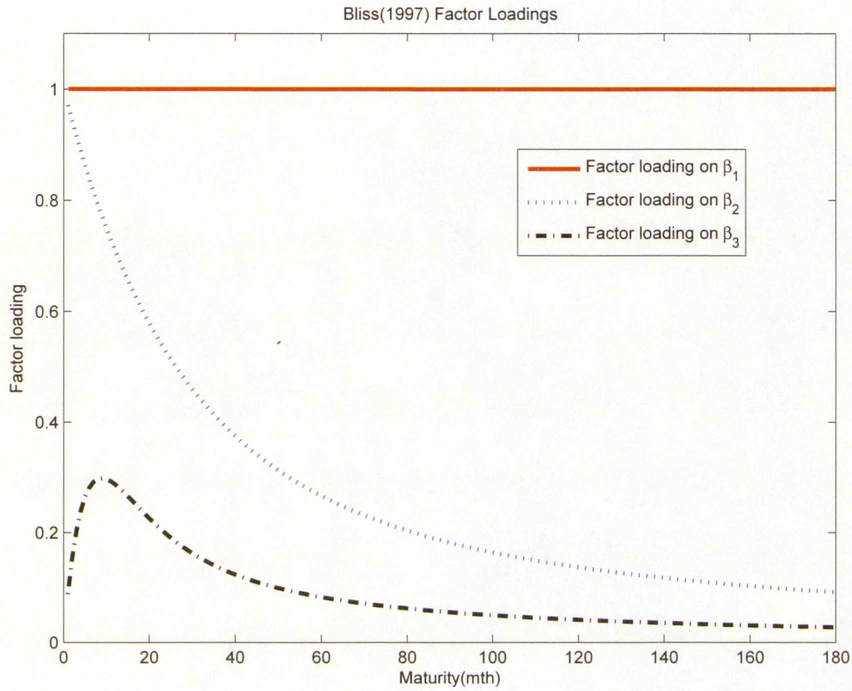
By incorporating a different λ, the author argues that the different decay parameters introduce more flexibility into the Nelson-Siegel framework. Intuitively, when $\lambda_1 = \lambda_2$, the Bliss model is identical to the original Nelson-Siegel model. The three factor loadings, with two fixed λs, are illustrated in Figure 2.3. Compared with Nelson and Siegel [1987] and Svensson [1994], the loading of β_1 remains the same. However, the slope and curvature factor loadings have different shapes because of the different decay parameters.

2.1.2.3 Bjork and Christensen (1999)

Björk and Christensen [1999] introduce a second slope factor and propose a four-factor Nelson-Siegel model,

$$f(t,T) = \beta_1 + \beta_2 e^{-\lambda_1 \tau} + \beta_3 \lambda_1 \tau e^{-\lambda_2 \tau} + \beta_4 e^{-\lambda_2 \tau}$$

Figure 2.3: The three factor loadings of the Bliss model ($\lambda_1 = 0.0609, \lambda_2 = 0.177$)

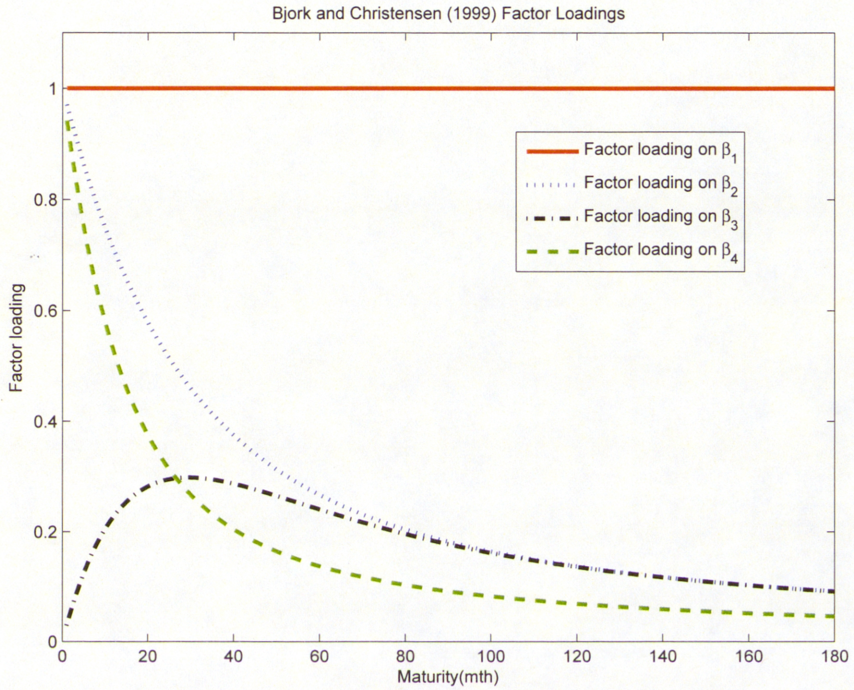


where $f(t,T)$ is the instantaneous forward rate and $\beta_1, \beta_2, \beta_3, \beta_4$ are level, curvature, first slope and second slope factors, respectively. The decay rate is governed by the summation of λ_1 and λ_2. The corresponding spot rate $y(t,T)$ is then given by,

$$\begin{aligned} y_t(\tau) = \beta_1 + \beta_2 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \right) + \beta_3 \left(\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right) \\ + \beta_4 \left(\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} \right) \end{aligned}$$

The slope of the yield curve can thus be explained by β_2 and β_4 instead of depending only on β_2. Similar to Svensson [1994], the different λs provide additional flexible to fit the interest rate curve. The four factor loadings, with two fixed

Figure 2.4: The four factor loadings of the Bjork and Christensen model ($\lambda_1 = 0.0609, \lambda_2 = 0.177$)



λs, are illustrated in Figure 2.4. Two slope parameters can explain more complicated shapes of the yield curve. According to Diebold, Rudebusch, and Aruoba [2006], the Bjork and Christensen Nelson-Siegel model provides a better in-sample fitting in terms of Root-Mean-Square-Error(RMSEs), compared with to the original Nelson-Siegel model using US government bond yields from January 1972 to December 2000.

2.1.2.4 Diebold Piazzesi and Rudebusch (2005)

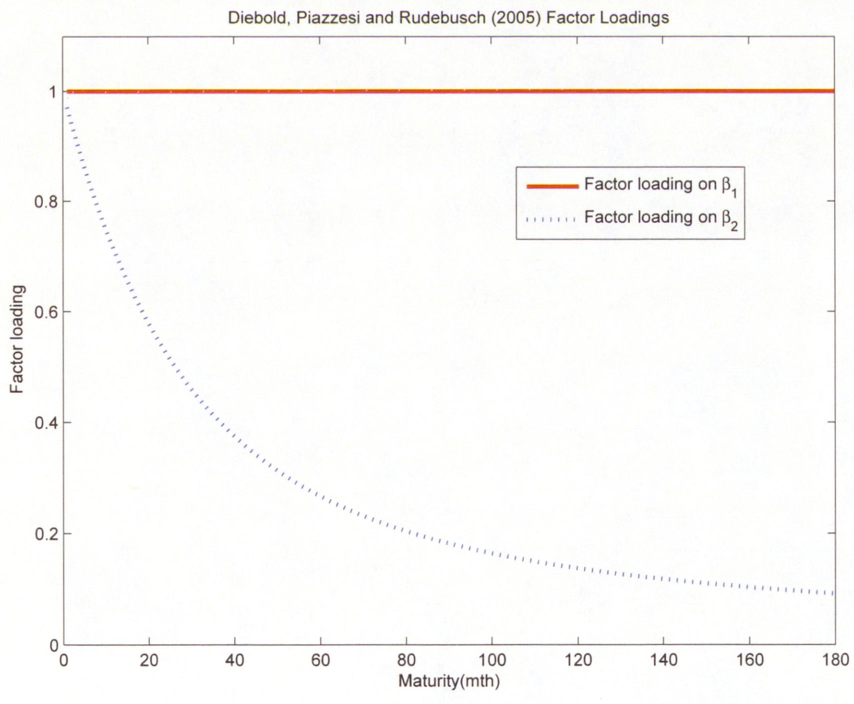
Diebold, Piazzesi, and Rudebusch [2005] introduce a two-factor Nelson-Siegel model. The instantaneous forward rate in Diebold, Piazzesi, and Rudebusch

[2005] is then given by,

$$f(t,T) = \beta_1 + \beta_2 e^{-\lambda\tau}$$

where β_1, β_2 are the level and slope factors. According to Litterman and Scheinkman [1991], most of the variation in term structure can be explained by three principal components, with the first two components explaining over 90%. Although Diebold, Piazzesi, and Rudebusch [2005] suggest the two-factor model is most likely not enough to fit the interest rate curve accurately, its parsimony may produce better out-of-sample forecasting. The corresponding spot rate $y(t,T)$ can

Figure 2.5: The two factor loadings of the Diebold Piazzesi and Rudebusch model ($\lambda = 0.0609$)



be derived by,

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

where the level and slope factor loadings are 1 and $\frac{1-e^{-\lambda\tau}}{\lambda\tau}$, which is illustrated in Figure 2.5 by fixing the decay parameter λ.

Table 2.1 summarises the Nelson-Siegel family of models introduced.

Table 2.1: Summary of Nelson-Siegel family of models

Model	Number of Factors	Level	Slope	Curvature
NS	3	1	1	1
Sv	4	1	1	2
Bl	3	1	1	1
BC	4	1	2	1
DPR	2	1	1	0

NS, Sv, Bl, BC and DRA represent Nelson and Siegel [1987], Svensson [1994], Bliss [1996], Björk and Christensen [1999] and Diebold, Piazzesi, and Rudebusch [2005]

2.1.3 Model calibration

2.1.3.1 State-space form

All the models listed in the previous section are variants of the original Nelson-Siegel model. This section captures the general model specification in a state-space framework. The variants of the Nelson-Siegel model in the previous section are special cases of equations (2.1) and (2.2) by selecting a different number of factors or a different value of the decay parameter.

The state-space model consists of a measurement equation and a transition equation. According to Harvey [1989] and Diebold, Rudebusch, and Aruoba [2006], the measurement equation for the Nelson-Siegel family is given by,

$$Y_t = H\boldsymbol{\beta}_t + \epsilon_t \tag{2.1}$$

where Y_t is the $N \times 1$ column vector of interest rates, which includes N different maturities. The matrix H is the $N \times K$ factor loading matrix and K is the number of the factors. For example, the value of K equals to three in the original Nelson-Siegel model. The time-varying factors $\boldsymbol{\beta}_t$ is the $K \times 1$ column vector of the factors, and $\boldsymbol{\epsilon}_t$ is a $N \times 1$ column vector of the measurement errors, which are assumed to be normally independent distributed.

The transition equation for the Nelson-Siegel family is given by,

$$\boldsymbol{\beta}_t = (I - \boldsymbol{\phi})\boldsymbol{\mu} + \boldsymbol{\phi}\boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t \tag{2.2}$$

where β_t is the $K \times 1$ column vector of the factors at time t. I is the $K \times K$ identity matrix and μ is the $K \times 1$ column vector, which represents the unconditional mean of the factors. Transition matrix ϕ is the $K \times K$ diagonal matrix since the factors are assumed to be independent. The error term η_t is a $K \times 1$ column vectors of the transition errors, which are also assumed to be normally independent distributed.

2.1.3.2 Estimation procedure

In general, two types of estimation methods are used to calibrate the state space model: two-step estimation and one-step estimation. As stated earlier, the two-step estimation procedure was introduced by Diebold and Li [2006] to calibrate the dynamic Nelson-Siegel model. In the first step, λ is assumed to be fixed which reduces the equations to a multi-linear regression problem. The three latent factors are then estimated by ordinary least squares from a cross-section of yields for each month t. In the second step, a time series process is employed to model three latent factors which is further used to fit and forecast the term structure of interest rates.

By using a AR(1) process, Diebold and Li [2006] show that the out-of sample forecasts outperform a random walk model, especially at longer forecast horizons. Similar conclusions have been obtained in forecasting the term structure of interest rates in Japan, Germany and the UK (Diebold, Li, and Yue [2008]), Brazil (Vicente and Tabak [2008]) and China (Luo, Han, and Zhang [2012]).

De Pooter [2007] and Diebold, Rudebusch, and Aruoba [2006] argue that the time series of the factors obtained from the Diebold and Li [2006] two-step estimation procedure can be potentially biased with outliers estimated in the first step. Rather than fitting the zero-coupon yield curve separately from the modeling of the factor dynamics, De Pooter [2007] estimates all parameters simultaneously with a one-step state-space approach. The parameters are estimated by the maximum likelihood function and Kalman filter used to obtain the optimal factor estimates. Although the number of parameters to estimate in a one-step approach is substantial, De Pooter [2007] argues that this method ensures that the dynamics of all parameters are taken into account simultaneously. His results illustrate better performance in out-of-sample forecasting of US data compared to the two-step approach.

2.1.4 The arbitrage-free Nelson-Siegel model

The Nelson-Siegel model could not exclude the probability of arbitrage. To the best of my knowledge, Björk and Christensen [1999] followed by Filipovic [1999] and Filipovic [2000] are the original articles establishing the inconsistency of Nelson-Siegel models with the absence of arbitrage. Sharef and Filipovic [2004] develop an arbitrage-free version of the Svensson [1994] extension to the Nelson-Siegel model and Krippner [2006] develops a special case of the arbitrage-free Nelson-Siegel model assuming constant market prices of risk and independent innovations. Christensen, Diebold, and Rudebusch [2009] also propose an extended arbitrage-free Nelson-Siegel model with additional slope and curvature components. Christensen, Diebold, and Rudebusch [2011] develop the general arbitrage-free Nelson-Siegel model by maintaining the Nelson-Siegel factor loading structure, in which the zero-coupon bond rate is given by,

$$y(t,T) = X_t^1 + \frac{1 - e^{-\lambda\tau}}{\lambda\tau} X_t^2 + \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) X_t^3 + L(t,T)$$

where X_t^1, X_t^2, and X_t^3 correspond to β_{1t}, β_{2t} and β_{3t}.[1] The model exactly matches the Nelson-Siegel factor loading structure, but there is an additional yield-adjustment term $L(t,T)$. One can also write the version of the affine arbitrage-free Nelson-Siegel model with respect to the instantaneous forward rate curve as follows,

$$f(t,T) = X_t^1 + X_t^2 e^{-\lambda\tau} + X_t^3 \lambda\tau e^{-\lambda\tau} + \frac{\partial L(t,T)\tau}{\partial\tau}$$

[1] The notation has been changed since the factors are now described by an stochastic differential equation instead of discrete time series models.

The probability of arbitrage is automatically excluded since the dynamic Nelson-Siegel model is incorporated into the framework of affine term structure models, see Duffie and Kan [1996] and Dai and Singleton [2000]. Krippner [2012] provides one form of theoretical foundation for Nelson-Siegel models, by relating them explicitly to the Gaussian affine term structure class of models, see Diebold and Rudebusch [2013]. Krippner [2008] also provides a theoretical basis for interpreting the arbitrage-free Nelson-Siegel components from a macroeconomic perspective.

2.2 General equilibrium model

Term structure modeling is one of the most active and sophisticated areas of financial theory applied to everyday business problems, ranging from managing the risk of a bond portfolio to the design and pricing of collateralized mortgage obligations. In this section, the affine term structure models are introduced. An understanding of these theoretical models is particularly useful in Chapter 4 and Chapter 5, where I derive the generalised arbitrage-free Nelson-Siegel model and the price of the interest rate derivatives.

2.2.1 Affine term structure model

The term structure of interest rates is defined as the relationship between the level of the interest rate and its corresponding time to maturity. A general version of the affine term structure model is introduced first, see Duffie and Kan [1996]. Affine models are developed in a general equilibrium setting, as originated by Vasicek [1977] and Cox, Ingersoll, and Ross [1985]. Duffie and Kan [1996] provide a characterization of the multifactor affine model, finally classified by Dai and Singleton [2000]. Simply put, in affine term structure models the bond price $P(t,T)$ maturing at time T is given by a simple exponential affine equation,

$$P(t,T) = e^{-A(t,T)-\mathbf{B}(t,T)'\mathbf{X}_t} \tag{2.3}$$

where $A(t,T)$ and $\mathbf{B}(t,T)$ are coefficients that depend on maturity τ and $\mathbf{X}_t$ satisfies a stochastic differential equation. Duffie and Kan [1996] suggest modeling the dynamics of interest rates with latent factors where the state vectors X_t cannot be obtained directly. Vasicek [1977] proposes the one-dimensional affine class. Additional dimensions are considered in Cox, Ingersoll, and Ross [1985] and Langetieg [1980]. Multifactor affine models include those of Longstaff and Schwartz [1993], Strickland [1996], Balduzzi, Das, and Foresi [1998] and Berardi and Esposito [1999]. A key contribution of the affine term structure models is the closed-form solution. Otherwise, numerical optimization, such as Monte Carlo, is required to solve these partial differential equations(PDEs).

The state variables X_t are assumed to follow a Markovian process and are further described by the following stochastic differential equation,

$$d\mathbf{X}_t = \boldsymbol{\mu}(t,\mathbf{X}_t)dt + \boldsymbol{\sigma}(t,\mathbf{X}_t)d\mathbf{W}_t \tag{2.4}$$

where $\boldsymbol{\mu}(t,\mathbf{X}_t) \in \mathbb{R}^k$, $\boldsymbol{\sigma}(t,\mathbf{X}_t)$ is a $k \times k$ matrix and $d\mathbf{W}_t$ is a k-dimensional Wiener process. Equations (2.3) and (2.4) together are the affine term structure framework. Applying Itô's lemma to equations (2.3) and (2.4), the dynamics of the bond price $P(t,T)$ can be derived by,

$$dP(t,T) = P(t,T)\Bigg(\Big[-A_t(t,T) - \mathbf{B}_t(t,T)\mathbf{X}_t - \boldsymbol{\mu}(t,\mathbf{X}_t)^\top\mathbf{B}(t,T) + \frac{1}{2}\sum_{i=1}^{k}\sum_{i=j}^{k}\mathbf{B}_i(t,T)\mathbf{B}_j(t,T)\boldsymbol{\sigma}_i(t,\mathbf{X}_t)\boldsymbol{\sigma}_j(t,\mathbf{X}_t)^\top\Big]dt - \mathbf{B}(t,T)^\top\boldsymbol{\sigma}(t,\mathbf{X}_t)d\mathbf{W}_t\Bigg)$$

If the market is free of arbitrage, there must exist a probability measure $\mathbb{P}^*$ to guarantee that the bond price $P(t,T)$ is a martingale,

$$dP(t,T) = r_t P(t,T)dt$$

where r_t is the risk-free rate. Given the bond price $P(t,T)$ is a martingale under the $\mathbb{P}^*$ measure, the drift term is then equal to the risk-free rate, which gives,

$$-A_t(t,T) - \mathbf{B}_t(t,T)\mathbf{X}_t - \boldsymbol{\mu}(t,\mathbf{X}_t)^\top\mathbf{B}(t,T) + \frac{1}{2}\sum_{i=1}^{k}\sum_{i=j}^{k}\mathbf{B}_i(t,T)\mathbf{B}_j(t,T)\boldsymbol{\sigma}_i(t,\mathbf{X}_t)\boldsymbol{\sigma}_j(t,\mathbf{X}_t)^\top\Big] - r_t = 0 \tag{2.5}$$

The drift condition in equation (2.5) provides general constraints between $A(t,T)$, $\mathbf{B}(t,T)$, $\boldsymbol{\mu}(t,\mathbf{X}_t)$ and $\boldsymbol{\sigma}(t,\mathbf{X}_t)$. There may be many solutions for $A(t,T)$, $\mathbf{B}(t,T)$, $\boldsymbol{\mu}(t,\mathbf{X}_t)$ and $\boldsymbol{\sigma}(t,\mathbf{X}_t)$. Duffie and Kan [1996] and Dai and Singleton [2000] provide the condition that is required to be placed on $\boldsymbol{\mu}(t,\mathbf{X}_t)$ and $\boldsymbol{\sigma}(t,\mathbf{X}_t)$ to guarantee the existence of $A(t,T)$ and $\mathbf{B}(t,T)$. The bond prices can then be calculated as solutions to a system of ordinary differential equations of $A(X_t,\tau)$ and $B(X_t,\tau)$. Because of their computational and empirical tractability, affine term structure models have been applied in finance for decades, including term structure model theory, pricing for interest rate derivatives and modeling of credit risk.

2.2.2 Short-rate models

Short-rate models are actually one-factor term structure models, because they treat the entire term structure of interest rates at any time as a function of a single state variable, the short rate of interest, see Merton [1974], Vasicek [1977], Cox, Ingersoll, and Ross [1985], Ho and Lee [1986], and Black, Derman, and Toy [1990]. Merton [1974] originates literature on the Gaussian short-rate model. The model is extended by Ho and Lee [1986]. They also develop the econometric procedure to calibrate the model to the current term structure. Black, Derman, and Toy [1990] further develop the calibration methodology in Ho and Lee [1986]. Alternatively, the framework of Cox, Ingersoll, and Ross [1985] is developed in a general equilibrium setting. In one-factor term structure models, the short rate r_t is given by a SDE of the following form:

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dB_t^Q \tag{2.6}$$

where $\mu : \mathbb{R} \times [0, T] \to \mathbb{R}$ and $\sigma : \mathbb{R} \times [0, T] \to \mathbb{R}^d$ satisfy technical conditions guaranteeing the existence of a solution to equation (2.6). The $\mu(r_t, t)$ and $\sigma(r_t, t)$ can be positive constants (time-homogenous short rate model) or deterministic functions of time (time-varying parameters short rate model).

2.2.2.1 Vasicek model

The framework of Vasicek [1977] is considered as one of the earliest models of interest rate dynamics. The author suggests that the short rate under the real world measure $\mathbb{P}$ evolves as an Ornstein-Uhlenbeck process,

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t \tag{2.7}$$

where κ, θ and σ are positive constants. The most important aspect of the Vasicek model is the mean reversion property, which implies that interest rates are always reverting towards a long-term equilibrium level θ. This property is both supported by economic theory and observed interest rate dynamics. Integrating equation (2.7) yields,

$$r_t = r_s e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) + \sigma \int_s^t e^{-\kappa(t-u)} dW_u$$

where r_t is normally distributed conditional on $\mathcal{F}_s$ with mean and variance,

$$\begin{aligned} \mathtt{E}(r_t|\mathcal{F}_s) &= r_s e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right) \\ \mathtt{Var}(r_t|\mathcal{F}_s) &= \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(t-s)}\right) \end{aligned}$$

The major drawback of this framework is that the short rate r_t can be negative with positive probability. However, without the normal distribution assumption, the analytical tractability of the interest rate model is rarely achieved, see Brigo and Mercurio [2007]. The framework of Vasicek [1977] is also notable because of its closed-form solution to a zero-coupon bond price $P(t,T)$, see equation (2.3).

$A(t,T)$ and $B(t,T)$ are given by,

$$A(t,T) = \frac{\sigma^2}{4\kappa}B(t,T)^2 - \left(\theta - \frac{\sigma^2}{2\kappa^2}\right)(B(t,T) - T + t)$$
$$B(t,T) = \frac{1}{\kappa}\left(1 - e^{-\kappa(T-t)}\right)$$

2.2.2.2 Cox, Ingersoll and Ross model

The framework of Cox, Ingersoll, and Ross [1985] has been a benchmark for many years because it overcomes the problem of negative interest rates by introducing a square root term into the diffusion coefficient. The dynamics of the short rate is given by,

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

where κ, θ and σ are positive constants. To prohibit the interest rate from being zero, the following condition is required to be placed on κ, θ and σ,

$$2\kappa\theta > \sigma^2$$

Following Cox, Ingersoll, and Ross [1985], the short rate process features a non-central chi-squared distribution under the risk-neutral measure $\mathbb{P}^*$. The mean and the variance of r_t conditional on $\mathcal{F}_s$ are therefore given by,

$$\mathtt{E}(r_t|\mathcal{F}_s) = r_s e^{-\kappa(t-s)} + \theta\left(1 - e^{-\kappa(t-s)}\right)$$
$$\mathtt{Var}(r_t|\mathcal{F}_s) = r_t\frac{\sigma^2}{\kappa}\left(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}\right) + \theta\frac{\sigma^2}{2\kappa}\left(1 - e^{-\kappa(t-s)}\right)^2$$

The zero-coupon bond price $P(t,T)$ is again given by equation (2.3), where,

$$A(t,T) = \exp\left(- \left[\frac{2he^{\frac{(\kappa+h)(T-t)}{2}}}{2h + (k+h)\left(e^{(T-t)h} - 1\right)}\right]^{\frac{2\kappa\theta}{\sigma^2}}\right)$$

$$B(t,T) = \frac{2\left(e^{(T-t)h} - 1\right)}{2h + (k+h)\left(e^{(T-t)h} -\right)}$$

$$h = \sqrt{\kappa^2 + 2\sigma^2}$$

The framework of Cox, Ingersoll, and Ross [1985] is significantly more complicated than Vasicek [1977] in terms of derivation.

2.2.2.3 Dothan Model

Dothan [1978] proposes a driftless geometric Brownian motion to model the dynamics of interest rates under the real world measure $\mathbb{P}$,

$$dr_t = \sigma r_t dW_t$$

where σ is a positive constant. By incorporating a constant market price of risk, the rate SDE of Dothan [1978] can be written in the following manner,

$$dr_t = ar_t + \sigma r_t dW_t \tag{2.8}$$

where a is a real constant. This is actually a continuous version of the framework of Rendleman and Bartter [1980]. Integrating equation 2.8,

$$r_t = r_s e^{\left(a - \frac{1}{2}\sigma^2\right)(t-s) + \sigma(W_t - W_s)}$$

The short rate r_t process features a lognormal distribution conditional on $\mathcal{F}_s$. The mean and the variance of r_t conditional on $\mathcal{F}_s$ are therefore given by,

$$\begin{aligned} \mathtt{E}(r_t|\mathcal{F}_s) &= r_s e^{a(t-s)} \\ \mathtt{Var}(r_t|\mathcal{F}_s) &= r_s^2 e^{2a(t-s)} \left(e^{\sigma^2(t-s)} - 1\right) \end{aligned}$$

This model also overcomes the main drawback in Vasicek [1977] since the short rate r_t is always positive because of the definition of a lognormal distribution. Unfortunately, the Dothan model is only a mean-reverting interest rate model when $a < 0$, see Brigo and Mercurio [2007]. The zero-coupon bond price of the Dothan model is derived in analytical form. However, the pricing formula can not be written in the manner of equation (2.3). Therefore, the Dothan model is a short rate model that is not included in the affine term structure framework.

