A No-Arbitrage Macro-Finance Approach to the Term Structure of Interest Rates

by

Phumza Thafeni

Thesis presented in partial fulfilment of the requirements for the degree of Master of Science in Mathematics in the Faculty of Science at Stellenbosch University

Department of Mathematical Sciences,
Mathematics Division,
University of Stellenbosch,
Private Bag X1, Matieland 7602, South Africa.

Supervisor: Dr. R. Ghomrasni

March 2015

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: ....
P. Thafeni

Date: 2014/08/28

Abstract

This work analysis the main macro-finance models of the term structure of interest rates that determines the joint dynamics of the term structure and the macroeconomic fundamentals under a no-arbitrage approach. There has been a long search during the past decades of trying to study the relationship between the term structure of interest rates and the economy, to the extent that much of recent research has combined elements of finance, monetary economics, and the macroeconomics to analyse the term structure.

The central interest of the thesis is based on two important notions. Firstly, it is picking up from the important work of Ang and Piazzesi (2003) model who suggested a joint macro-finance strategy in a discrete time affine setting, by also imposing the classical Taylor (1993) rule to determine the association between yields and macroeconomic variables through monetary policy. There is a strong intuition from the Taylor rule literature that suggests that such macroeconomic variables as inflation and real activity should matter for the interest rate, which is the monetary policy instrument. Since from this important framework, no-arbitrage macro-finance approach to the term structure of interest rates has become an active field of cross-disciplinary research between financial economics and macroeconomics.

Secondly, the importance of forecasting the yield curve using the variations on the Nelson and Siegel (1987) exponential components framework to capture the dynamics of the entire yield curve into three dimensional parameters evolving dynamically. Nelson-Siegel approach is a convenient and parsimonious approximation method which has been trusted to work best for fitting and forecasting the yield curve. The work that has caught quite much of interest under this framework is the generalized arbitrage-free Nelson-Siegel macro-finance term structure model with macroeconomic fundamentals, (Li *et al.* (2012)), that characterises the joint dynamic interaction between yields and the macroeconomy and the dynamic relationship between bond risk-premia and the economy. According to Li *et al.* (2012), risk-premia is found to be closely linked to macroeconomic activities and its variations can be analysed. The approach improves the estimation and the challenges on identification of risk parameters that has been faced in recent macro-finance literature.

Keywords: Yield curve, term-structure of interest rates, macroeconomic fundamentals and financial factors.

Opsomming

Hierdie werk ontleed die makro-finansiese modelle van die term struktuur van rentekoers pryse wat die gesamentlike dinamika bepaal van die term struktuur en die makroekonomiese fundamentele faktore in 'n geen arbitrage wêreld. Daar was 'n lang gesoek in afgelope dekades gewees wat probeer om die verhouding tussen die term struktuur van rentekoerse en die ekonomie te bestudeer, tot die gevolg dat baie onlangse navorsing elemente van finansies, monetêre ekonomie en die makroekonomie gekombineer het om die term struktuur te analiseer.

Die sentrale belang van hierdie proefskrif is gebaseer op twee belangrike begrippe. Eerstens, dit tel op by die belangrike werk van die Ang and Piazzesi (2003) model wat 'n gesamentlike makro-finansiering strategie voorstel in 'n diskrete tyd affiene ligging, deur ook die klassieke Taylor (1993) reël om assosiasie te bepaal tussen opbrengste en makroekonomiese veranderlikes deur middel van monetêre beleid te imposeer. Daar is 'n sterk aanvoeling van die Taylor reël literatuur wat daarop dui dat sodanige makroekonomiese veranderlikes soos inflasie en die werklike aktiwiteit moet saak maak vir die rentekoers, wat die monetêre beleid instrument is. Sedert hierdie belangrike raamwerk, het geen-arbitrage makro-finansies benadering tot term struktuur van rentekoerse 'n aktiewe gebied van kruis-dissiplinêre navorsing tussen finansiële ekonomie en makroekonomie geword.

Tweedens, die belangrikheid van voorspelling van opbrengskromme met behulp van variasies op die Nelson and Siegel (1987) eksponensiële komponente raamwerk om dinamika van die hele opbrengskromme te vang in drie dimensionele parameters wat dinamies ontwikkel. Die Nelson-Siegel benadering is 'n gerieflike en spaarsamige benaderingsmetode wat reeds vertrou word om die beste pas te bewerkstellig en voorspelling van die opbrengskromme. Die werk wat nogal baie belangstelling ontvang het onder hierdie raamwerk is die algemene arbitrage-vrye Nelson-Siegel makro-finansiele term struktuur model met makroekonomiese grondbeginsels, (Li *et al.* (2012)), wat kenmerkend van die gesamentlike dinamiese interaksie tussen die opbrengs en die makroekonomie en die dinamiese verhouding tussen band risiko-premies en die ekonomie is. Volgens Li *et al.* (2012), word risiko-premies bevind om nou gekoppel te wees aan makroekonomiese aktiwiteite en wat se variasies ontleed kan word. Die benadering verbeter die skatting en die uitdagings van identifisering van risiko parameters wat teegekom is in die afgelope makro-finansiese literatuur.

Sleutelwoorde: Opbrengskromme, termyn-struktuur van rentekoerse, makroekonomiese grondbeginsels en finansiële faktore.

Acknowledgements

I would like to express my heartfelt gratitude to Dr. Raouf Ghomrasni for his unwavering commitment and supervision towards my master thesis. The journey was not easy and this dissertation would not have been possible without his guidance and persistent help.

I am grateful to all of my dear friends which are innumerous to mention, who helped me to move a little bit further on this thesis, so that I could see a little bit higher. They provided me with a strong love shield that always surrounds me and never let my sadness enter inside. Your friendship gave my life a wonderful experience.

I wish to also thank the African Institute for Mathematical Science's family, for providing facilities and atmosphere that is incredibly collaborative to conduct this work. I feel very fortunate that I was part of the community. Your mentorship throughout the years I have spent at AIMS, gave me an invaluable tool to have as my career moves forward.

To my parents, I am so thankful for your endless love, support and encouragement, you were always giving me that little hope that all things are possible through Christ who strengthens us. In the same spirit, I would also like to extend my gratitude to my entire family as a whole and all my friends in Christ for their prayers, inspiring motivations and support.

Lastly, all the glory be to God, who made this possible for me.

Dedication

To the most precious gift I have on earth, my parents,
Raymond and Sinah,
your unconditional love and support, have motivated me to set higher targets.

Contents

List of Figures

List of Tables

Chapter 1

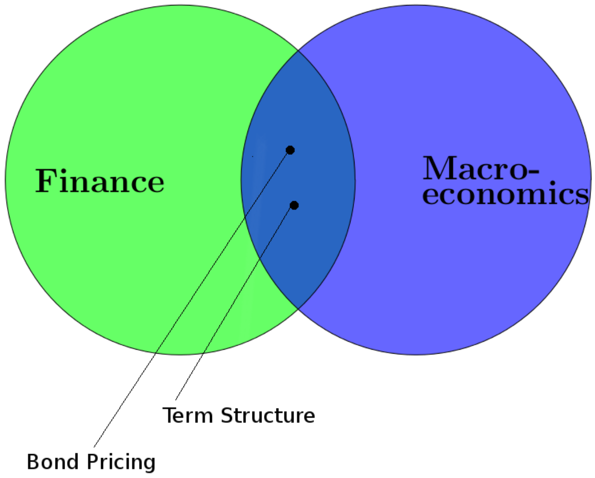
Introduction

The relation of the short-term to long-term interest rate yields has often been an extremely significant area of interest that has been deeply analysed by macroeconomists, financial economists and composes an issue of special relevance for practitioners in financial markets, as it reflects expectations of market participants about future changes in interest rate. This relationship usually referred to as the term structure of interest rates, and plays a major role on the analysis of different aspects of the economy. An understanding of the stochastic behaviour of term structure of interest rates in the light of macroeconomic evolution is important for understanding the transmission mechanism of monetary policy. There has been a long-standing separation between macro and finance literature in the modelling of the term structure of interest rates until a joint macro-finance model introduced by Ang and Piazzesi (2003), which examines the relationship between the term structure of interest rates and the economy. The joint macro-finance approach has since then become an active field of cross-disciplinary research between the two disciplines.

In the canonical finance models, the short-term interest rates is expressed as a linear function of few latent factors of the yield curve, but with no fundamental economic interpretation of the underlying latent factors. The long-term interest rates are attributable to these latent factors, and the movement in long-term yields reflects the changes in risk premiums, which also depend on these latent factors. Contrary to macro literature, the short-term interest rates is controlled by the central banks through several mechanisms to realize a powerful monetary policy. Long term interest rates are determined by the expectations of future short rates. Since short rates are determined by expectations of the macro variables and changes in risk premiums are often ignored in this case, then the expectations hypothesis of the term structure holds.

The finance and macroeconomic literature have developed to a remarkable extent of mutual interest that connects the two literatures through the short term interest rates. From a finance perspective, the short-term interest rate is a fundamental building block for yields of other maturities, because long-term yields are risk adjusted averages of expected future short rates. From a macro perspective, the short-term interest rate is a key policy instrument that is controlled by the central bank through several mechanisms to realize a powerful monetary policy. Monetary authority decides how to conduct the monetary policy by adjusting the rates to achieve the goals of economic stabilization. Connecting the edges, a joint macro-finance strategy would suggest that understanding the way central banks move the short rate in response to fundamental macroeconomic shocks, will provide the most extensive understanding of the movements in the short end of the term structure of interest rates with the help of macroeconomic variables. Moreover, the consistency between short and long-term rates is enforced by no-arbitrage assumption.

Figure 1.1: Macro-finance representation



In modern finance, different methods and techniques have been developed with the aim of determining a fair price of any tradable asset. As seen in the groundbreaking work of Black and Scholes (1973), Merton (1973), that gives a theoretical estimate of the price of option over time, which includes risk neutral and martingale pricing, see Harrison and Pliska (1983), under arbitrage-free assumptions. Not only did this specify the first successful options pricing formula, but it also described a general framework for pricing other derivative instruments such as interest rate derivatives. Constructing a model of the term structure of interest rates which combines macroeconomics and finance models will enable us to do bond pricing, which is viewed as a special case of asset pricing. The main focus in this research is on the modelling of the term structure of interest rates using recently developed no-arbitrage macro-finance models and also, to compare the variety of joint models of the term structure and the macroeconomy to study the impact of different macroeconomic variables and how they drive expected future interest rates under discrete time setting.

1.1 Macro-Finance Background Theory

The term structure of interest rates has long been of fundamental importance to analyse different aspects of the economy. Modelling the yield curve has been essential to researchers, that is to achieve goodness of fit, forecasting the economy and derivative pricing. Empirical studies reveal that more than 95% of the movements of various treasury bond yields are captured by three factors, as shown in the seminal paper of Litterman and Scheinkman (1991). Extracting these factors rely on the yield curve information that can be classified in two categories, the no-arbitrage approach based on financial theory and exploiting the cross-equation restrictions of bond pricing. The other approach is through statistical interpolation such as the Nelson and Siegel (1987) interpolation, regarded as reduced-form model. Different methods and techniques have been developed with the aim of modelling the dynamics of the yield curve.

Macroeconomics have developed term structure models to determine the relationship between interest rates, monetary policy and macroeconomic fundamentals without any considerations about the absence of arbitrage and based on expectation hypothesis. The aim was on the effect of macroeconomic shocks on the short rate of the yield curve, where the set of yields are estimated using unrestricted vector autoregressive (VAR) estimation, see Evans and Marshall (2007), Amisano and Giannini (1997). This is done by combining macroeconomic variables (inflation and economic activity indicator) and the bond yields of different maturities in a standard VAR process to estimate the exogenous impulses to monetary policy effect on bond yields of various maturities. The results are based on the type of yields chosen.

Short-rate plays a core building block in most of the empirical studies found in monetary policy. Amongst those we have Hamilton and Jordá (2002), where in most related papers short rate process is estimated using data on short rates.

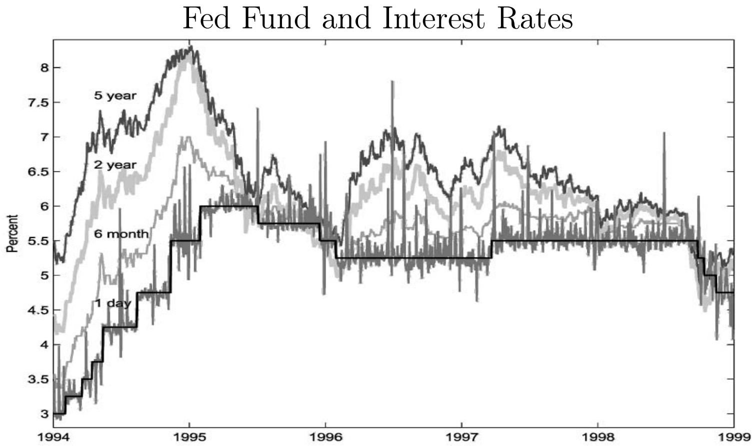


Figure 1.2: Daily data on target (step function), federal funds rate (one-day), LIBOR (six month), and swap yields (two and five years). 1994-1999

According to expectations theory, long term rates are related to short term rates through market expectations of future short rates, see Rudebusch (1995), which cannot match the long end of the yield curve, because the estimation involves only short-end data. In addition, there is strong evidence against the expectations hypothesis that "the expectations hypothesis that long yields are the average of future expected short yields", see Fama and Bliss (1987); Campbell and Shiller (1991). To compute the expected federal fund target, most literatures, for example Kuttner (2001) and others, make use of federal funds futures data and again the expectations hypothesis, which has been the most cited theory to explain the determinants of the shape of the term structure of interest rates.

The disadvantages of the macro-models are that they require a large number of coefficients to be estimated when dealing with a broad range of yields maturities and did not rule out theoretical inconsistencies due to the presence of arbitrage opportunities along the yield curve. However, as pointed out by Duffee (2002), term structure models that take on arbitrage opportunities do not say much about dynamics of interest rates and they are mainly focused on fitting the curve at one point in time and are not suitable for forecasting. These macroeconomic models could not find the relation between term structure dynamics and macroeconomy.

In the case of the empirical finance literature, the no-arbitrage approach is a centre of interest for the analysis of yield curves and it provides tractability and consistency in bond pricing. Financial economists have mainly focused on forecasting the yield curves and pricing interest rate related securities using the affine term structure models (ATSMs) which provides a good fit, based on the assumptions of no-arbitrage framework. In this research, yields are modeled as a linear function which is explained by few unobservable factors under no-arbitrage assumptions that enforces the consistency of evolution of yields over time with cross-section of yields of different maturities, as discussed in Duffie and Kan (1996) and Dai and Singleton (2000).

Popular affine no-arbitrage models provide useful statistical descriptions of term structure dynamics. However, the underlying factors on bond yields have no direct economic nature and do not provide any information on macroeconomic forces behind the movement of the yield curve. The pioneers of this literature are stemming from Vasicek (1977) and Cox *et al.* (1985), who introduced an analytically tractable special class of term structure models derived in continuous time under no-arbitrage approach. In this approach, dynamics of short-term interest rates are taken as underlying state factors.

Affine models were then popularised by Duffie and Kan (1996), on derivation of an unobservable three factor ATSM, whose formalization encompasses earlier ATSMs. Dai and Singleton (2000) made a thorough specification analysis of the unobservable three factor ATSM. Specifically, empirical studies reveal, by use of principal component analysis, that some of the movements of bond yields are captured by three state factors, which are often called level, slope and curvature as shown in the seminal paper by Litterman and Scheinkman (1991), but typically left unspecified the relationship between the term structure and the fundamental macroeconomic variables.