2.2.2.4 Hull and White model

Vasicek [1977], Dothan [1978] and Cox, Ingersoll, and Ross [1985] are time-homogeneous short-rate models that provide a poor fit to the initial term structure of interest rates, see Ho and Lee [1986] and Hull and White [1990]. Hull and White [1990] propose a time-varying parameter in the Vasicek model to accurately fit the currently-observed yield curve,

$$dr_t = (\theta_t - a_t r_t)dt + \sigma_t dW_t$$

where θ_t, a_t and σ_t are deterministic functions of time. Although it fits the current yield curve perfectly, the perfect fitting of a volatility term structure can be rather misleading because of the liquidity issues of financial markets, see Hull and White [1995]. Hull and White [1994] propose a different framework,

$$dr_t = (\theta_t - ar_t)dt + \sigma dW_t \tag{2.9}$$

where a and σ are positive constants and θ_t is a deterministic function of time. Integrating equation (2.9),

$$r_t = r_s e^{-a(t-s)} + \alpha_t - \alpha_s e^{-a(t-s)} + \sigma \int_s^t e^{-a(t-u)} dW_u$$

where

$$\alpha_t = f(0,t) + \frac{\sigma^2}{2a^2}(1 - e^{-at})^2$$
$$f(0,t) = -\frac{\partial \ln P(0,t)}{\partial t}$$

$f(0,t)$ is the instantaneous forward rate at time zero for the maturity t. The mean and the variance of r_t conditional on $\mathcal{F}_s$ are therefore given by,

$$\mathtt{E}(r_t|\mathcal{F}_s) = r_s e^{-a(t-s)} + \alpha_t - \alpha_s e^{-a(t-s)}$$
$$\mathtt{Var}(r_t|\mathcal{F}_s) = \frac{\sigma^2}{2a}\left(1 - e^{-2a(t-s)}\right)$$

Again, the dynamics of r_t feature a normal distribution conditional on $\mathcal{F}_s$. The zero-coupon bond price can be classified into equation (2.3), where $A(t,T)$ and $B(t,T)$ are given by,

$$A(t,T) = \left(\frac{\sigma^2}{4a}(1 - e^{-2at})B(t,T)^2 - B(t,T)f(0,t)\right) e^{\frac{P(0,T)}{P(0,t)}}$$
$$B(t,T) = \frac{1}{a}\left(1 - e^{-a(T-t)}\right)$$

The framework of Hull and White [1994] provides an exact fit of interest rates by setting a time-varying θ_t. The Gaussian distribution of the short rate r_t allows for the derivation of closed-form solutions of interest rate contracts. However, similar to Vasicek [1977], the short rate r_t may have negative values with positive probability.

Table 2.2 summarises the short-rate models introduced.[1]

Table 2.2: Summary of short-rate models

Model	Dynamics	$r > 0$	Distribution
V	$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t$	N	normal
CIR	$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t$	Y	noncentral chi-squared
D	$dr_t = ar_t + \sigma r_t dW_t$	Y	lognormal
HW	$dr_t = (\theta_t - ar_t)dt + \sigma dW_t$	N	normal

V, CIR, D, and HW represent Vasicek [1977], Cox, Ingersoll, and Ross [1985], Dothan [1978] and Hull and White [1994], N and Y represent No and Yes.

2.3 HJM framework and its no-arbitrage condition

Heath, Jarrow, and Morton [1992] provide a framework to directly model the dynamics of instantaneous forward rate instead of the short rate in section 2.2.2, which is defined by,

$$f(t,T) = -\frac{\partial \ln P(t,T)}{\partial T} \tag{2.10}$$

[1]This summary is similar to Table 3.2 of Brigo and Mercurio [2007]

By integrating both sides of equation (2.10),

$$\int_t^T f(t,u)du = -\int_t^T \frac{\partial \ln P(t,u)}{\partial u} du$$
$$= -\ln P(t,T)$$

The bond price $P(t,T)$ can be derived as follows,

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

Intuitively, the instantaneous forward rate is the fundamental building block in the HJM model. The stochastic differential equation of the instantaneous rate is therefore assumed explicitly under real world measure,

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)^\top dW_t$$

where W_t is a d-dimensional standard Brownian motion. $\alpha(t,T)$ and $\sigma(t,T)$ are $\mathfrak{F}_t$-adapted processes for all $T > 0$. Thus, according to the proof in Heath, Jarrow, and Morton [1992], the no-arbitrage condition is guaranteed by the particular relationship between the drift $\alpha(t,T)$ and the volatility $\sigma(t,T)$.

$$\alpha(t,T) = \sigma(t,T)^\top \int_t^T \sigma(t,u)du$$

This is the key result in the HJM model and is used as my no-arbitrage condition for the generalised arbitrage-free Nelson-Siegel model. However, HJM models are non-Markovian in general. It is well known that interest rate derivatives cannot be priced in a non-Markovian models by using a PDE-based approach. Since the unspanned stochastic volatility arbitrage-free Nelson-Siegel model is derived in the HJM framework, it has to facilitate a Markovian version to be able to price interest rate derivatives. Related studies are Miltersen [1994] and Frachot [1995]. The Markovian decomposition method used in this thesis comes from Chiarella and Kwon [2001] and Chiarella and Kwon [2003].

Chapter 3

Modeling and forecasting the government bond term structure in Australia

This chapter looks at the performance of interest rate models in fitting and forecasting the government bond yields in Australia. The Diebold and Li interest rate model is improved by utilizing a more powerful and robust state-space framework estimated via a Kalman filter. I show that this approach outperforms a random walk and the original Diebold and Li model in forecasting the Australian government term structure across various forecasting horizons.

3.1 Introduction

A vast amount of research has been devoted to modeling the term structure of interest rates and pricing of fixed-income derivatives. Significant advances have been achieved in equilibrium term structure models, summarized by Duffie and Kan [1996], where future yield dynamics depend on several state variables (see for example, Vasicek [1977]; Cox, Ingersoll, and Ross [1985]) and no-arbitrage models, focusing on the cross-sectional fitting of the interest rates (Ho and Lee [1986]; Hull and White [1990]). The ability to predict the yield curve is of equal importance in many areas of finance and of great interest to both academics and market practitioners. However, literature on forecasting the term structure of Australian interest rates is sparse considering Australian government bonds are playing an increasing role in international markets. Largely due to the governments good credit rating, the Australian market has attracted new international investors such as central banks and sovereign wealth funds. According to the Reserve Bank of Australias statement on monetary policy, as at the end of 2011, 75 per cent of Commonwealth government securities were held by non-residents, up from 60 per cent five years earlier. Interest rates are generally forecast with statistical models that express yields as a vector time series process (for example, Diebold and Li [2006]). These models generally outperform theoretical models at forecasting the yield curve, yet they still rarely outperform even the most basic random walk benchmark. Recent improvements in the forecasting performance of statistical models (see for example, De Pooter [2007]) have been driven by improvements in the estimation methodology.

This chapter focuses on statistical methods, forecasting the term structure of Australian Government bonds using a dynamic factor model proposed by Diebold and Li [2006] estimated with a State-Space model of De Pooter [2007]. Using an out-of-sample period from August 2001 to August 2010 I conclude that more efficient estimation produces forecasts that are superior to both the Diebold and

Li [2006] model and the Random Walk model.

The remainder of this paper is organized as follows: The next section presents the original Nelson and Siegel model framework and its recent extensions. In section 3.2, I introduce the State-Space representation of the model and the Kalman filter estimation procedure. Section 3.3 presents the dataset while section 3.4 presents the forecasting performance of the state-space Diebold and Li model in predicting Australian zero coupon government bond term structure. Section 3.5 summaries and concludes.

3.2 State space representation and the Kalman filter

The time series of the factors obtained from the Diebold and Li two-step estimation procedure can be potentially biased with outliers estimated in the first step, see De Pooter [2007] and Diebold, Rudebusch, and Aruoba [2006]. Rather than fitting the zero-coupon yield curve separately from the modeling of the factor dynamics, De Pooter [2007] suggests a one-step state-space approach via Kalman filter to calibrate all parameters simultaneously. The one-step state-space method outperforms the two-step approach in out-of-sample forecasting of US government bond yields. The Kalman filter has also been used to analyze the affine interest rates models, see Duffee and Stanton [2004] and general HJM models, see Trolle and Schwartz [2009a]. Consistent with these papers, all parameters are estimated simultaneously using information from both the cross-section and time-series of yields with a linear state space model specification estimated by a Kalman filter, in order to obtain more efficient estimates of βs.

A state-space model is specified by a measurement and a transition equation. In my case, the measurement equation represents the relationship between the market zero-coupon rate and the underlying state variables while the transition equation describes the time series dynamics of the underlying state variables. The optimal parameters are estimated via a Kalman filter by maximising the log-likelihood function, which is derived recursively from the state-space formulation. According to Duan and Simonato [1999] and Duffee and Stanton [2004], the Kalman filter is considered an efficient approach for modelling state variables that are not directly observed. Consistent with Diebold and Li [2006], λ is set to a fixed number 0.0609. [1] To clarify the notation, I substitute β_1, β_2, β_3 with L_t, S_t, C_t to represent the level, slope and curvature of the yield curve, respectively. In a linear Gaussian state space model, the transition equation is given by:

$$\begin{pmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{pmatrix} + \begin{pmatrix} \eta_t(L) \\ \eta_t(S) \\ \eta_t(c) \end{pmatrix}$$

I define

$$F = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

[1] λ does not necessarily have to be at Diebold and Li [2006] value. A different value may prove more suitable for different markets or data sets, see Krippner and Thorsrud [2009]. However, many authors have simply adopted the Diebold and Li [2006] value for their application, see De Pooter [2007], Vicente and Tabak [2008], the fixed value 0.0609 is therefore applied to consistent with these studies.

where $\eta_t|\mathcal{F}_{t-1} \sim N(0,Q)$. Since the factors are assumed to be independent, the Q matrix is therefore a diagonal matrix. μ_L, μ_S and μ_C are the unconditional mean of the three factors. The measurement equation can be written according to Diebold and Li [2006],

$$Y_t = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} + \begin{pmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_N) \end{pmatrix}$$

Where

$$Y_t = \begin{pmatrix} y(t,Y_1) \\ y(t,Y_2) \\ \vdots \\ y(t,Y_N) \end{pmatrix}$$

where $y(t,Y_i)$ indicates the zero-coupon rate maturating at time Y_i with time to maturity $\tau_i = Y_i - t$. I also define

$$H = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix}$$

where $\epsilon_t|\mathcal{F}_{t-1} \sim N(0,R)$ and matrix R is diagonal, thus I assume that the error terms in the measurement equation are normally and independently distributed.

The Kalman filter estimation technique is applied as follows.

Step 1 The Kalman filter is initialized with the unconditional mean and variance of the underlying unobserved state variable using historical data preceding the start of the sample.

$$E(X_{t_0}|\mathcal{F}_0) = \begin{pmatrix} L_t & S_t & C_t \end{pmatrix}^\top$$

Step 2 The recursive procedure begins by forecasting the measurement equation. The conditional forecast of the measurement equation is given by,

$$E(Y_{t_i}|\mathcal{F}_{t_{i-1}}) = HE(X_{t_i}|\mathcal{F}_{t_{i-1}})$$

Where the associated conditional variance has the following form,

$$Var(Y_{t_i}|\mathcal{F}_{t_{i-1}}) = HVar(X_{t_i}|\mathcal{F}_{t_{i-1}})H^\top + R$$

Step 3 Once the true value of the observed zero-coupon rate Y_{t_i} enters the recursive system, the error in the conditional prediction in Step 2 can be obtained from,

$$\zeta_{t_i} = Y_{t_i} - E(Y_{t_i}|\mathcal{F}_{t_{i-1}})$$

The prediction error is used to update the dynamics of transition equation in the Kalman filter technique. The updating equations are given by,

$$E(X_{t_i}|\mathcal{F}_{t_i}) = E(X_{t_i}|\mathcal{F}_{t_{i-1}}) + K_{t_i}\zeta_{t_i}$$

$$Var(X_{t_i}|\mathfrak{F}_{t_i}) = (I - K_{t_i}H)Var(X_{t_i}|\mathfrak{F}_{t_{i-1}})$$

where K_{t_i} is the Kalman gain matrix, which determines the weight given to the new observation in the updated state system forecast. It can be denoted as,

$$K_{t_i} = Var(X_{t_i}|\mathfrak{F}_{t_{i-1}})H^{\top}Var(Y_{t_i}|\mathfrak{F}_{t_{i-1}})$$

Step 4 In the recursive loop, the next time period state variables are forecasted conditioning on the information from 3 steps above. According to the transition equation, the conditional expectation of the underlying state variables during the next period is given by:

$$E(X_{t_{i+1}}|\mathfrak{F}_{t_i}) = (I - F)\mu + FE(X_{t_i}|\mathfrak{F}_{t_i})$$

and the conditional variance has the following form:

$$Var(X_{t_{i+1}}|\mathfrak{F}_{t_i}) = Var(X_{t_i}|\mathfrak{F}_{t_i}) + FVar(X_{t_i}|\mathfrak{F}_{t_i})F^{\top} + Q$$

Step 5 The likelihood function is constructed by repeating the previous four steps in the data sample. At each step a measurement-system prediction error ζ_{t_i} and a prediction error covariance matrix $Var(X_{t_i}|\mathfrak{F}_{t_{i-1}})$ is generated. Under the assumption that measurement-system prediction errors are Gaussian, the log-likelihood function can be given as follow:

$$\ell(\theta) = -\frac{nNln(2\pi)}{2} - \frac{1}{2}\sum_{i=1}^{N}[ln(det(Var(X_{t_i}|\mathfrak{F}_{t_{i-1}}))) + \zeta_{t_i}^{\top}Var(X_{t_i}|\mathfrak{F}_{t_{i-1}})^{-1}\zeta_{t_i}]$$

To find the optimal parameter set, the preceding algorithm is treated as the objective function and non-linear numerical optimization techniques are used to find the maximum.

3.3 Data and methodology

As zero-coupon bonds are not issued in Australia, I bootstrap the zero-coupon rate from Australian government coupon-bearing bonds using the Fama and Bliss [1987] assumption (constant forward rates between maturities of the yield curve data). I utilize 14 different maturities; 1, 3 and 6 months, and 1, 2, 3, 4, 5, 7, 10, 12 and 15 years observed at the monthly frequency from January 1994 to August 2010 for a total of 200 monthly observations. The period from January 1994 to July 2000 is used to estimate the model and then roll the calibration period ahead by one month to estimate a new set of parameters. A rolling forecast window ensures estimates are constructed from more recent data. Similar to all types of multi-factors term structure models, forecasting future yield curve is equivalent to forecasting the factors. I implement the Nelson-Siegel model preferred by Diebold and Li [2006], in which each of the three latent factors follows an AR(1) process. The factors are recursively estimated out-of-sample at a monthly interval for the 11-year period between August 2000 and August 2010 and construct point forecasts across the twelve maturities for four different horizons, 1-, 3-, 6- and 12-months ahead. For maturities of less than one year, I use Treasury notes or Overnight Indexed Swaps (OIS), if Treasury notes are not available. All data is supplied by the Reserve Bank of Australia. The descriptive statistics, summarized in 3.1, support the stylized facts of term structure, including predominantly upward sloping yield curve and high persistence. Persistence tends to slightly decrease with maturity while yield volatility generally tends to increase with maturity.

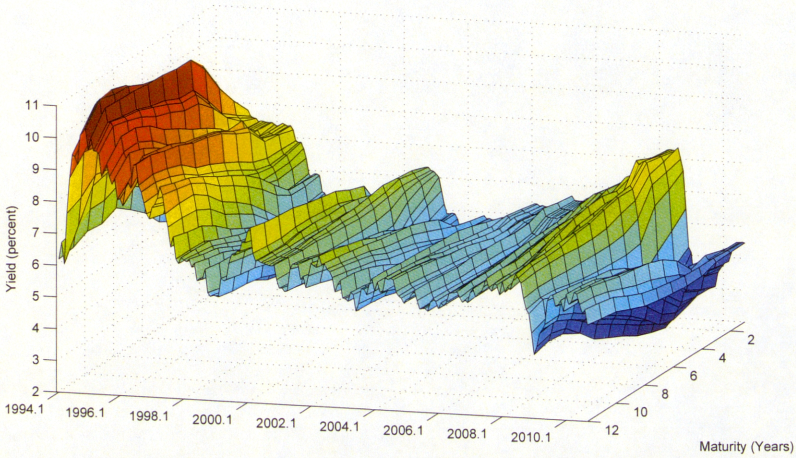
Table 3.1: Descriptive Statistics for Australian Government bonds

Maturity	μ	σ	Max	Min	$\rho(1)$	$\rho(12)$	$\rho(24)$
1	5.439	1.120	7.780	2.890	0.979	0.342	-0.189
3	5.464	1.174	8.350	2.590	0.977	0.336	-0.166
6	5.521	1.245	9.160	2.295	0.972	0.325	-0.134
12	5.638	1.311	10.023	2.430	0.965	0.441	-0.012
24	5.807	1.363	10.497	2.605	0.964	0.469	0.089
36	5.966	1.369	10.501	2.888	0.965	0.521	0.163
48	6.088	1.382	10.583	3.124	0.967	0.559	0.211
60	6.189	1.401	10.594	3.365	0.969	0.598	0.254
84	6.340	1.436	10.826	3.824	0.972	0.633	0.305
120	6.451	1.449	10.981	4.097	0.973	0.662	0.345
144	6.513	1.438	10.975	4.153	0.971	0.655	0.344
180	6.472	1.287	10.333	4.016	0.962	0.632	0.367

Descriptive statistics across the various bond maturities expressed in months. μ and σ, Max, Min represent mean, standard deviation, maximum and minimum yield on the zero coupon bond. $\rho(.)$ shows the level of autocorrelation with the order (in months) specified in the brackets. All autocorrelation coefficients are significant at the 5 per cent level of significance.

A 3-dimensional plot of the Australian zero-coupon rate is shown in Figure 3.1. It clearly demonstrates that the large amount of temporal variation in the Australian data comes from the interest rate level from the figure.

Figure 3.1: 3-dimension Yield Curves



3.4 Results

3.4.1 In sample performance

The above estimation procedure is used to forecast the term structure of the Australian Government bonds. The descriptive statistics for the Diebold and Li [2006] model are reported in Table 3.2.

Table 3.2: Descriptive Statistics for Diebold and Li model

Diebold and Li model			
Lt	St	Ct	μ
0.973	0	0	6.194
0	0.951	0	-0.943
0	0	0.869	-1.037

The parameter estimates reported in Table 3.2 are similar to the results obtained by Diebold and Li [2006] using US data. The level factor displays the highest persistency while the curvature factor has the fastest rate of mean-reversion.

The summary statistics of the fitted errors are shown in Table 3.3. The average error across all maturities is close to zero indicating a good in-sample fit.

Table 3.3: Descriptive Statistics for the fitted errors

Maturity	μ	σ	Max	Min	$\rho(1)$	$\rho(12)$	$\rho(24)$
1	-0.43	10.64	-50.17	38.25	0.81	-0.07	0.08
3	0.62	3.19	-8.85	16.29	0.45	0.02	0.13
6	0.17	12.55	-35.01	41.45	0.85	0.05	0.23
12	-1.12	12.80	-45.93	38.99	0.75	0.16	0.06
24	0.93	7.58	-21.70	42.04	0.78	-0.18	0.13
36	0.61	5.96	-21.42	22.62	0.82	0.16	0.00
48	0.71	6.59	-22.91	25.71	0.80	0.05	-0.04
60	0.14	7.48	-47.48	18.73	0.75	-0.08	-0.04
84	-1.86	6.80	-56.90	5.70	0.71	0.29	0.18
120	-1.75	9.19	-65.19	8.55	0.81	0.39	0.15
144	-3.37	7.89	-33.20	15.79	0.88	0.19	-0.19
180	5.36	15.95	-12.09	48.88	0.85	0.45	0.19

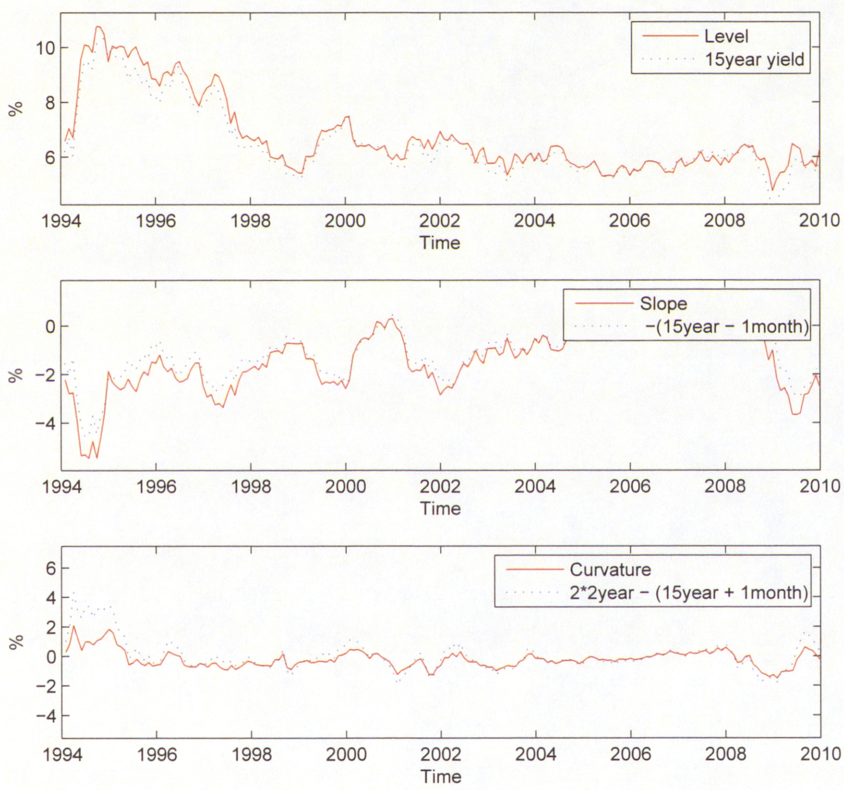
Descriptive statistics across the various bond maturities expressed in months. μ and σ, Max, Min represent mean, standard deviation, maximum and minimum yield on the zero coupon bond. (in bps) $\rho(.)$ shows the level of autocorrelation with the order (in months) specified in the brackets.

Figure 3.2 plots the actual and empirical level, slope, and curvature factors. The figure confirms the in-sample performance of the model is quite good with all three factors closely fitting the actual values for the Australian government bond zeros data.

3.4.2 Out of sample evaluation

This section investigates the performance of the State Space Diebold and Li model. There are numerous evaluation metrics available but ultimately I am interested in the model forecasting ability. The accuracy of the state space Diebold and Li model out-of sample forecasts is compared with that of the no change

Figure 3.2: Evaluation of the actual and empirical DL factors



forecast which is equivalent to a Random Walk process. I use the root mean squared prediction error (RMSPE) criteria and the Diebold and Mariano [1995] test for comparing predictive accuracy. The Random Walk model forecasts are constructed by using previous observations as forecasts of future rates. Although the Random Walk is one of the most elementary benchmarks, Duffee [2002] and Ang and Piazzesi [2003] all conclude that few term structure forecasting models have outperformed it. I also compare it to the Diebold and Li [2006] dynamic Nelson-Siegel model. The dynamic Nelson-Siegel model has superior forecasting power when compared with an affine term structure model and can produce better forecasts than the Random Walk model when forecasting yield maturities less than 5 years at the 12-month ahead horizon. The RMSPE in basis points (bps) and the Diebold-Mariano statistic across the 14 maturities at the 1-, 3-, 6- and 12-month horizons are displayed in Table 3.4 and 3.5. I disaggregate the forecast evaluation by its maturity.

The table 3.4 and 3.5 show that with an average RMSPE of 18.4 bps the 1-month maturity bonds estimated with the state space Diebold and Li model outperforms both the Random Walk and the dynamic Nelson-Siegel models at the 1-month ahead horizon. Its outperformance over the Random Walk is significant at the 5 per cent level of significance. However, neither model systematically outperforms the basic RW model benchmark at the 1-month horizon. The Random Walk model exhibits a marginally lower RMSPE of 26.6 bps compared to 27.2 bps and 27.6 bps, for the state space Diebold and Li model and two-step Diebold and Li model, respectively across all the maturities at the 1-month horizon.

However, the state space Diebold and Li model dominates the Diebold and Li

Table 3.4: Out-of-sample RMSE forecast evaluation for 1-, 3-, 6-, and 12-month ahead horizons (in bps), Part A

Maturity	Horizon	RW	DL	SSDL	SS vs RW	SS vs DL
1	1	22.088	18.619	18.399	-2.282	-1.517
	3	55.121	61.701	51.022	-1.570	-2.555
	6	96.117	112.424	90.695	-1.087	-1.321
	12	144.946	123.581	133.817	-0.879	0.460
3	1	23.098	23.955	23.626	0.884	-2.267
	3	60.089	71.832	59.623	-0.261	-2.910
	6	103.868	118.727	99.597	-0.862	-1.206
	12	152.145	127.787	138.712	-0.926	0.482
6	1	25.994	31.086	30.708	2.814	-2.516
	3	65.713	81.160	66.994	0.382	-3.084
	6	112.006	124.496	106.770	-0.783	-1.223
	12	159.830	131.545	142.289	-0.952	0.479
12	1	27.609	33.624	33.182	2.269	-2.386
	3	63.627	76.682	66.897	0.580	-1.977
	6	101.092	111.395	99.972	-0.237	-0.942
	12	137.114	114.047	125.981	-1.474	0.555
24	1	29.292	29.887	29.517	0.337	-1.963
	3	64.231	72.645	63.233	-0.320	-1.680
	6	96.884	103.210	92.258	-1.365	-1.323
	12	126.249	101.483	110.560	-1.740	0.529
36	1	29.921	29.450	29.083	-1.82	-2.012
	3	62.920	69.664	60.938	-0.826	-1.569
	6	91.326	97.057	86.358	-1.320	-1.524
	12	115.123	92.348	99.364	-1.898	0.518
48	1	28.836	29.058	28.667	-0.309	-2.201
	3	60.124	67.090	58.538	-0.680	-1.633
	6	85.577	92.696	81.105	-1.140	-1.692
	12	106.280	86.034	91.271	-1.980	0.477
60	1	27.767	28.220	27.815	0.087	-2.341
	3	56.572	63.422	55.191	-0.630	-1.671
	6	79.357	88.176	75.463	-1.059	-1.785
	12	95.489	80.076	83.282	-2.082	0.343
84	1	26.079	26.700	26.400	0.770	-1.780
	3	50.072	57.127	49.150	-0.458	-1.832
	6	69.994	81.802	66.271	-1.289	-2.157
	12	83.080	73.518	73.482	-2.167	-0.005
120	1	25.750	26.406	26.013	0.599	-2.334
	3	45.421	53.305	44.764	-0.370	-2.184
	6	61.674	78.233	58.995	-1.165	-2.496
	12	71.385	69.637	64.959	-1.872	-0.645

Table 3.5: Out-of-sample RMSE forecast evaluation for 1-, 3-, 6-, and 12-month ahead horizons (in bps), Part B

Maturity	Horizon	RW	DL	SSDL	SSDL vs RW	SSDL vs DL
144	1	26.123	26.061	25.857	-1.325	-1.216
	3	44.265	50.410	43.157	-0.897	-1.898
	6	59.212	74.560	55.993	-1.825	-2.425
	12	67.571	66.532	60.948	-2.400	-0.819
180	1	26.970	27.661	27.464	0.427	-1.232
	3	45.668	52.329	44.735	-0.326	-2.079
	6	61.410	76.846	57.557	-1.203	-2.654
	12	68.681	69.195	61.814	-2.005	-1.136

RW represents the Random Walk, DL represent the Diebold and Li two-step estimation via AR(1), SSDL represents the Diebold and Li one-step estimation via Kalman filter. The last two columns tabulate the D-M statistic. Under the null of equal predictive accuracy $D-M \sim N(0,1)$. Negative values indicate the state space Diebold and Li model is preferred to the competing model. We reject the null of predictive accuracy at the 5% level if $|D-M| > 1.96$

model across the whole yield curve at the 1-month ahead forecasts. I attribute the results to a more efficient recursive estimation technique combined with high persistence in the yields in the Australian data. As the forecast horizon increases to 3-months ahead, the State space Diebold and Li continues to dominate the Diebold and Li model and also begins to outperform the Random Walk highlighting the advantage of the recursive Kalman filter estimation technique. In the 3-month ahead forecast the State space Diebold and Li is superior to the Random Walk in 10 out of the 12 maturities. As the forecast horizon is further increased to 6-months ahead, both the Diebold and Li and the State space Diebold and Li begin to outperform the Random Walk and the advantage of the recursive Kalman filter estimation technique over the simple AR(1) process diminishes. Nevertheless, overall, the State space Diebold and Li model is still the preferred model outperforming the Random Walk and the Diebold and Li in all maturities, respectively. Interestingly the State space Diebold and Li forecasts perform worse than the Diebold and Li forecasts at the 12-month horizon, although they still outperform the Random Walk across all maturities. Overall, it can be concluded that although the State space Diebold and Li tends to outperform its rival models, its performance depends on the maturity and forecast horizon. I can summarize its performance into two parts. Firstly, the State space Diebold and Li model consistently outperforms the Diebold and Li model across all maturities in forecasts up to 6-months ahead and majority of the outperformance is significant at the 5 per cent level of significance. At longer horizons the recursive nature of the Kalman filter offers little benefit over an AR(1) process and thus it is not unexpected that the Diebold and Li model offers slightly better forecasts at the 12-month horizon, particularly at shorter maturities. Secondly, the State space Diebold and Li outperforms the Random Walk benchmark at maturities medium (3-months) to long term-maturities but the outperformance is only statistically significant at longer forecasting horizons. To check that these results are robust, I vary the starting point of the out-of-sample period between August 2000 (80th observation) to December 2003 (120th observation). I find that relative results do not change through time and are not affected by the financial crisis.

3.5 Conclusion

Superior yield curve forecasting provides a better base to understand fixed-income securities derivative pricing, portfolio allocation and risk management. Therefore, I evaluate the most successful empirical term structure model, Diebold and Li model, using data from the Australian fixed-income market. In this chapter, it is demonstrated that, based on Australian government bond data Diebold and

Li approach still fits and forecasts well, like in US data according to Diebold and Li [2006]. And I take a step further by incorporating Diebold and Li framework into a State-space model. Accordingly, the Kalman filter can be used to estimate parameters of my model instead of a simple OLS in Diebold and Li [2006]. My forecast results conclude that Kalman filter estimation procedure outperforms the original methodology of the Diebold and Li model, especially in one month ahead forecasting. The previous literature has shown that introducing dynamic coefficients into the Nelson-Siegel model not only provides good in sample fit but also outperformance over other benchmarks models in out-of sample forecasting, especially at the longer forecast horizon. I show that improved single-step estimation of the three parameters leads to an improvement of the out-of sample performance of the Nelson-Siegel model in forecasting the Australian government bond rates. The results confirm that the dynamic Nelson-Siegel factor model not only provides robust in sample fitting but also outperforms the other competitors on the out-of-sample forecast accuracy.

Chapter 4

A generalised arbitrage-free Nelson and Siegel model and pricing of interest rate derivatives

According to the results shown in chapter 3 and Diebold and Li [2006], the dynamic Nelson-Siegel model provides exceptional in-sample fitting and out-of-sample forecasting of interest rates. However, the lack of theoretical background on this model has been criticized by academics and practitioners. Most notably, the model does not exclude the possibility of arbitrage. In this chapter, I develop a general arbitrage-free Nelson-Siegel model under the HJM framework. The model maintains the Nelson-Siegel factor loading structure and features unspanned stochastic volatility factors. The corresponding market price of risk is derived based on Duffie, Pan, and Singleton [2000], Trolle and Schwartz [2009a], and Christensen, Diebold, and Rudebusch [2011] and the semi-closed form interest rate derivative prices are then derived from the Fourier transform techniques of Carr and Madan [1999], Duffie, Pan, and Singleton [2000] and Schrager and Pelser [2006].

4.1 Introduction

Derivative prices reflect investors' expectations of future movements in the underlying assets and hence contain the forward looking information relevant to these assets. In addition, interest rate derivatives, such as caps and swaptions, are among the most liquid securities being traded in financial markets. Thus, the information obtained from interest rate derivative prices may improve the performance of the dynamic Nelson-Siegel model. To extract this additional information from interest rate derivatives I require a framework which jointly models the interest rate dynamics and their derivative prices in the Nelson-Siegel structure.

However, the Nelson-Siegel family is not arbitrage-free, see Björk and Christensen [1999], Filipovic [1999], Filipovic [2000], Krippner [2006], Christensen, Diebold, and Rudebusch [2009], Krippner [2009], Christensen, Lopez, and Rudebusch [2010] and Christensen, Diebold, and Rudebusch [2011]. Without arbitrage-free constraints, it is impossible to jointly model interest rates and their derivatives. Krippner [2006] addresses that deficiency by deriving an intertemporally consistent and arbitrage-free version of the Nelson-Siegel model. His result is further extended by Christensen, Diebold, and Rudebusch [2011] to propose an arbitrage-free Nelson-Siegel model. By adding an arbitrage-free adjustment term, the structure of Nelson-Siegel models has been incorporated into classical affine term structure models in the spirit of Duffie and Kan [1996] and Dai and Singleton [2000]. The framework of Christensen, Diebold, and Rudebusch [2011] is thus arbitrage-free since it is an affine term structure model. However, the arbitrage-free Nelson-Siegel model may not be suitable for derivative pricing due to its constant volatility. Interest rates are clearly stochastic; this stochastic volatility contains important unspanned parts, see Collin-Dufresne and Goldstein [2002] and Li and Zhao [2006]. Therefore, it is important to incorporate the unspanned stochastic volatility into the Nelson-Siegel framework in order to price interest rate derivatives accurately.

A number of studies have been proposed for pricing interest rate derivatives. Singleton and Umantsev [2002] introduce a method that is based on an approximation of the exercise region in the space of underlying factors by line segments. This reduces the exercise probability of the swaption to the form of a caplet. A transform inversion is introduced to calculate the required exercise probabilities, (see Duffie, Pan, and Singleton [2000]). This method requires a simplifying assumption to bring the framework back to the affine term structure models, which is similar to Schrager and Pelser [2006]. Collin-Dufresne and Goldstein [2002] propose an Edgeworth expansion of the coupon bond density to price the swaption approximately. The moments of the coupon bond are approximated by a sequence of zero-coupon bonds, which are available in closed form. Munk [1999] introduces the stochastic duration of the coupon-bond to approximately price the coupon-bond options through zero-bond options. Schrager and Pelser [2006] and Trolle and Schwartz [2009a] introduce a similar approximation procedure in their derivation of swaption prices. Schrager and Pelser [2006] extend the literature on both speed and accuracy. The pricing structure has been incorporated into the affine term structure models using an approximation similar to Munk [1999].

In this chapter, I develop an unspanned stochastic volatility arbitrage-free

Nelson-Siegel model to relax the constant volatility assumptions in Christensen, Diebold, and Rudebusch [2011] while maintaining the Nelson-Siegel factor loading structures. In contrast to Christensen, Diebold, and Rudebusch [2011], which is derived from the affine term structure framework, I present the unspanned stochastic volatility arbitrage-free Nelson-Siegel model by incorporating HJM no-arbitrage constraint into the Nelson-Siegel framework, similar to the work of Devin, Hanzon, and Ribarits [2010]. Indeed, the arbitrage-free Nelson-Siegel model of Christensen, Diebold, and Rudebusch [2011] is a special case of the model introduced in this chapter.[1] Furthermore, I derive a Markovian representation (see Björk and Svensson [2001] and Chiarella and Kwon [2001]) of the additional term in Christensen, Diebold, and Rudebusch [2011], which is assumed to be locally riskless in my model. Thus, the unspanned stochastic volatility arbitrage-free Nelson-Siegel model is actually a Markovian factor term structure model, which enjoys the advantage of the factor model in estimation and the HJM models' advantage of being theoretically arbitrage-free. Finally, according to Duffie, Pan, and Singleton [2000], Cheridito, Filipović, and Kimmel [2007], Trolle and Schwartz [2009a], and Christensen, Diebold, and Rudebusch [2011], I derive the market price of risk in the unspanned stochastic volatility arbitrage-free Nelson-Siegel model. The Markovian representation method applied in this chapter is applicable to the entire family of Nelson-Siegel models discussed in chapter 2.

The prices of the zero-coupon option and interest rate caps are derived according to Duffie, Pan, and Singleton [2000] and Trolle and Schwartz [2009a].

[1] There are two Gaussian affine term structure models with unspanned/stochastic volatility that are conceptually similar to the model developed in this chapter, Fong and Vasicek [1991] and Balduzzi, Das, Foresi, and Sundaram [1996]. Of course, the proposed model in this chapter extends all of the models noted in these references.

The swaption prices are also derived by applying the framework of Schrager and Pelser [2006]. Although there are a number of studies on pricing interest rate contingent claims, as far as the author is aware, this is the first study to price interest derivatives in a Nelson-Siegel framework. This method provides a numerically accurate approximation to the prices of interest rate derivatives, which is substantially furthered the understanding for pricing interest rate securities.

The remainder of this chapter is organized as follows. In section 4.2, I derive the unspanned stochastic volatility arbitrage-free Nelson-Siegel model and the market price of risk specifications. In section 4.3, I derive the interest rate contingent claims under the framework of section 4.2. Section 4.4 summarises and concludes.

4.2 The model

4.2.1 Framework

Define $f(t,T)$ to be the time t instantaneous forward rate with maturity T, $T > t$ and let τ be the corresponding time to maturity, $\tau = T - t$. The proposed model for $f(t,T)$ is given by,

$$f(t,T) = \boldsymbol{\alpha}(t,T)^\top \mathbf{X}_t + Y_t \quad , \tag{4.1}$$

where $\mathbf{X}_t = (X_t^1, X_t^2, X_t^3)^\top$ is assumed to be a 3-dimensional stochastic process defined on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The factor loading $\boldsymbol{\alpha}(\cdot,\cdot)$ is a 3-dimensional deterministic function of the present time t and maturity T, which maintains the specific structure of the Nelson-Siegel model,

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1(t,T) \\ \alpha_2(t,T) \\ \alpha_3(t,T) \end{bmatrix} = \begin{bmatrix} 1 \\ e^{-\lambda(T-t)} \\ \lambda(T-t)e^{-\lambda(T-t)} \end{bmatrix} \tag{4.2}$$

The additional term Y_t ensures the model is arbitrage-free. This function is allowed to fluctuate slightly and is assumed to be locally riskless. The arbitrage-free term Y_t must have some risk (i.e. an unanticipated/stochastic component) because Y_t essentially accounts for the effect that volatility has on expected returns, and volatility is explicitly stochastic in the model specification. However, explicitly allowing for the risk in Y_t due to the stochastic volatility would introduce a substantial amount of additional complexity into the model, and that complexity can be mitigated with the simplifying approximation. It is worth noting that the proposed model exactly matches the dynamic Nelson-Siegel model of Diebold and Li [2006] without this additional term. Meanwhile, the zero-coupon rate $y(t,T)$ can be presented as follows,

$$y(t,T) = \frac{\left(\int_t^T \boldsymbol{\alpha}(t,u)^\top du\right)\mathbf{X}_t + \int_t^T Y(t,u)du}{T-t}.$$

Lemma 4.2.1. $\bar{\boldsymbol{\alpha}}(t,T) = \begin{bmatrix} \frac{1}{\lambda}\frac{\alpha_3(t,T)}{\alpha_2(t,T)} \\ \frac{1-\alpha_2(t,T)}{\lambda} \\ \frac{1-\alpha_2(t,T)-\alpha_3(t,T)}{\lambda} \end{bmatrix}$, *where* $\bar{\boldsymbol{\alpha}}(t,T) = \int_t^T \boldsymbol{\alpha}(t,u)^\top du$

Proof. Given $\bar{\boldsymbol{\alpha}}(t,T) = \int_t^T \boldsymbol{\alpha}(t,u)^\top du$ and equation (4.2),

$$\bar{\boldsymbol{\alpha}}(t,T) = \begin{bmatrix} \int_t^T 1dt \\ \int_t^T e^{-\lambda(T-t)}dt \\ \int_t^T \lambda(T-t)e^{-\lambda(T-t)}dt \end{bmatrix} = \begin{bmatrix} T-t \\ \frac{1-e^{-\lambda(T-t)}}{\lambda} \\ \frac{1-e^{-\lambda(T-t)}-\lambda(T-t)e^{-\lambda(T-t)}}{\lambda} \end{bmatrix}.$$

and recall my equation (4.2), I may write,

$$\bar{\boldsymbol{\alpha}}(t,T) = \begin{bmatrix} \frac{1}{\lambda}\frac{\alpha_3(t,T)}{\alpha_2(t,T)} \\ \frac{1-\alpha_2(t,T)}{\lambda} \\ \frac{1-\alpha_2(t,T)-\alpha_3(t,T)}{\lambda} \end{bmatrix}.$$

□

Assume $(\mathbf{X}_t^\top, v_t)$ to be a unique strong solution of the following stochastic differential equation under the risk-neutral measure $\mathbb{P}^*$,

$$d\begin{bmatrix} \mathbf{X}_t \\ v_t \end{bmatrix} = \begin{bmatrix} \mathbf{K} & 0 \\ 0 & k_v \end{bmatrix}\left(\begin{bmatrix} \boldsymbol{\theta} \\ \theta_v \end{bmatrix} - \begin{bmatrix} \mathbf{X}_t \\ v_t \end{bmatrix}\right) dt + \sqrt{v_t}\begin{bmatrix} \boldsymbol{\Sigma} & \mathbf{0} \\ \sigma_v^1 \; \sigma_v^2 \; \sigma_v^3 & \sigma_v^4 \end{bmatrix}\begin{bmatrix} dW_1^* \\ dW_2^* \\ dW_3^* \\ dW_4^* \end{bmatrix} \tag{4.3}$$

where $(W_t^{*1}, W_t^{*2}, W_t^{*3}, W_t^{*4})^\top$ is a 4 dimensional standard Brownian motion under the risk-neutral measure $\mathbb{P}^*$. The 3×3 matrix of parameters $\mathbf{K}$ is the reversion factor, while the 3×1 vector $\boldsymbol{\theta}$ is the central tendency factor. The 3×3 matrix of parameters $\boldsymbol{\Sigma}$ represents the constant part of volatility for the state variable $\mathbf{X}_t$. The parameter v_t is a stochastic volatility factor described by the long term mean k_v, the mean reversion factor θ_v, and the volatilities of the volatility $\boldsymbol{\sigma}_v = (\sigma_v^1, \sigma_v^2, \sigma_v^3, \sigma_v^4)$. It is worth noting that the unspanned stochastic volatility σ_v^4

is modeled into the unspanned stochastic volatility arbitrage-free Nelson-Siegel structure.

To obtain the SDE for the instantaneous forward rate, I apply Itô's lemma to equation (4.1),

$$\begin{aligned} df(t,T) &= d\left[\boldsymbol{\alpha}(t,T)^\top \mathbf{X}_t + Y_t\right] \\ &= \boldsymbol{\alpha}(t,T)^\top d\mathbf{X}_t + \frac{\partial \boldsymbol{\alpha}(t,T)^\top}{\partial t}\mathbf{X}_t dt + dY_t \\ &= \left(\boldsymbol{\alpha}^\top \mathbf{K}(\boldsymbol{\theta} - \mathbf{X}_t) + \frac{\partial \boldsymbol{\alpha}^\top}{\partial t}\mathbf{X}_t\right) dt + dY_t + \sqrt{v_t}\boldsymbol{\alpha}^\top \Sigma dW_t^*, \end{aligned} \tag{4.4}$$

where $\mathbf{W}_t^*$ is the vector containing three Brownian motions.

$$\mathbf{W}_t^* = (W_t^{*1}, W_t^{*2}, W_t^{*3})^\top \quad .$$

The 4th Brownian motion $\mathbf{W}_t^{*4}$ is used to model the unspanned stochastic volatility, therefore it does not appear in the SDE of $f(t,T)$ in equation (4.4). Note that, the arbitrage-free term Y_t is a deterministic function in Christensen, Diebold, and Rudebusch [2011], however it is assumed to be locally riskless in my model to allow the stochastic volatility. This can be viewed as the sacrifice for relaxing the constant volatility assumption. The additional term Y_t can therefore be written into the drift term of equation (4.4),

$$df(t,T) = \left(\boldsymbol{\alpha}^\top \mathbf{K}(\boldsymbol{\theta} - \mathbf{X}_t) + \frac{\partial \boldsymbol{\alpha}^\top}{\partial t}\mathbf{X}_t + \frac{dY_t}{dt}\right) dt + \sqrt{v_t}\boldsymbol{\alpha}^\top \boldsymbol{\Sigma} dW_t^* \quad .$$

The drift of $f(t,T)$ under the risk-neutral measure is thus given by,

$$\mu^*(t,T) = \boldsymbol{\alpha}^\top \mathbf{K}(\boldsymbol{\theta} - \mathbf{X}_t) + \frac{\partial \boldsymbol{\alpha}^\top}{\partial t}\mathbf{X}_t + \frac{dY_t}{dt} \tag{4.5}$$

4.2.2 Implementation of HJM no-arbitrage restriction

According to Heath, Jarrow, and Morton [1992], to guarantee that the proposed model is arbitrage-free, the drift term under the risk-neutral measure should be determined only by its volatility. In addition, according to Girsanov's theorem, the volatilities are identical under the risk-neutral measure $\mathbb{P}^*$ and real world measure $\mathbb{P}$. Thus, the drift under the risk-neutral measure $\mu^*(t,T)$ can be derived from equation (4.4),

$$\begin{aligned}
\mu^*(t,T) &= \left(\sqrt{v_t}\boldsymbol{\alpha}(t,T)^\top \boldsymbol{\Sigma}\right)\left(\int_t^T \sqrt{v_t}\boldsymbol{\alpha}(t,u)^\top \boldsymbol{\Sigma} du\right)^\top \\
&= \left(\sqrt{v_t}\boldsymbol{\alpha}(t,T)^\top \boldsymbol{\Sigma}\right)\boldsymbol{\Sigma}^\top \sqrt{v_t}\int_t^T \boldsymbol{\alpha}(t,u)du \\
&= v_t \boldsymbol{\alpha}(t,T)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}}(t,T)
\end{aligned} \tag{4.6}$$

Thus, to guarantee the proposed model to be arbitrage-free, the risk neutral drift $\mu^*(t,T)$ in equations (4.5) and (4.6) should match,

$$\begin{aligned}
\mu^*(t,T) &= \boldsymbol{\alpha}^\top(t,T)\mathbf{K}(\boldsymbol{\theta}-\mathbf{X}_t) + \frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t}\mathbf{X}_t + \frac{dY_t}{dt} \\
&= v_t \boldsymbol{\alpha}^\top(t,T)\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}}(t,T)
\end{aligned}$$

which provides the expression for the additional arbitrage-free term Y_t as,

$$\begin{aligned}
dY_t = \Big(& v_t \boldsymbol{\alpha}^\top(t,T)\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}}(t,T) \\
&- \boldsymbol{\alpha}^\top(t,T)\mathbf{K}(\boldsymbol{\theta}-\mathbf{X}_t) - \frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t}X_t\Big)dt
\end{aligned} \tag{4.7}$$

Note that there are no extra diffusion terms entering the HJM drift constraint in equation (4.6) since Y_t is assumed to be locally riskless process.

Lemma 4.2.2. *Assume the matrix of the reversion factor* $\mathbf{K}$ *is given as*

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}$$

The equation (4.7) *can be further reduced to* $dY_t = \left(v_t \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^\top \mathbf{K}\boldsymbol{\theta}\right) dt.$[1]

Proof. The reversion factor matrix $\mathbf{K}$ takes the following form:

$$\mathbf{K} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}$$

Following equation (4.7), dY_t can be written as,

$$dY_t = \Big(v_t \boldsymbol{\alpha}^\top(t,T)\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}}(t,T) - \boldsymbol{\alpha}^\top(t,T)\mathbf{K}\boldsymbol{\theta} \\ + \boldsymbol{\alpha}^\top(t,T)\mathbf{K}\mathbf{X}_t - \frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t}\mathbf{X}_t\Big) dt \ .$$

Intuitively, the main fluctuations of Y_t come from the $\mathbf{X}_t$ term. Substitute matrix $\mathbf{K}$ into the $\mathbf{X}_t$ terms in this equation, which gives,

$$\frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t} = \begin{bmatrix} 0 \\ \lambda e^{-\lambda\tau} \\ \lambda^2 \tau e^{-\lambda\tau} - \lambda\tau e^{-\lambda\tau} \end{bmatrix}$$

[1]This is identical to the reversion factor matrix $\mathbf{K}$ in Christensen, Diebold, and Rudebusch [2011]

and

$$\boldsymbol{\alpha}^\top(t,T)\mathbf{K} = \begin{bmatrix}1, e^{-\lambda\tau}, \lambda\tau e^{-\lambda\tau}\end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \lambda e^{-\lambda\tau} \\ \lambda^2\tau e^{-\lambda\tau} - \lambda\tau e^{-\lambda\tau} \end{bmatrix}$$

Therefore

$$\boldsymbol{\alpha}^\top(t,T)\mathbf{K}\mathbf{X}_t - \frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t}\mathbf{X}_t = \left(\boldsymbol{\alpha}^\top(t,T)\mathbf{K} - \frac{\partial \boldsymbol{\alpha}^\top(t,T)}{\partial t}\right)\mathbf{X}_t = 0$$

which implies that the locally riskless term Y_t thus becomes,

$$dY_t = \left(v_t \boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^\top \mathbf{K}\boldsymbol{\theta}\right) dt \quad .$$

□

Note that the main fluctuations due to the factors $\mathbf{X}_t$ vanish under the setting of the reversion matrix $\mathbf{K}$. The only source of fluctuation now comes from the stochastic volatility term v_t. The form of Y_t presented in Christensen, Diebold, and Rudebusch [2011] may be recovered by setting v_t equal to the constant $v_0 = 1$ without loss of generality,

$$dY_t = \left(\boldsymbol{\alpha}^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \bar{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^\top \mathbf{K}\boldsymbol{\theta}\right) dt \quad .$$

Thus, the arbitrage-free Nelson-Siegel model of Christensen, Diebold, and Rudebusch [2011] arises as a special case in a more general HJM framework. The additional arbitrage-free term can be further divided into two components:

a volatility and a drift contribution. In Christensen, Diebold, and Rudebusch [2011], both components are deterministic and consequently can be analysed together as a merged term. The volatility contribution $v_t\boldsymbol{\alpha}^\top(t,T)\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\bar{\boldsymbol{\alpha}}(t,T)$ is however locally riskless in my model and can not be analysed together with the drift component $-\alpha^\top(t,T)\mathbf{K}\boldsymbol{\theta}$ as in Christensen, Diebold, and Rudebusch [2011].

Without loss of generality, assume that,

$$Y_t = M(t,T) + D(t,T)$$

where $D(t,T)$ is the drift contribution component,

$$\begin{aligned} D(t,T) &= \int_t^T \boldsymbol{\alpha}^\top(s,T)\mathbf{K}\boldsymbol{\theta}ds \\ &= \int_t^T \alpha^\top(s,T)ds\mathbf{K}\boldsymbol{\theta} \\ &= \bar{\boldsymbol{\alpha}}^\top(t,T)\mathbf{K}\boldsymbol{\theta} \ . \end{aligned}$$

Note that at maturity T, $D(T,T) = 0$. The volatility contribution component $M(t,T)$ is defined by,

$$dM(t,T) = v_t\boldsymbol{\alpha}^\top(t,T)\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\bar{\boldsymbol{\alpha}}(t,T)dt \tag{4.8}$$

with initial value $M(0,T)$, which is assumed to be determined by the market data.[1] The Markovian HJM models, see Chiarella and Kwon [2001][2], are used to analyse the volatility contribution part $M(t,T)$. The difference is that in Chiarella

[1]This is similar to the initial value of the instantaneous forward rate, $f(0,\tau)$, in HJM framework. However, for the HJM case, $f(0,\tau)$ is assumed to be given, whereas $M(0,T)$ is hidden and needs to be modeled in this thesis.