To bridge the gap between the two literatures, the seminal work by Ang and Piazzesi (2003), suggests a discrete-time affine model, and impose a no-arbitrage restrictions on a classical Taylor rule VAR model. Taylor rule specification implies that central banks aim at stabilising inflation around its target level and output around its potential, by adjusting short-term interest rates in response to movements in inflation and real activity. Ang and Piazzesi (2003) work is attempting to find the connection between macroeconomic and the financial description of the term structure of interest rates to improve the model performance. The analysis, however, introduced a special case of the Gaussian affine set-up corresponding to a discrete time multi-dimensional Vasicek model. The use of the discrete-time framework has become an active field of cross-disciplinary research between finance and macroeconomics in macro-finance literature over the past decade.

For forecasting purpose affine models turn out to show a poor empirical performance, as shown by Duffee (2002). This is further confirmed in the work of Hamilton and Wu (2012), claiming that, because of the tremendous numerical difficulties encountered in estimating the necessary parameters of the affine term structure models in the literature, has led to the establishment that, three popular parametrizations of ATSMs are unidentified and propose to estimate the reduced form representation of the Gaussian affine term structure models.

Alternatively, the reduced-form statistical approach, is popular among markets and central bank practitioners, namely the three factor interpolation Nelson and Siegel (1987) model. It is an empirically motivated approach due to the relative parsimony, its goodness of fit of their model which is usually given in discrete-time settings. According to the dynamic extensions made by Diebold and Li (2006), the main findings are that this dynamic Nelson-Siegel extension (denoted DNS) exhibits good empirical performance and the model factors are extracted from principal component analysis but lacks a solid theatrical foundation.

Recently, Christensen *et al.* (2011) introduced a merged, affine arbitrage-free Nelson-Siegel (AFNS) model, in discrete time, by combining the affine arbitrage-free assumption and Nelson-Siegel interpolation to address the empirical problem of ATSMs. The no-arbitrage restriction on the reduced-form DNS model impose additional constant term to the yield equations in the AFNS model. However, even though the existence of the solutions for both AFNS models and the more general ATSMs is successfully proven, the uniqueness of the solution is left out. To derive both the existence and uniqueness of solution, Linlin and Gengming (2012) developed a simple and fast procedure for estimating the AFNS model with reduced dimension optimization and a multi-step embedded regression.

1.2 Motivation on Macro-Finance Theory

The interaction between the science of macroeconomics and the science of macro-finance in finance under the subject of the term structure of interest rates is a relatively new branch of research that combines models of the term structure of interest rates in financial literature with simple macroeconomic variables or models thereof. Little attention has been paid on the relationship between macro and finance literature, until since early 2000s where economists tried to find the interaction between the dynamics of the term structure with the help of macroeconomic variables. The important work of Ang and Piazzesi (2003) showed some success, and macro-finance approach to model the term structure of interest rates has under no-arbitrage setting become an active field of cross-disciplinary research between financial economics and macroeconomics.

The main motivation of studying such model-structure is to understand both the influence of macroeconomic variables on yield curve and the information value of yield curve with respect to the macroeconomy. Imposition of no-arbitrage assumptions on macro-finance models add macroeconomic variables to the vector of unobservable yield factors. The term structure is then fitted at a point in time to eliminate the arbitrage possibilities along yield curve. Term structure modelling gives the most valuable information to forecasters and policy makers. The information reflects expectations of market participants about future changes in interest rate which in turn help determine what actually happens in the future.

1.3 Thesis Structure

Thesis is organized as follows: Chapter 2 first reviews the preliminaries of various interest rates related modelling approaches of the yield curve. Chapter 3 provides the most popular approaches for term structure modelling, reduced-form models and no-arbitrage approach, where the types of these two approaches are presented to give thorough explanations about the yield curve. A detailed description about the newly developed standard affine-Gaussian setup of macro-finance models with other existing related frameworks and the original macro-finance models of the term structure is provided in Chapter 4. Chapter 5 proceeds to discuss a model by Li *et al.* (2012), which is divided into two sub-classes: the unspanned versus spanned model, that describes the interaction between yield factors and macroeconomic factors, depending on whether the macro factors span the yield curve. The unspanned model will also be favoured as it enjoys more parsimony of the Nelson-Siegel model. Chapter 6 concludes.

Chapter 2

Yield Curve Preliminaries

Numerous statistical and financial economist literatures attempt to address the important conceptual, descriptive and theoretical considerations regarding the information conveyed on nominal government bond yield curves by elaborating on particular approaches to yield curve modelling. Conceptually, this chapter focuses on what is being measured and the simple way of understanding the behaviour of bond yields at different maturities over time. Descriptively, how do yield curves tend to behave over time. Furthermore, can a simple yet accurate dynamic characterizations and forecasting be obtained. Theoretically, what governs and restricts the shape and the behaviour of the yield curve. Moreover, the relationship between the yield curve, macroeconomic fundamentals and the central bank behaviour. The first section of this chapter traces back the origin of the yield curve modelling framework and the rest of the chapter touches upon the basic bond pricing terminology that will be later used when valuing yield curves.

2.1 Specifications of Interest Rate Curves.

The yield curve is described as the representation of the relationship between interest rates and different maturities, which is known to be the term structure of interest rates. There are numerous representations of the term structure of interest rates, but this section elaborates more on three key theoretical bond market constructs and the relationships between them, notably, the pure discount curve, forward rate curve, and the yield curve. The three representations convey the same information, however, in practice, none of these representations are directly observable, instead, they are estimated from observed bond prices.

Let the price of a pure discount bond with maturity τ, be denoted by $P_t(\tau)$ for $\tau \in \{\tau_0, \tau_1, \tau_2, ..., \tau_N\}$, as a function of the current time t that pays 1 unit at maturity τ. If $y_t(\tau)$ is a continuously compounded rate of return anticipated on a bond up to maturity, then the price of a bond at time t is given in terms of yield to maturity as

$$P_t(\tau) = e^{-\tau y_t(\tau)}. \tag{2.1.1}$$

Hence, the discount function shows a one-to-one correspondence between bond price and its yield to maturity. If a bond's price is known, or assuming we are holding the term constant, the bond's yield can easily be computed. The yield curve is formally denoted as a function that presents the relationship between the yields and their maturities at a fixed point in time. This relation is determined by

$$y_t(\tau) = -\frac{1}{\tau} \ln P_t(\tau). \tag{2.1.2}$$

Likewise, knowledge of a bond's yield (holding the term constant), enables one to calculate its price.

The instantaneous (short-term) interest rate, r_t, is the annualized risk-free rate on a one period bond denoted by the limit of yield to maturity[1]. That is,

$$r_t \equiv y_t(0) = \lim_{\tau \to 0} y_t(\tau).$$

The discount function is an exponential decay curve, whose rate of decay is the instantaneous forward interest rate. That is, the future interest rate that one would obtain today for over a specified period in the future. The forward rate $f_t(\tau, T)$, is the future interest rate obtainable at time t for over a specified period in the future. That is, the $T-$year future forward rate, beginning at time $[\tau, T]$ years, hence is expressed as

$$f_t(\tau, T) \quad = \quad \frac{1}{T-\tau} \ln \left[\frac{P_t(\tau)}{P_t(T)} \right] \qquad \text{for} \quad 0 < t \leq \tau \leq T. \tag{2.1.3}$$

Taking the limit of (2.1.3) as $T \to \tau$, gives the instantaneous forward rate $f_t(\tau)$, of τ years ahead[2], which represents the continuously compounded in-

[1]In reality, instantaneous short term rates does not exist, it is just a theoretical construct used to model interest rates, as cited in most traditional stochastic interest rate models

[2] The instantaneous forward rate curve is a very important theoretical construct. However, its value for a single maturity τ is of little practical concern, because it is prohibitively expensive in terms of transactions cost to make a forward contract between two points in the distant future if these points are only a small distance apart.

stantaneous return for a future date that one would require at time t:

$$\begin{aligned} f_t(\tau) &= \lim_{T\to\tau} f_t(\tau,T) \\ &= -\lim_{T\to\tau} \frac{\ln P_t(T) - \ln P_t(\tau)}{T-\tau} \\ &= -\lim_{\varepsilon\to 0} \frac{\ln P_t(\tau+\varepsilon) - \ln P_t(\tau)}{\varepsilon} \qquad (2.1.4) \\ &= -\frac{\partial \ln P_t(\tau)}{\partial \tau} \\ &= -\frac{P_t'(\tau)}{P_t(\tau)} \quad \text{"forward rate curve."} \qquad (2.1.5) \end{aligned}$$

Equation (2.1.5)[3] shows that forward rates are closely related to discount function since they can be obtained in terms of bond prices. The discount function is assumed to be continuously differentiable, this implies that bond prices can be obtained in terms of forward rates as

$$P_t(\tau) = \exp[-\int_0^\tau f_t(x)dx].$$

Moreover, using Equation (2.1.2), the relationship between yield curve and the forward curve, is given by

$$y_t(\tau) = \frac{1}{\tau}\int_0^\tau f_t(x)dx \quad \text{"yield curve",} \qquad (2.1.6)$$

which states that the yield curve may be interpreted as a weighted average of forward rates over the interval $[0,\tau]$. Given the knowledge of the discount function, enables one to calculate the bond yields, so as the valuation of the forward rates. However one of the curves can be used to construct the term structure, since they are effectively interchangeable. The core interest of both academic and industry practice is based on the use of the yield curve $y(\tau)$, which is the most common method amongst literatures.

2.2 Zero-Coupon Bond Yields

In practice, constructing yield curves, discount curves and forward curves is challenging, since these yields are not directly observed, instead, they are estimated from observed bond prices. Throughout this thesis, following much of the literature, yields that will be considered are estimated from observed zero-coupon bond prices with various amounts of time to maturity. To achieve

[3]Only the average of $f_t(\tau)$, the mean forward interest rate $f_t(\tau,T)$ over a considerable interval in the future is of practical concern.

a meaningful yield curve, the bond yields that are being compared must have similar risk and payout characteristics. This is central to various term structure literatures to model asset prices specialized on zero-coupon bonds, because they represent the fundamental discount rates embedded in fixed-income products that make payments through time.

The background of the yield curve construction results from the two historically popular approaches that estimate a smooth discount curve to observations on bond prices with varying maturities and then converting to yields at the relevant maturities by making use of the formulas denoted by Equation (2.1.5) and Equation (2.1.6) above. Among the predecessors of yield curve estimation, the first discount curve approach is due to, McCulloch (1975), who modeled the discount curve by proposing the use of polynomial splines. The fitted discount curve, instead of converging to zero, it however, tends to diverge at long maturities due to the polynomial structure. This results to the corresponding yield curve inheriting the same manner. The method turns out, however, to poorly fit yields that flatten out with maturity, as noted by Shea (1984).

Later, Vasicek and Fong (1982) provide an improvement of the discount curve approach to yield curve construction, by modelling the discount curve using exponential splines. The method introduces the use of a negative transformation of maturity rather than maturity itself, to ensure that forward rates and zero-coupon yields converge to a fixed limit as maturity increases. However, considering this clever technique, the Vasicek-Fong approach appears much more accurate at fitting yield curves with flat long ends. Nevertheless, the discount curve approach in Vasicek and Fong (1982) suffers from some drawbacks, because the estimation requires iterative non-linear optimization which make it almost impossible to restrict the implied forward rates to be always positive.

An alternative and a very popular approach to yield curve construction that seems to out-stand the various standard benchmark issues, is due to Fama and Bliss (1987). Yields are constructed from estimated forward rates at the observed maturities, which is contrary to the use of estimated discount curve. The constructed forward rates are necessary to price long-maturity bonds and are often called "unsmoothed Fama-Bliss" forward rates, see Diebold and Li (2006). The unsmoothed Fama-Bliss forward rates are transformed to unsmoothed Fama-Bliss yields[4] by appropriate averaging, using Equation (2.1.6) above.

At any time, different yields corresponding to different bond maturities

[4]The unsmoothed Fama-Bliss yields exactly price the included bonds. Numerous studies make use of this estimation technique to fit empirical yield curves (i.e Nelson-Siegel family), as this is to be discussed in details throughout this thesis.

evolve dynamically by shifting among different shapes. This, however shows that, for a given set of bonds, the slope of the yield curve reflects an expectation about the movement of interest rates over time. Different shapes can be interpreted as "Downward sloping", where long-term borrowing costs are lower than short-term borrowing costs. "Upward sloping"[5], where short-term borrowing costs are lower than long-term borrowing costs. Lastly, the "Flat yield curve" where borrowing costs are relatively similar for short and long term loans.

The construction of the term-structure conveys an important information that can be useful for various aspects such as conducting monetary policy, forecasting future interest rate path of the economy, the financing of public debt, risk management of a portfolio of securities and the pricing of interest rates derivatives. Last but not the least, the formation of expectations about real economic activity and inflation.

2.3 Holding Period Return

Another important tool for the valuation of the term structure of interest rates is the *holding period return*. Firstly, to forecast the holding period returns, we determine the log forward rate at time t for loans between time $t + \tau - m$ and $t + \tau$ as,

$$f_t(\tau) = p_t(\tau - m) - p_t(\tau).$$

The holding period return is the return on τ-year zero-coupon bond at time t, and then selling it as an $(\tau - m)$-year zero-coupon bond at time $t + m$, which generates a log holding period return of

$$r_{t+m}(\tau) = p_{t+m}(\tau - m) - p_t(\tau).$$

The holding period is usually random because it depends on the resale value of the bond $P_{t+m}(\tau - m)$, which is generally not known at time t. Since the holding period m cannot exceed time to maturity τ, the resale value equals its payoff when the bond matures so that holding a bond until maturity $m = \tau$ generate a return which is known at time t. The per-period holding period return in this case is the yield to maturity

$$y_t(\tau) = \frac{r_{t+m}(\tau)}{\tau} = -\frac{\ln P_t(\tau)}{\tau},$$

where the short rate $r_t = \lim_{\tau \downarrow 0} y_t(\tau)$. The difference between the log holding period return and the m-year yield gives the log excess holding period return,

$$rx^{\tau}_{t+m} = r_{t+m}(\tau) - y_t(\tau),$$

[5]The upward-sloping yield curve has been the most prevalent historically.

these are returns made in excess of riskless return over the holding period. This is an important feature of the data that the term structure models have to match.

2.4 Risk Neutral Pricing

Bonds are usually priced using "risk-neutral probability measure" $\mathbb{Q}$, under no-arbitrage assumptions. The principle of no arbitrage allows one to develop a consistent pricing and hedging theory. Following from Ross (1976) and Harrison and Kreps (1979),[6] the most important implication of the arbitrage-free assumption is the existence of a positive stochastic process known as *Pricing Kernel*, $M(t)$, which enables one to compute prices of bonds of any maturity under probability measure $\mathbb{P}$. However, the price of any asset that promises a payoff of $S(T)$ at time T results in asset prices being the expected values of their discounted future payoffs. Pricing relations at time t, are therefore given by,

$$S(t) = \mathbb{E}_t^{\mathbb{P}} \left[\frac{M(T)}{M(t)} S(T) \right]. \tag{2.4.1}$$

Stock prices usually over perform money market under the actual measure, however, when it comes to derivative pricing, the right measure to work with is the risk-neutral measure. One key characteristic of risk-neutral measure is that under this measure, every discounted price process is a martingale. This is not true under the actual measure. Therefore, in order to price derivatives, we need to change the measure from the actual to the risk-neutral.