[2]The method of Chiarella and Kwon [2001] is used in my model, however, Björk and Svensson [2001] provides a general treatment of Markovian realization.

and Kwon [2001], the Markovian realization is applied to the entire forward rate structure. However, only the volatility component $M(t,T)$ is decomposed by applying the Markovian realization in this study.[1] This makes my model a hybrid of the affine short rate models of Duffie and Kan [1996], Dai and Singleton [2000] and the Markovian HJM framework of Chiarella and Kwon [2001].

4.2.3 Markovian factor loading structures

Proposition 4.2.3. *The volatility contribution part $M(t,T)$ can be written in a Markovian factor loading form,*

$$M(t,T) = \sum_{i=1}^{8} E^i(t,T)\phi_t^i + M(0,T)$$

where $E^i(t,T)$, $i \in \{1,\ldots,8\}$ and $\boldsymbol{\phi}_t = (\phi_t^1,\ldots,\phi_t^8)^\top$ are given by,

$$\begin{aligned}
E^1(t,T) &= B_{11}^3 & d\phi_t^1 &= \phi_t^8 dt \\
E^2(t,T) &= B_{21}^3 + B_{13}^3 + \lambda B_3^4 & d\phi_t^2 &= (\phi_t^4 - \lambda\phi_t^2)dt \\
E^3(t,T) &= B_{31}^3 & d\phi_t^3 &= (2\lambda\phi_t^2 - \lambda\phi_t^3)dt \\
E^4(t,T) &= B_{12}^3 + B_2^4 & d\phi_t^4 &= (\nu_t - \lambda\phi_t^4)dt \\
E^5(t,T) &= B_{22}^3 & d\phi_t^5 &= (\nu_t - 2\lambda\phi_t^5)dt \\
E^6(t,T) &= B_{32}^3 + \frac{B_{23}^3}{\lambda} & d\phi_t^6 &= (\lambda\phi_t^5 - 2\lambda\phi_t^6)dt \\
E^7(t,T) &= B_{33}^3 & d\phi_t^7 &= (2\phi_t^6 - 2\lambda\phi_t^7)dt \\
E^8(t,T) &= B_1^4 & d\phi_t^8 &= \nu_t dt
\end{aligned}$$

[1]Only $M(0,T)$ requires the Markovian treatment because, being a Gaussian affine specification, the other components are already Markovian

with initial values for all ϕ_t^i equal to 0.

In Proposition 4.2.3, $M(0,T)$ is the initial value of the volatility contribution, which generates various shapes to fit the initial forward rate curve. This should hence be determined from the market data, similar to HJM models.

The proof of Proposition 4.2.3 is non trivial and so it is divided into several lemmata and definitions.

Lemma 4.2.4. *For $0 < s \le t \le T$, the Nelson-Siegel factor loading structure $\boldsymbol{\alpha}(s,T)$ is equal to $\mathbf{B}^2(t,T)\boldsymbol{\alpha}(s,t)$, where*

$$\mathbf{B}^2(t,T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha_2(t,T) & 0 \\ 0 & \alpha_3(t,T) & \alpha_2(t,T) \end{bmatrix}$$

.

Proof. Given equation (4.2),

$$\begin{aligned}
\mathbf{B}^2(t,T)\boldsymbol{\alpha}(s,t) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha_2(t,T) & 0 \\ 0 & \alpha_3(t,T) & \alpha_2(t,T) \end{bmatrix} \begin{bmatrix} 1 \\ e^{-\lambda(T-s)} \\ \lambda(T-s)e^{-\lambda(T-s)} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda(T-t)} & 0 \\ 0 & \lambda(T-t)e^{-\lambda(T-t)} & e^{-\lambda(T-t)} \end{bmatrix} \begin{bmatrix} 1 \\ e^{-\lambda(t-s)} \\ \lambda(t-s)e^{-\lambda(t-s)} \end{bmatrix} \\
&= \begin{bmatrix} 1 \\ e^{-\lambda(T-s)} \\ \lambda(T-s)e^{-\lambda(T-s)} \end{bmatrix}
\end{aligned}$$

$$= \boldsymbol{\alpha}(s,T)$$

□

Lemma 4.2.5. *For $0 < s \le t \le T$, the factor loading of zero rate $\bar{\boldsymbol{\alpha}}(s,T)$ where*

$$\bar{\boldsymbol{\alpha}}(s,T) = \int_s^T \boldsymbol{\alpha}(s,u)du$$

is equal to $\mathbf{B}^0(t,T)+\mathbf{B}^1(t,T)\widetilde{\boldsymbol{\alpha}}(s,t)$, where $\mathbf{B}^0(t,T)$, $\widetilde{\boldsymbol{\alpha}}(s,t)$ and $\mathbf{B}^1(t,T)$ are give as follows

$$\mathbf{B}^0(t,T) = \begin{bmatrix} T-t \\ \frac{1}{\lambda} \\ \frac{1}{\lambda} \end{bmatrix}, \qquad \widetilde{\boldsymbol{\alpha}}(s,t) = \begin{bmatrix} t-s \\ e^{-\lambda(t-s)} \\ (t-s)e^{-\lambda(t-s)} \end{bmatrix}$$

$$\mathbf{B}^1(t,T) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{\lambda}e^{-\lambda(T-t)} & 0 \\ 0 & \left(-\frac{1}{\lambda} - (T-t)\right)e^{-\lambda(T-t)} & -e^{-\lambda(T-t)} \end{bmatrix}$$

Proof. Recall Lemma 4.2.1,

$$\begin{aligned} &\mathbf{B}^0(t,T) + \mathbf{B}^1(t,T)\widetilde{\boldsymbol{\alpha}}(s,t) \\ =& \begin{bmatrix} T-t \\ \frac{1}{\lambda} \\ \frac{1}{\lambda} \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{\lambda}e^{-\lambda(T-t)} & 0 \\ 0 & \left(-\frac{1}{\lambda} - (T-t)\right)e^{-\lambda(T-t)} & -e^{-\lambda(T-t)} \end{bmatrix} \begin{bmatrix} t-s \\ e^{-\lambda(t-s)} \\ (t-s)e^{-\lambda(t-s)} \end{bmatrix} \end{aligned}$$

$$= \begin{bmatrix} T-s \\ \frac{1-e^{-\lambda(T-s)}}{\lambda} \\ \frac{1-e^{-\lambda(T-s)}-\lambda(T-s)e^{-\lambda(T-s)}}{\lambda} \end{bmatrix} = \bar{\boldsymbol{\alpha}}(s,T)$$

□

Definition 4.2.6. $\mathbf{B}^3(t,T)$ *is defined as the following* 3×3 *matrix*

$$\mathbf{B}^3(t,T) = \mathbf{B}^2(t,T)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \mathbf{B}^1(t,T) = \begin{bmatrix} B^3_{11} & B^3_{12} & B^3_{13} \\ B^3_{21} & B^3_{22} & B^3_{23} \\ B^3_{31} & B^3_{32} & B^3_{33} \end{bmatrix} .$$

Definition 4.2.7. $\mathbf{B}^4(t,T)$ *is defined as the following* 3×1 *vector*

$$\mathbf{B}^4(t,T) = \mathbf{B}^2(t,T)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \mathbf{B}^0(t,T) = \begin{bmatrix} B^4_1 \\ B^4_2 \\ B^4_3 \end{bmatrix} .$$

Using Lemma 4.2.4 and 4.2.5, Definition 4.2.6 and 4.2.7, and equation (4.8), proposition 4.2.3 may be demonstrated as follows,

Proof. $\mathbf{M}(t,T)$ can be given by,

$$\mathbf{M}(t,T) = \mathbf{M}(0,T) + \int_0^t \boldsymbol{\alpha}(s,t)^\top v_s ds \mathbf{B}^2(t,T)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \mathbf{B}^0(t,T) \tag{4.9}$$
$$+ \int_0^t \boldsymbol{\alpha}(s,t)^\top \mathbf{B}^3(T-t)\tilde{\boldsymbol{\alpha}}(s,t) v_s ds \ .$$

The second term in equation (4.9) can be written as,

$$\begin{aligned}&\int_0^t \boldsymbol{\alpha}(s,t)^\top v_s ds \mathbf{B}^2(t,T)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \mathbf{B}^0(t,T)\\&= \int_0^t \alpha_1 v_s ds B_1^4 + \int_0^t \alpha_2 v_s ds B_2^4 + \int_0^t \alpha_3 v_s ds B_3^4 \quad .\end{aligned}$$

The third term in equation (4.9) can thus be obtained as follows,

$$\begin{aligned}&\int_0^t \boldsymbol{\alpha}(s,t)^\top \mathbf{B}^3(t,T)\widetilde{\boldsymbol{\alpha}}(s,t) v_s ds\\&= \int_0^t \alpha_1\widetilde{\alpha_1} v_s ds B_{11}^3 + \int_0^t \alpha_2\widetilde{\alpha_1} v_s ds B_{21}^3 + \int_0^t \alpha_3\widetilde{\alpha_1} v_s ds B_{31}^3\\&+ \int_0^t \alpha_1\widetilde{\alpha_2} v_s ds B_{12}^3 + \int_0^t \alpha_2\widetilde{\alpha_2} v_s ds B_{22}^3 + \int_0^t \alpha_3\widetilde{\alpha_2} v_s ds B_{32}^3\\&+ \int_0^t \alpha_1\widetilde{\alpha_3} v_s ds B_{13}^3 + \int_0^t \alpha_2\widetilde{\alpha_3} v_s ds B_{23}^3 + \int_0^t \alpha_3\widetilde{\alpha_3} v_s ds B_{33}^3 \quad .\end{aligned}$$

Therefore, the volatility contribution component $\mathbf{M}(t,T)$ can be separated into the following twelve terms.

$$\begin{aligned}\int_0^t \alpha_1\widetilde{\alpha_1} v_s ds B_{11}^3 &= B_{11}^3 \int_0^t (t-s) v_s ds\\\int_0^t \alpha_2\widetilde{\alpha_1} v_s ds B_{21}^3 &= B_{21}^3 \int_0^t e^{-\lambda(t-s)} v_s ds\\\int_0^t \alpha_3\widetilde{\alpha_1} v_s ds B_{31}^3 &= B_{31}^3 \int_0^t \lambda(t-s)^2 e^{-\lambda(t-s)} v_s ds\\\int_0^t \alpha_1\widetilde{\alpha_2} v_s ds B_{12}^3 &= B_{12}^3 \int_0^t e^{-\lambda(t-s)} v_s ds\\\int_0^t \alpha_2\widetilde{\alpha_2} v_s ds B_{22}^3 &= B_{22}^3 \int_0^t e^{-2\lambda(t-s)} v_s ds\\\int_0^t \alpha_3\widetilde{\alpha_2} v_s ds B_{32}^3 &= B_{32}^3 \int_0^t \lambda(t-s) e^{-2\lambda(t-s)} v_s ds\\\int_0^t \alpha_1\widetilde{\alpha_3} v_s ds B_{13}^3 &= B_{13}^3 \int_0^t (t-s) e^{-\lambda(t-s)} v_s ds\end{aligned}$$

$$
\begin{aligned}
\int_0^t \alpha_2\widetilde{\alpha_3}v_s ds B_{23}^3 &= \frac{B_{23}^3}{\lambda}\int_0^t \lambda(t-s)e^{-2\lambda(t-s)}v_s ds \\
\int_0^t \alpha_3\widetilde{\alpha_3}v_s ds B_{33}^3 &= B_{33}^3\int_0^t \lambda(t-s)^2 e^{-2\lambda(t-s)}v_s ds \\
\int_0^t \alpha_1 v_s ds B_1^4 &= B_1^4\int_0^t v_s ds \\
\int_0^t \alpha_2 v_s ds B_2^4 &= B_2^4\int_0^t e^{-\lambda(t-s)}v_s ds \\
\int_0^t \alpha_3 v_s ds B_3^4 &= \lambda B_3^4\int_0^t (t-s)e^{-\lambda(t-s)}v_s ds
\end{aligned}
$$

It is clear that B_{21}^3, B_{13}^3 and λB_3^4, B_{12}^3 and B_2^4, B_{32}^3 and $\frac{B_{23}^3}{\lambda}$ have identical Markovian loadings. Therefore the volatility contribution component therefore is further derived in the following nine terms,

$$
M(t,T) = \sum_{i=1}^{8} E^i(t,T)\phi_t^i + M(0,T)
$$

where $E^i(t,T)$, $i \in \{1,\ldots,8\}$ and the state variables $\boldsymbol{\phi}_t = (\phi_t^1,\ldots,\phi_t^8)^\top$ are given by,

$$
\begin{aligned}
E^1(t,T) &= B_{11}^3 & d\phi_t^1 &= \phi_t^8 dt \\
E^2(t,T) &= B_{21}^3 + B_{13}^3 + \lambda B_3^4 & d\phi_t^2 &= (\phi_t^4 - \lambda\phi_t^2)dt \\
E^3(t,T) &= B_{31}^3 & d\phi_t^3 &= (2\lambda\phi_t^2 - \lambda\phi_t^3)dt \\
E^4(t,T) &= B_{12}^3 + B_2^4 & d\phi_t^4 &= (\nu_t - \lambda\phi_t^4)dt \\
E^5(t,T) &= B_{22}^3 & d\phi_t^5 &= (\nu_t - 2\lambda\phi_t^5)dt \\
E^6(t,T) &= B_{32}^3 + \frac{B_{23}^3}{\lambda} & d\phi_t^6 &= (\lambda\phi_t^5 - 2\lambda\phi_t^6)dt \\
E^7(t,T) &= B_{33}^3 & d\phi_t^7 &= (2\phi_t^6 - 2\lambda\phi_t^7)dt \\
E^8(t,T) &= B_1^4 & d\phi_t^8 &= \nu_t dt
\end{aligned}
$$

with initial values for all ϕ_t^i equal to 0. □

4.2.4 Zero-coupon bond price

The derivation of the unspanned stochastic volatility arbitrage-free Nelson-Siegel model leads to the following proposition:

Proposition 4.2.8. *Assume the instantaneous forward rate is given by*

$$f(t,T) = \begin{bmatrix} \boldsymbol{\alpha}(t,T) \\ \mathbf{E}(t,T) \end{bmatrix}^{\top} \begin{bmatrix} \mathbf{X}_t \\ \boldsymbol{\phi}_t \end{bmatrix} + D(t,T) + m$$

where the factors $\mathbf{X}_t$ and $\boldsymbol{\phi}_t$ are defined in equation (4.3) and Proposition 4.2.3 under the risk-neutral measure and the loadings $\boldsymbol{\alpha}(t,T)$ and $\mathbf{E}(t,T)$ are defined in equation (4.2) and Proposition 4.2.3. The model then becomes arbitrage-free and the zero coupon bond price is given by

$$P(t,T) = \exp\left\{-\int_t^T f(t,u)du\right\}$$

and the zero coupon bond yields are given by,

$$\begin{aligned} y(t,T) &= X_t^1 + X_t^2\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau}\right) + X_t^3\left(\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \\ &\quad + \bar{\mathbf{E}}(t,T)^{\top}\boldsymbol{\phi}_t + \bar{D}(t,T) \end{aligned} \tag{4.10}$$

where τ is the time to maturity and $\bar{\mathbf{E}}(t,T)$ represents the factor loadings of the locally riskless part $M(t,T)$, such that,

$$\bar{\mathbf{E}}(t,T) = \frac{\int_t^T \mathbf{E}(t,u)du}{T-t}$$

while $\bar{D}(t,T)$ is a deterministic drift term of the form,

$$\bar{D}(t,T) = \frac{\int_t^T D(t,u)du}{T-t} + M(0,T)$$

Note that the first three factor loadings for $y(t,T)$ in equation (4.10) exactly match the dynamic Nelson-Siegel model, but there are the additional terms $\bar{\mathbf{E}}(t,T)^\top \boldsymbol{\phi}_t$ and $\bar{D}(t,T)$. Assuming a constant volatility, the model reduces exactly to that of Christensen, Diebold, and Rudebusch [2011]. In this case, since the volatility term v_t is allowed to fluctuate, $\bar{\mathbf{E}}(t,T)^\top \boldsymbol{\phi}_t$ is now a locally riskless random variable, the corresponding fluctuation of which is negligible for order dt.

4.2.5 Market price of risk specifications

The Wiener process in the dynamics of the factors $\widetilde{\mathbf{X}}_t = (X_t^1, X_t^2, X_t^3, v_t)^\top$ in the SDE given by (4.3) is under the risk neutral measure $\mathbb{P}^*$. The dynamics of $\widetilde{\mathbf{X}}_t$ under the real world measure $\mathbb{P}$ are required for parameter estimation. As a result, I need to specify the market price of risk $\boldsymbol{\Lambda}_t$ to link $\mathbb{P}$ and $\mathbb{P}^*$ by applying Girsanov's theorem,

$$d\mathbf{W}_t^* = d\mathbf{W}_t + \boldsymbol{\Lambda}_t dt \ .$$

Furthermore, for computational simplicity, a specific market price of risk process $\boldsymbol{\Lambda}_t$ is required to be specified such that (4.3), when converted to the real world measure $\mathbb{P}$, is still affine in the sense of Duffie, Pan, and Singleton [2000]. The market price of risk specification is derived using the setting of Trolle and Schwartz [2009a] and Christensen, Diebold, and Rudebusch [2011], $\boldsymbol{\Lambda}_t = (\lambda_t^1, \lambda_t^2, \lambda_t^3, \lambda_t^4)^\top$

where

$$\begin{bmatrix} \lambda_t^1 \\ \lambda_t^2 \\ \lambda_t^3 \\ \lambda_t^4 \end{bmatrix} = \frac{1}{\sqrt{v_t}} \left(\begin{bmatrix} \lambda_1^0 \\ \lambda_2^0 \\ \lambda_3^0 \\ \lambda_4^0 \end{bmatrix} + \begin{bmatrix} \lambda_{11}^1 & \lambda_{12}^1 & \lambda_{13}^1 & \lambda_{14}^1 \\ \lambda_{21}^1 & \lambda_{22}^1 & \lambda_{23}^1 & \lambda_{24}^1 \\ \lambda_{31}^1 & \lambda_{32}^1 & \lambda_{33}^1 & \lambda_{34}^1 \\ \lambda_{41}^1 & \lambda_{42}^1 & \lambda_{43}^1 & \lambda_{44}^1 \end{bmatrix} \begin{bmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \\ v_t \end{bmatrix} \right) , \tag{4.11}$$

and for $i \in \{1, 2, 3\}$,

$$\lambda_{4i}^1 = -\frac{1}{\sigma_v^4} \sum_{j=1}^{3} \sigma_v^1 \lambda_{ji}^1 \ . \tag{4.12}$$

Here $\mathbf{\Lambda^0} = (\lambda_i^0)_{i=1}^4$ is a 4×1 vector of constants and $\mathbf{\Lambda^1} = (\lambda_{ij}^0)_{i,j=1}^{4,4}$ is a 4×4 matrix of constants. The reason for the asymmetry for λ_{41}^1, λ_{42}^1 and λ_{43}^1 is that the latent factors X_t^1, X_t^2 and X_t^3 can, at times, be very negative. To guarantee the strict positivity of the volatility process, X_t^1, X_t^2 and X_t^3 should not be incorporated into the drift of v_t. Furthermore, we would like to guarantee that the volatility process v_t also forms a square-root process under the measure $\mathbb{P}$, which serves to simplify the calibration.

Note that in equation (4.11), there exists the inverse of the square-root of v_t. As a result, according to Trolle and Schwartz [2009a], to prevent $\mathbf{\Lambda}_t$ from exploding, as well as to guarantee that v_t does not exit its state space under both $\mathbb{P}$ and $\mathbb{P}^*$, we need to prevent v_t from reaching zero under both $\mathbb{P}$ and $\mathbb{P}^*$ by applying the following Feller condition,

$$\frac{2k_v\theta_v}{\sum_{i=1}^4 (\sigma_v^i)^2} \geq 1$$
$$\frac{2k_v^{\mathbb{P}}\theta_v^{\mathbb{P}}}{\sum_{i=1}^4 (\sigma_v^i)^2} \geq 1$$

where $k_v^{\mathbb{P}}\theta_v^{\mathbb{P}} = \widetilde{\mathbf{\Psi}}_{41}^{\mathbb{P}} = k_v\theta_v + \mathbf{\Sigma}_v^\top \mathbf{\Lambda}^0$. The dynamics of $\widetilde{\mathbf{X}}_t = (X_t^1, X_t^2, X_t^3, v_t)^\top$

under the real world measure $\mathbb{P}$ can then be expressed as follows,

$$d\widetilde{\mathbf{X}}_t = (\widetilde{\mathbf{\Psi}}^{\mathbb{P}} - \mathbf{K}^{\mathbb{P}}\widetilde{\mathbf{X}}_t)dt + \sqrt{v_t}\widetilde{\mathbf{\Sigma}}^{\mathbb{P}} d\mathbf{W}_t^{\mathbb{P}} \tag{4.13}$$

where $\mathbf{K}^{\mathbb{P}} = \mathbf{K}^{\mathbb{P}^*} - \mathbf{\Sigma}^{\mathbb{P}^*}\mathbf{\Lambda}^1$. The SDE (4.13) under the real world measure $\mathbb{P}$ is still affine, see Christensen, Diebold, and Rudebusch [2011]. The dynamics of the associated factors $\boldsymbol{\phi}_t = (\phi_t^1, \ldots, \phi_t^8)$ are identical under both the risk-neutral measure and the real world measure, since they are assumed to be locally riskless.

4.3 Pricing of interest rate derivatives

Interest rate derivatives, such as caps and swaptions, are among the most liquid derivatives traded in financial markets. It is also widely believed that the price of interest rate derivatives contains forward looking information about the risk-adjusted averages of expected rates, see Almeida, Graveline, and Joslin [2011]. In this section, I derive the price of interest rate derivatives using the framework of Carr and Madan [1999], Duffie, Pan, and Singleton [2000] and Schrager and Pelser [2006].

4.3.1 Zero-coupon bond options

Call and put options written on zero-coupon bonds are fundamental elements in pricing interest rate derivatives. More complicated interest rate contingent claims may be calculated based on the prices of these simple bond options.

Denote $\mathbb{P}^*$ as the risk-neutral measure, $C_t^{T_0,T_1}(K)$ as a call option written on a zero-coupon bond $P(T_0, T_1)$ with strike price K, maturity at T_0 is defined by,

$$C_t^{T_0,T_1}(K) = E_t^* \left(e^{-\int_t^{T_0} r_s ds} \left[P(T_0, T_1) - K \right]^+ \right) \quad . \tag{4.14}$$

Denote $B(t)$ as the bank saving account,

$$B(t) = \int_0^t r_s ds \quad , \tag{4.15}$$

where r_t is the short rate at time t. Substitute equation (4.15) into equation (4.14), this expression may be rewritten as,

$$C_t^{T_0,T_1}(K) = E_t^* \left(\frac{B(t)}{B(T_0)} \left[P(T_0, T_1) - K \right]^+ \right) \quad .$$

The Radon-Nikodym derivative is defined as,

$$\frac{d\mathbb{P}^{T_0}}{d\mathbb{P}} = \frac{B(s)}{P(s, T_0)}$$

which changes equation 4.14 to the forward measure $\mathbb{P}^{T_0}$ to yield,

$$\begin{aligned} C_t^{T_0,T_1}(K) &= E_t^{T_0} \left(\frac{\frac{B(T_0)}{P(T_0,T_0)}}{\frac{B(t)}{P(t,T_0)}} \left[P(T_0, T_1) - K \right]^+ \right) \\ &= P(t, T_0) E_t^{T_0} \left(\left[P(T_0, T_1) - K \right]^+ \right) \end{aligned} \tag{4.16}$$

where the conditional expectation $E_t^{T_0}$ is taken under the T_0th forward measure. The zero-coupon bond price $P(t, T_0)$ is derived from equation (4.10). The problem thus reduces to solving the conditional expectation term under the T_0th forward measure. To this end, I apply the transform inversion technique, proposed by

Duffie, Pan, and Singleton [2000] to solve this conditional expectation term.

Define $\phi_{T_0}(u)$[1] to be the conditional characteristic function of the $\ln P(T_0, T_1)$ under the T_0th forward measure,

$$\begin{aligned}\phi_{T_0}(u) &= E_t^{T_0}\left(e^{iu \ln P(T_0,T_1)}\right) \\ &= \int_{-\infty}^{+\infty} e^{ius} q^{T_0}(s) ds\end{aligned}$$

where $q^{T_0}(s)$ is the density of $\ln P(T_0, T_1)$ under the T_0th forward measure. Further define k to be the natural logarithm of the strike price K,

$$k = \ln K \quad .$$

The conditional expectation term in equation (4.16) can hence be further simplified to,

$$\begin{aligned}\Psi(k) &= E_t^{T_0}\left([P(T_0, T_1) - K]^+\right) \\ &= E_t^{T_0}\left(\left[e^{\ln P(T_0,T_1)} - e^k\right]^+\right) \\ &= \int_k^{+\infty} (e^s - e^k) q^{T_0}(s) ds \quad ,\end{aligned}$$

where the limits of integration are reduced to eliminate the plus function $[\quad]^+$. The next step is to apply the Fourier transform as in Carr and Madan [1999]. This requires square integrability and it is straightforward to show the functions used satisfy this property. Define the modified conditional expectation term $\overline{\Psi}(k)$

[1]To be consistent with the notation in Carr and Madan [1999], I use ϕ to represent the conditional characteristic function. This should not be confused with the factors in proposition 4.2.3

by,

$$\overline{\Psi}(k) := e^{\alpha k}\Psi(k) \quad ,$$

where $\alpha > 1$ is a constant. Now denote $\psi_{T_0}(v)$ as the Fourier transform of $\overline{\Psi}(k)$, which gives,

$$\psi_{T_0}(v) = \int_{-\infty}^{+\infty} e^{ivk}\overline{\Psi}(k)dk \quad .$$

The conditional expectation $\Psi(k)$ may then be obtained through the Fourier inversion,

$$\Psi(k) = \frac{e^{-\alpha k}}{2\pi}\int_{-\infty}^{+\infty} e^{-ivk}\psi_{T_0}(v)dv \quad .$$

Similar to Carr and Madan [1999], the function $\psi_{T_0}(v)$ is odd in its imaginary part and even in its real part, which gives,

$$\Psi(k) = \frac{e^{-\alpha k}}{\pi}\int_{0}^{+\infty} \texttt{Real}\{e^{-ivk}\psi_{T_0}(v)\}dv$$

Lemma 4.3.1. *The expression, $\psi_{T_0}(v)$, equals $\frac{\phi_{T_0}(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v}$ where ϕ_{T_0} is the conditional characteristic function of* $\ln P(T_0, T_1)$

Proof. The proof is identical to Carr and Madan [1999]. The expression for $\psi_{T_0}(v)$ can be derived as follows,

$$\psi_{T_0}(v) = \int_{-\infty}^{+\infty} e^{ivk}\overline{\Psi}(k)dk$$

$$= \int_{-\infty}^{+\infty} e^{ivk} \int_{k}^{+\infty} e^{\alpha k}(e^s - e^k) q^{T_0}(s) ds dk$$

using Fubini's theorem

$$\begin{aligned}
\psi_{T_0}(v) &= \int_{-\infty}^{+\infty} q^{T_0}(s) \int_{-\infty}^{s} e^{ivk} \left(e^{s+\alpha k} - e^{(1+\alpha)k}\right) dk ds \\
&= \int_{-\infty}^{+\infty} q^{T_0}(s) \left[\frac{e^{(\alpha+1+iv)s}}{\alpha + iv} - \frac{e^{(\alpha+1+iv)s}}{\alpha + 1 + iv}\right] ds \\
&= \int_{-\infty}^{+\infty} q^{T_0}(s) e^{(\alpha+1+iv)s} \frac{1}{(\alpha + iv)(\alpha + 1 + iv)} ds \\
&= \frac{\phi_{T_0}(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v} \quad .
\end{aligned}$$

□

Finally, the conditional characteristic function is given by,

$$\begin{aligned}
\phi_{T_0}(u) &= E_t^{T_0}\left(e^{iu \ln P(T_0,T_1)}\right) \\
&= E_t^{*}\left(\frac{B(t)}{B(T_0)} \frac{1}{P(t,T_0)} e^{iu \ln P(T_0,T_1)}\right) \quad .
\end{aligned} \tag{4.17}$$

The probability measure has been changed back from T_0th forward to a risk-neutral measure $\mathbb{P}^*$ to allow the use of the results from Duffie, Pan, and Singleton [2000]. An exponential affine solution for the corresponding characteristic function can therefore be obtained under the risk-neutral measure.

Proposition 4.3.2. *The conditional characteristic function introduced in* (4.17) *is given by,*[1]

$$E_t^*\left(\frac{B(t)}{B(T_0)}\frac{1}{P(t,T_0)}e^{iu\ln P(T_0,T_1)}\right)$$
$$= e^{M(\tau)+N(\tau)v_t+iu\ln\frac{P(t,T_1)}{P(t,T_0)}}$$

where $M(\tau)$ and $N(\tau)$ satisfy the following system of Riccati equations

$$\frac{dM(\tau)}{d\tau} = N(\tau)k_v\theta_v$$
$$M(0) = 0$$
$$\frac{dN(\tau)}{d\tau} = -\frac{1}{2}(1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\boldsymbol{\alpha}(t,T_0) - \frac{1}{2}iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\boldsymbol{\alpha}(t,T_1)$$
$$+\left(N(\tau)\boldsymbol{\Sigma}_v^\top - (1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma} - iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\right)$$
$$\left(N(\tau)\boldsymbol{\Sigma}_v^\top - (1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma} - iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\right)^\top$$
$$- N(\tau)k_v + \frac{1}{2}N(\tau)^2(\sigma_v^4)^2$$
$$N(0) = 0$$

where $\boldsymbol{\Sigma}_v = (\sigma_v^1, \sigma_v^2, \sigma_v^3)^\top$.

The proof of this statement is non trivial and requires a prior result presented in lemma 4.3.3.

Lemma 4.3.3. *if $x_t = e^{-\int_0^t r_s ds}e^{M(\tau)+N(\tau)v_t+\beta\ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$ is a martingale, proposition 4.3.2 holds true.*

Proof. It is easy to show that $M(0) = 0$ and $N(0) = 0$ at maturity time $T_0 (\tau = 0)$.

[1] I use $M(.)$ to represent the Riccati equation. This should not be confused with the volatility contribution component $M(t,T)$ in section 4.2.3

if the process x_t

$$x_t = e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$$

is a martingale, therefore the equation $E(X_T) = x_t$ must hold which gives,

$$E_t^* \left(e^{-\int_0^{T_0} r_s ds} e^{M(0)+N(0)v_{T_0}+\beta \ln P(T_0,T_1)+(1-\beta)\ln P(T_0,T_0)} \right)$$
$$= e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$$

Recalling $M(0) = 0$, $N(0) = 0$ and $P(T_0, T_0) = 1$, which leads to,

$$E_t^* \left(e^{-\int_0^{T_0} r_s ds} e^{M(0)+N(0)v_{T_0}+\beta \ln P(T_0,T_1)+(1-\beta)\ln P(T_0,T_0)} \right)$$
$$= E_t^* \left(e^{-\int_0^{T_0} r_s ds+\beta \ln P(T_0,T_1)} \right)$$
$$= e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$$

and given $e^{\ln P(t,T_0)} = P(t, T_0)$,

$$E_t^* \left(e^{-\int_0^{T_0} r_s ds} e^{M(0)+N(0)v_{T_0}+\beta \ln P(T_0,T_1)+(1-\beta)\ln P(T_0,T_0)} \right)$$
$$= P(t,T_0) e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{P(t,T_1)}{P(t,T_0)}}$$

rearranging the equation by moving the known price at time t, $P(t, T_0)$ and $e^{-\int_0^t r_s ds}$ to the lefthand side gives,

$$E_t^* \left(\frac{1}{P(t,T_0)} e^{-\int_0^{T_0} r_s ds+\beta \ln P(T_0,T_1)+\int_0^t r_s ds} \right)$$
$$= e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{P(t,T_1)}{P(t,T_0)}}$$

and as the bank savings account $B(t) = \int_0^t r_s ds$, I finally arrive at,

$$E_t^* \left(\frac{B(t)}{B(T_0)} \frac{1}{P(t,T_0)} e^{iu \ln P(T_0,T_1)} \right)$$
$$= e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{P(t,T_1)}{P(t,T_0)}} \quad .$$

By simply specifying $\beta = iu$, the proposition can thus be proved. □

I start to prove the proposition 4.3.2 here,

Proof. if I can prove x_t is a martingale, the proposition can be proved according to lemma 4.3.3. Recalling the definition of a martingale is that the drift term must be equal to zero. I therefore need to specify the ODEs for $M(\tau)$ and $N(\tau)$ to guarantee that the drift of,

$$de^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$$

is zero. The SDE of the bond price $P(t,T_i)$ in the generalised arbitrage-free Nelson-Siegel model is,

$$dP(t,T_i) = P(t,T_i) \left(r_t - \sqrt{v_t} \boldsymbol{\alpha}(t,T_i)^\top \boldsymbol{\Sigma} d\mathbf{W}_t^* \right)$$

Using Itô's formula, the $d \ln P(t,T_i)$ can be derived as,

$$\begin{aligned} d\ln P(t,T_i) &= \left(r_t - \frac{1}{2} v_t \boldsymbol{\alpha}(t,T_i)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \boldsymbol{\alpha}(t,T_i) \right) dt \\ &\quad - \sqrt{v_t} \boldsymbol{\alpha}(t,T_i)^\top \boldsymbol{\Sigma} d\mathbf{W}_t^* \quad . \end{aligned} \tag{4.18}$$

Further, define

$$z = -\int_0^t r_s ds + M(\tau) + N(\tau)v_t + \beta \ln P(t, T_1) + (1-\beta)\ln P(t, T_0) \quad . \quad (4.19)$$

Again, using the Itô formula, given $Y = e^z$, I have,

$$de^z = e^z dz + \frac{1}{2}e^z (dz)^2 \quad . \quad (4.20)$$

By substituting equations (4.18) and (4.19) into equation (4.20) and considering only the drift term,

$$\begin{aligned} dz = &-r_t dt + dM(\tau) + v_t dN(\tau) + N(\tau)dv_t \\ &+ \beta\left(r_t - \frac{1}{2}v_t \boldsymbol{\alpha}(t, T_1)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \boldsymbol{\alpha}(t, T_1)\right) dt \\ &+ (1-\beta)\left(r_t - \frac{1}{2}v_t \boldsymbol{\alpha}(t, T_0)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \boldsymbol{\alpha}(t, T_0)\right) dt \quad . \end{aligned}$$

Note that all terms containing r_t may be canceled, which leads to,

$$\begin{aligned} dz = &\, dM(\tau) + v_t dN(\tau) + N(\tau)dv_t + \beta\left(-\frac{1}{2}v_t \boldsymbol{\alpha}(t, T_1)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \boldsymbol{\alpha}(t, T_1)\right) dt \\ &+ (1-\beta)\left(-\frac{1}{2}v_t \boldsymbol{\alpha}(t, T_0)^\top \boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top \boldsymbol{\alpha}(t, T_0)\right) dt \end{aligned}$$

where as stated in equation (4.3), the dynamics of v_t are given by,

$$dv_t = k_v(\theta_v - v_t)dt + \sqrt{v_t}\boldsymbol{\sigma}_v dW_t^* \quad .$$

The drift of $(dz)^2$ is given by,

$$(dz)^2 = \frac{1}{2}N(\tau)^2 v_t \sigma_{v_4}^2 dt + \frac{1}{2}v_t \left(N(\tau)\boldsymbol{\sigma}_{v_{1,2,3}}^\top - \beta\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma} - (1-\beta)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma}\right)$$
$$\left(N(\tau)\boldsymbol{\sigma}_{v_{1,2,3}}^\top - \beta\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma} - (1-\beta)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma}\right)^\top dt$$

By specifying $\beta = iu$ and rearranging the terms in v_t which are all equal to zero by the properties of martingales, the proposition has been proved. □

The semi closed-form solution of zero-bond calls has thus been derived. The corresponding put option prices may be obtained via put-call parity.

4.3.2 Caps and floors

An interest rate cap is a contract that can be viewed as a payer interest-rate swap where each exchange payment is executed only if it has a positive value. The discounted payoff for a cap $\bar{H}_T^{\text{cap}}$ is given by

$$\bar{H}_T^{\text{cap}} = \sum_{i=2}^{n} e^{-\int_t^{T_i} r_s ds} N\tau_i (L(T_{i-1}, T_i) - K)^+$$

where N is the face value, usually assumed to be 1. Define the variable τ_i as the scaled time interval such that $\tau_i = T_i - T_{i-1}$ and $L(T_{i-1}, T_i)$ is the LIBOR rate, given by

$$L(T_{i-1}, T_i) = F(T_{i-1}, T_{i-1}, T_i) \quad ,$$

where $F(t,S,T)$ is the forward rate defined as

$$F(t,S,T) = \frac{1}{T-S}\left(\frac{P(t,S)}{P(t,T)} - 1\right) .$$

The discounted payoff and thus the price of a cap can be represented by the sum of a sequence of caplets, where each caplet has its own discounted payoff $\bar{H}_t^{\text{Caplet},i}$ where

$$\bar{H}_t^{\text{Caplet},i} = e^{-\int_t^{T_i} r_s ds} N\tau_i(L(T_{i-1},T_i) - K)^+ \tag{4.21}$$

Lemma 4.3.4. *The price of a caplet with discounted payoff* $\bar{H}_t^{Caplet,i}$ *given in* (4.21) *is equivalent to* $(1+\tau_i K)$ *European put options of zero coupon bond* $P(T_{i-1},T_i)$ *with strike* $\frac{1}{1+\tau_i K}$ *at maturity time* T_{i-1}, *such that*

$$V^{Caplet}(t,T_{i-1},T_i,K) = (1+\tau_i K)P_t^{T_{i-1},T_i}\left(\frac{1}{1+\tau_i K}\right)$$

Proof. With the payoff $\bar{H}_t^{\text{Caplet},i}$ given in (4.21), the price of the caplet equals

$$\begin{aligned}
V^{\text{Caplet}}(t,T_{i-1},T_i,K) &= E_t^*\left(\bar{H}_t^{\text{Caplet},i}\right)\\
&= E_t^*\left(e^{-\int_t^{T_i} r_s ds} N\tau_i(L(T_{i-1},T_i) - K)^+\right)\\
&= N\tau_i E_t^*\left(e^{-\int_t^{T_i} r_s ds}\left(\frac{1}{\tau_i P(T_{i-1},T_i)} - \left(\frac{1}{\tau_i} + K\right)\right)^+\right)\\
&= N(1+K\tau_i)E_t^*\left(\frac{e^{-\int_t^{T_i} r_s ds}}{P(T_{i-1},T_i)}\left(\frac{1}{1+K\tau_i} - P(T_{i-1},T_i)\right)^+\right)
\end{aligned}$$

by applying iterated expectations

$$
\begin{aligned}
&= N(1+K\tau_i)E_t^*\left(E_{T_{i-1}}^*\left(\frac{e^{-\int_t^{T_i} r_s ds}\left(\frac{1}{1+K\tau_i} - P(T_{i-1},T_i)\right)^+}{P(T_{i-1},T_i)}\right)\right)\\
&= N(1+K\tau_i)E_t^*\left(\frac{P(T_{i-1},T_i)e^{-\int_t^{T_{i-1}} r_s ds}\left(\frac{1}{1+K\tau_i} - P(T_{i-1},T_i)\right)^+}{P(T_{i-1},T_i)}\right)\\
&= N(1+K\tau_i)E_t^*\left(e^{-\int_t^{T_{i-1}} r_s ds}\left(\frac{1}{1+K\tau_i} - P(T_{i-1},T_i)\right)^+\right)\\
&= N(1+K\tau_i)P_t^{T_{i-1},T_i}\left(\frac{1}{1+K\tau_i}\right)
\end{aligned}
$$

□

Analogously, a floor is equivalent to a receiver interest-rate swap where each exchange payment is executed only if it has positive value, which can also be interpreted by a sequence of floorlets. The corresponding floorlet can also be calculated from the caplet via put-call parity.

The price of interest rate caps may now be easily computed by simply taking the sum of the prices of sequential caplets,

$$
V^{\text{Cap}}(t,T_1,T_n,K) = \sum_{i=2}^{n} V^{\text{Caplet}}(t,T_{i-1},T_i,K)
$$

where i is a prespecified set of dates T_2 ... T_n and τ_i is a year fraction between T_{i-1} and T_i. As discussed previously, the price of the caplet at time T_{i-1} equals $1+\tau_i K$ zero-bond puts with strike $\frac{1}{1+\tau_i K}$. The caplet price thus can be calculated according to the zero-bond options that have been derived in section 4.3.1.

4.3.3 Swaptions

Now I derive the price of a swaption via an approximation method proposed by Schrager and Pelser [2006]. A swaption can be viewed as a European option on a coupon bond as follows. Given a set of dates, T_i, $i = n+1, \dots, N$ at which swap payments are to be made, the forward par swap rate $y_{n,N}(t)$ is given as

$$y_{n,N}(t) = \frac{P(t,T_n) - P(t,T_N)}{P_{n+1,N}(t)} \tag{4.22}$$

where

$$P_{n+1,N}(t) = \sum_{i=n+1}^{N} \Delta_{i-1}^{Y} P(t,T_i)$$

is the present value of a basis point, which is used as the numéraire for the $n+1, N$th swap measure and Δ_{i-1}^{Y} represents the year fraction between T_{i-1} and T_i. The swaption price, $\mathtt{PS}_t(K)$ is then given by,

$$\begin{aligned} \mathtt{PS}_t(K) &= B_t E_t^* \left(\frac{P_{n+1,N}(T_n)}{B_{T_n}} [y_{n+1,N}(T_n) - K]^+ \right) \\ &= P_{n+1,N}(t) E_t^{n+1,N} \left([y_{n+1,N}(T_n) - K]^+ \right) \end{aligned}$$

where the expectation $E_t^{n+1,N}(\cdot)$ is taken under the $n+1, N$th swap measure.

To solve the conditional expectation under the $n+1, N$th swap measure $\Psi_t(K) := E_t^{n+1,N}([y_{n+1,N}(T_n) - K]^+)$, the Carr and Madan [1999]'s approach is again used. Assume $\phi(u, t, T_n)$ to be the conditional characteristic function of

$y_{n,N}(T_n)$ under the $n+1, N$th swap measure, such that

$$\begin{aligned}\phi(\mu, t, T_n) &= E_t^{n+1,N}\left(e^{iuy_{n,N}(T_n)}\right)\\ &= \int_{-\infty}^{+\infty} e^{ius} g(s) ds\end{aligned}$$

with the density of $y_{n,N}(T_n)$ under the $n+1, N$th swap measure given by $g(s)$. Now define

$$\overline{\Psi}_t(K) = e^{\alpha K}\Psi_t(K)$$

and the Fourier transform

$$\psi(v, t, T_n) = \int_{-\infty}^{+\infty} e^{ivK}\overline{\Psi}_t(K) dK \quad .$$

The conditional expectation $\Psi_t(K)$ can thus be obtained as the corresponding inversion

$$\begin{aligned}\Psi_t(K) &= \frac{e^{-\alpha K}}{2\pi}\int_{-\infty}^{+\infty} e^{ivk}\psi_T(v) dv\\ &= \frac{e^{-\alpha K}}{\pi}\int_{0}^{+\infty} \texttt{Real}\left\{e^{ivk}\psi_T(v)\right\} dv\end{aligned}$$

Lemma 4.3.5. *The expression $\psi_{T_n}(v)$ is given by*

$$\psi_{T_n}(v) = \frac{\phi(v - i\alpha, t, T_n)}{(iv+\alpha)^2}$$

where ϕ is the conditional characteristic function of $y_{n+1,N}(T_n)$ under the $n+1, N$th swap measure.

Proof. The proof of this Lemma is similar to Lee [2004]. The expression of $\psi_{T_n}(v)$ can be derived by

$$\begin{aligned}\psi_{T_n}(v) &= \int_{-\infty}^{+\infty} e^{ivK}\overline{\Psi}_t(K)dK \\ &= \int_{-\infty}^{+\infty} e^{ivK}\int_{K}^{+\infty} e^{\alpha K}(s-K)q(s)dsdK\end{aligned}$$

applying Fubini's theorem,

$$\begin{aligned}\psi_{T_n}(v) &= \int_{-\infty}^{+\infty} q(s)\int_{-\infty}^{s} e^{ivK}(se^{\alpha K}-Ke^{\alpha K})dKds \\ &= \int_{-\infty}^{+\infty} q(s)\left[\frac{se^{s(iv+\alpha)}}{iv+\alpha}-\frac{e^{s(iv+\alpha)}(s(iv+\alpha)-1)}{(iv+\alpha)^2}\right]ds \\ &= \frac{\phi(v-i\alpha,t,T_n)}{(iv+\alpha)^2}\end{aligned}$$

This completes the proof. □

The proposition 4.3.6 incorporates the entire structure into the affine framework.

Proposition 4.3.6. *Denote by*

$$\begin{cases} q_n^Y(t) = -\frac{P(t,T_n)}{P_{n+1,N}} \\ q_N^Y = \left(1+\Delta_{i-1}^Y y_{n,N}(t)\right)\frac{P(t,T_i)}{P_{n+1,N}(t)} \\ q_i^Y(t) = \Delta_{i-1}^Y y_{n,N}(t)\frac{P(t,T_i)}{P_{n+1,N}(t)} \quad i=n+1,\dots,N-1 \\ p^i(t) = \Delta_{i-1}^Y\frac{P(t,T_i)}{P_{n+1,N}} \end{cases} \tag{4.23}$$

Assume that p_t^i and q_t^i can be approximated by their initial values p_0^i and q_0^i and the conditional characteristic function of $y_{n,N}(T)$ can be obtained in closed form, *then*

$$E_t^*\left(e^{iuy_{n,N}(T)+\ln \bar{P}_{n+1,N}(T)}\right)=e^{\widetilde{M}(\tau)+\widetilde{\mathbf{N}}(\tau)^\top \widetilde{\mathbf{X}}_t}$$

where $\widetilde{M}(\tau)$ *and* $\widetilde{\mathbf{N}}(\tau)$ *satisfies the following system of Riccati equations*

$$\begin{aligned}\frac{\partial \widetilde{M}(\tau)}{\partial \tau}&=\widetilde{\mathbf{N}}(\tau)^\top \widetilde{\mathbf{K}}(\tau)\widetilde{\boldsymbol{\theta}}(\tau)\\ \widetilde{M}(0)&=0\\ \frac{\partial \widetilde{\mathbf{N}}(\tau)^\top}{\partial \tau}&=-\widetilde{\mathbf{N}}(\tau)^\top \widetilde{\mathbf{K}}(\tau)+\frac{1}{2}\widetilde{\mathbf{N}}(\tau)^\top \widetilde{\boldsymbol{\Sigma}}\widetilde{\boldsymbol{\Sigma}}^\top \widetilde{\mathbf{N}}(\tau)\mathbf{S}_{6,6}\\ \widetilde{\mathbf{N}}(0)^\top&=[iu,\mathbf{S}_{5,1}]\end{aligned} \tag{4.24}$$

where $\mathbf{S}_{i,j}$ *is a length* i *row vector with all elements zero, except for position* j *which equals* 1.

Proof. The conditional characteristic function $\phi(u,t,T_n)$ under the $n+1,N$th swap measure is given by,

$$\phi(u,t,T_n)=\frac{B_t}{P_{n+1,N}(t)}E_t^*\left(e^{iuy_{n,N}(T_n)+\ln \bar{P}_{n+1,N}(T_n)}\right) .$$

The swap rate $y_{n,N}(t)$, see equation (4.22), is a martingale under the $n+1,N$th swap measure with SDE

$$dy_{n,N}(t)=\sqrt{v_t}\left[\sum_{i=n}^{N} q_i^Y \boldsymbol{\alpha}(t,T_i)^\top\right]\boldsymbol{\Sigma} d\mathbf{W}_t^{n+1,N}$$

where $q_n^Y(t)$, q_N^Y and $q_i^Y(t), i=n+1,\ldots,N-1$ are defined in equation (4.23) and $\mathbf{W}_t^{n+1,N}$ is the 3-dimension standard Brownian Motion under the $n+1,N$th swap measure, such that

$$d\mathbf{W}_t^{n+1,N} = \sqrt{v_t}\mathbf{\Sigma}^\top \left[\sum_{i=n+1}^{N} \Delta_{i-1}^Y \frac{P(t,T_i)}{P_{n+1,N}(t)} \boldsymbol{\alpha}(t,T_i) \right] dt + d\mathbf{W}_t^* \ .$$

By applying Itô's formula, the SDE for $y_{n,N}(T_n)$ and $\ln \bar{P}_{n+1,N}(T_n)$ may be obtained as follows,

$$\begin{aligned} dy_{n,N}(t) =& v_t \left(\sum_{i=n}^{N} q_t^i \boldsymbol{\alpha}(t,T_i)^\top \right) \mathbf{\Sigma}\mathbf{\Sigma}^\top \left(\sum_{i=n+1}^{N} p_t^i \boldsymbol{\alpha}(t,T_i) \right) dt \\ & + \sqrt{v_t} \left(\sum_{i=n}^{N} q_t^i \boldsymbol{\alpha}(t,T_i)^\top \right) \mathbf{\Sigma} d\mathbf{W}_t^* \end{aligned}$$

and

$$\begin{aligned} d\ln \bar{P}_{n+1,N}(t) =& -\frac{1}{2} v_t \left(\sum_{i=n+1}^{N} p_t^i \boldsymbol{\alpha}(t,T_i)^\top \right) \mathbf{\Sigma}\mathbf{\Sigma}^\top \left(\sum_{i=n+1}^{N} p_t^i \boldsymbol{\alpha}(t,T_i) \right) dt \\ & + \sqrt{v_t} \left(\sum_{i=n+1}^{N} p_t^i \boldsymbol{\alpha}(t,T_i)^\top \right) \mathbf{\Sigma} d\mathbf{W}_t^* \end{aligned}$$

where q_t^i is given in equation (4.23) and p_t^i is defined as,

$$p_t^i = \Delta_{i-1}^Y \frac{P(t,T_i)}{P_{n+1,N}(t)} \ .$$

Since p_t^i and q_t^i are low volatility martingales, which could be approximated by their initial value p_0^i and q_0^i respectively, see Schrager and Pelser [2006].