The no-arbitrage assumption guarantees the existence of an equivalent martingale measure $\mathbb{Q}$ or risk-neutral measure. When investors are risk-neutral, this pricing results applies under the data-generating measure $\mathbb{P}$. In general, the risk-neutral probability measure $\mathbb{Q}$ will be different from $\mathbb{P}$. The change of the two measures are connected through the Radon-Nikodým derivative Z,

$$Z = \frac{d\mathbb{Q}}{d\mathbb{P}} > 0.$$

In order to price derivatives at any time $t \in [0, T]$, we need conditional expectation under $\mathbb{Q}$, and we have the Radon-Nikodým process to handle it:

$$Z_t = \mathbb{E}_t(Z) = \frac{d\mathbb{Q}}{d\mathbb{P}}\bigg|_{\mathcal{F}_t} \equiv \exp\left(-\frac{1}{2} \int_t^T \lambda'(s)\lambda(s)ds - \int_t^T \lambda(s)dz(s) \right), \tag{2.4.2}$$

where $\lambda(t)$ is the market price of risk and the shocks of the economy $z(t)$ following a standard Brownian motion at time t. This implies that, under

[6] Ross (1976) and Harrison and Kreps (1979) explore the principle of no-arbitrage in frictionless markets to obtain a convenient representation of pricing operators.

probability measure $\mathbb{P}$, the pricing kernel $M(t)$ is given by,

$$\frac{dM(t)}{M(t)} = -r(t)dt - \lambda(t)'dz(t), \tag{2.4.3}$$

where $r(t)$ is the short rate. An application of Itô's lemma gives

$$d\ln M(t) = [-r(t) - \frac{1}{2}\lambda^{'}(t)\lambda(t)]dt - \lambda(t)dz(t). \tag{2.4.4}$$

Applying an integration on Equation (2.4.4),

$$\frac{M(T)}{M(t)} = \exp\left(-\int_t^T r(s)ds - \frac{1}{2}\int_t^T \lambda^{'}(s)\lambda(s)ds - \int_t^T \lambda(s)dz(s)\right).$$

The pricing equation takes the form,

$$S(t) = \mathbb{E}_t^{\mathbb{P}}\left[\frac{M(T)}{M(t)}S(T)\right] = \mathbb{E}_t^{\mathbb{Q}}[e^{-\int_t^T r(s)ds}S(T)].$$

Easy manipulations of the above equation can prove that under $\mathbb{Q}$, the instantaneous expected returns for all assets are equal to the risk-free rate. For this reason the measure $\mathbb{Q}$ is also called risk-neutral measure. Specializing the above equation for a zero-coupon bond that matures at time T and promises a payoff of 1-unit gives,

$$\begin{aligned} P_t(T) &= \mathbb{E}_t^{\mathbb{P}}\left[\frac{M(T)}{M(t)}\right] \\ &= \mathbb{E}_t^{\mathbb{P}}[\underbrace{\exp\left(-\frac{1}{2}\int_t^T \lambda^{'}(s)\lambda(s)ds - \int_t^T \lambda(s)dz(s)\right)}_{\frac{d\mathbb{Q}}{d\mathbb{P}}}\exp\left(-\int_t^T r(s)ds\right)]. \end{aligned}$$

This gives the specification of the Radon-Nikodým derivative, and leads to the risk neutral $\mathbb{Q}-$formula:

$$P_t(T) = \mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\exp(-\int_t^T r(s)ds)\right], \tag{2.4.5}$$

then the price of a zero-coupon bond becomes,

$$P_t(T) = \mathbb{E}_t^{\mathbb{Q}}[e^{-\int_t^T r(s)ds}],$$

where $\mathbb{E}_t^{\mathbb{Q}}$ denotes the conditional expectation under measure $\mathbb{Q}$. Standard results found in Duffie and Lando (2001) show that if there exists a risk-neutral probability measure $\mathbb{Q}$, a system of asset prices is arbitrage-free. Moreover, the uniqueness of probability measure $\mathbb{Q}$ is equivalent to market completeness. From the above pricing equations we observe that the key variables that govern the bond prices dynamics are the interest rate $r(t)$, and the time varying price of risk $\lambda(t)$. Under risk-neutral measure, expected excess returns on bonds are zero, the expected rate of return on a long bond equals the risk-free rate.

2.5 Expectations Hypothesis

The *Expectations Hypothesis* (EH) is the benchmark term structure model that provides a theoretical link between yields, returns on bonds and forward rates of different maturities. Most related literatures on this framework are based on market expectations hypothesis which are mainly developed for understanding the returns and yields on long versus short term bonds, and the movement of the term structure. The hypothesis suggests that long-term yields are equal to the average of expected short-term interest rates until the maturity date, and allows for a maturity specific constant term premium of the long yield over the average expected short-term interest rates. Three implications are then suggested by the hypothesis, that is, yield term premia and forward term premia are constant, expected excess returns are constant over time. Another illustration of the expectation hypothesis states that, investors price all bonds as though they were risk-neutral. That is to say, investors are concerned about the expected outcomes and they are indifferent between two investments strategies. Consider the two following strategies: when buying R100 of a 1-period bond, when it matures, every year re-invest the remains in another 1-period bond for τ-periods or buy a R100 of τ-period bond and hold it. Hence the expected returns on these bonds of different maturities are equal. This shows the relationship between short-term yields with long-term yields,

$$y_t^{EH}(\tau) = \frac{1}{\tau}\sum_{i=0}^{\tau-1} \mathbb{E}_t y_{t+i}^{EH}(1), \tag{2.5.1}$$

where $y_t(\tau)$ is the τ-period yield of discounted bond that matures at time t. The expression in Equation (2.5.1) postulates that the movements in long yields are due to movements in expected future 1-period maturity yield or short-term yields. The expectations hypothesis argue that the current term structure gives information about investors expectations regarding future path of interest rates. At any time, different yields corresponding to different bond maturities evolve dynamically by shifting among different shapes. This, however shows that, for a given set of bonds, the slope of the curve reflects an expectation about the movement of interest rates over time. Different shapes on yield curves can be interpreted as "Downward sloping", where expected future short rates are falling, and this results to long-term borrowing costs being lower than short-term borrowing costs. Secondly, "Upward sloping", where expected future short rates are expected to rise and short-term borrowing costs are lower than long-term borrowing costs in this scenario. Lastly, the "Flat yield curve", where borrowing costs are relatively similar for short-term and long-term loans. The upward-sloping yield curve has been the most prevalent historically.

Another alternative scenario regarding the upward sloping yield curve gives uncertainty about future long term bond yields which makes them systematically less attractive to lenders than the short term bonds. This concludes that expectations hypothesis is exactly true in a certain world by the force of no-arbitrage, but when interest rates are stochastic, uncertainty causes systematic distortion of expectations hypothesis.

2.5.1 Expectations Hypothesis and Risk Premia

Yield curve relationships are captured in three ways, that is, the long-term yield is an average of expected future short-term rates plus a risk premium, expressed as

$$y_t(\tau) = \frac{1}{\tau}\mathbb{E}_t(y_t(1) + y_{t+1}(1) + \ldots + y_{t+\tau-1}(1)) + rpy_t(\tau). \tag{2.5.2}$$

Secondly, the forward rate is the expected future short rate plus risk premia, given by

$$f_t(\tau) = \mathbb{E}_t(y_{t+\tau-1}(1)) + rpf_t(\tau). \tag{2.5.3}$$

Lastly, the expected one period return on long-term bonds equals the expected return on short-term bonds plus a risk premium. That is,

$$\mathbb{E}_t(r_{t+1}(\tau)) = y_t(1) + rpr_t(\tau). \tag{2.5.4}$$

All the three equations can be taken as definitions of the risk premia, which are equivalent such that if one equation holds with $rp = 0$ or constant, then all other equations hold with $rp = 0$ or constant over time.

2.6 The Taylor Rule (1993) Specification

Taylor (1993) introduced a rule for guiding and assessing monetary policy performance. Taylor rule is a simple monetary policy technique linking the levels of the policy rate mechanically to deviations of inflation from its target and output from its potential (the output gap). According to the policy rule recommended by Taylor (1993), movements in the short rate r_t are traced to movements in contemporaneous observable macro variables and a component which is not explained by macro variables, an orthogonal shock η_t. The Taylor's rule original specification considers two macro variables as factors in observable variables such as the annual inflation rate, and the output gap. The Taylor rule also implies that central banks aim at stabilising inflation around its target level and output around its potential, by adjusting short-term interest rates in response to movements in inflation and real activity, (Svensson (1997)). Now, the monetary authority is assumed to set the short term interest rate according to this simple aspect of the Taylor rule given by:

$$i_t = \pi_t + r_t^* + a_\pi(\pi_t - \pi_t^*) + a_y(y_t - \bar{y}_t) + \eta_t.$$

The central bank reacts to high inflation and to deviations of output from its trend. The parameter i_t describes the nominal policy rate, π_t^* is the central bank's time-varying desired rate of inflation, while π_t is the current period inflation rate as measured by the gross domestic product (DGP) deflator. The long run or assumed equilibrium real rate of interest is given by r_t^*, y_t is the current period output gap, while $\bar{y}_t$ gives the potential output. The constant a_π measures the response of the central bank to inflation, while a_y describes its reaction to output gap fluctuations. Lastly, the monetary policy shock η_t. This monetary policy rule specifies that the rate of the central bank should change the nominal interest rate in relation to changes in inflation, output gap, or other economic conditions. It emphasizes the importance of adjusting policy rates more than one-for-one in response to these economic conditions. In particular, the rule stipulates that for each one-percent increase in inflation, the central bank should raise the nominal interest rate by more than one percentage point.

The rule recommends a relatively high interest rate when inflation is more than its target or when output is more than its full-employment level, in order to reduce inflationary pressure. It recommends a relatively low interest rate in the opposite situation, to stimulate output. During the period of rising inflation and falling output, monetary policy goals may conflict, when inflation is above its target while output is below full employment. In such a situation, specifically the great inflation of American central bank during the $1970'$s where policy rates were below the level implied by this benchmark, a Taylor rule specifies the relative weights given to reducing inflation versus increasing output. This is done to cool the economy when inflation increases. The positive or negative deviations of inflation and output gap from their target or potential level is associated with a tightening or loosening of monetary policy. The Taylor rule benchmarks for the global aggregate as well as the aggregate of advanced and emerging markets over a period from first quarter of 1995 to the first quarter of 2012 is computed in Figure 2.1 to obtain a range of possible Taylor rule implied rates for all combinations of four measures of inflation (headline, core, GDP deflator and consensus headline forecasts) and measures of the output gap obtained from three different statistical ways to compute potential output (HP filter, segmented linear trend and unobserved components).

The Taylor (1993) rule and Policy Rates

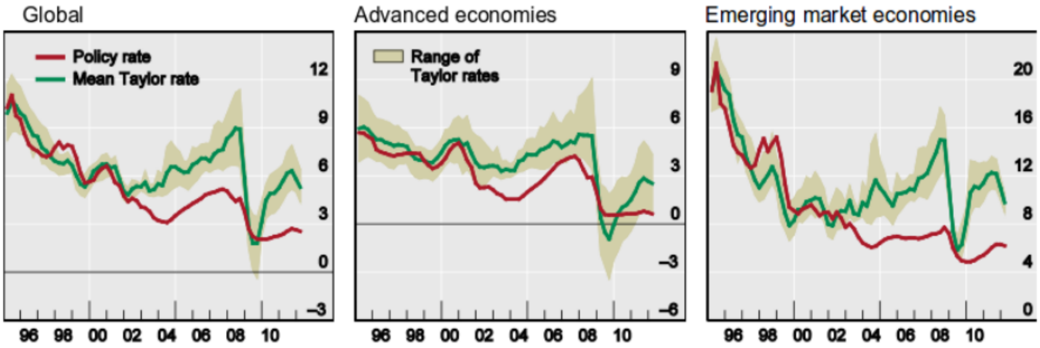


Figure 2.1: Taylor rates are computed for all combinations of four measures of inflation and three measures of the output gap.

Sources: IMF, International Financial Statistics and World Economic Outlook; Bloomberg; CEIC; Consensus Economics; Data stream; national data; authors' calculations.

Figure 2.1, left-hand panel results reveal the systematic deviation of policy rates[7] from the Taylor rule since the early 2000s. This further shows that policy rates were consistent with the levels implied by the Taylor rule up until the early years of the new millennium, a systematic deviation emerged thereafter. Since 2003, global policy rates have almost always been below the levels indicated by Taylor rules. Only during the Great Recession of 2009 where policy rates briefly inside the Taylor rule range, there after remained low while the global economy recovered, the gap opened up again reflecting the recent weakening of the global economy. However, the deviation narrowed somewhat in the first quarter of 2012. In the advanced economies, policy rates have been below the range of Taylor rule rates since around 2001 to 2009, but the deviation is smaller as compared to Figure 2.1, left and right-hand panel. In the Great Recession, the Taylor rule would on average have suggested negative policy rates for a short period of time, but actual policy rates were still well inside the range. In 2011, the spectrum of Taylor rates shifted back to positive levels and policy rates have been at the lower bound of the range since then.

[7] "Global" comprises the economies listed here. Advanced economies: Australia, Canada, Denmark, the euro area, Japan, New Zealand, Norway, Sweden, Switzerland, the United Kingdom and the United States. Emerging market economies: Argentina, Brazil, China, Chinese Taipei, the Czech Republic, Hong Kong SAR, Hungary, India, Indonesia, Korea, Malaysia, Mexico, Peru, Poland, Singapore, South Africa and Thailand

The Taylor (1999) rule and policy rates

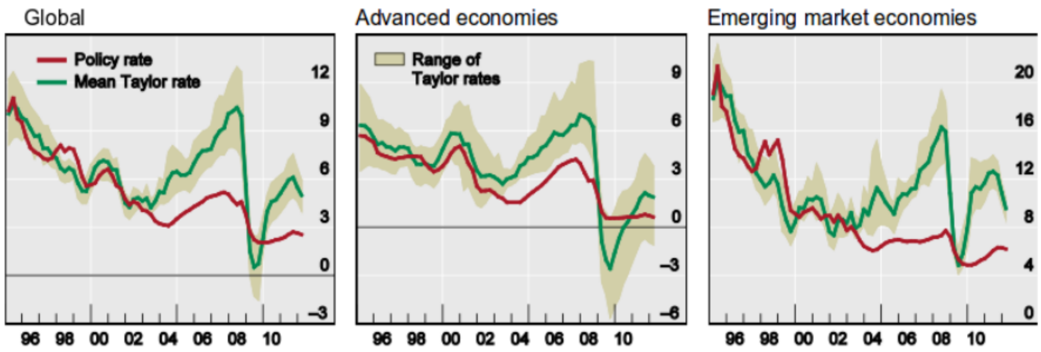


Figure 2.2: Taylor rates are computed for all combinations of four measures of inflation and three measures of the output gap.