The previous approximation makes the model structure affine such that

$$\begin{cases} dy_{n,N}(t) = v_t A_t^q dt + \sqrt{v_t}\left(B_t^q\right)^\top \Sigma d\mathbf{W}_t^* \\ d\ln \bar{P}_{n+1,N}(t) = -\frac{1}{2} v_t A_t^p dt + \sqrt{v_t}\left(B_t^p\right)^\top \Sigma d\mathbf{W}_t^* \\ d\mathbf{X}_t = \mathbf{K}\left(\boldsymbol{\theta} - \mathbf{X}_t\right) dt + \sqrt{v_t}\Sigma d\mathbf{W}_t^* \\ dv_t = k_v\left(\theta_v - v_t\right) dt + \sqrt{v_t}\Sigma_v^\top d\mathbf{W}_t^* \end{cases} \tag{4.25}$$

where A_t^q, $\mathbf{B}_t^q$, A_t^p and $\mathbf{B}_t^p$ are deterministic functions of time, given by,

$$\begin{cases} A_t^q = \left(\sum_{i=n}^N q_t^i \boldsymbol{\alpha}(t,T_i)^\top\right) \Sigma\Sigma^\top \left(\sum_{i=n+1}^N p_t^i \boldsymbol{\alpha}(t,T_i)\right) \\ \mathbf{B}_t^q = \sum_{i=n}^N q_t^i \boldsymbol{\alpha}(t,T_i) \\ A_t^p = -\frac{1}{2}\left(\sum_{i=n+1}^N p_t^i \boldsymbol{\alpha}(t,T_i)^\top\right) \Sigma\Sigma^\top \left(\sum_{i=n+1}^N p_t^i \boldsymbol{\alpha}(t,T_i)\right) \\ \mathbf{B}_t^p = \sum_{i=n+1}^N p_t^i \boldsymbol{\alpha}(t,T_i) \end{cases}$$

Define $\widetilde{\mathbf{X}}_t = \left[y_{n,N}(t), \ln \bar{P}_{n+1,N}(t), \overline{\mathbf{X}}_t^\top\right]^\top$ to be the factors under the enlarged state space. Equation (4.25) can be written in the following manner,

$$d\widetilde{\mathbf{X}}_t = \widetilde{\mathbf{K}}(t)\left(\widetilde{\boldsymbol{\theta}}(t) - \widetilde{\mathbf{X}}_t\right) dt + \sqrt{v_t}\widetilde{\Sigma}(t) d\mathbf{W}_t^* \ . \tag{4.26}$$

Thus, according to Duffie and Kan [1996], one has the systems of the Riccati equations (4.24), which completes the proof. □

Swaptions prices can thus be obtained via,

$$\text{PS}_t(K) = P_{n+1,N}(t)\Psi_t(K) \ .$$

Although the methodology applied here is based on Schrager and Pelser [2006], their original approach was based on the affine term structure framework. My method on the other hand derives from the HJM framework and thus represents an original contribution.

4.4 Conclusion

Although most factor models of term structure perform well both in-sample and out-of-sample, criticism largely focuses on the lack of a theoretical background, especially the arbitrage-free assumption. In this chapter, I present a methodology, based on the HJM no-arbitrage condition, to connect empirical term structure models with theoretical interest rate models. The presented framework enjoys the advantages of both factor models as well as theoretical models. Additionally, I introduce the stochastic volatility into this model to relax the constant volatility assumption in Christensen, Diebold, and Rudebusch [2011], which is inconsistent with the empirical findings from financial markets. The market price of risk is derived according to Duffie, Pan, and Singleton [2000], Trolle and Schwartz [2009a], and Christensen, Diebold, and Rudebusch [2011]. I have derived the solutions to these contingent claims under my arbitrage-free Nelson-Siegel model by employing the method of Carr and Madan [1999], Duffie, Pan, and Singleton [2000] and Schrager and Pelser [2006]. This framework allows me to exploit as much information as possible for modelling the interest rates and their derivatives.

Chapter 5

On the joint calibration of LIBOR/Swap and interest rate derivatives

The consistent pricing framework of interest rate caps and swaptions is derived from the studies of Carr and Madan [1999] and Duffie, Pan, and Singleton [2000]. The model is calibrated using extensive panel data, including Libor/Swap rates, ATM caps and swaptions. By casting the entire model into a state-space form, an extended Kalman filter is employed to estimate the model. The results show that the model prices both interest rates and their derivatives accurately.

5.1 Introduction

Interest rate contracts are amongst the most liquid instruments in the financial market. There are numerous term structure models that are proposed in academic literature for pricing interest rate derivatives. Examples are Hull and White [1990], Carr and Madan [1999], Singleton and Umantsev [2002], Chacko and Das [2002], Collin-Dufresne and Goldstein [2003] and Schrager and Pelser [2006]. The closed-form solutions also exploit the potential to jointly model interest rates and their derivatives, for example, the general stochastic volatility model of Trolle and Schwartz [2009a], as well as stochastic volatility Libor market models of Han [2007] and Jarrow, Li, and Zhao [2007]. Trolle and Schwartz [2009a] estimate their model using extensive data, including the US Libor/Swap rate, caps and swaptions, whereas Han [2007] estimates his model using swaption data and Jarrow, Li, and Zhao [2007], using cap skew data.

The pricing method of Schrager and Pelser [2006] used in section 4.3 is mathematically elegant, however from an empirical perspective, there are two disadvantages. Firstly, by using an extended Kalman filter, it is computationally inefficient as fourteen Riccati equations must be solved in every loop of the Kalman filter. Secondly it is also inefficient in that the pricing method for swaptions and caps are different in chapter 4. The pricing and hedging of interest rate derivatives not only needs an accurate price, but also requires the results to be calculated in real time. The computationally efficient therefore turns into a big problem for trading interest rate derivatives. To overcome these problems, an approximated scheme is introduced to price the coupon bond options.[1] Using this scheme, the Riccati equations only have to be solved once for each Kalman loop. The zero-coupon bond options can also be priced as a special case of this method which implies that swaptions and caps are priced under a consistent framework.

[1]Similar to the zero-bond options in pricing the caps, it is widely agreed that the swaption prices can be given by a specialized coupon-bond option contract.

This joint framework is then estimated using Libor/Swap rates, ATM caps and ATM swaptions from the US market. Similar to Trolle and Schwartz [2009a], the model is estimated using a quasi maximum likelihood via the extended Kalman filter. The results show that the three factor model provides a very good fit to both interest rates and interest rate derivatives, which is consistent with the literature that three factors are required to explain the variations in the term structure, see Litterman and Scheinkman [1991]. The additional unspanned stochastic volatility factors are necessary to capture the variation of interest rate derivatives.

The remainder of this chapter is organized as follows. Section 5.2 introduces the consistent pricing framework of interest rate contracts. Section 5.3 describes in detail the estimation framework of the model in previous section, such as the state-space form and the extended Kalman filter. In Section 5.4, the data and the estimation results are presented. Section 5.5 concludes the chapter.

5.2 Pricing of interest rate derivatives

5.2.1 Bond options

Denote $P_t^{T_1,T_m}$ as the coupon-bond price,

$$P_t^{T_1,T_m} = \sum_{i=1}^{m} c_i P(t, T_i)$$

where c_i is the coupon payments at time T_i, $i \in \{1, \ldots, m\}$ and $P(t, T_i)$ is the zero-coupon bond price maturity at time T_i.

$C_t^{T_0,T_1}(K)$ is the European call option written on a coupon bond $P_t^{T_1,T_m}$ with strike price K, maturity T_0, $T_0 < T_1$ such that

$$C_t^{T_0,T_1}(K) = E_t^*\left(e^{-\int_t^{T_0} r_s ds}[P_t^{T_1,T_m} - K]^+\right) \quad ,$$

where the conditional expectations $E^*(\cdot)$ is taken under the risk neutral measure. Recall that the Radon-Nikodym derivative is given by,

$$\frac{d\mathbb{P}^{T_0}}{d\mathbb{P}} = \frac{B(s)}{P(s,T_0)}$$

Changing to the forward measure $\mathbb{P}^{T_0}$,

$$\begin{aligned} C_t^{T_0,T_1}(K) &= E_t^{T_0}\left(\frac{\frac{B(T_0)}{P(T_0,T_0)}}{\frac{B(t)}{P(t,T_0)}}[P_t^{T_1,T_m} - K]^+\right) \\ &= P(t,T_0)E_t^{T_0}\left([P_t^{T_1,T_m} - K]^+\right) \end{aligned} \tag{5.1}$$

where the conditional expectation $E_t^{T_0}$ is taken under the T_0th forward measure $\mathbb{P}^{T_0}$. The zero-coupon bond price $P(T_0, T_i)$ is calculated by equation 4.10. The transform inversion techniques in Duffie, Pan, and Singleton [2000] and Carr and Madan [1999] are applied to solve the conditional expectations in equation 5.1. Basically, an approximate form for the conditional characteristic function of the logarithm coupon-bond price $\ln P_t^{T_1,T_m}$ is derived. The Fourier inversion has been applied to the approximate conditional characteristic function to solve the coupon-bond option prices numerically.

Denote $\phi_{T_0}(u)$ as the conditional characteristic function of the logarithm of the coupon-bond price under the T_0th forward measure $\mathbb{P}^{T_0}$,

$$\begin{aligned} \phi_{T_0}(u) &= E_t^{P^{T_0}}\left(e^{iu \ln \sum_{i=1}^{m} c_i P(T_0,T_i)}\right) \\ &= \int_{-\infty}^{+\infty} e^{ius} q_c^{T_0}(s) ds \end{aligned} \tag{5.2}$$

where $q_c^{T_0}(s)ds$ is the density of $\ln P_t^{T_1,T_m}$ under the T_0th forward measure. Further define k as the natural logarithm of the strike price K,

$$k = \ln K$$

$\Psi(k)$ is the conditional expectation in equation (5.1),

$$\begin{aligned} \Psi(k) &= E_t^{T_0}\left([P_t^{T_1,T_m} - K]^+\right) \\ &= E_t^{T_0}\left([e^{\ln P_t^{T_1,T_m}} - e^k]^+\right) \end{aligned}$$

Using equation 5.2,

$$\Psi(k) = \int_k^{+\infty} (e^s - e^k) q_c^{T_0}(s) ds$$

To guarantee square integrability, define the modified conditional expectation term $\overline{\Psi}(k)$ by,

$$\overline{\Psi}(k) = e^{\alpha k} \Psi(k)$$

where α is any real number and $\alpha > 1$. Similar to the previous pricing framework, the Fourier transform of $\overline{\Psi}(k)$ is given by,

$$\psi_{T_0}(v) = \int_{-\infty}^{+\infty} e^{ivk}\overline{\Psi}(k)dk$$

Reversing the Fourier transform gives,

$$\overline{\Psi}(k) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-ivk}\psi_{T_0}(v)dv$$

Recall that $\overline{\Psi}(k) = e^{\alpha k}\Psi(k)$, $\Psi(k)$ can thus be derived by,

$$\Psi(k) = \frac{e^{-\alpha k}}{2\pi}\int_{-\infty}^{+\infty} e^{-ivk}\psi_{T_0}(v)dv$$

$\Psi(k)$ is a real number since it is the price of the coupon-bond options. The function $\psi_{T_0}(v)$ is therefore odd in its imaginary part and even in its real part, which gives,

$$\Psi(k) = \frac{e^{-\alpha k}}{\pi}\int_{0}^{+\infty} Re\{e^{-ivk}\psi_{T_0}(v)\}dv$$

$\psi_{T_0}(v)$ is given in lemma 4.3.1,

$$\psi_{T_0}(v) = \frac{\phi_{T_0}(v-(\alpha+1)i)}{\alpha^2+\alpha-v^2+i(2\alpha+1)v} \ .$$

where $\phi_{T_0}(u)$ is given by,

$$\begin{aligned}\phi_{T_0}(u) &= E_t^{T_0}\left(e^{iu\ln\sum_{i=1}^{m} c_iP(T_0,T_i)}\right)\\ &= E_t^{*}\left(\frac{B(t)}{B(T_0)}\frac{1}{P(t,T_0)}e^{iu\ln\sum_{i=1}^{m} c_iP(T_0,T_i)}\right)\end{aligned}$$

However, in the case of the coupon-bond options, there are no closed-form solutions for the conditional characteristic function $\phi_{T_0}(u)$, see Brace, Gatarek, and Musiela [1997], Singleton and Umantsev [2002] and Schrager and Pelser [2006]. Approximate methods are needed in this situation.

Assumption 5.2.1. *Assume the fraction of the jth coupon payment, $c_j P(t, T_j)$, $j \in \{1, \ldots, m\}$, with respect to the entire coupon bond price $\sum_{i=1}^m c_i P(t, T_i)$*

$$w_j(t) = \frac{c_j P(t, T_j)}{\sum_{i=1}^m c_i P(t, T_i)}$$

for $t > 0$ can be approximated by its initial value $W_j(0)$.

This assumption is similar to Brace, Gatarek, and Musiela [1997] and Schrager and Pelser [2006].

Proposition 5.2.2. *Given assumption 5.2.1, the conditional characteristic function of the coupon bond is given by*

$$\begin{aligned}\phi_{T_0}(u) &= E_t^*\left(\frac{B(t)}{B(T_0)}\frac{1}{P(t, T_0)}e^{iu \ln \sum_{i=1}^m c_i P(T_0, T_i)}\right)\\ &= e^{M(\tau)+N(\tau)v_t+iu\ln\frac{\sum_{i=1}^m c_i P(t,T_i)}{P(t,T_0)}}\end{aligned}$$

where $M(\tau)$ and $N(\tau)$ satisfies the following system of Riccati equations

$$\begin{aligned}&\frac{dM}{d\tau} = N k_v \theta_v\\ &M(0) = 0\\ &\frac{dN}{d\tau} = -N k_v - \frac{1}{2} iu \left[\sum_{i=1}^m w_i(t) \boldsymbol{\alpha}_i^\top\right] \Sigma\Sigma^\top \left[\sum_{i=1}^m w_i(t) \boldsymbol{\alpha}_i\right]\\ &\quad + \frac{1}{2}\left[N \Sigma_{v_{1-3}}^\top - iu \sum_{i=1}^m \boldsymbol{\alpha}_i^\top \Sigma - (1 - iu)\boldsymbol{\alpha}_0^\top \Sigma\right]\end{aligned} \tag{5.3}$$

$$\left[N\Sigma_{v_{1-3}}^{\top} - iu\sum_{i=1}^{m}\boldsymbol{\alpha}_i^{\top}\Sigma - (1-iu)\boldsymbol{\alpha}_0^{\top}\Sigma\right]^{\top}$$

$$-\frac{1}{2}(1-iu)\boldsymbol{\alpha}_0^{\top}\Sigma\Sigma^{\top}\boldsymbol{\alpha}_0 + \frac{1}{2}N^2\sigma_{v_4}^2$$

$$N(0) = 0$$

The proof of this statement is not obvious, so I first introduce the lemma 5.2.3.

Lemma 5.2.3. *Assume* $x_t = e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta\ln\sum_{i=1}^m c_i P(t,T_i)+(1-\beta)\ln P(t,T_0)}$ *is a martingale, the proposition 5.2.2 can thus be proved.*

Proof. if the process x_t,

$$x_t = e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta\ln\sum_{i=1}^m c_i P(t,T_i)+(1-\beta)\ln P(t,T_0)}$$

is a martingale, then

$$E_t^*\left(e^{-\int_0^{T_0} r_s ds} e^{M(0)+N(0)v_t+\beta\ln\sum_{i=1}^m c_i P(T_0,T_i)+(1-\beta)\ln P(T_0,T_0)}\right)$$

given $M(0) = 0$, $N(0) = 0$ and $P(T_0, T_0) = 1$,

$$= E_t^*\left(e^{-\int_0^{T_0} r_s ds} + \beta\ln\sum_{i=1}^m c_i P(T_0, T_i)\right)$$

$$= e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta\ln\sum_{i=1}^m c_i P(t,T_i)+(1-\beta)\ln P(t,T_0)}$$

taking $\ln P(t,T_0)$ out of the exponential term,

$$= P(t,T_0)e^{-\int_0^t r_s ds}e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{\sum_{i=1}^m c_i P(t,T_i)}{P(t,T_0)}}$$

rearranging the equation,

$$E_t^*\left(\frac{1}{P(t,T_0)}e^{-\int_0^{T_0} r_s ds+\beta \ln \sum_{i=1}^m c_i P(T_0,T_i) - e^{-\int_0^t r_s ds}}\right)$$
$$= e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{\sum_{i=1}^m c_i P(t,T_i)}{P(t,T_0)}}$$

using the dynamics of the bank saving accout $B(t) = \int_0^t r_s ds$,

$$E_t^*\left(\frac{B(t)}{B(T_0)}\frac{1}{P(t,T_0)}e^{iu \ln \sum_{i=1}^m c_i P(T_0,T_i)}\right)$$
$$= e^{M(\tau)+N(\tau)v_t+\beta \ln \frac{\sum_{i=1}^m c_i P(t,T_i)}{P(t,T_0)}}$$

The theorem can then be proved by specifying $\beta = iu$. □

Proof. At the maturity time T_0,

$$E_t^*\left(\frac{B(T_0)}{B(T_0)}\frac{1}{P(T_0,T_0)}e^{iu \ln \sum_{i=1}^m c_i P(T_0,T_i)}\right) = e^{M(0)+N(0)v_t+iu \ln \frac{\sum_{i=1}^m c_i P(T_0,T_i)}{P(T_0,T_0)}}$$

where $M(0)$ and $N(0)$ are all zero at maturity time T_0 ($\tau = 0$) as well as $P(T_0,T_0) = 1$,

$$E_t^*\left(e^{iu \ln \sum_{i=1}^m c_i P(T_0,T_i)}\right) = e^{iu \ln \sum_{i=1}^m c_i P(T_0,T_i)}$$

at the maturity time T_0, the $e^{iu \ln \sum_{i=1}^m c_i P(T_0,T_i)}$ can be taken out of the expectation

E_t^* since it is known at time T_0. Thus,

$$e^{iu \ln \sum_{i=1}^{m} c_i P(T_0,T_i)} = e^{iu \ln \sum_{i=1}^{m} c_i P(T_0,T_i)}$$

The theorem has been proved at the maturity time T_0.

In the generalised situation, following lemma 5.2.3, the only task left is to prove that x_t is a martingale. According to properties of the martingale, the ODEs for $M(\tau)$ and $N(\tau)$ are required to guarantee that the drift term of dx_t is zero. Recalling the SDE of $\ln \sum_{i=1}^{m} c_i P(T_0, T_i)$ and 4.3.1, the $M(\tau)$ and $N(\tau)$ are required to satisfy the ODEs in 5.2.2 to guarantee that the x_t is a martingale. □

This model is different from Schrager and Pelser [2006] since the framework of Heston [1993] has been applied. The logarithm of the characteristic function is calculated by the transform inversion techniques of Duffie, Pan, and Singleton [2000]. Only two Riccati equations are required to be solved in this case. If one simply applies Schrager and Pelser [2006], more Riccati equations are needed to solve by using the affine term structure models from Duffie and Kan [1996] and Dai and Singleton [2000]. There are fourteen Riccati equations must be solved by the method from section 4.3.3.

5.2.2 Caps and floors

If the coupon-bond pays one unit of the currency at time T_m ($c_m = 1$) and zero at the other paying dates $c_i = 0$, $i \in \{1, \dots, m-1\}$. This coupon-bond is reduced to a zero-coupon bond $P(T_0, T_m)$. The conditional characteristic function $\phi_{T_0}(u)$ can be obtained in closed form from the proof in 4.3.1. The conditional characteristic function introduced in (4.17) is given by

$$E_t^*\left(\frac{B(t)}{B(T_0)}\frac{1}{P(t,T_0)}e^{iu\ln P(T_0,T_1)}\right)$$
$$= e^{M(\tau)+N(\tau)v_t+iu\ln\frac{P(t,T_1)}{P(t,T_0)}}$$

where $M(\tau)$ and $N(\tau)$ satisfy the following system of Riccati equations

$$\begin{aligned}
\frac{dM(\tau)}{d\tau} &= N(\tau)k_v\theta_v \\
M(0) &= 0 \\
\frac{dN(\tau)}{d\tau} &= -\frac{1}{2}(1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\boldsymbol{\alpha}(t,T_0) - \frac{1}{2}iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\boldsymbol{\Sigma}^\top\boldsymbol{\alpha}(t,T_1) \\
&+ \left(N(\tau)\boldsymbol{\Sigma}_v^\top - (1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma} - iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\right) \qquad (5.4) \\
&\left(N(\tau)\boldsymbol{\Sigma}_v^\top - (1-iu)\boldsymbol{\alpha}(t,T_0)^\top\boldsymbol{\Sigma} - iu\boldsymbol{\alpha}(t,T_1)^\top\boldsymbol{\Sigma}\right)^\top \\
&- N(\tau)k_v + \frac{1}{2}N(\tau)^2(\sigma_v^4)^2 \\
N(0) &= 0
\end{aligned}$$

where $\boldsymbol{\Sigma}_v = (\sigma_v^1, \sigma_v^2, \sigma_v^3)^\top$.

Proof. The proof is similar to proposition 5.2.2. According to Duffie, Pan, and Singleton [2000], Collin-Dufresne and Goldstein [2003], and Trolle and Schwartz [2009a],

$$E_t^*\left(e^{-\int_t^T r_s ds}e^{\beta\ln P(T_0,T_1)}\right) = e^{M(\tau)+N(\tau)v_t+\beta\ln P(t,T_1)+(1-\beta)\ln P(t,T_0)} \qquad (5.5)$$

where $M(\tau)$ and $N(\tau)$ satisfy the system of Riccati equations. This can be shown if one assumes that the process

$$x_t := e^{-\int_0^t r_s ds} e^{M(\tau)+N(\tau)v_t+\beta \ln P(t,T_1)+(1-\beta)\ln P(t,T_0)}$$

is a martingale. Thus, equation (5.5) would hold by applying the law of iteration. If one specifies $M(\tau)$ and $N(\tau)$ to make x_t as a martingale, then specifying $\beta = iu$ will give equation (5.5). □

According to section 4.3.2, the prices of the interest rate caps can be calculated by summing the prices of the sequential caplets. The price of the corresponding caplets can be further solved by a special zero-coupon bond put option, which is written on a $1 + \tau_i K$ zero-coupon bond with a strike $\frac{1}{1+\tau_i K}$.

5.2.3 Swaptions

A swaption can be seen as an European option on a coupon bond as follows. Given a set of dates, T_i, $i = n+1, \ldots, N$ at which swap payments are to be made, it is well known that a forward par swap rate $y_{n,N}(t)$ is given as

$$y_{n,N}(t) = \frac{P(t,T_n) - P(t,T_N)}{\sum_{i=n+1}^{N} \Delta_{i-1}^{Y} P(t,T_i)} \quad .$$

Furthermore, the swaptions price is given by

$$\begin{aligned}
\mathsf{PS}_t(K) &= E_t^*\left(e^{-\int_t^{T_n} r_s ds} \sum_{i=n+1}^{N} \Delta_{i-1}^{Y} P(T_n,T_i)\,[y_{n+1,N}(T_n) - K]^+\right) \\
&= E_t^*\left(e^{-\int_t^{T_n} r_s ds}\left[1 - P(T_0,T_N) - \sum_{i=n+1}^{N} \Delta_i K P(T_n,T_i)\right]^+\right)
\end{aligned}$$

which can be viewed as a European put option of a coupon bond

$$\text{PS}_t(K) = E_t^* \left(e^{-\int_t^{T_n} r_s ds} \left[1 - P_{T_0}^{T_{n+1},T_N}(T_n) \right]^+ \right)$$

where the coupon payments are $c_i = \Delta_i K$ for $i \in \{1, \ldots, N-1\}$, $c_N = 1 + \Delta_N K$ and the strike price equals 1. Again, the prices of the swaptions can be calculated by this specialised put option written on the coupon-bond by proposition 5.2.2.

5.3 Estimation approach

5.3.1 Model framework

The Libor and Swap rates can be calculated from the corresponding zero-coupon bond prices, which are given by proposition 4.2.8. According to Shreve [2004] and Brigo and Mercurio [2007], the time-t Libor rate for the period t to T is given by,

$$L(t,T) = \frac{1 - P(t,T)}{(T-t)P(t,T)} \tag{5.6}$$

and the time-t swap rate for the period t to T_n and fixed-leg payments at T_1, ..., T_n is given by,

$$S(t,T_n) = \frac{1 - P(t,T_n)}{\nu \sum_{j=1}^{n} P(t,T_j)} \tag{5.7}$$

The US Dollar swap has a semiannual fixed-leg and a quarterly floating-leg. The ν is thus defined by $\nu = \frac{1}{2}$ with regard to the market convention, see Trolle and Schwartz [2009a]. Since the zero-coupon bond prices are derived in closed-form in proposition 4.2.8, the Libor and swap rates can then be calculated based on the formulas.

An interest rate cap contract is designed to provide insurance against the risk of interest rate rising, which is offered by financial institutions in the over-the-counter market, see Hull [2009]. As discussed in section 4.3.2, the cap contract can be priced by the sum of a group of the corresponding caplets. To price interest rate caps therefore reduces in order to pricing the corresponding caplets. According to Brigo and Mercurio [2007], the caplet can be manipulated to $(1+\nu K)$ European put options with strike $\frac{1}{1+\nu K}$. The price of European put options is derived in section 5.2. Therefore, if the Riccati equations in 5.4 can be solved numerically, the price of the corresponding European call option can be calculated by integrals. Put-Call parity is employed to price European put options,

$$\texttt{Call} = E_t^*\left(e^{-\int_t^{T_0} r_s ds}\left[P(T_0,T_1)-K\right]^+\right)$$

Similarly, the price of a European put option is defined by,

$$\texttt{Put} = E_t^*\left(e^{-\int_t^{T_0} r_s ds}\left[K-P(T_0,T_1)\right]^+\right)$$

Intuitively,

$$\begin{aligned}
-\texttt{Put}+\texttt{Call} &= E_t^*\left(e^{-\int_t^{T_0} r_s ds}\left[K-P(T_0,T_1)\right]\right)\\
&= P(t,T_1)-P(t,T_0)K
\end{aligned}$$

Swaptions are options on interest rate swaps. Similar to zero-coupon bonds in pricing the interest rate cap contracts, swaptions can be priced by manipulating coupon-bond options. Specifically, according to Trolle and Schwartz [2009a] the swaption price can be viewed as an European put option on a coupon bond, which has the coupon payments $c_i = \nu_i K$ for $i \in \{1, \ldots, N-1\}$, $c_N = 1 + \nu_N K$ and the strike price 1. The price of a coupon-bond call option is derived in section 5.2.