Figure 2.2 uses an alternative calibration of the Taylor rule considered by Taylor (1999) for the same aggregates. The only difference from the original calibration Taylor (1993) is a larger output reaction coefficient. In addition to this larger output weight, the range of the aggregate of advanced economies shifts down for the period since the Great Recession, indicating negative policy rates for a longer period and putting policy rates well inside the Taylor rule range at the end of the sample period. This popular gauge for assessments of the monetary policy stance has since been subject to considerable attention, as seen in the work of Ang and Piazzesi (2003), for assessment of monetary stance both in advance economies and emerging market economies as a prescription for desirable policy. The instantaneous short rate equation that arise in the affine term structure literature can be extended with macroeconomic variables to a Taylor (1993) rule as shown in Ang and Piazzesi (2003) approach. The basic insight of the approach is that a well known Taylor rule specification of the short rate also has an affine form:

$$r_t = \rho_\pi \pi_t^Y + \rho_g \mathrm{gap}_t + \mathrm{constant},$$

where π_t^Y is the annual inflation and gap_t is the gross domestic product gap. Therefore, using inflation and GDP gap variables as part of the state vector in the affine Gaussian model set up, that is,

$$x_t = [\pi_t^Y, \mathrm{gap}_t, ...]',$$

provides a system in which bond yields are linked to macroeconomics variables. A high order of VAR process can be written as a VAR(1) process with an expanded state vector that includes lags of these variables, that is,

$$x_t = [\pi_t^Y, \pi_{t-1}^Y, ..., \mathrm{gap}_t, \mathrm{gap}_{t-1}, ...]'.$$

This combination set up permits the expansion of affine term structure models to explicitly include observable macroeconomic outcomes.

2.7 Term Structure of Interest Rate Modelling

This section reviews the general structure and the importance of the term structure of interest rate models. Term structure of interest rates is also known as yield curve, plays a central role both theoretically and practically in the economy. Firstly, understanding the stochastic behaviour of yields and what moves bond yields is important for different motives. One of these reasons would be for conducting monetary policy. Model of the yield curve at any given state of the economy, helps to understand the relationship between movements at the short end into longer-term yields. This involves adjustments of the short rate by the central banks to conduct monetary policy and how the transmission mechanism works. The expectations hypothesis (EH) is commonly cited to model the term structure in most previous research in this area.

Interest rate forecasting is another crucial aspect of understanding the future path of the economy. The literature on expectations hypothesis states that yields on long-maturity bonds are expected values of average future short yields, at least after an adjustment for risk. This implies that the current yield curve contains information about the future path of the economy, as shown in Hicks (1946). Yield spreads have been useful for forecasting not only future short yields, but also inflation and real activity, (Ang *et al.* (2006)), even though the forecasting may tend to give unstable results.

Public debt policy and the risk management of a portfolio of interest rate sensitive securities, gives the third reason. When issuing new debt, governments need to decide about the maturity of the new bonds,[8] this helps to study how various strategies behave under different interest rate outcomes, while minimizing the risk that short-term interest rates spike. Lastly, Derivative pricing and hedging provide a fourth reason. Governments need to take good care of risk management and derivative pricing on bonds,[9] and also to compute hedging strategies[10] in response to the price of derivative securities which are depending on the changing state of the economy.

There are various early models attempted to model the term-structure of interest rates. Amongst the most popular, is the adopted no-arbitrage approach[11] pioneered by Vasicek (1977) and Cox *et al.* (1985). This model is known to be the one factor interest rate model that considers short rate as the

[8] The administration of John Kennedy persuaded the Federal Reserve to co-operate on selling short maturity debts and buying long maturity notes, using the idea that came to be known as "Operation Twist".

[9] To manage the risk of paying short-term interest rates on deposits while receiving long-term interest rates on loans

[10] Hedging strategies involve contracts that are contingent on future short rates, such as swap contracts.

[11] The formal definition of no-arbitrage condition and its applications are given in Harrison and Kreps (1979).

basis for modelling the term-structure of interest rate, typically using affine models. In the case of the empirical finance literature, no-arbitrage approach is a centre of interest for the analysis of yield curves and it provides tractability and consistency in bond pricing.

Term-structure models of this nature have essentially the same form such that some useful key assumptions are involved in the modelling process. The first assumption is that the interest rate system is the function of some set of state variables, (i.e., unobservable variables), as demonstrated in the work of Litterman and Scheinkman (1991) or observable factors, such as macroeconomic variables. The second assumption involves the dynamics of the state variable vector. In a discrete time setting, the dynamics of state variables follow VAR process to describe evolution of the state variable vector. This is the discrete-time analogue of the Ornstein-Uhlenbeck process. The third, and final, assumption relates to the mapping between the state variables and the term structure of interest rates. Most of the models in the finance literature ensure that the mapping excludes arbitrage opportunities. This is commonly applicable to the most popular affine-class of the term structure model, Duffie and Kan (1996).

There are, however, other alternative mappings. This includes the work introduced in Diebold and Li (2006), where a mapping between a set of discrete time, continuous-value state variables that does not exclude arbitrage. Even though the model does not consider arbitrage-free restrictions, empirical results have shown that these models still lead to superior success and they are parsimonious models as noted by Diebold and Li (2006).

2.8 Arbitrage-free Affine Term-Structure Models

Since bonds usually trade in deep and well organised markets, the theoretical restrictions that eliminate opportunities for riskless arbitrage across maturities and over time hold powerful appeal, and they provide the foundation for a large finance literature on arbitrage free models. The pioneers of the literature of the arbitrage-free term structure models are from the popular work of Vasicek (1977) and Cox *et al.* (1985), where a single-factor model is introduced as a short rate, which determines the bond prices. Both models are discussed in discrete time setting, to specify the risk neutral evolution of the underlying yield curve factors as well as the dynamics of risk premia.

2.8.1 Vasicek Model

Vasicek (1977) presentation under risk neutral assumptions continues directly from the work of Black and Scholes (1973) that introduced a breakthrough on the asset pricing theory. Vasicek model describes the evolution of interest rates[12] that incorporates mean reversion and a risk premium. The short-rate is given by

$$r_t = (1-\phi)r_{t-1} + \phi\tau + z_t,$$

where $z_t \sim N(0, \sigma^2)$, with the three following parameters. The parameter τ is the long-term steady-state rate to which the process is reverting. The coefficient ϕ is the strength of mean-reversion[13], and σ is the volatility parameter of the stochastic process. The short-term rate follows a normal distribution of the form,

$$r_t \sim N(\tau, \frac{\sigma^2}{1-\phi^2}).$$

The normal distribution of rates makes it impossible to avoid negative interest rate scenarios, because of this fact, it is possible to develop a closed form solution for the parameters that will exactly meet the key calibration points.[14] The single state variable z, that follows a first-order autoregressive is given by,

$$z_{t+1} = (1-\phi)\mu + \phi z_t + \sigma\epsilon_{t+1}, \tag{2.8.1}$$

where $\epsilon_{t+1} \overset{\text{i.i.d.}}{\sim} N(0, 1)$. The other parameters give the following interpretation; μ is the mean of z, σ^2 is the conditional variance, and the unconditional variance is given by $\frac{\sigma^2}{1-\phi^2}$. The pricing kernel is given by

$$-\ln M_{t+1} = \delta + z_t + \lambda\epsilon_{t+1}, \tag{2.8.2}$$

where the market price of risk, λ, determines the covariance between shocks to kernel M, and state variable z, risk characteristics and the corresponding bonds. Following the property of log-normal random variable, from Equation (2.8.2), the log-normal kernel, $\ln M_{t+1}$ has a conditional mean $-(\delta + z_t)$, and a conditional variance of λ^2. This gives the one period bond price that satisfies

$$\ln P_t^1 = -\delta - z_t + \frac{\lambda^2}{2} = -z_t. \tag{2.8.3}$$

The short rate is therefore,

$$r_t = -\ln P_t^1 = z_t. \tag{2.8.4}$$

Then, the price of an $n-$period bond can be assumed as,

$$-\ln P_t^n = A_n + B_n z_t. \tag{2.8.5}$$

[12] Hull and White (1994) adapt this specification to introduce a discrete-time continuous state form of the Vasicek (1977) model to value the short-term interest rate.

[13] The parameter ϕ controls mean reversion and must be between 0 and 1. A zero value would results in no mean-reversion while a one value would results in full reversion in the next period.

[14] The 10^{th} and the 97.5^{th} percentiles. Other calibration points are not likely to be a constraint for the Vasicek model form.

Given the bond price coefficients for maturity n, to evaluate the bond price of an $(n+1)-$period ahead, and using the fact that bond prices satisfy the following,

$$\begin{aligned}
P_t^{n+1} &= \mathbb{E}_t(M_{t+1}P_{t+1}^n) \\
\ln P_t^{n+1} &= \mathbb{E}_t[\ln M_{t+1} + \ln P_{t+1}^n] + \frac{1}{2}\mathrm{Var}_t(\ln M_{t+1} + \ln P_{t+1}^n) \\
&= -[\delta + A_n + B_n(1-\phi)\mu] - (1 + B_n\phi)z_t \\
&\quad + \frac{1}{2}(\lambda + B_n\sigma)^2\mathrm{Var}_t(\epsilon_{t+1}).
\end{aligned}$$

Equating the coefficients with Equation (2.8.5), gives the recursions

$$A_{n+1} = A_n + \delta + B_n(1-\phi)\mu - \frac{1}{2}(\lambda + B_n\sigma)^2\mathrm{Var}_t(\epsilon_{t+1})], \quad (2.8.6)$$

$$B_{n+1} = 1 + B_n\phi. \quad (2.8.7)$$

Using the parameters, $(\mu, \phi, \sigma, \lambda)$, the coefficients can be easily computed. Forward rates in this model take a particularly simple form,

$$f_t^n = (1-\phi)\mu + \frac{1}{2}\left[\lambda^2 - \left(\lambda + \frac{1-\phi^n}{1-\phi}\sigma\right)^2\right] + \phi^n z_t. \quad (2.8.8)$$

The forward rate expression illustrates the impact of short rates on long forwards and the form of the risk premium, which give insight into the relation between forward rates and expected future short rates. If $\lambda = \delta = 0$, the forward rate, $f_t^n = \mathbb{E}_t r_{t+n}$, forms a version of expectations hypothesis. When $\lambda = 0$ and $\sigma \neq 0$, the non-linearity of the prices relation results in a downward sloping average forward rate curve. If $\lambda \geq 0$, this effect is reversed and average forward rates, such as yields, increase with maturity.

2.8.2 Cox, Ingersoll and Ross (CIR) Model

Cox *et al.* (1985) introduced an alternative approach to Vasicek model. In this model, the stochastic process is scaled by the square root of interest rate. This ensures that interest rate does not become negative, because the closer the interest rate r_t gets toward zero, the closer the formulation becomes to a mean reversion process with no stochastic term. This model is represented by

$$r_t = (1-\phi)r_{t-1} + \phi\tau + \sqrt{r_{t-1}}z_t,$$

where $z_t \sim N(0, \sigma^2)$. It is possible to determine the parameters for this model using a set of historical data, with a constrained MLE approach. However, in this case, the steady state rate does not follow a normal distribution. Simulations can be used to calculate the needed percentiles. If such an approach is used, then it is necessary to fix a reasonable duration at which steady-state is presumed.

In the case of the discrete time version, the CIR model has also a similar structure as the Vasicek, but the difference lies in the behaviour of the state variable z, which in this case obeys the "square root process", and also takes on the first order autoregression. More precisely, we have

$$z_{t+1} = (1-\phi)\mu + \phi z_t + \sigma z_t^{1/2}\epsilon_{t+1}, \tag{2.8.9}$$

with $0 < \phi < 1$ and $\mu > 0$. The conditional variance is,

$$\mathrm{Var}_t(z_{t+1}) = z_t\sigma^2, \tag{2.8.10}$$

which has a mean of $\mu\sigma^2$, and the unconditional variance is $\mathrm{Var}(z) = \frac{\mu\sigma^2}{1-\phi^2}$. Equation (2.8.9) ensures that the interest rate does not become negative. If the time interval is small with the square-root process, the conditional variance gets small as z approaches zero, this results on the limited chances of getting a negative value. Since ϵ is distributed normally, there is still a positive probability that z_{t+1} is negative, but the probability falls to zero as the time interval shrinks. The conditionally log-normal pricing kernel for a discrete time version of CIR is

$$-\ln M_{t+1} = (1+\lambda^2/2)z_t + \lambda z_t^{1/2}\epsilon_{t+1}. \tag{2.8.11}$$

Adding Equation (2.8.11) with the log-price, we get

$$\begin{aligned}
\ln M_{t+1} + \ln P^n_{t+1} &= -(1+\lambda^2/2)z_t - \lambda z_t^{1/2}\epsilon_{t+1} - A_n - B_n z_{t+1} \\
&= -[A_n + B_n(1-\phi)\mu] - [(1+\lambda^2/2) + B_n\phi]z_t \\
&\quad -[\lambda z_t^{1/2} + B_n\sigma z_t^{1/2}]\epsilon_{t+1}.
\end{aligned}$$

The expected log-normal pricing becomes,

$$\mathbb{E}_t(\ln M_{t+1} + \ln P^n_{t+1}) = -[A_n + B_n(1-\phi)\mu] - [(1+\lambda^2/2) + B_n\phi]z_t$$

and,

$$\mathrm{Var}_t(\ln M_{t+1} + \ln P^n_{t+1}) = (\lambda + B_n\sigma)^2 z_t.$$

The bond price becomes,

$$P^n_{t+1} = A_n + B_n(1-\phi)\mu + [\frac{(\lambda + B_n\phi)^2}{2} + (1+\lambda^2/2) + B_n\phi]z_t,$$

where the coefficients of the log-linear bond price formula satisfy the recursion

$$\begin{aligned}
A_{n+1} &= A_n + B_n(1-\phi)\mu \\
B_{n+1} &= 1 + \lambda^2/2 + B_n\phi + (\lambda + B_n\sigma)^2/2,
\end{aligned}$$

starting with $A_0 = B_0 = 0$. Since $A_1 = 0$ and $B_1 = 1$, z is the short rate. Financial economists have mainly focused on forecasting yield curves and pricing interest rate related securities. They have therefore developed powerful models (e.g. ATSM) based on the assumption of absence of arbitrage opportunities, but typically left unspecified the relationship between the term structure and other economic variables.

Chapter 3

Macro-Finance (Yields-Macro) Approach

Macro-finance modelling framework focuses on the joint dynamics of term structure variables and macroeconomic variables. The link between these two literatures is a new branch of research that has been of perennial interest to researchers, to understand the behaviour of the yield curve concerning economic variables. As noted in Ang *et al.* (2006), that macroeconomics focuses on the overall movement and trends in the economy and the economic approach typically works with vector autoregressive, which may give inconsistent results such as arbitrage opportunity. The no-arbitrage vector autoregressive, mostly used in macro-finance model is introduced to overcome these features. In this approach, macroeconomic variables and latent factors are used in the term structure estimation of interest rates.

3.1 Linking Macro and Yield Curve Variables

Both macroeconomic and finance are linked through the short-term interest rate. In finance, long-term interest rates are risk adjusted averages of expected future short rates, and for this reason, short rates serve as a fundamental building block for rates of other maturities. From a macro perspective, central bank controls the short rates, by adjusting the rates in order to achieve the economic stabilization goals of monetary policy. The two perspectives suggest that the short-term interest rate is a critical point of intersection because understanding the manner in which central banks move the short rate in response to fundamental macroeconomic shocks should explain movements in the short end of the yield curve.

A joint macro-finance connects the macroeconomic variables with the yield factors. Term structure models constructed in these two distinct branches have different aims. Financial economists develop affine models based on the absence of arbitrage and manly focused on forecasting and pricing. But models of this nature could not specify the relationship between the term structure and other economic variables. On the other hand, macroeconomists have developed term structure models to determine the relation between interest rates monetary policy and macroeconomic fundamentals. The macroeconomic models are based on the expectation hypothesis and could not provide the interactions between macroeconomy and the term structure dynamics.

Understanding the stochastic behaviour of the yield curve in response to macroeconomic shocks is crucial for central banks since yield curves provide a useful information about underlying expectations of inflation and output over a number of different horizons. Number of researchers have tried different frameworks to determined the joint macro-finance characterization of the term structure of interest rates with the help of macroeconomic fundamentals, but some could not provide an explicit relation between the determinants of the yield curve shape and macroeconomic factors. The work by Ang and Piazzesi (2003) is the break through to bridge the gap in macro-finance literature that offers a vector autoregressive framework to capture the joint dynamics of macro factors and yield factors under no-arbitrage restrictions where macroeconomic factors are measures of inflation and real activity.

3.2 Advantages of Macro-Finance Modelling

Most macro-finance models in the literature are based on the "affine-Gaussian" model, and the no-arbitrage key assumptions of their affine model under simple vector autoregressive process that captures bond yield movements over time. However, Gaussian macro-dynamic term structure models are further developed, in which bond yields follow a dimensional factor structure and the historical distribution of bond yields. The joint macro-finance model offers a number of advantages over both pure term structure finance models and pure macroeconomic models to understand the behaviour of the yield curve.

The first advantage is that term structure models relate the yield curve to current and past interest rates whereas the macro-finance studies the bidirectional interactions betweens interest rates and macroeconomic variables, since the joint approach recognizes that interest rates and macroeconomic variables evolve jointly over time. A second advantage of macro-finance models is that they allow the behaviour of risk premiums to depend explicitly on macroeconomic conditions but term structure models determine risk premium by covariance of asset returns with the marginal utility consumption. Moreover, there is a strong relationship between economic activity and excess return in bond markets as stated in Cochrane and Piazzesi (2005). A third advantage of macro-finance models is that a substantial component of observed bond yields reacts with the evolution of time varying term of risk premiums. Furthermore, term premiums evolve according to the estimated dynamics of the model to maintain consistency in arbitrage-free efficient markets. The remainder of the chapter proceeds by describing the distinctive macro-finance modelling frameworks offered by Ang and Piazzesi (2003) and Diebold *et al.* (2006) approach and their estimation.

3.3 The Ang and Piazzesi (2003) Approach

The model represents the no-arbitrage term structure approach that determines the movements of the yields of all maturities in response to the movements of the underlying state factors but could not identify the sources of those movements in state factors. On the other hand, macroeconomic VAR model could give a good interpretation of the economic sources of movements in the state variables for a set of given yields, but could not say anything about the response of the entire yield curve against those movements.

The formulation of Ang and Piazzesi (2003) combines the two frameworks to derive the movements of the yield curve and present a no-arbitrage vector autoregressive model. This joint approach captures the joint dynamics of macroeconomic and bond yield factors based on the absence of arbitrage opportunities. The three latent factors interpreted as level, slope and curvature and the two measures of macroeconomic factors such as inflation and output gap are included in the model as state variables. The macroeconomic factors are derived from the principal component of a set of large collection of candidate macroeconomic series of inflation, real activity and latent factors ending up measuring from yields.

The Ang and Piazzesi (2003) model focuses on the joint macro-finance model in a discrete-time affine settings. An affine model is used in this setting to find a connection between macroeconomics and financial descriptions of the term structure of interest rates. The first assumption in this approach is the combination of the classical Taylor (1993) rule and short rate equation used in the affine term structure. The affine term structure model, is often used in different ways. In this work, it will be used to describe any arbitrage-free model, in which bond yields are affine (meaning they are constant and linear) functions of some state vector X_t. The use of "affine" refers to the way yield depend on the state variables. Affine models are thus a special class of term structure models.

Ang and Piazzesi (2003) model demonstrates how the combination of macroeconomic factors into a term-structure model can actually improve model performance. Affine term structure model has three components, namely a collection of observable and/or unobservable state variables, a description of the dynamics of these state variables and a mapping between these state variables and the term structure of interest rate. Two assumptions are maintained in this framework, the independence of latent and macroeconomic factors and in discrete time setting, both latent and macroeconomic factors follow VAR process to describe the state variable vector.

3.3.1 State Dynamics

Based on macro dynamics and short rate discussed before, this section focuses on the unobservable and observable macroeconomic variable which are involved in the collection of inflation π_t and output gap g_t. More specifically on how macroeconomic variables can be linked to the zero coupon bonds price dynamics, by using a simple discrete time Gaussian affine set-up based on the work of Ang and Piazzesi (2003). There are basically three steps to consider when setting up the joint model. The first step consist of specifying the separate models for the macro-variable and for the bond prices or yields. We now define Gaussian Vector Autoregressive (VAR) dynamics for the state variables

$$\begin{bmatrix} \text{output gap} \\ \text{inflation} \\ \text{short rate} \end{bmatrix} = \begin{bmatrix} g_t \\ \pi_t \\ r_t \end{bmatrix} \equiv X_t = \left(X_t^{o'}, X_t^{u'} \right)'.$$

Now, the dynamics of the entire state vector of N risk factors X_t, follows a first-order Gaussian Vector Autoregressive under the historical measure $\mathbb{P}$ given by

$$X_t = \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{P}},$$

where the errors $\epsilon_t^{\mathbb{P}}$ represent the vector of shocks to the observable and unobservable factors respectively and are normally distributed with zero mean and $\mathbb{E}^{\mathbb{P}}[\epsilon_t^{\mathbb{P}} \epsilon_t^{\mathbb{P}'}] = \Sigma_X \Sigma_X'$. The parameters $\mu_X^{\mathbb{P}}$ and $\Phi_X^{\mathbb{P}}$ contain the appropriate conditional means and autocorrelation (block) matrices. This brings us to the derivation of the term-structure model, which uses the state-variable vector to construct a discrete time affine term structure model.

3.3.2 Short-rate Equation

Based on the policy rule recommended by Taylor (1993), the dynamics in r_t are described due to movements in observable macro variable f_t^i and an orthogonal monetary policy shock u_t such that the short-term interest rate is an affine function of the factors:

$$r_t = a_0 + a_1' f_t^0 + u_t, \tag{3.3.1}$$

where the coefficients of the short rate, a_0 and a_1 represents a scaler and a $K \times 1$ vector respectively. Another policy rule proposed by Clarida *et al.* (2000), is a forward-looking version of the Taylor rule, according to this rule, the central bank reacts to expected inflation and expected output gap. This simply means that any variable that forecasts inflation or output gap will enter the right-hand side of the short rate equation. To capture all the information of the underlying macro forecasts, we add lagged macro variable, which are written as

$$X_t^o = (f_t^o, f_{t-1}^o, ..., f_{t-p-1}^o)',$$

for some length p including the lags as arguments in the policy rule. The state vector $X_t^o = [g_t \quad \pi_t]'$ represents the recommended observable macroeconomic state variables, this gives,

$$r_t = b_0 + b_1' X_t^o + u_t. \tag{3.3.2}$$

The purpose of this forward-looking rule is to include forecast errors $f_{t+1}^o - \mathbb{E}_t(f_{t+1}^o)$ in the shock term u_t. This allows for the use of future values of macro variables f_{t+1}^o as arguments on the short rate equation. Affine term structure models are based on short rate Equation (3.3.1) and the assumption on risk premia, see Duffie and Kan (1996). The short rate dynamics is an affine function of latent factors X_t^u which are given by

$$r_t = c_0 + c_1' X_t^u, \tag{3.3.3}$$

where X_t^u is the set of unobservable yield-curve state variables which follow the affine process. By combining Equation (3.3.2) and Equation (3.3.3), we get

$$r_t = \underbrace{\delta_0 + \delta_1' X_t^o}_{\text{Taylor rule}} + \delta_2' X_t^u. \tag{3.3.4}$$

The latent factors X_t^u, are assumed to be independent of the macro factors X_t^o. In this case, the combined short rate dynamics of the term structure model can be interpreted as a version of the Taylor (1993) rule with the errors

$$u_t = \delta_2' X_t^u,$$

being unobserved factors. This is due to using the restrictions from no-arbitrage to separately identify the individual latent factors. The one period short rate r_t is assumed to be an affine function of pricing factors, for all state variables:

$$\begin{aligned} r_t &= \delta_0 + \delta_1' X_t^o + \delta_2' X_t^u, \\ &= \delta_0 + \begin{bmatrix} \delta_1' & \delta_2' \end{bmatrix} \begin{bmatrix} X_t^o \\ \\ X_t^u \end{bmatrix}, \\ &= \delta_0 + \delta_1' X_t, \end{aligned} \tag{3.3.5}$$

with δ_0 a scalar and δ_1 a $K \times 1$ vector. By constraining the coefficient δ_1 to depend only on contemporaneous factor values, we obtain Taylor rule (3.3.1), which consists of macroeconomic variables. We also consider the case where δ_1 is unconstrained, which corresponds to the forward looking Taylor rule incorporating lags. This can be referred to as "Macro Lag Model", because it uses lags of macro variables in the short rate equation. The second key term-structure modelling assumption relates to the states variable dynamics. A VAR(p) specification for X_t^o and VAR(1) approach for X_t^u. To implement this model, we consider only the contemporaneous macroeconomic state variable, VAR(1). The actual state variable dynamics has the form:

$$\begin{bmatrix} X_t^o \\ \\ X_t^u \end{bmatrix} = \begin{bmatrix} \mu_X^o \\ \\ \mu_X^u \end{bmatrix} + \begin{bmatrix} \Phi_X^o & 0_{2\times 3} \\ \\ 0_{3\times 2} & \Phi_X^u \end{bmatrix} \begin{bmatrix} X_{t-1}^o \\ \\ X_{t-1}^u \end{bmatrix} + \begin{bmatrix} \epsilon_t^o \\ \\ \epsilon_t^u \end{bmatrix},$$

$$X_t = \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_{t-1} + \epsilon_t^{\mathbb{P}},$$

where $\mu_X^u \equiv 0$ and,

$$\epsilon_t = \begin{bmatrix} \epsilon_t^o \\ \\ \epsilon_t^u \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \Omega^o & 0_{2\times 3} \\ \\ 0_{3\times 2} & \mathcal{I}_{3\times 3} \end{bmatrix} = \Omega \right). \tag{3.3.6}$$

Equation (3.3.6) shows that the dynamics of the latent state variables are determined independently of the macroeconomic variables. The current model does not permit the macro economic variable to improve the description of the latent yield curve state variables. This can lead to different expectations that are associated with flat term-structure with low inflation and high output than a flat term-structure with high inflation and low output. This shows that the influence of the macroeconomic factors on the term structure arises elsewhere in the model.

3.3.3 The Market Price of Risk

We consider the final third term structure modelling assumption. The market price of risk, which is arising because of the uncertainty $\epsilon_t^{\mathbb{P}}$. The link between the risk-neutral distribution $\mathbb{Q}$ and the physical distribution $\mathbb{P}$ is given by the time varying price of risk λ_t. The $N-$dimensional vector of risk prices is parametrized as affine process in the pricing factors:

$$\lambda_t = \lambda_0 + \lambda_1 X_t, \tag{3.3.7}$$

with parameters in $N-$vector λ_0 and $N \times N$ matrix λ_1 which corresponds to lagged macro variables, are set to zero. If λ_0 and λ_1 takes any value other than zero, we will then allow for constant or time varying risk premia. This specification is introduced in Duffee (2002). The physical innovations $\epsilon_t^{\mathbb{P}}$, is a vector martingale difference sequence (m.d.s.) under measure $\mathbb{P}$, are related to the innovations under measure $\mathbb{Q}$ by

$$\epsilon_t^{\mathbb{Q}} = \epsilon_t^{\mathbb{P}} + \lambda_{t-1}.$$

The risk-neutral innovations, while being m.d.s. under measure $\mathbb{Q}$, can have non-zero mean and be predictable under measure $\mathbb{P}$, depending on the risk price specification.

3.3.4 The Pricing Kernel

To develop the term structure model under arbitrage-free considerations that guarantees the existence of the equivalent martingale measure $\mathbb{Q}$, the price of an asset V_t that does not pay any dividends in future, satisfies

$$V_t = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t)V_{t+1}],$$

where expectation is taken under risk neutral measure $\mathbb{Q}$. Ang and Piazzesi (2003) further assumed that the existence of the pricing kernel is of the form:

$$M_{t+1} = e^{-r_t}[\frac{\varphi_{t+1}}{\varphi_t}], \tag{3.3.8}$$

where r_t is the short rate, φ_t is a positive martingale under probability measure $\mathbb{Q}$. The Radon-Nikodým derivative, which converts the risk neutral measure to the data generating measure, is denoted by φ_{t+1}. Thus, we have for any random variable Z_{t+1} that

$$\mathbb{E}_t^{\mathbb{Q}}(Z_{t+1}) = \mathbb{E}_t^{\mathbb{P}}\{Z_{t+1}[\frac{\varphi_{t+1}}{\varphi_t}]\}.$$

According to Girsanov's theorem, this derivative

$$\mathbb{E}_t^{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = \prod_{j=1}^{n} \exp(-\frac{1}{2}\lambda'_{t+j-1}\lambda_{t+j-1} - \lambda'_{t+j-1}\epsilon_{t+j}^{\mathbb{P}}),$$

where φ_{t+1} follows the log-normal process and λ_t is the time varying price of risk associated with the uncertainty of the errors $\epsilon_t^{\mathbb{P}}$, which are normally distributed with zero mean, $\mathbb{E}[\epsilon_t^{\mathbb{P}}] = 0$ and $\mathbb{E}[\epsilon_t^{\mathbb{P}}\epsilon_t^{\mathbb{P}\prime}] = \Sigma_X\Sigma_X'$. It follows that

$$\frac{\varphi_{t+1}}{\varphi_t} = \exp(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}), \tag{3.3.9}$$

substituting both the short rate (3.3.5) and Equation (3.3.9) to Equation (3.3.8), we have

$$M_{t+1} = \exp(-\frac{1}{2}\lambda_t'\lambda_t - \delta_0 - \delta_1'X_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}), \tag{3.3.10}$$

where $\epsilon_{t+1}^{\mathbb{P}} \sim N(0, I)$.