Put-Call parity is used to solve for the coupon-bond put option. The call option price is defined by,

$$\texttt{Call} = E_t^* \left(e^{-\int_t^{T_0} r_s ds} \left[P_{T_0}^{T_1, T_m} - K \right]^+ \right)$$

where the put option price is defined by,

$$\texttt{Put} = E_t^* \left(e^{-\int_t^{T_0} r_s ds} \left[K - P_{T_0}^{T_1, T_m} \right]^+ \right)$$

The portfolio of long one unit call option and short one unit put option is given by,

$$\begin{aligned} -\texttt{Put} + \texttt{Call} &= E_t^* \left(e^{-\int_t^{T_0} r_s ds} \left[P_{T_0}^{T_1, T_m} - K \right] \right) \\ &= E_t^* \left(e^{-\int_t^{T_0} r_s ds} \left[\sum_{i=1}^{m} c_i P(T_0, T_i) - K \right] \right) \\ &= \sum_{i=1}^{m} c_i P(t, T_i) - K P(t, T_0) \end{aligned}$$

5.3.2 State space form

The state-space representations of the model are presented below. Let $\mathbf{X}_t = (X_t^{i=1:3}, \phi_t^{i=1:8}, v_t)$ denote the vector of state variables, the SDE of the factor $\mathbf{X}_t$

under the physical measure $\mathbb{P}$ is given by,

$$d\mathbf{X}_t = (\boldsymbol{\Psi} - \mathbf{K}\mathbf{X}_t)\, dt + \sqrt{v_t}\boldsymbol{\Sigma} d\mathbf{W}_t^P$$

Applying Itô's Lemma to $e^{\boldsymbol{K}t}\mathbf{X}_t$, we obtain the conditional mean of the factor $\mathbf{X}_t$,

$$E_t(\mathbf{X}_{t+1}) = \int_t^{t+1} e^{-\mathbf{K}(t+1-u)}\boldsymbol{\Psi} du + e^{-\mathbf{K}\Delta t}\mathbf{X}_t \tag{5.8}$$

and the conditional covariance matrix of the factor $\mathbf{X}_t$ is given by,

$$\begin{aligned} Cov_t(\mathbf{X}_{t+1}) = & \int_t^{t+1} \left(1 - e^{-k_v^{\mathbb{P}}(t+1-u)}\right) \theta_v^{\mathbb{P}} e^{-\mathbf{K}(t+1-u)} \Sigma\Sigma^\top e^{-\mathbf{K}^\top(t+1-u)} du \\ & + \left(\int_t^{t+1} e^{-k_v^{\mathbb{P}}(t+1-u)} e^{-\mathbf{K}(t+1-u)} \Sigma\Sigma^\top e^{-\mathbf{K}^\top(t+1-u)} du\right) v_t \end{aligned} \tag{5.9}$$

The state transition equation is thus,

$$\begin{aligned} \mathbf{X}_{t+1} &= \phi_0 + \phi_x \mathbf{X}_t + \omega_{t+1}, \qquad \omega_{t+1} \sim \mathcal{N}(0, Q_{t+1}) \\ Q_{t+1} &= Q_0 + Q_v v_t \end{aligned}$$

where ϕ_0, ϕ_x, and Q_{t+1} are defined based on 5.8 and 5.9. Let $\mathbf{Y}_t$ be the vector of observable quantities and h is the pricing function for the price of interest rate derivatives, the observation equation is given by,

$$Y_t = h(\mathbf{X}_t) + u_t \tag{5.10}$$

Note that $h(\mathbf{X}_t)$ is a non-linear function. However, the standard Kalman filter is designed for linear Gaussian state-space models. This model requires to be calibrated by the extended Kalman filter, which is based on Trolle and Schwartz [2009a].[1]

$$Y_t = \left(h(\hat{\mathbf{X}}_{t|t-1}) - H_t'\hat{\mathbf{X}}_{t|t-1}\right) + H_t'\mathbf{X}_t + u_t, \qquad u_t \sim \mathcal{N}(0, S)$$

where

$$H_t' = \left.\frac{\partial h(\mathbf{X}_t)}{\partial \mathbf{X}_t'}\right|_{\mathbf{X}_t = \hat{\mathbf{X}}_{t|t-1}}$$

The model is now casted into a state-space form. The next step is to estimate the parameters using the extended Kalman filter framework.

5.3.3 Extended Kalman filter

The extended Kalman filter[2] is widely used in engineering to solve nonlinear systems based on the first-order linearization, see Ljung [1979]. By comparing the finite-sample properties of some standard techniques used to estimate term structure models. Duffee and Stanton [2004] conclude that the Kalman filter is the most accurate estimation technique when maximum likelihood is unavailable.

According to Trolle and Schwartz [2009a], this model is cast into a state-space framework that can be estimated by quasi maximum-likelihood (QML)

[1]The QML estimation is in fact not consistent in the present linearized state-space model. However, in Trolle and Schwartz [2009a], the small-sample properties of the QML kalman filter method, in terms of estimating multi-factor stochastic term structure models, is investigated. The results show that there are virtually no biases in the estimation of parameters under both risk-neutral and physical measures.

[2]Particle filter is another possible way to calibrate the model, see Chib, Nardari, and Shephard [2002]. However, in order to compare my model with Trolle and Schwartz [2009a] model, identical estimation method, the extended Kalman filter is employed in this thesis.

in conjunction with the extended Kalman filter.[1] The extended Kalman filter is initialised with the unconditional mean and variance of the underlying unobserved state variable.[2] The recursive process is started by forecasting the state factors where the conditional forecast of the state factors, which are given by,

$$\mathbf{X}_{t+1} = \phi_0 + \phi_x \mathbf{X}_t + \omega_{t+1}, \qquad \omega_{t+1} \sim \mathcal{N}(0, Q_{t+1})$$
$$Q_{t+1} = Q_0 + Q_v v_t$$

The associated conditional variance can be written as follows,

$$P_{t+1|t} = \phi_x P_t \phi_x^{\top} + Q_t \quad .$$

The model forecast can therefore be calculated using the measurement equation,

$$Y_t = \left(h(\hat{\mathbf{X}}_{t|t-1}) - H_t' \hat{\mathbf{X}}_{t|t-1} \right) + H_t' \mathbf{X}_t + u_t, \qquad u_t \sim \mathcal{N}(0, S)$$

The true value of the observed data enters the recursive system and then the predicted errors can be obtained from,

$$\epsilon_t = Y_t - h(X_{t+1|t})$$

[1] By comparing several estimation methods in terms of estimating term structure models, namely Efficient Method of Moments (EMM), Simulated Maximum likelihood (SML), and Quasi-Maximum likelihood (QML), in conjunction with the Kalman filter, Duffee and Stanton [2004] conclude that the QML/Kalman filter approach is preferable because of its better finite-sample properties.

[2] Note that, several different initial values are employed in the estimation procedure to find the global optimal set.

The predicted error is used to update the dynamics of the transition equation,

$$\hat{X}_{t+1} = \hat{X}_{t+1|t} + P_{t+1|t}H_t^{'}F_t^{-1}\epsilon_t$$

The corresponding conditional variance is updated by,

$$P_{t+1} = P_{t+1|t} - P_{t+1|t}H_t^{'}F_t^{-1}H_tP_{t+1|t}$$

where the F_t is given by,

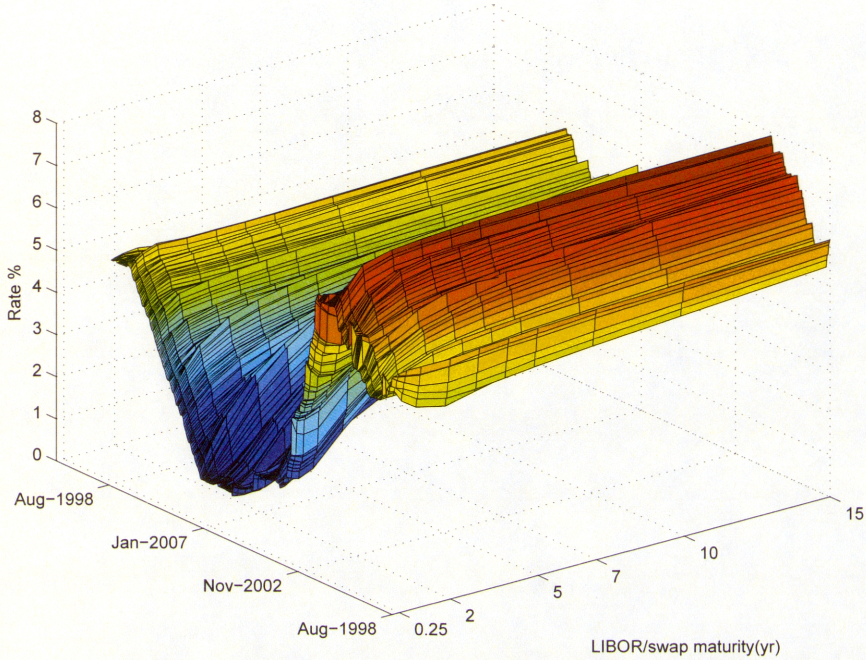
$$F_t = H_tP_{t+1|t}H_t^{'} + S\Omega$$

In the recursive loop, the following time period state variables are forecasted conditional on the previous information. Therefore, the likelihood function can be constructed by repeating the above steps. Under the assumption that the measurement errors are Gaussian, the log-likelihood function is given by,

$$\log L = -\frac{1}{2}\log 2\pi\sum_{i=1}^{T}N_t - \frac{1}{2}\sum_{i=1}^{T}\log|F_t| - \frac{1}{2}\sum_{i=1}^{T}\epsilon_t^{'}F_t^{-1}\epsilon_t$$

The preceding algorithm is treated as the objective function to find the optimal parameter set using non-linear numerical optimization techniques to find the maximum.

Figure 5.1: The term structure of LIBOR/Swap rates



5.4 Empirical results

5.4.1 Data

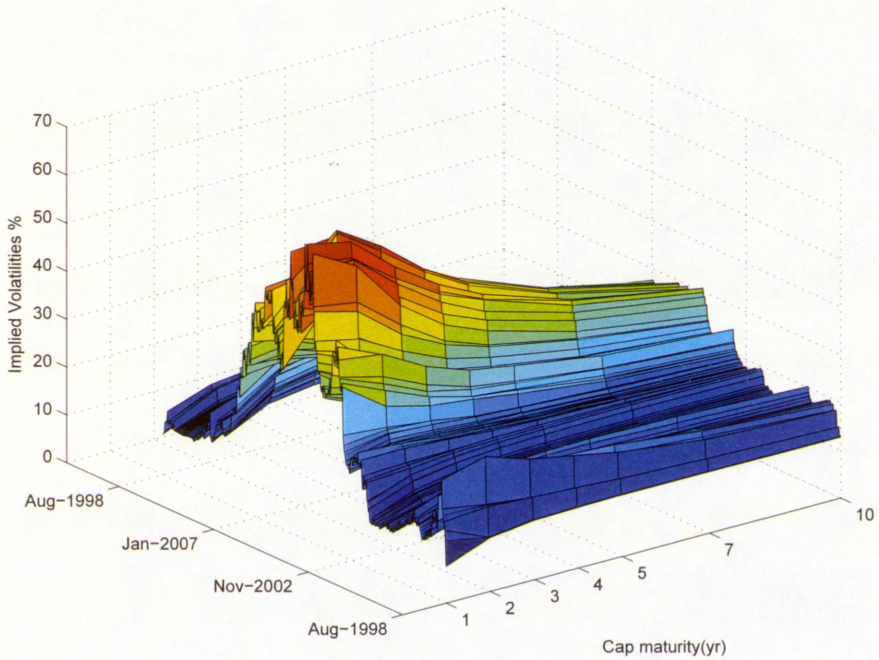
The generalised arbitrage-free Nelson-Siegel model and its corresponding derivative prices are presented in section 5.3.1. The next step is to calibrate the parameters of the model with financial market data.

The models are estimated using the weekly Libor and swap rates and ATM swaptions and caps implied volatilities from August 21, 1998 to January 26, 2007.[1] This data is obtained from Bloomberg and all observations are closing midquotes on Friday. Tables 5.1 and 5.2 present the descriptive statistics of the

[1]There are 441 observations in total. The data period used in this study is identical to Trolle and Schwartz [2009a]

entire sample. The average Libor and swap rates increases with maturity and the volatilities of these rates decrease with maturity.

Figure 5.2: The term structure of interest rate caps data



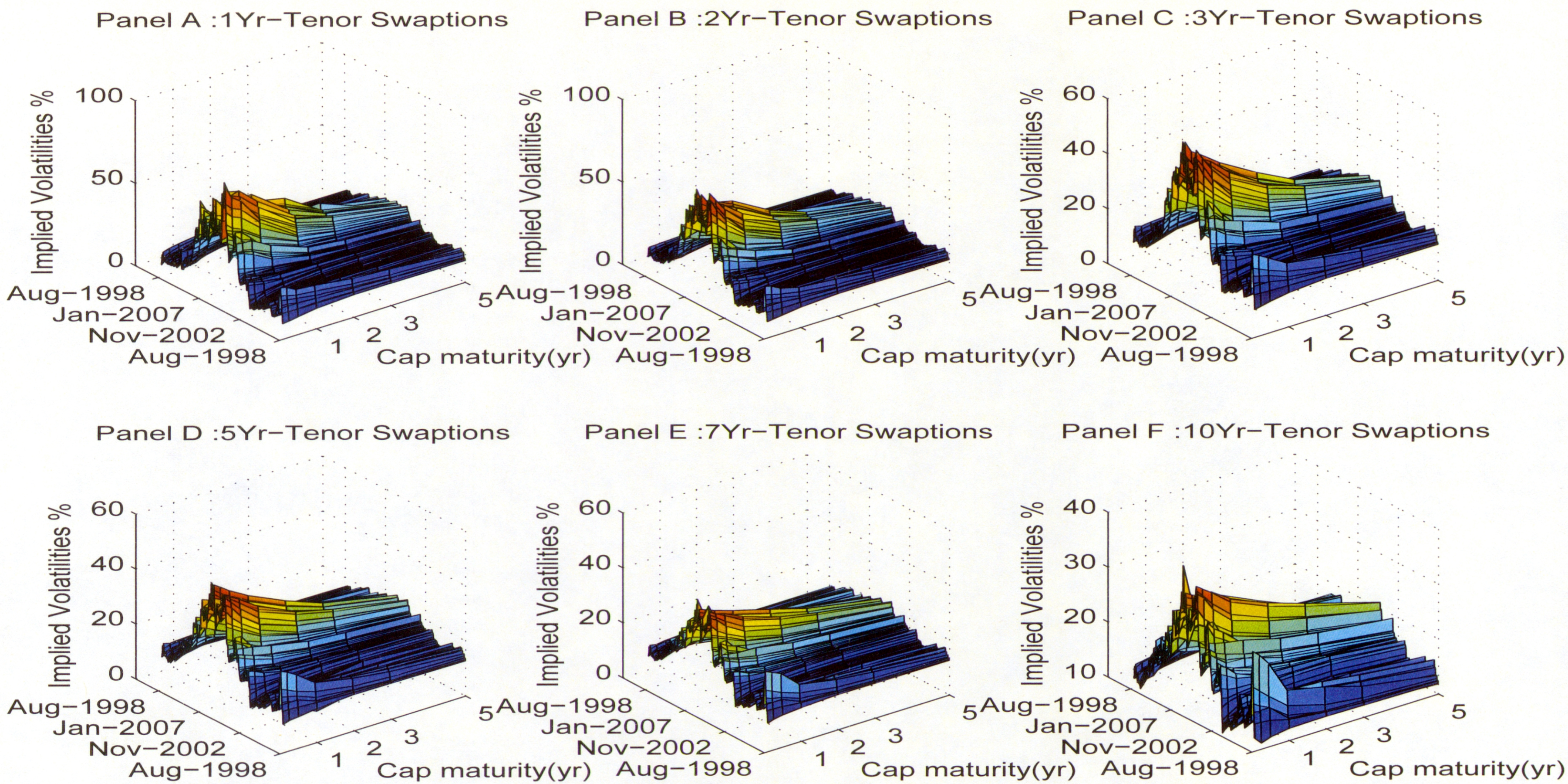
The average implied volatilities of caps and several tenors of implied volatility of swaptions present a humped shape. Following Trolle and Schwartz [2009a], Libor, Swap, Swaption, and Cap markets are assumed to be of homogeneous credit quality because of the identical discount factors for cash-flows. Note that the data used here experiences considerable volatility during long term capital management crisis in 1998 and the dot com bubble in 2001.

The Libor and Swap term structures are constructed from the end-of-week Libor rates at three maturities: 1, 3 and 6 months and swap rate at seven maturities: 1, 2, 3, 5, 7, 10 and 15 years. Figure 5.1 displays the Libor and Swap data.

In Figure 5.3, I show the implied volatilities of ATM swaptions across tenors and time. Swaptions with different tenors at 1, 2, 3, 5, 7 and 10 years are used. The maturities are at 1 month, 3 months, 6 months, 1 year, 2 years, 3 years and 5 years. Thus, the total number of the ATM swaptions is 42. In Figure 5.2, the US dollar caps are graphed across maturity and time. The caps used in this study have maturities of 1, 2, 3, 4, 5, 7 and 10 years.

In order to calibrate the model, several off-line processes are needed. First, the entire forward curve on each observation is calibrated using a cubic spline interpolation. The zero-coupon rate term structure is calculated using the forward curve by using the price formula for Libor and Swap rates, see equations (5.6) and (5.7). Based on the zero-coupon rate term structure, the swaption and cap prices can be calculated as in Black [1976].

Figure 5.3: The term structure of interest rate swaptions data[a]



[a]Note that the "z" scales on the figures are different

Table 5.1: Summary Statistics of Libor and Swap Rates and Implied Volatilities of Caps

Maturity	Mean	Std Dev	Skewness	Kurtosis	Auto	Maturity	Mean	Std Dev	Skewness	Kurtosis	Auto
Libor Rate						Implied Volatilities					
1m	3.660	1.900	0.002	1.491	0.997	1y	24.037	14.362	0.738	2.144	0.984
3m	3.731	1.917	-0.003	1.505	0.997	2y	26.733	14.108	0.886	2.361	0.989
6m	3.813	1.911	-0.015	1.550	0.997	3y	25.906	11.431	0.915	2.461	0.989
1y	3.998	1.866	-0.010	1.695	0.997	4y	24.776	9.257	0.904	2.520	0.988
Swap Rate						5y	23.719	7.611	0.856	2.506	0.986
2y	4.349	1.635	0.028	1.991	0.996	7y	22.035	5.556	0.733	2.438	0.983
3y	4.626	1.449	0.101	2.171	0.995	10y	20.247	3.953	0.592	2.439	0.972
4y	4.834	1.311	0.187	2.285	0.994						
5y	5.000	1.209	0.283	2.352	0.993						
7y	5.244	1.069	0.433	2.429	0.991						
10y	5.483	0.951	0.571	2.467	0.990						
15y	5.729	0.844	0.650	2.467	0.989						

Entries are summary statistics of Libor and Swap rates and implied volatilities of Caps, Mean, Std Dev, Skewness, Kurtosis, and Auto denote the sample estimates of the mean, standard deviation, skewness, excess kurtosis, and first-order autocorrelations, respectively. In the maturity column, Libor maturities are in months(m) and years(y), Swap and Cap maturities are in years(y). The data, are weekly closing quotes from Bloomberg from August 21, 1998 to January 26, 2007. (441 observations in total)

Table 5.2: Summary Statistics of Implied Volatilities of Swaptions

Maturity	Mean	Std Dev	Skewness	Kurtosis	Auto	Maturity	Mean	Std Dev	Skewness	Kurtosis	Auto
1-year tenor						5-year tenor					
1m	26.219	16.142	0.838	2.647	0.977	1m	23.268	10.008	0.732	2.359	0.980
3m	26.425	15.446	0.743	2.266	0.982	3m	23.445	9.527	0.697	2.213	0.987
6m	27.368	15.589	0.813	2.276	0.986	6m	22.952	8.648	0.729	2.286	0.987
1y	26.871	13.300	0.959	2.542	0.988	1y	21.832	7.057	0.755	2.403	0.986
2y	23.944	8.413	1.003	2.842	0.986	2y	20.296	5.051	0.661	2.419	0.983
3y	21.933	5.669	0.845	2.660	0.976	3y	19.202	3.821	0.499	2.344	0.978
5y	19.441	3.162	0.450	2.336	0.971	5y	17.332	2.525	0.159	2.286	0.969
2-year tenor						7-year tenor					
1m	27.635	16.153	0.849	2.403	0.982	1m	21.328	7.755	0.614	2.288	0.975
3m	27.608	15.200	0.790	2.189	0.987	3m	21.566	7.473	0.592	2.156	0.982
6m	26.806	13.727	0.842	2.305	0.988	6m	21.267	6.826	0.612	2.174	0.985
1y	25.192	10.838	0.939	2.604	0.988	1y	20.502	5.739	0.648	2.276	0.986
2y	22.679	7.092	0.880	2.612	0.987	2y	19.352	4.292	0.569	2.352	0.982
3y	20.999	5.007	0.711	2.471	0.982	3y	18.374	3.343	0.393	2.310	0.976
5y	18.763	3.010	0.334	2.235	0.972	5y	16.620	2.274	0.041	2.318	0.965
3-year tenor						10-year tenor					
1m	25.635	13.159	0.798	2.353	0.984	1m	19.540	6.117	0.540	2.355	0.966
3m	25.629	12.350	0.733	2.143	0.988	3m	19.873	5.812	0.474	2.122	0.978
6m	24.990	11.142	0.778	2.250	0.988	6m	19.730	5.345	0.505	2.153	0.982
1y	23.657	8.989	0.858	2.476	0.988	1y	19.146	4.547	0.534	2.223	0.982
2y	21.700	6.143	0.798	2.527	0.985	2y	18.237	3.508	0.452	2.310	0.976
3y	20.298	4.501	0.625	2.395	0.981	3y	17.400	2.770	0.258	2.348	0.970
5y	18.261	2.823	0.277	2.255	0.972	5y	15.768	1.980	-0.144	2.462	0.960

Entries are summary statistics of swaptions implied volatilities, Mean, Std Dev, Skewness, Kurtosis, and Auto denote the sample estimates of the mean, standard deviation, skewness, excess kurtosis, and first-order autocorrelations, respectively. In the maturity column, swaption maturities are in months(m) and years(y) and tenors are in years(y). The data, are weekly closing quotes from Bloomberg from August 21, 1998 to January 26, 2007. (441 observations in total)

5.4.2 Estimation results

The joint framework is estimated on the entire dataset from August 21, 1998 to January 26, 2007, which consists of 441 weekly observations. The parameter estimates are given in Table 5.3. All parameters in matrix λ^1 are negative, which guarantees that the mean reversion of the model under the real world measure $\mathbb{P}$. The unspanned stochastic volatility term σ_v^4 is positive and statistically significant, which confirms the existence of the unspanned stochastic volatility.

Table 5.3: Parameter estimates

m	0.0243	σ_v^4	0.3016	λ_3^0	0.0295	σ_3	0.2057	k_θ^3	2.3688
	(0.3641)		(0.1215)		(0.1386)		(0.3687)		(0.8693)
α	1.1496	k_v	0.3672	λ_4^0	0.1135	σ_v^1	0.2117	λ_1^0	0.2536
	(0.0234)		(0.1037)		(0.3678)		(0.0011)		(0.0013)
λ	0.4663	θ_v	1.5443	λ_1^1	-0.3372	λ_4^1	-0.3096	σ_v^3	0.0499
	(0.0091)		(0.5564)		(0.0023)		(0.2358)		(0.0236)
σ_1	0.4394	k_θ^1	0.3694	λ_2^1	-0.0295	σ_v^2	0.2044	λ_2^0	0.3972
	(0.2423)		(0.1283)		(0.0136)		(0.3687)		(0.2861)
σ_2	0.5975	k_θ^2	0.4408	λ_3^1	-0.0096	σ_{CS}	0.0236	σ_{LS}	0.0008
	(0.0346)		(0.0783)		(0.1078)		(0.0000)		(0.0000)
							Loglikelihood		-43786.37

Standard errors are in parentheses, σ_{LS} denotes the standard deviation of interest rate measurement errors and σ_{CS} denotes the standard deviation of swaption and cap price measurement errors.

This finding is also reported by Collin-Dufresne and Goldstein [2002] and Li and Zhao [2006]. The parameter α is also statistically significantly greater than 1. According to the restriction of α in section 4.3.1, the square integrability of the model is guaranteed, see Carr and Madan [1999]. The standard deviation of swaption and cap price measurement errors σ_{cs} and σ_{ls} are small and significant, which implies the model accurately fits the data.

Table 5.4: Summary statistics for Libor/Swap fitting errors

Libor			Maturity				
	1mth	3mth	6mth				
	0.72	-0.63**	1.02***				
	(0.54)	(-1.72)	(3.15)				
Swap			Maturity				
	1yr	2yr	3yr	5yr	7yr	10yr	15yr
	1.77***	-2.13***	-0.67***	0.58	-0.23	-0.18	0.17
	(4.72)	(-3.09)	(-2.78)	(0.69)	(-0.37)	(-0.23)	(0.12)

The table reports the mean pricing errors for the Libor/Swap rates. The pricing errors are defined as the difference between fitted rates and actual rates (in bps), which are reported in percentages. t-statistics are in parentheses, which are calculated by Newey and West [1987] standard error with 12 lags. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 5.4 reports the mean valuation errors and associate t-statistics for Libor and Swap rates. Overall, the fitting errors of Libor and Swap rates are very small and range from -2.13 to 1.77 in basis points. The average fitted error is less than 1 basis point. These results are consistent with the results from chapter 3, which show that the variations of interest rates can be explained by the three-factor model.

Table 5.5: Summary statistics for caps fitting errors

Cap			Maturity			
	1yr	2yr	3yr	5yr	7yr	10yr
	1.13	-3.16**	2.87**	3.64***	-2.37	-1.06
	(1.02)	(-2.13)	(2.36)	(4.58)	(-1.14)	(-0.79)

The table reports the mean pricing errors for interest rate caps. The pricing errors are defined as the difference between fitted cap prices and actual cap prices divided by the actual prices, which are reported in percentages. t-statistics are in parentheses, which are calculated by Newey and West [1987] standard error with 12 lags. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Further inspection shows that the model fits 6-month Libor, 1-year and 2-year Swap rates relatively worse. This may indicate the structure difference between the Libor and Swap market, which is also documented by Longstaff, Santa-Clara, and Schwartz [2001] and Heidari and Wu [2009].

The pricing errors and associated t-statistics for interest rate caps are presented in Table 5.5. The average ATM cap errors range from -3.16% to 3.64%, which are slightly better than the results reported in Jarrow, Li, and Zhao [2007](around 5% during the same period). The pricing errors for 2yr, 3yr and 5yr are relatively large and statistically significant, which implies that these caps fit relatively worse.