3.3.5 Bond Prices

The term structure modelling determines the price of zero-coupon bonds. These bonds pay a terminal payoff, usually normalised to 1-unit, without risk of default. The pricing kernel between bonds of different maturities links the n^{th} period bond price with that of a bond of maturity $n-1$ of the next period. This is given as

$$P_t^n = \mathbb{E}_t^{\mathbb{P}}\left[M_{t+1}P_{t+1}^{n-1}\right], \tag{3.3.11}$$

which is the current price of the zero-coupon bond, expressed as the conditional expectation of discounted future price of the bond. The factor M_{t+1} is given as the function of the short rate and the risk perceived by the market,

$$M_{t+1} = \exp(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}). \tag{3.3.12}$$

By iterating Equation (3.3.11) to the maturity n of the bond price, leads to

$$P_t^n = \mathbb{E}_t^{\mathbb{P}}\left[M_{t+1}M_{t+2}...M_{t+n}P_{t+n}^0\right] = \mathbb{E}_t^{\mathbb{P}}[M_{t+1}M_{t+2}...M_{t+n}\cdot \underbrace{1}_{P_{t+n}^0=1}].$$

Here, $P_{t+n}^0 = 1$ for all t, because the price of a zero coupon bond at maturity is unity. We will make use of the randomness ϵ_{t+1} that arises in the dynamics of the Radon-Nikodým derivative and the state-variable system in the derivation of the recursion relation used to represent the price of an arbitrary pure-discount bond P_t^n. The idea is that the pure-discount bond is an exponential-affine function of the state variables. Using the pricing kernel which fixes the Radon-Nikodým derivative and the market price of risk to determine the pure-discount bond price. Thus, we have that,

$$P_t^n = \mathbb{E}_t^{\mathbb{P}}\{\exp[-\sum_{i=0}^{n-1}(r_{t+i} + \frac{1}{2}\lambda_{t+i}'\lambda_{t+i} + \lambda_{t+i}'\epsilon_{t+1+i})]\}. \tag{3.3.13}$$

Since our attention is restricted to affine term structure, the no-arbitrage recursive relations is derived such that the price of bonds are exponential affine functions of state variables, denoted as

$$P_t^n = \exp(A_n + B_n'X_t). \tag{3.3.14}$$

The price of an $(n+1)-$period zero coupon bond can be computed recursively by using the pricing kernel under physical measure $\mathbb{P}$. This becomes,

$$\begin{aligned}
P_t^{n+1} &= \mathbb{E}_t^{\mathbb{P}}[P_{t+1}^n M_{t+1}] \\
&= \mathbb{E}_t^{\mathbb{P}}\left[\exp\{A_n + B_n' X_{t+1}\}\exp\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}\}\right] \\
&= \mathbb{E}_t^{\mathbb{P}}\left[\exp\{A_n + B_n'(\mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t + \Sigma_X \epsilon_{t+1}^{\mathbb{P}})\}\exp\{-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}\}\right] \\
&= \exp\left[A_n - \delta_0 + B_n'\mu_X^{\mathbb{P}} + (B_n'\Phi_X^{\mathbb{P}} - \delta_1')X_t - \frac{1}{2}\lambda_t'\lambda_t\right] \\
&\times \mathbb{E}_t^{\mathbb{P}}\left[\exp((B_n'\Sigma_X - \lambda_t')\epsilon_{t+1}^{\mathbb{P}})\right]. \qquad (3.3.15)
\end{aligned}$$

We used the properties of a log-normal random variable $\epsilon_{t+1}^{\mathbb{P}}$, to solve the conditional expectation part of the pure discount price. Since the random variable $\epsilon_{t+1}^{\mathbb{P}}$ is independent of $\mathcal{F}_t$, then the expectation term becomes

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{P}}\left\{\exp[(B_n'\Sigma_X - \lambda_t')\epsilon_{t+1}^{\mathbb{P}}]\right\} &= \mathbb{E}_t^{\mathbb{P}}\left[\exp\{\langle B_n\Sigma_X', \epsilon_{t+1}^{\mathbb{P}}\rangle - \langle\lambda_t, \epsilon_{t+1}^{\mathbb{P}}\rangle\}\right] \\
&= \mathbb{E}_t^{\mathbb{P}}\left[\exp\langle B_n\Sigma_X' - \lambda_t, \epsilon_{t+1}^{\mathbb{P}}\rangle\right] \\
&= \exp[\frac{1}{2}\langle B_n\Sigma_X' - \lambda_t, B_n\Sigma_X' - \lambda_t\rangle] \\
&= \exp[\frac{1}{2}\langle B_n\Sigma_X', B_n\Sigma_X'\rangle + \frac{1}{2}\langle\lambda_t, \lambda_t\rangle - \langle B_n\Sigma_X', \lambda_t\rangle] \\
&= \exp[\frac{1}{2}(B_n'\Sigma_X\Sigma_X' B_n - 2B_n'\Sigma_X\lambda_t + \lambda_t'\lambda_t)]. \\
& \qquad (3.3.16)
\end{aligned}$$

If we substitute the above expectation Equation (3.3.16) to the discount bond price Equation (3.3.15) to solve the recursive form of the pure discount bond price,

$$\begin{aligned}
P_t^{n+1} &= \exp\left[A_n - \delta_0 + B_n'\mu_X^{\mathbb{P}} + (B_n'\Phi_X^{\mathbb{P}} - \delta_1')X_t - \frac{1}{2}\lambda_t'\lambda_t\right] \\
&\quad \times \exp\left[\frac{1}{2}(B_n'\Sigma_X\Sigma_X' B_n - 2B_n'\Sigma_X\lambda_t + \lambda_t'\lambda_t)\right] \\
&= \exp\left[A_n - \delta_0 + B_n'\mu_X^{\mathbb{P}} + \frac{1}{2}B_n'\Sigma_X\Sigma_X' B_n - B_n'\Sigma_X\lambda_0\right] \\
&\quad \times \exp\left[(B_n'(\Phi_X^{\mathbb{P}} - \Sigma_X\lambda_1) - \delta_1')X_t\right]. \qquad (3.3.17)
\end{aligned}$$

This expression allows us to use the current value of the state variables as well as our model parameters to recursively construct the term structure of interest rate. Equation (3.3.15) illustrates how the macroeconomics variables influence the term structure of interest rates. The variables describe the cross section of the term structure at a specific point in time. Their influence on the yield curve occurs through mapping. The no-arbitrage imposes the following structure on the coefficients of the measurement equation, for $n > 1$:

$$\begin{aligned} A_{n+1} &= A_n + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_n'\Sigma_X\Sigma_X'B_n + A_1 && (3.3.18) \\ B_{n+1}' &= B_n'(\Phi_X^{\mathbb{P}} - \Sigma_X\lambda_1) + B_1' && (3.3.19) \\ &= B_1'\sum_{i=0}^{n}(\Phi_X^{\mathbb{P}} - \Sigma_X\lambda_1)^i, && (3.3.20) \end{aligned}$$

with $A_1 = -\delta_0$ and $B_1 = -\delta_1$. These difference equations can be derived by recursive form using Equation (3.3.11), as shown in Appendix (A.2). On the basis of the affine term-structure assumptions, it can be shown that affine functions of the state variables for the continuously compounded yields y_t^n, on an n-period zero coupon bond is also dependent on the three factors, which are the short rate, the risk price and the variance of the factors. That is, in terms of bond yields, this relationship is equivalent to,

$$\begin{aligned} y_t^n &= -\frac{\ln P_t^n}{n} \\ &= -\frac{1}{n}\ln \mathbb{E}_t^{\mathbb{P}}\{\exp[-\sum_{i=0}^{n-1} r_{t+i}]\} \\ &= -\frac{1}{n}\ln \mathbb{E}_t^{\mathbb{P}}\{\exp[-\sum_{i=0}^{n-1}(\delta_0 + \delta_1' X_{t+i})]\} \\ &= \bar{A}_n + \bar{B}_n' X_t, && (3.3.21) \end{aligned}$$

where $\bar{A}_n = -\frac{1}{n}A_n(\lambda_0, \lambda_1, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, \Sigma_X, n)$ and $\bar{B}_n = -\frac{1}{n}B_n(\lambda_0, \lambda_1, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, \Sigma_X, n)$ solve recursive equations under the actual measure. The non-linear functions of $\bar{A}_n$ and $\bar{B}_n$ make the yield equations consistent with each other for different values of t. The functions also make the yield equations consistent with the state dynamics. Zero-coupon bond yields are therefore affine functions of the state variable X_t, which are also taken to be the observation equation of the state space system given by observable variables. Once the coefficients, $(A_1 = -\delta_0, B_1 = -\delta_1)$ on the short rate equation are fixed, all the other coefficients for longer maturity yields are determined by the parameters in the state equation and the risk pricing equation:

$$B_{n+1} = B_1 + (\Phi_X^{\mathbb{P}} - \lambda_1\Sigma_X)' B_n = [\sum_{i=0}^{n}(\Phi_X^{\mathbb{P}\prime} - \lambda_1'\Sigma_X)^i]B_1.$$

$$A_{n+1} = (n+1)A_1 + \sum_{i=0}^{n} B^{(i)}, \quad \text{where} \quad B^{(i)} = B_i'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_i'\Sigma_X\Sigma_X' B_i,$$

as proved in (A.2.1) and (A.2.2). It is worth pointing out that Equation (3.3.21) is not the expectation of future interest rates because of a convexity term. The convexity term arises because yields are the expectations of a nonlinear function of future interest rates, which are stochastic, and given by:

$$\begin{aligned} c_t^n &= y_{t,n}^{\mathbb{P}} - y_{t,n}^{EH} \\ &= -\frac{1}{n}\ln \mathbb{E}_t^{\mathbb{P}}\{\exp[-\sum_{i=0}^{n-1} r_{t+i}]\} - \frac{1}{n}\ln \mathbb{E}_t^{\mathbb{P}}\{\sum_{i=0}^{n-1} r_{t+i}\} \\ &= A_n^{\mathbb{P}} - A_n^{EH} + (B_n^{\mathbb{P}} - B_n^{EH})X_t. \end{aligned}$$

The "Expectation Hypothesis" (EH) yields are equivalent to treating interest rates as non-stochastic, which can be done by simply imposing $\Sigma = 0$ in the same recursions:

$$A_n^{EH} = -\frac{1}{n}A_n(\lambda_0, \lambda_1, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, 0, n) \tag{3.3.22}$$

$$B_n^{EH} = -\frac{1}{n}B_n(\lambda_0, \lambda_1, \mu^{\mathbb{Q}}, \Phi^{\mathbb{Q}}, 0, n). \tag{3.3.23}$$

From Equation (3.3.22) and Equation (3.3.23), spot term premia tP_t, are easily calculated as an affine function of the state as well,

$$tP_t^n = y_{t,n}^{\mathbb{Q}} - y_{t,n}^{\mathbb{P}} \tag{3.3.24}$$

$$= (A_n^{\mathbb{Q}} - A_n^{\mathbb{P}}) + (B_n^{\mathbb{Q}} - B_n^{\mathbb{P}})X_t. \tag{3.3.25}$$

We can also easily calculate forward rates starting in n periods in the future maturing in $m > n$ periods,

$$y_t^{m-n} = (m-n)^{-1}(my_t^m - ny_t^n),$$

which becomes,

$$y_t^{m-n} = A_{m-n}^{\mathbb{Q}} + B_{m-n}^{\mathbb{Q}}X_t,$$

with

$$\begin{aligned} A_{m-n}^{\mathbb{Q}} &= (m-n)^{-1}[mA_m^{\mathbb{Q}} - nA_n^{\mathbb{Q}}] \\ B_{m-n}^{\mathbb{Q}} &= (m-n)^{-1}[mB_m^{\mathbb{Q}} - nB_n^{\mathbb{Q}}]. \end{aligned}$$

Forward term premia can be calculated by substituting the corresponding forward yield in Equation (3.3.24). We define the excess return

$$xr_{t+1}^n = \ln P_{t+1}^{n-1} - \ln P_t^n - r_t,$$

as the return on buying an $n-$period real bond at time t for P_t^n and selling it at time $t+1$ for P_{t+1}^{n-1} in excess of the short rate. Now, the log expected excess returns for a bond of maturity n is given by

$$\begin{aligned} \ln \mathbb{E}_t^{\mathbb{P}}[xr_{t+1}^n] &\equiv \ln \mathbb{E}_t^{\mathbb{P}} P_{t+1}^{n-1} - \ln \mathbb{E}_t^{\mathbb{P}} P_t^n - \ln \mathbb{E}_t^{\mathbb{P}} r_t \\ &= A_{n-1} + B_{n-1}\left(\mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t\right) + \underbrace{1/2 B_{n-1}\Sigma_X \Sigma_X' B_{n-1}'}_{\mathbb{E}_t^{\mathbb{P}}(B_{n-1}\Sigma_X \epsilon_{t+1}^{\mathbb{P}})} \\ &\quad -(A_n + B_n X_t) - r_t. \end{aligned}$$

Substituting the pricing coefficients A_n and B_n in the expression, gives the following,

$$\begin{aligned}
\ln \mathbb{E}_t^{\mathbb{P}}[xr_{t+1}^n] &= A_{n-1} + B_{n-1}(\mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t) + \frac{1}{2} B_{n-1}\Sigma_X \Sigma_X' B_{n-1}^{'} \\
&\quad -(-\delta_0 + A_{n-1} + \frac{1}{2} B_{n-1}\Sigma_X \Sigma_X' B_{n-1}^{'} + B_{n-1}\mu_X^{\mathbb{Q}}) \\
&\quad -(-\delta_0 + B_{n-1}\Phi_X^{\mathbb{Q}})X_t - r_t \\
&= B_{n-1}[(\mu_X^{\mathbb{P}} - \mu_X^{\mathbb{Q}}) + (\Phi_X^{\mathbb{P}} - \Phi_X^{\mathbb{Q}})X_t] \\
&= B_{n-1}\Sigma_X(\lambda_0 + \lambda_1 X_t) \\
&= B_{n-1}\Sigma_X \lambda_t.
\end{aligned}$$

The excess return measures the prices of risk λ_t with the quantity of risk exposure on bonds. The loading $B_{n-1}\Sigma_X$ represents the amount of exposure to each factor risk.

3.3.6 Cross-Equation Restrictions

When estimating a system of yield equations in a way that ensures no-arbitrage, the cross-equation restrictions have to be derived from parameters that describe the state dynamics and risk premia. Although the model is affine in the state vector X_t, but the parameter functions, A_n and B_n are non-linear functions. Using ordinary least squares (OLS) or maximum likelihood to estimate is not possible because the density of yields is not available in closed form. New econometric methods have been developed to solve these estimation problems as shown in Bekaert and Hodrick (2001).

3.3.7 Advantages of Cross-Equation Restrictions

Cross-equation restrictions ensure that the yield dynamics are consistent. The parameters A_n and B_n make yield curve equations consistent with each other in the cross section and the time series.

Term structure models allow us to separate risk premia from expectations about future short-rates. These models are therefore, key to understanding to what extent investors think of long bonds as safe investments. Sargent (1979) is amongst the ones who explore the expectation hypothesis, under which expected excess bond returns are zero. The extension of expectation hypothesis have been developed under which expected excess returns are constant. The comparison of the two tests about the ratio of the likelihood function with or without restrictions implied by the expectation hypothesis shows that expected returns on long bonds are on average higher than on short bonds and that they are time varying equations, this leads to the necessity of cross-equations to model these risk premia.

Unrestricted regression imply that the number of variables needed to describe the yield curve, corresponds with the number of yields in the regression. From the work of Litterman and Scheinkman (1991), based on factor decompositions of variance-covariance matrix of yield changes reveals that over 97% of variance is attributable to just three principal components, namely, level, slope and curvature according to how shocks to these factors affect the yield curve.

The number of estimated parameters in unrestricted regressions is usually large, imposing no-arbitrage restrictions improves the efficiency of these estimates and out-of-sampling forecasting of yields, as shown in Ang and Piazzesi (2003).