Table 5.6: Summary statistics for swaptions fitting errors

Swaption			Maturity				
Tenor	1mth	3mth	6mth	1yr	2yr	3yr	5yr
1yr	-2.05***	-0.62***	1.82***	1.61***	1.37***	2.35***	0.63
	(-4.86)	(-2.67)	(5.13)	(4.88)	(6.32)	(5.28)	(1.13)
2yr	0.33	1.96*	0.47***	1.68***	2.12***	1.28**	2.63
	(0.96)	(1.78)	(4.23)	(6.78)	(9.32)	(2.01)	(0.58)
3yr	0.65	1.39	0.73***	2.43***	0.82	0.24	1.19
	(1.21)	(0.86)	(2.97)	(5.38)	(0.87)	(0.55)	(1.28)
5yr	1.85	2.13	0.40	1.72	1.26	0.36	0.74
	(0.54)	(0.68)	(0.99)	(0.86)	(0.56)	(0.93)	(0.48)
7yr	-1.29***	-0.58***	-0.93**	-1.41	-1.02	-0.83	-0.67
	(-3.76)	(-2.84)	(-1.96)	(-0.87)	(-1.23)	(-0.56)	(-0.43)
10yr	-0.29**	-0.97***	-0.02	-1.29	-0.76	-0.88	-0.83
	(-2.24)	(-3.68)	(-1.07)	(-0.86)	(-0.56)	(-0.76)	(-0.54)

The table reports the mean pricing errors for interest rate swaptions. The pricing errors are defined as the difference between fitted swaption prices and actual cap prices divided by the actual prices, which are reported in percentages. t-statistics are in parentheses, which are calculated by Newey and West [1987] standard error with 12 lags. *, **, and *** denote significance at the 10%, 5% and 1% levels, respectively.

Table 5.6 presents the average pricing errors and associated t-statistics for swaptions. The average pricing error ranges from -2.05% to 2.43%. Several pricing errors are statistically significant, which implies the model prices interest rate swaption accurately. These results are similar to Han [2007] and Trolle and Schwartz [2009a]. In addition, most of significant fitting errors are found in 1-year and 2-year tenors. These swaptions are on the 'edges' of the volatility surface, which are usually more difficult to fit. On the other hand, the model fits 5-year tenor swaptions almost perfectly. Overall, according to tables 5.4, 5.5 and 5.6, the joint framework provides a very good in-sample fitting of Libor/Swap rates and their derivatives.

5.5 Conclusion

A consistent pricing framework is proposed under the arbitrage-free Nelson-Siegel model from chapter 4. As discussed in the introduction, this method is computational efficient in terms of empirical calibration because there are only two Riccati equations must be solved in every loop of Kalman filter(fourteen Riccati equations require to be solved in the framework of Schrager and Pelser [2006] and section 4.3.3. The semi-closed solutions also provide a accurate way to calculate the Greeks for interest rate derivatives. The consistent framework is then cast into a state-space form, which is further estimated via an extended Kalman filter using the US Libor/Swap rates, ATM caps and swaptions. The results show that the proposed joint framework prices the data both accurately and efficiently.

Chapter 6

Conclusions

This thesis investigates the empirical performance of the Nelson-Siegel model in Australian government bond rates, which are similar to studies on US government bond rates, see Diebold and Li [2006]. The advanced estimation procedure, which uses a state-space model calibrated via a Kalman filter is introduced to improve the results further. As discussed in chapter 4, because the Nelson-Siegel model is not arbitrage-free, I therefore introduce the generalized arbitrage-free Nelson-Siegel model to overcome this problem. In particular, the methodology I proposed here not only makes Nelson and Siegel [1987] arbitrage-free, but also applies to the entire family of Nelson-Siegel models, including Svensson [1994], Bliss [1996], Björk and Christensen [1999] and Diebold, Piazzesi, and Rudebusch [2005]. The derivatives pricing framework based on the generalized arbitrage-free Nelson-Siegel model is also proposed in this thesis. Finally, I calibrate the arbitrage-free Nelson-Siegel model and the derivative pricing framework using US financial data, including Libor and swap rates, interest rate caps and swaptions. In conclusion, I summarize the main contributions of this thesis and then discuss potential future studies.

The three main chapters extend the Nelson-Siegel family of models both empirically and theoretically. Chapter 3 models and forecasts the government bond term structure in Australia. This empirical study uses the dynamic Nelson-Siegel model to model Australian government bond rates. The question is whether the dynamic Nelson-Siegel model can deliver a robust empirical performance in the Australian fixed-income market. I have also cast the dynamic Nelson-Siegel mode into a state-space form, which can be estimated robustly via a Kalman filter. The estimation results confirm that the dynamic Nelson-Siegel model still performs well in the Australian financial market. Moreover, by incorporating the state-space form of the Nelson-Siegel model, the out-of-sample forecasts can be further improved compared to the two-step estimation procedure in Diebold and Li [2006]. The result suggests that the dynamic Nelson-Siegel model via one-step estimation procedure should be applied in Australian financial markets. In particular, its prominent out-of-sample forecasting could provide a basis for investment decisions of the financial industry and policy decisions of Reserve Bank of Australia.

The research question that naturally arises is whether the dynamic Nelson-Siegel model can be further extended. The first possible extension is that because the dynamic Nelson-Siegel model is not arbitrage-free, Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011] present an arbitrage-free version of the Nelson-Siegel model, which incorporates the Nelson-Siegel structure into the affine term structure framework by adding an additional arbitrage-free term. However, interest rate volatility is assumed to be constant in Christensen, Diebold, and Rudebusch [2009] and Christensen, Diebold, and Rudebusch [2011], which is not an appropriate assumption. According to Li and Zhao [2006] and Trolle and Schwartz [2009a], interest rates are clearly stochastic, and stochastic volatility contains important unspanned parts. In addition, unspanned stochastic volatility is difficult to incorporate into the affine term structure model, which has been widely proved throughout the literature, see Trolle and Schwartz [2009a]. Therefore in chapter 4, I introduce a generalized arbitrage-free Nelson-Siegel model, which is derived from a more generalized framework, the HJM framework. It is widely believed that stochastic volatility can be incorporated much more easily in HJM models, see Trolle and Schwartz [2009a]. The generalized arbitrage-free Nelson-Siegel model is introduced to relax the constant volatility assumptions in Christensen, Diebold, and Rudebusch [2011]. I introduce a Markovian transformation for the additional terms, which is assumed to be locally riskless. The model can be finally classified into the Markovian HJM model class. By assuming constant volatility, the framework of Christensen, Diebold, and Rudebusch [2011] can also be derived based on my model framework as a special case.

When generalized arbitrage-free model is built, the second extension, which comprises pricing and hedging the interest rate contingent claims under the new model, arises naturally. According to Chacko and Das [2002], pricing the interest rate caps and swaptions is actually to price the bond options, both zero-coupon bonds and coupon-bearing bonds, because the interest rate caps and swaptions can be manipulated into particular bond options with specialized strike prices. The prices of interest rate contracts are derived based on Schrager and Pelser [2006] scheme. Their original approach is based on the affine term structure framework. My method on the other hand derives from the HJM framework and thus also represents an original contribution. Nonetheless, this method has two disadvantages during the estimation procedure. The first problem is that the pricing procedure for interest rate caps and swaptions is not consistent; the second drawback is that the estimation of this model is computationally inefficient. The computational efficient matters in the financial markets, especially for pricing and hedging interest rate derivatives.

Therefore, in chapter 5, I introduce a second pricing methodology based on the generalized arbitrage-free Nelson-Siegel model to overcome this problem. The model provides a consistent and efficient way to price interest rate caps and swaptions by using the transform inversion techniques in Carr and Madan [1999] and Duffie, Pan, and Singleton [2000]. In addition, the consistent framework for pricing interest rate derivatives also exploits the potential to jointly model the interest rates and their derivatives. The framework is calibrated via an extended Kalman filter using the US Libor/Swap rates, caps and swaptions. The results show that the model prices both interest rates and their derivatives accurately. This consistent framework dramatically improves the program efficiency, which is important for real time interest rate derivative trading and fixed-income portfolio management. The semi-closed form solutions also provide an accurate and efficient way to calculate Greeks for interest rate derivatives, which are vital tools in risk management.

The results found in this thesis indicate that several potential extensions warrant future research.

- **Economic meaning of no-arbitrage constraints.** The generalized arbitrage-free Nelson-Siegel model is designed to make the Nelson-Siegel model consistent with no-arbitrage pricing. However, similar to Coroneo, Nyholm, and Vidova-Koleva [2011], whether the arbitrage-free model is necessary to produce better fits and forecasts needs further research.

- **Comparison with other models based on Australian data.** Several papers also conclude that their out-of-sample forecasts are better than the Random Walk benchmark, for instance Bowsher and Meeks [2008], Chua, Foster, Ramaswamy, and Stine [2008]. The comparison of the empirical performances of these models using Australian data shows potential for Australian financial markets.

- **Risk-neutral dynamics of the generalized arbitrage-free Nelson-Siegel model.** Given the closed-form risk-neutral dynamics of the forward rate in the generalised arbitrage-free Nelson-Siegel model, the more complicated interest rate derivatives can be derived by Monte Carlo simulations, see Longstaff and Schwartz [2001].

- **Extended to commodity futures and options pricing.** Similar to Trolle and Schwartz [2009b], the framework of the generalised arbitrage-free Nelson-Siegel model can be applied to value derivatives of other assets, such as commodity futures and options.

References

ALMEIDA, C., J. J. GRAVELINE, AND S. JOSLIN (2011): "Do interest rate options contain information about excess returns?," *Journal of Econometrics*, 164(1), 35–44, Annals Issue on Forecasting.

ANG, A., AND M. PIAZZESI (2003): "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics*, 50(4), 745–787.

BALDUZZI, P., S. R. DAS, AND S. FORESI (1998): "The central tendency: A second factor in bond yields," *The Review of Economics and Statistics*, 80(1), 62–72.

BALDUZZI, P., S. R. DAS, S. FORESI, AND R. K. SUNDARAM (1996): "A simple approach to three-factor affine term structure models," *The Journal of Fixed Income*, 6(3), 43–53.

BALI, T. (2003): "Modeling the stochastic behavior of short-term interest rates: Pricing implications for discount bonds," *Journal of banking and finance*, 27(2), 201–228.

Bank for International Settlements, . (2005): "Zero-coupon yild curves: Technical documentation," *BIS Papers*, 25.

Berardi, A., and M. Esposito (1999): "A base model for multifactor specifications of the term structure," *Economic Notes*, 28(2), 145.

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

Björk, T., and L. Svensson (2001): "On the existence of Finite-Dimensional realizations for nonlinear forward rate models," *Mathematical Finance*, 11(2), 205–243.

Black, F. (1976): "The pricing of commodity contracts," *Journal of financial economics*, 3(1), 167–179.

Black, F., E. Derman, and W. Toy (1990): "A one-factor model of interest rates and tts application to Treasury bond options.," *Financial Analysts Journal*, 46(1), 33–39.

Bliss, R. (1996): *Testing term structure estimation methods.* Federal Reserve Bank of Atlanta.

Bowsher, C. G., and R. Meeks (2008): "The dynamics of economic functions: Modeling and forecasting the yield curve," *Journal of the American Statistical Association*, 103(484), 1419–1437.

Brace, A., D. Gatarek, and M. Musiela (1997): "The market model of interest rate dynamics," *Mathematical Finance*, 7(2), 127–147.

BRACE, A., AND M. MUSIELA (1995): "Swap derivatives in a Gaussian HJM framework," *Working Paper.*

BRIGO, D., AND F. MERCURIO (2007): *Interest rate models-theory and practice: with smile, inflation and credit.* Springer.

BROWN, R., AND S. SCHAEFER (1994): "The term structure of real interest rates and the Cox, Ingersoll, and Ross model," *Journal of Financial Economics*, 35(1), 3–42.

CARR, P., AND D. MADAN (1999): "Option valuation using the fast Fourier transform," *Journal of Computational Finance*, 2(4), 61–73.

CARVERHILL, A., AND K. PANG (1995): "Efficient and flexible bond option valuation in the Heath, Jarrow and Morton framework," *Journal of Fixed Income*, 5(2), 70–77.

CHACKO, G., AND S. DAS (2002): "Pricing interest rate derivatives: A general approach," *The Review of Financial Studies*, 15(1), 195–241.

CHAN, K. C., G. A. KAROLYI, F. A. LONGSTAFF, AND A. B. SANDERS (1992): "An empirical comparison of alternative models of the short-term interest rate," *The Journal of Finance*, 47(3), 1209–1227.

CHEN, L. (1996): *Stochastic Mean and Stochastic Volatility: A Three-factor Model of the Term Structure of Interest Rates and Its Applications in Derivatives Pricing and Risk Management.* Blackwell publishers.

CHEN, R., AND L. SCOTT (1992): "Pricing interest rate options in a two-factor

Cox-Ingersoll-Ross model of the term structure," *Review of Financial Studies*, 5(4), 613–636.

——— (1993): "Maximum likelihood estimation for a multifactor equilibrium model of the term structure of interest rates," *The Journal of Fixed Income*, 3(3), 14–31.

CHERIDITO, P., D. FILIPOVIĆ, AND R. KIMMEL (2007): "Market price of risk specifications for affine models: Theory and evidence," *Journal of Financial Economics*, 83(1), 123–170.

CHIARELLA, C., AND O. KWON (2000): *A class of Heath-Jarrow-Morton term structure models with stochastic volatility.* Citeseer.

——— (2001): "Forward rate dependent Markovian transformations of the Heath-Jarrow-Morton term structure model," *Finance and Stochastics*, 5, 237–257.

——— (2003): "Finite dimensional affine realisations of HJM models in terms of forward rates and yields," *Review of Derivatives Research*, 6(2), 129–155.

CHIB, S., F. NARDARI, AND N. SHEPHARD (2002): "Markov chain Monte Carlo methods for stochastic volatility models," *Journal of Econometrics*, 108(2), 281–316.

CHRISTENSEN, J., J. LOPEZ, AND G. RUDEBUSCH (2010): "Stochastic volatility in arbitrage-free Nelson-Siegel models of the term structure," *working paper.*

CHRISTENSEN, J. H., F. X. DIEBOLD, AND G. D. RUDEBUSCH (2009): "An arbitrage-free generalized Nelson-Siegel term structure model," *The Econometrics Journal*, 12(32), 33–64.

——— (2011): "The affine arbitrage-free class of Nelson-Siegel term structure models," *Journal of Econometrics*, 164(1), 4–20, Annals Issue on Forecasting.

Chua, C. T., D. Foster, K. Ramaswamy, and R. Stine (2008): "A dynamic model for the forward curve," *Review of Financial Studies*, 21(1), 265–310.

Collin-Dufresne, P., and R. S. Goldstein (2002): "Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility," *The Journal of Finance*, 57(4), 1685–1730.

——— (2003): "Generalizing the affine framework to HJM and random field models," *Available at SSRN*.

Coroneo, L., K. Nyholm, and R. Vidova-Koleva (2011): "How arbitrage-free is the Nelson–Siegel model?," *Journal of Empirical Finance*, 18(3), 393–407.

Cox, J. C., J. E. Ingersoll, and S. A. Ross (1985): "A theory of the term structure of interest rates," *Econometrica*, 53(2), 385–407.

Dai, Q., and K. J. Singleton (2000): "Specification analysis of affine term structure models," *Journal of Finance*, 55(36), 1943–1978.

——— (2002): "Expectation puzzles, time-varying risk premia, and affine models of the term structure," *Journal of Financial Economics*, 63(3), 415–441.

DE POOTER, M. (2007): "Examining the Nelson-Siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance," *Available at SSRN.*

DEMUNNIK, J. (1992): *The valuation of interest rate derivative securities.* Amsterdam: Tinbergen Researh Institute.

DEVIN, S., B. HANZON, AND T. RIBARITS (2010): "A finite-dimensional HJM model: How important is arbitrage-free evolution?," *International Journal of Theoretical and Applied Finance*, 13(08), 1241–1263.

DIEBOLD, F. X., AND C. LI (2006): "Forecasting the term structure of government bond yields," *Journal of Econometrics*, 130(2), 337–364.

DIEBOLD, F. X., C. LI, AND V. Z. YUE (2008): "Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach," *Journal of Econometrics*, 146(2), 351–363.

DIEBOLD, F. X., AND R. MARIANO (1995): "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13(3), 253–263.

DIEBOLD, F. X., M. PIAZZESI, AND G. D. RUDEBUSCH (2005): "Modeling bond yields in finance and macroeconomics," *American Economic Review*, 95(2), 415–420.

DIEBOLD, F. X., AND G. D. RUDEBUSCH (2013): *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach.* Princeton University Press.

DIEBOLD, F. X., G. D. RUDEBUSCH, AND S. B. ARUOBA (2006): "The macroeconomy and the yield curve: A dynamic latent factor approach," *Journal of Econometrics*, 131(1-2), 309–338.

DOTHAN, L. (1978): "On the term structure of interest rates," *Journal of Financial Economics*, 6(1), 59–69.

DUAN, J.-C., AND J.-G. SIMONATO (1993): "Multiplicity of solutions in maximum likelihood factor analysis," *Journal of Statistical Computation and Simulation*, 47(1-2), 37–47.

——— (1999): "Estimating and testing exponential-affine term structure models by Kalman filter," *Review of Quantitative Finance and Accounting*, 13, 111–135.

DUFFEE, G. R. (2002): "Term premia and interest rate forecasts in affine models," *The Journal of Finance*, 57(39), 405–443.

DUFFEE, G. R., AND R. H. STANTON (2004): "Estimation of dynamic term structure models," *Working Paper.*

DUFFIE, D., AND R. KAN (1996): "A yield factor model of interest rates," *Mathematical Finance*, 6(4), 379–406.

DUFFIE, D., J. PAN, AND K. SINGLETON (2000): "Transform analysis and asset pricing for affine Jump-Diffusions," *Econometrica*, 68(6), 1343–1376.

DUFFIE, D., AND K. J. SINGLETON (1997): "An econometric model of the term structure of Interest-Rate swap yields," *The Journal of Finance*, 52(4), 1287–1321.

El Karoui, N., and V. Lacoste (1992): "Multifactor analysis of the term structure of interest rates," *Proceedings, AFFI.*

Fama, E. F., and R. R. Bliss (1987): "The information in long-maturity forward rates," *American Economic Review*, 77(4), 680–692.

Filipovic, D. (1999): "A note on the Nelson-Siegel family," *Mathematical Finance*, 9(4), 349–359.

——— (2000): "Exponential-polynomial families and the term structure of interest rates," *Bernoulli*, 6(6), 1081–1107.

Fong, H. G., and O. A. Vasicek (1991): "Fixed-income volatility management," *The Journal of Portfolio Management*, 17(4), 41–46.

Frachot, A. (1995): "Factor models of domestic and foreign interest rates with stochastic volatilities," *Mathematical Finance*, 5(2), 167–185.

Frachot, A., D. Janci, and V. Lacoste (1992): *Factor analysis of the term structure: a probabilistic approach.* Banque de France.

Han, B. (2007): "Stochastic volatilities and correlations of bond yields," *The Journal of Finance*, 62(3), 1491–1524.

Harvey, A. (1989): *Forecasting, structural time series models and the Kalman filter.* Cambridge university press.

Heath, D., R. Jarrow, and A. Morton (1992): "Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation," *Econometrica*, 60(1), 77–105.

HEIDARI, M., AND L. WU (2009): "A joint framework for consistently pricing interest rates and interest rate derivatives," *Journal of Financial and Quantitative Analysis*, 44(03), 517–550.

HESTON, S. L. (1993): "A Closed-Form solution for options with stochastic volatility with applications to bond and currency options," *The Review of Financial Studies*, 6(2), 327–343.

HO, T. S. Y., AND S. B. LEE (1986): "Term structure movements and pricing interest rate contingent claims," *Journal of Finance*, 41(5), 1011–1029.

HULL, J. (2009): *Options, futures, and other derivatives.* Pearson.

HULL, J., AND A. WHITE (1990): "Pricing interest-rate-derivative securities," *Review of Financial Studies*, 3(4), 573–592.

——— (1994): "Branching out," *Risk*, 7(7), 34–37.

——— (1995): "A note on the models of Hull and White for pricing options on the term structure," *The Journal of Fixed Income*, 5(2), 97–102.

JARROW, R., H. LI, AND F. ZHAO (2007): "Interest rate caps "Smile" too! But can the LIBOR barket bodels capture the smile?," *The Journal of Finance*, 62(1), 345–382.

JEFFREY, A. (1995): "Single factor Heath-Jarrow-Morton term structure models based on Markov spot interest rate dynamics," *Journal of Financial and Quantitative Analysis*, 30(4), 619–642.

KOOPMAN, S. J., M. I. P. MALLEE, AND M. VAN DER WEL (2010): "Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters," *Journal of Business and Economic Statistics*, 28(3), 329–343.

KRIPPNER, L. (2006): "A theoretically consistent version of the Nelson and Siegel class of yield curve models," *Applied Mathematical Finance*, 13(1), 39–59.

——— (2008): "A macroeconomic foundation for the Nelson and Siegel class of yield curve models," Discussion paper, Reserve Bank of New Zealand.

——— (2009): "A theoretical foundation for the Nelson and Siegel class of yield curve models," Reserve Bank of New Zealand Discussion Paper Series DP2009/10, Reserve Bank of New Zealand.

——— (2012): "A theoretical foundation for the Nelson and Siegel class of yield curve models," Discussion paper, Centre for Applied Macroeconomic Analysis.

KRIPPNER, L., AND L. A. THORSRUD (2009): "Forecasting New Zealand's economic growth using yield curve information," Discussion paper, Reserve Bank of New Zealand.

LANGETIEG, T. C. (1980): "A multivariate model of the term structure," *Journal of Finance*, 35(1), 71–97.

LEE, R. (2004): "Option pricing by transform methods: Extensions, unification, and error control," *Journal of Computational Finance*, 7(3), 51–86.

LI, H., AND F. ZHAO (2006): "Unspanned stochastic volatility: Evidence from hedging interest rate derivatives," *The Journal of Finance*, 61(1), 341–378.

LITTERMAN, R. B., AND J. SCHEINKMAN (1991): "Common factors affecting bond returns," *Journal of Fixed Income*, 1(1), 54–61.

LIU, J., J. PAN, AND L. PEDERSEN (1999): "Density-based inference in affine jump-diffusions," Discussion paper, Tech. report, Graduate School of Business, Stanford University.

LJUNG, L. (1979): "Asymptotic behavior of the extended Kalman filter as a parameter estimator for linear systems," *Automatic Control, IEEE Transactions on*, 24(1), 36–50.

LONGSTAFF, F., P. SANTA-CLARA, AND E. SCHWARTZ (2001): "The relative valuation of caps and swaptions: Theory and empirical evidence," *The Journal of Finance*, 56(6), 2067–2109.

LONGSTAFF, F., AND E. SCHWARTZ (1993): "Interest rate volatility and bond prices," *Financial Analysts Journal*, 49(4), 70–74.

——— (2001): "Valuing American options by simulation: A simple least-squares approach," *Review of Financial studies*, 14(1), 113–147.

LUO, X., H. HAN, AND J. E. ZHANG (2012): "Forecasting the term structure of Chinese Treasury yields," *Pacific-Basin Finance Journal*, 20(5), 639–659.

MERTON, R. C. (1974): "On the pricing of corporate debt: The risk structure of interest rates," *The Journal of Finance*, 29(2), 449–470.

MILTERSEN, K., AND S. PERSSON (1999): "Pricing rate of return guarantees in a Heath–Jarrow–Morton framework," *Insurance: Mathematics and Economics*, 25(3), 307–325.

MILTERSEN, K. R. (1994): "An arbitrage theory of the term structure of interest rates," *The Annals of Applied Probability*, 4(4), 953–967.

MUNK, C. (1999): "Stochastic duration and fast coupon bond option pricing in multi-factor models," *Review of Derivatives Research*, 3(2), 157–181.

NELSON, C. R., AND A. F. SIEGEL (1987): "Parsimonious modeling of yield curves," *Journal of Business*, 60(4), 473–489.

NEWEY, W., AND K. WEST (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix," *Econometrica*, 55(3), 703–708.

RENDLEMAN, R., AND B. BARTTER (1980): "The pricing of options on debt securities," *Journal of Financial and Quantitative Analysis*, 15(1), 11–24.

RICHARD, S. (1978): "An arbitrage model of the term structure of interest rates," *Journal of Financial Economics*, 6(1), 33–57.

RUTKOWSKI, M. (1996): "Valuation and hedging of contingent claims in the HJM model with deterministic volatilities," *Applied Mathematical Finance*, 3(3), 237–267.

SCHRAGER, D. F., AND A. A. J. PELSER (2006): "Pricing swaptions and coupon bond options in affine term structure models," *Mathematical Finance*, 16(4), 673–694.

SHAREF, E., AND D. FILIPOVIC (2004): "Conditions for consistent exponential-polynomial forward rate processes with multiple nontrivial factors," *International Journal of Theoretical and Applied Finance*, 7(06), 685–700.

SHREVE, S. E. (2004): *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer Finance. Springer.

SINGLETON, K. (2001): "Estimation of affine asset pricing models using the empirical characteristic function," *Journal of Econometrics*, 102(1), 111–141.

SINGLETON, K. J., AND L. UMANTSEV (2002): "Prcing coupon-bond options and swaptions in affine term structure models," *Mathematical Finance*, 12(4), 427–446.

STRICKLAND, C. (1996): "A comparison of diffusion models of the term structure," *The European Journal of Finance*, 2(1), 103–123.

SVENSSON, L. E. (1994): "Estimating and interpreting forward interest rates: Sweden 1992 - 1994," NBER Working Papers 4871, National Bureau of Economic Research, Inc.

TROLLE, A., AND E. SCHWARTZ (2009a): "A general stochastic volatility model for the pricing of interest rate derivatives," *Review of Financial Studies*, 22(5), 2007–2057.

——— (2009b): "Unspanned stochastic volatility and the pricing of commodity derivatives," *Review of Financial Studies*, 22(11), 4423–4461.

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

VICENTE, J., AND B. M. TABAK (2008): "Forecasting bond yields in the Brazilian fixed income market," *International Journal of Forecasting*, 24(3), 490–497.

YU, W.-C., AND D. M. SALYARDS (2009): "Parsimonious modeling and forecasting of corporate yield curve," *Journal of Forecasting*, 28(1), 73–88.