3.3.8 Change of Measure - Risk Neutral Dynamics

The results in this section show what kind of term structure factors follow under risk neutral probability measure $\mathbb{Q}$, which is assumed to be equivalent to the physical measure $\mathbb{P}$[1]. Following the assumptions of the term structure, the risk neutral dynamics are given as follows:

Suppose the transition equation for an $r \times 1$ vector of risk factors X_t under risk neutral measure $\mathbb{Q}$ follows a first-order homoskedastic Gaussian VAR(1):

$$X_{t+1} = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t + \Sigma_X \epsilon_t^{\mathbb{Q}},$$

where $\epsilon_t^{\mathbb{Q}} \overset{\text{i.i.d}}{\sim} N(0, I_r)$, under the risk adjusted probabilities and $\mathbb{E}^{\mathbb{Q}}(\epsilon_m^{\mathbb{Q}} \epsilon_n^{\mathbb{Q}\prime}) = 0$, $m \neq n$. The lower triangular Σ_X summarizes unpredictability of factors and risk premia is a function of Σ_X. For any current information set Ω_t,

$$X_{t+1}|\Omega_t \quad \overset{\mathbb{Q}}{\sim} \quad N(\mu_t^{\mathbb{Q}}, \Sigma_X \Sigma_X'),$$

where the dynamics of the state vector of factors, X_t, are characterised by the risk neutral vector of constant $\mu_X^{\mathbb{Q}}$ and the autoregressive matrix $\Phi_X^{\mathbb{Q}}$. VAR parameters that are describing physical and risk neutral dynamics are linked to each other through the relationship:

$$\mu_X^{\mathbb{Q}} = \mu_X^{\mathbb{P}} - \Sigma_X \lambda_0, \tag{3.3.26}$$

$$\Phi_X^{\mathbb{Q}} = \Phi_X^{\mathbb{P}} - \Sigma_X \lambda_1. \tag{3.3.27}$$

An asset that pays $X_{i,t+1}$ dollars in future, for each of $i = 1, ..., r$ assets today, the price of a risk neutral investor would be $e^{-r_t}\mu_t^{\mathbb{P}}$, and if risk averse the price would be $e^{-r_t}\mu_t^{\mathbb{Q}}$. The mean parameters of both probability measures

$$\begin{aligned} \underbrace{\mu_t^{\mathbb{Q}}}_{(r\times1)} &= \underbrace{\mu_t^{\mathbb{P}}}_{(r\times1)} - \underbrace{\Sigma_X \lambda_t}_{(r\times r)(r\times1)}, \text{ and} \\ \mu_t^{\mathbb{P}} &= \mathbb{E}_t^{\mathbb{P}}(X_{t+1}) = \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t, \end{aligned}$$

[1] Two measures $\mathbb{P}$ and $\mathbb{Q}$ are said to be equivalent if they agree on sets of probability zero.

with the parameter λ_{it}, giving the market price of factor i risk. When implementing the model under the actual measure and risk-neutral measure, the stochastic discount factor M_{t+1}, defines the change of probability measure between the two measures, in which under risk-neutral measure this implies

$$\begin{aligned} e^{-r_t}\mu_t^{\mathbb{Q}} &= \mathbb{E}_t(M_{t+1}X_{t+1}) \\ &= \int M_{t+1}X_{t+1}\psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')dX_{t+1}. \end{aligned}$$

If $M_{t+1}\psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X') = e^{-r_t}\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X')$, then by using the multivariate normal distribution the expression becomes,

$$\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X') = \frac{|\Sigma_X|^{-1}}{(2\pi)^{r/2}}\exp\left[-\frac{(X_{t+1}-\mu_t^{\mathbb{Q}})'(\Sigma_X\Sigma_X')^{-1}(X_{t+1}-\mu_t^{\mathbb{Q}})}{2}\right],$$

where,

$$\begin{aligned} (X_{t+1}-\mu_t^{\mathbb{Q}})'(\Sigma_X\Sigma_X')^{-1}(X_{t+1}-\mu_t^{\mathbb{Q}}) &= (X_{t+1}-\mu_t^{\mathbb{P}}+\Sigma_X\lambda_t)'(\Sigma_X\Sigma_X')^{-1} \\ &\quad \times(X_{t+1}-\mu_t^{\mathbb{P}}+\Sigma_X\lambda_t) \\ &= (X_{t+1}-\mu_t^{\mathbb{P}})'(\Sigma_X\Sigma_X')^{-1}(X_{t+1}-\mu_t^{\mathbb{P}}) \\ &\quad +\lambda_t'\lambda_t + 2\lambda_t'\Sigma_X^{-1}(X_{t+1}-\mu_t^{\mathbb{P}}). \end{aligned}$$

Since

$$X_{t+1} = \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}}X_t + \Sigma_X\epsilon_{t+1}^{\mathbb{P}}, \tag{3.3.28}$$

$$\Sigma_X^{-1}(X_{t+1}-\mu_t^{\mathbb{P}}) = \epsilon_{t+1}^{\mathbb{P}} \tag{3.3.29}$$

then, this concludes with the following:

$$\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X') = \psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')\exp[-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}].$$

The Radon-Nikodým derivative, connects the densities under the actual measure and risk-neutral measure, and takes the following form:

$$\frac{\psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')}{\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X')} = \left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right](X_{t+1};\lambda_t) = \exp(\frac{1}{2}\lambda_t'\lambda_t + \lambda_t'\epsilon_{t+1}^{\mathbb{P}}),$$

where the change of measure gives the one period stochastic discount factor of the form:

$$\begin{aligned} M_{t+1} &= \frac{e^{-r_t}\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X')}{\psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')} \\ &= \exp[-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}]. \end{aligned} \tag{3.3.30}$$

The risk-neutral conditional Laplace transform X_{t+1}, under measure $\mathbb{Q}$ is given by

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}\{\exp(u'X_{t+1})\} &= \int \exp(u'X_{t+1})\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X')dX_{t+1} \\
&= \int \exp(u'X_{t+1})\frac{\psi(X_{t+1};\mu_t^{\mathbb{Q}},\Sigma_X\Sigma_X')dX_{t+1}}{\psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')dX_{t+1}} \\
&\quad \times \psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')dX_{t+1} \\
&= \int \exp(u'X_{t+1})\exp(-\frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}}) \\
&\quad \times \psi(X_{t+1};\mu_t^{\mathbb{P}},\Sigma_X\Sigma_X')dX_{t+1} \\
&= \mathbb{E}_t^{\mathbb{P}}[\exp(u'\mu_X^{\mathbb{P}} + u'\Phi_X^{\mathbb{P}}X_t + u'\Sigma_X\epsilon_{t+1}^{\mathbb{P}} - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\epsilon_{t+1}^{\mathbb{P}})] \\
&= \exp(u'\mu_X^{\mathbb{P}} + u'\Phi_X^{\mathbb{P}}X_t - \frac{1}{2}\lambda_t'\lambda_t)\mathbb{E}_t^{\mathbb{P}}\left[\exp(u'\Sigma_X\epsilon_{t+1}^{\mathbb{P}} - \lambda_t'\epsilon_{t+1}^{\mathbb{P}})\right].
\end{aligned}$$

Since ϵ_{t+1} is independent of $\mathcal{F}_t$

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{P}}\left\{\exp(u'\Sigma_X\epsilon_{t+1}^{\mathbb{P}} - \lambda_t'\epsilon_{t+1}^{\mathbb{P}})\right\} &= \mathbb{E}_t^{\mathbb{P}}\left[\exp\{\langle\Sigma_X'u,\epsilon_{t+1}^{\mathbb{P}}\rangle - \langle\lambda_t,\epsilon_{t+1}^{\mathbb{P}}\rangle\}\right] \\
&= \mathbb{E}_t^{\mathbb{P}}\left[\exp\langle\Sigma_X'u - \lambda_t,\epsilon_{t+1}^{\mathbb{P}}\rangle\right] \\
&= \exp\left[\frac{1}{2}\langle\Sigma_X'u - \lambda_t,\Sigma_X'u - \lambda_t\rangle\right] \\
&= \exp\left[\frac{1}{2}\langle\Sigma_X'u,\Sigma_X'u\rangle + \frac{1}{2}\langle\lambda_t,\lambda_t\rangle - \langle\Sigma_X'u,\lambda_t\rangle\right] \\
&= \exp[\frac{1}{2}u'\Sigma_X\Sigma_X'u + \frac{1}{2}\lambda_t^2 - u'\Sigma_X\lambda_t].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}}\left\{\exp(u'X_{t+1})\right\} &= \exp[u'\mu_X^{\mathbb{P}} + u'\Phi_X^{\mathbb{P}}X_t]\exp[\frac{1}{2}u'\Sigma_X\Sigma_X'u - u'\Sigma_X\lambda_t] \\
&= \exp[u'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_t + \Phi_X^{\mathbb{P}}X_t) + \frac{1}{2}u'\Sigma_X\Sigma_X'u],
\end{aligned}$$

which is said to have a conditional moment-generating function of a multivariate normal distribution with the underlying mean and variance respectively,

$$\mu_t^{\mathbb{P}} = \mu_X^{\mathbb{P}} - \Sigma_X\lambda_t + \Phi_X^{\mathbb{P}}X_t \qquad \text{and} \qquad \sigma^2 = \Sigma_X\Sigma_X'.$$

Zero-coupon bonds are evaluated under no-arbitrage assumptions, this implies the existence of a risk-neutral measure $\mathbb{Q}$, equivalent to $\mathbb{P}$. If risk neutrality holds, this means that the price of risk $\lambda_0 = \lambda_1 = 0$, and equation (3.3.30) would simply reduce to $M_t = e^{-r_t}$ and the price of a bond would become,

$$\begin{aligned}
P_t^n &= \mathbb{E}_t^{\mathbb{Q}}\left[\exp(-r_t)P_{t+1}^{n-1}\right] \\
&= \mathbb{E}_t^{\mathbb{Q}}[\exp(-\sum_{i=0}^{n-1} r_{t+i})].
\end{aligned}$$

Investors act the same way a risk neutral investor would who thought the actors follow the $\mathbb{Q}-$measure distribution, that is, the no-arbitrage bond prices under risk-neutral measure,

$$P_t^n = \mathbb{E}_t^{\mathbb{Q}}\left[\exp(-r_t)P_{t+1}^{n-1}\right] = \mathbb{E}_t^{\mathbb{P}}\left[M_{t+1}P_{t+1}^{n-1}\right]. \tag{3.3.31}$$

Assume that the nominal bond prices are exponential affine functions of the state variable,

$$P_t^n = \exp\left(A_n + B_n' X_t\right). \tag{3.3.32}$$

Recall that if,

$$z \sim N(\mu, \Sigma_X^2)$$

then,

$$\mathbb{E}(e^z) = \exp(\mu + \Sigma_X^2/2).$$

Thus, we obtain the recursion:

$$\begin{aligned}
\ln P_t^n &= \ln \mathbb{E}_t^{\mathbb{Q}}\left[\exp(r_t + A_{n-1} + B_{n-1}'X_{t+1})\right] \\
&= \ln \mathbb{E}_t^{\mathbb{Q}}\left[\exp\{-(\delta_0 + \delta_1 X_t) + A_{n-1} + B_{n-1}'(\mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t + \Sigma_X \epsilon_{t+1}^{\mathbb{Q}})\}\right] \\
&= -(\delta_0 + \delta_1 X_t) + A_{n-1} + B_{n-1}'(\mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t) \\
&\quad + \ln \mathbb{E}_t^{\mathbb{Q}}[\exp\{B_{n-1}'\Sigma_X \epsilon_{t+1}^{\mathbb{Q}}\}] \\
A_n + B_n' X_t &= -(\delta_0 + \delta_1 X_t) + A_{n-1} + B_{n-1}'(\mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t) + \frac{1}{2} B_{n-1}'\Sigma_X \Sigma_X' B_{n-1} \\
&= -\delta_0 + A_{n-1} + B_{n-1}'\mu_X^{\mathbb{Q}} + \frac{1}{2} B_{n-1}'\Sigma_X \Sigma_X' B_{n-1} \\
&\quad + (-\delta_1 + B_{n-1}'\Phi_X^{\mathbb{Q}})X_t.
\end{aligned}$$

The pricing coefficients are given by the difference equations,

$$\begin{aligned}
A_n &= -\delta + A_{n-1} + B_{n-1}'\mu_X^{\mathbb{Q}} + \frac{1}{2} B_{n-1}'\Sigma_X \Sigma_X' B_{n-1} \\
B_n' &= -\delta_1 + B_{n-1}'\Phi_X^{\mathbb{Q}},
\end{aligned}$$

with the initial conditions $A_0 = B_0 = 0$. Following the assumptions that zero-coupon bonds are linear in the state variable, the nominal risk-free yield is given by

$$\begin{aligned}
y_t^n &= A_n^{\mathbb{Q}} + B_n^{\mathbb{Q}'} X_t \\
&= A_n^{\mathbb{Q}} + \{B_n^{\mathbb{Q}'}\}_z z_t + \{B_n^{\mathbb{Q}}\}_x x_t & (3.3.33) \\
&\triangleq \underbrace{A_n^{\mathbb{P}} + B_n^{\mathbb{P}'} X_t}_{\text{Short rate expectations}} + \underbrace{A_n^{TP} + B_n^{TP'} X_t}_{\text{Term premium}}, & (3.3.34)
\end{aligned}$$

where the coefficients A_n and B_n solve the recursive equations with boundary conditions $A_1 = -\delta_0$ and $B_1 = -\delta_1$, with the short rate equation given by $y_t^1 = r_t$. The behaviour of the yields and bond spreads is captured by the state vector $X_t = (z'_t, x'_t)$. The vector of macroeconomic variables contains Gross Domestic Product (GDP) growth and inflation, is given by $z_t = (g_t, \pi_t)'$. The variable x_t, contains the vector of latent factors in the model and can contain lags of z_t, any other macro variables not explicitly modeled, or of the unknown variables. In this case, X_t, fully reflects the available information at time t. The pricing iterations, following from Equation (3.3.18) and Equation (3.3.20) can therefore be restated in a similar manner in terms of the risk neutral VAR parameters as

$$\begin{aligned} A_{n+1} &= A_n + B'_n \mu_X^{\mathbb{Q}} + \frac{1}{2} B'_n \Sigma_X \Sigma'_X B_n + A_1 && (3.3.35) \\ B'_{n+1} &= B'_n \Phi_X^{\mathbb{Q}} + B'_1 \\ &= B'_1 \sum_{i=0}^{n} (\Phi_X^{\mathbb{Q}})^i. && (3.3.36) \end{aligned}$$

The market price of risk λ_t, can be written as a function of two sets of VAR parameters:

$$\lambda_t = \lambda_0 + \lambda_1 X_t = \Sigma_X^{-1}[(\mu_X^{\mathbb{P}} - \mu_X^{\mathbb{Q}}) + (\Phi_X^{\mathbb{P}} - \Phi^{\mathbb{Q}})X_t].$$

The model has a large number of potential parameters to estimate and calculate the log price of any bond P_t^n. The vector of parameters is given in terms of

$$\theta = \{\delta_0, \delta_1, \Phi_X^{\mathbb{P}}, \Phi_X^{\mathbb{Q}}, \mu_X^{\mathbb{Q}}, \mu_X^{\mathbb{P}}, \Sigma_X, \lambda_0, \lambda_1\},$$

where Φ_X and Σ_X are sparse block matrices. If $\mu_X^{\mathbb{Q}} = \mu_X^{\mathbb{P}}$ and $\Phi_X^{\mathbb{Q}} = \Phi_X^{\mathbb{P}}$ this would correspond to the expectations hypothesis of the term structure. Suppose there is an $r =$ (level, slope, curvature) $\times$ 1, vector X_t of possibly unobservable factors that captures every effects for determining interest rates. Assume that the price of risk is also affine function of the form:

$$\underbrace{\lambda_t}_{(r\times 1)} = \underbrace{\lambda_0}_{(r\times 1)} + \underbrace{\lambda_1 X_t}_{(r\times r)(r\times 1)},$$

then,

$$\begin{aligned} \mu_t^{\mathbb{Q}} &= \mu_t^{\mathbb{P}} - \Sigma_X \lambda_t \\ &= \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t - \Sigma_X \lambda_0 - \Sigma_X \lambda_1 X_{t+1} \\ &= \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t \\ \mu_X^{\mathbb{Q}} &= \mu_X^{\mathbb{P}} - \Sigma_X \lambda_0 \\ \Phi_X^{\mathbb{Q}} &= \Phi_X^{\mathbb{P}} - \Sigma_X \lambda_1. \end{aligned}$$

The $\mathbb{P}-$measure dynamics:

$$\begin{aligned} X_{t+1} &= \mu_X^{\mathbb{P}} + \Phi_X^{\mathbb{P}} X_t + \Sigma_X \epsilon_{t+1}^{\mathbb{P}} \\ \epsilon_{t+1}^{\mathbb{P}} &\overset{\mathbb{P}}{\sim} N(0, I_r). \end{aligned}$$

The $\mathbb{Q}-$measure dynamics:

$$\begin{aligned} X_{t+1} &= \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t + \Sigma_X \epsilon_{t+1}^{\mathbb{Q}} \\ \epsilon_{t+1}^{\mathbb{Q}} &\overset{\mathbb{Q}}{\sim} N(0, I_r). \end{aligned}$$

3.3.9 Forward Rates

The forward rates can be used to price bonds in response to future outcome of the short rate. Yields are derived by focusing on the expected future short rate under measure $\mathbb{Q}$. The rates $f_t^{\mathbb{Q}}(n) = \mathbb{E}_t^{\mathbb{Q}}(r_{t+n})$, is denoted as a forward rate received at time t for over a loan beginning at time $[t+1, t+n+1]$, which is contrary to the well known bond pricing literature of Ang and Piazzesi (2003). The risk adjusted expectations of the future short-term interest rate are expressed as follows

$$\begin{aligned} f_t^{\mathbb{Q}}(n) = \mathbb{E}_t^{\mathbb{Q}}(r_{t+n}) &= \mathbb{E}_t^{\mathbb{Q}}(\delta_0 + \delta_1' X_{t+n}) \\ &= \delta_0 + \delta_1' \left[\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i \mu_X^{\mathbb{Q}} + (\Phi_X^{\mathbb{Q}})^n X_t \right] \\ &= A_n^{\mathbb{Q}} + B_n^{\mathbb{Q}\prime} X_t. \end{aligned}$$

The corresponding loadings are given by,

$$A_n^{\mathbb{Q}} = \delta_0 + \delta_1' \left[\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i \mu_X^{\mathbb{Q}} \right], \quad B_n^{\mathbb{Q}\prime} = \delta_1' (\Phi_X^{\mathbb{Q}})^n,$$

which corresponds to the one-period forward rates in the bond pricing literature, as shown in Cochrane and Piazzesi (2008). The forward risk premium is given by $f_t^{\mathbb{P}}(n) - f_t^{\mathbb{Q}}(n)$, with the one period change $f_{t+1}^{\mathbb{P}}(n) - f_t^{\mathbb{P}}(n+1)$, notably as an excess return, because its risk-neutral expected return equals to zero. The one-period rate change can be written as

$$\begin{aligned} f_{t+1}^{\mathbb{P}}(n) - f_t^{\mathbb{P}}(n+1) &= A_n + B_n' X_{t+1} - A_{n+1} - B_{n+1}' X_t \\ &= B_n'(\mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_t + \Sigma_X \epsilon_{t+1}^{\mathbb{Q}}) - \delta_1' (\Phi_X^{\mathbb{Q}})^n \mu_X^{\mathbb{Q}} - B_n' \Phi_X^{\mathbb{Q}} X_t \\ &= B_n' \Sigma_X \epsilon_{t+1}^{\mathbb{Q}} \\ &= B_n' \Sigma_X \epsilon_{t+1}^{\mathbb{P}} + B_n' \Sigma_X \lambda_t \\ &= B_n^{\mathbb{Q}\prime} \Sigma_X \epsilon_{t+1}^{\mathbb{P}} + (B_n - B_n^{\mathbb{Q}})' \Sigma_X \epsilon_{t+1}^{\mathbb{P}} + B_n' \Sigma_X (\lambda_0 + \lambda_1 X_t). \end{aligned}$$

Three components that can be drawn from this interest rate change.

1. Revisions to short rate expectation - that results to the change in expected future short rate for time $t+n+1$,

$$\begin{aligned} (\mathbb{E}_{t+1} - \mathbb{E}_t) r_{t+n+1} &= \delta_1' (\mathbb{E}_{t+1} X_{t+n+1} - \mathbb{E}_t X_{t+n+1}) \\ &= \delta_1' \Phi_X^n \Sigma_X \epsilon_{t+1}^{\mathbb{P}} \\ &= B_n^{\mathbb{Q}\prime} \Sigma_X \epsilon_{t+1}^{\mathbb{P}}, \end{aligned}$$

The change in short rate expectation corresponds to the change in risk neutral rates $f_{t+1}^{\mathbb{Q}}(n) - f_t^{\mathbb{Q}}(n+1)$, that captures expectations of market participants on future monetary policy between dates $[t, t+1]$.

2. Surprise changes in the forward risk premium - this equals to the unexpected change in the forward risk premium given by

$$\begin{aligned} f_{t+1}^{\mathbb{P}}(n) - \mathbb{E}_t f_{t+1}^{\mathbb{P}}(n) &= (B_n - B_n^{\mathbb{Q}})'(X_{t+1} - \mathbb{E}_t^{\mathbb{Q}} X_{t+1}) \\ &= (B_n - B_n^{\mathbb{Q}})'\Sigma_X \epsilon_{t+1}. \end{aligned}$$

3. Expected returns - this component equals the expected change in the forward risk premium:

$$\begin{aligned} \mathbb{E}_t f_{t+1}^{\mathbb{P}}(n) - f_t^{\mathbb{P}}(n+1) &= A_n + B_n'\mathbb{E}_t X_{t+1} - A_{n+1} - B_{n+1}' X_t \\ &= B_n'\mathbb{E}_t \epsilon_{t+1}^{\mathbb{Q}} = B_n'\Sigma_X \lambda_t \\ &= B_n'\Sigma_X(\lambda_0 + \lambda_1 X_t), \end{aligned}$$

that captures the predictable part of the return and corresponds to the return risk premium. According to Bauer *et al.* (2011), decomposing rate changes shows that both short rate expectations and term premia can contribute to unpredictable changes in interest rates.

3.3.10 An Observed-Affine Model

Observed affine is the special case of the model. One of the assumptions mentioned before was the independence between the macroeconomic and latent yield-curve state variables. By relaxing the assumption of their independence allow macroeconomic variables to provide some information on yield curve dynamics, for example, following Collin-Dufresne *et al.* (2005). Their motivation was to find an invariant transformation of the latent variables in an affine model.

The idea is to use the first three principal components from an eigenvalue decomposition of zero-coupon interest rates. This will allow us to treat the yield curve related variables, and all state variables such as observable. We now adjust the latent state variables to lightly fit the term structure of interest rates. The state variable vector has the form,

$$X_t = [g_t \quad \pi_t \quad r_t \quad l_t \quad s_t \quad c_t],$$

where l_t, s_t, and c_t are the level, slope and curvature factors respectively, from the eigenvalue decomposition which can be written as X_t^e. The variables g_t, π_t and r_t are the macroeconomic factors representing the output gap, consumer price inflation and the short rate. These variables will be denoted by X_t^o. The instantaneous short rate is given by

$$r_t = [0 \quad 0 \quad 1 \quad 0 \quad 0 \quad 0]X_t = \mathbb{I}_r X_t.$$

For the second yield curve key assumption, we consider an unrestricted VAR(2) specification for the state variable dynamics, which are described as

$$\begin{bmatrix} X_t^o \\ \\ X_t^e \end{bmatrix} = \begin{bmatrix} \mu_X^o \\ \\ \mu_X^e \end{bmatrix} + \sum_{k=1}^{2} \begin{bmatrix} \Theta_k^o & \Theta_k^{eo} \\ \\ \Theta_k^{oe} & \Theta_k^e \end{bmatrix} \begin{bmatrix} X_{t-k}^o \\ \\ X_{t-k}^e \end{bmatrix} + \epsilon_t,$$

$$X_t = \mu_X^{\mathbb{P}} + \sum_{k=1}^{2} \Theta_k^{\mathbb{P}} X_{t-1} + \epsilon_t^{\mathbb{P}},$$

where $\mu^e \neq 0$ and $\epsilon_t \sim \mathcal{N}(0, \Omega)$. In this case, the interaction is permitted between the yield curve and macroeconomic state variables. It implies that the future zero-coupon term structure dynamics can now also vary depending on the state of the macroeconomy, which gives us the corresponding price of an $(n+1)$-period zero-coupon bond or the mapping between the state variable and the zero-coupon curve. That is,

$$P_t^{n+1} = \exp[A_n + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X \lambda_0) + \frac{1}{2} B_n' \Sigma_X \Sigma_X' B_n + \left(B_n'(\mu_X^{\mathbb{P}} - \Sigma_X \lambda_1) - \mathbb{I}_r\right) X_t],$$

where

$$\begin{aligned} A_{n+1} &= A_n + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X \lambda_0) + \frac{1}{2} B_n' \Sigma_X \Sigma_X' B_n, \\ B_{n+1}' &= B_n'(\Theta_X^{\mathbb{P}} - \Sigma_X \lambda_1) - \mathbb{I}_r, \end{aligned}$$

by recursive relations we get the following results, $A_1 = 0$ and $B_1' = -\mathbb{I}_r$. According to the Ang and Piazzesi (2003) model, the zero-coupon bond is a linear combination of yield curve and macroeconomic state variables. Thus, with this approach, the macroeconomic variables can influence both dynamics of yield curve state variables and the mapping between the state variable vector and pure-discount bond prices.

3.4 Main Findings of Ang Piazzesi Model

Ang and Piazzesi (2003) have inaugurated a prolific literature which presents a Gaussian VAR model of yields and macro variables by imposing modern no-arbitrage pricing models to study the joint dynamics of macro variables and bond prices in a factor model of the term structure. The incorporation of both macroeconomic variable and traditional yield variables play a significant role such that macro variables are found to explain mostly the short and middle parts of the yield curve as compared to the variations at the long end of the yield curve. The long term rates are described by the latent factors. Significant portion of level and slope factors are attributed to macro factors, particularly to inflation. Using only Treasury yields, Ang *et al.* (2006) demonstrate that a macro-finance term structure model leads to more efficient and accurate forecasts compared to those obtained by the standard approach using unrestricted OLS regressions. The term structure forecasts also outperform a number of alternative predictors. Moreover, Cochrane and Piazzesi (2005) show that, relative to standard latent-factor models, using macroeconomic information can substantially lower pricing errors. Ang and Piazzesi (2003) prove that the volatilities of short-term and medium-term yields are driven by macroeconomic factors. Their inflation and economic activity factors explain up to 85% of the long-horizon variance of shorter-term yields, but have a much smaller effect on long-term yields.

Earlier studies investigating the relation between the yield curve and macroeconomic variables, like Fama (1990), Mishkin and Jorion (1991), Estrella and Mishkin (1996) and Evans and Marshall (2007), did not consider no-arbitrage relations among yields and did not model bond pricing. As a consequence, they were able to make predictions only about the yields explicitly analysed (typically no more than three), they did not rule out theoretical inconsistencies due to the presence of arbitrage opportunities along the yield curve and, more importantly, they made no predictions about risk premia and their dynamics. Risk premia in Ang and Piazzesi (2003) model are time varying and depend on both macroeconomic and latent factors.

3.4.1 Limitations of Ang and Piazzesi (2003) model

As the size of state vector increase, the curse of dimensionality imposes severe estimation challenge and results in some undesirable restrictions in the structure, such that as unidirectional macro to yield linkage, have to be made. The identification of risk price is also problematic. Although Gaussian ATSMs have provided important insights, they can not capture conditional heteroskedasticity. The following remark presents recursive bond prices adapted to the VAR(2) model.

Remark 3.1. *The use of VAR model is found to be of importance when valuing the impact of individual house price risk model on pricing of equity release products. The method is employed to explain price variations by heterogeneous characteristics, with the adaptation of VAR model that incorporate exogenous variables to project the overall house price index and risk adjusted discount factors.*

In the work of Ang and Piazzesi (2003), the zero-coupon bond prices are assumed to be exponential linear functions of state variable. Following the same assumption to project state variables using VAR(2) model, zero-coupon bond prices are expressed as exponential linear functions of contemporaneous and 1-quarter lagged state variable given by

$$p_t^n = \exp\{A_n + B_n' X_t + C_n' X_{t-1}\}. \tag{3.4.1}$$

The difference equations of the parameters A_n, B_n and C_n are recursively derived from bond prices. The price of a 1-quarter zero coupon bond is:

$$p_t^1 = \exp\{-y_t^1\} = \exp\{A_1 + B_1' X_t + C_1' X_{t-1}\}, \tag{3.4.2}$$

where the initial coefficients are given by $A_1 = 0$, $B_1 = e_1$ and $C_1 = 0$. The price of an n-quarter zero-coupon bond is:

$$\begin{aligned}
p_t^n &= \mathbb{E}_t[M_{t+1} p_{t+1}^{n-1}] \\
&= \mathbb{E}_t[\exp\{-y_t^1 - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\xi_{t+1} + A_{n-1} + B_{n-1}'X_{t+1} + C_{n-1}'X_t\}] \\
&= \exp\{-y_t^1 - \frac{1}{2}\lambda_t'\lambda_t + A_{n-1} + C_{n-1}'X_t\}\mathbb{E}_t[\exp\{-\lambda_t'\xi_{t+1} + B_{n-1}'X_{t+1}\}] \\
&= \exp\{-\delta_0 - \delta_1'X_t - \frac{1}{2}\lambda_t'\lambda_t + A_{n-1} + C_{n-1}'X_t\} \\
&\quad \times \mathbb{E}_t\left[\exp\{-\lambda_t'\xi_{t+1} + B_{n-1}'(k + \phi_1 X_t + \phi_2 X_{t-1} + \Sigma_X \xi_{t+1})\}\right] \\
&= \exp\{-\delta_0 - \delta_1'X_t - \frac{1}{2}\lambda_t'\lambda_t + A_{n-1} + C_{n-1}'X_t + B_{n-1}'(k + \phi_1 X_t + \phi_2 X_{t-1})\} \\
&\quad \times \mathbb{E}_t\left[\exp\{(B_{n-1}'\Sigma_X - \lambda_t')\xi_{t+1}\}\right] \\
&= \exp\{-\frac{1}{2}\lambda_t'\lambda_t + B_{n-1}'k + A_{n-1} - \delta_0 + (C_{n-1}' + B_{n-1}'\phi_1 - \delta_1')X_t\} \\
&\quad \times \exp\{B_{n-1}'\phi_2 X_{t-1}\}\exp\{\frac{1}{2}(B_{n-1}'\Sigma_X - \lambda_t')(B_{n-1}'\Sigma_X - \lambda_t')'\} \\
&= \exp\{-\frac{1}{2}\lambda_t'\lambda_t + B_{n-1}'k + A_{n-1} - \delta_0 + (C_{n-1}' + B_{n-1}'\phi_1 - \delta_1')X_t\} \\
&\quad \times \exp\{B_{n-1}'\phi_2 X_{t-1}\}\exp\{\frac{1}{2}(B_{n-1}'\Sigma_X'\Sigma_X B_{n-1} - 2B_{n-1}'\Sigma_X\lambda_t + \lambda_t'\lambda_t)\} \\
&= \exp\{B_{n-1}'k + A_{n-1} - \delta_0 + (C_{n-1}' + B_{n-1}'\phi_1 - \delta_1')X_t + B_{n-1}'\phi_2 X_{t-1}\} \\
&\quad \times \exp\{B_{n-1}'\Sigma_X'\Sigma_X B_{n-1} - B_{n-1}'\Sigma_X(\lambda_0 + \lambda_1 X_t)\} \\
&= \exp\{B_{n-1}'(k - \Sigma_X\lambda_0) + \frac{1}{2}B_{n-1}'\Sigma_X'\Sigma_X B_{n-1} + A_{n-1} - \delta_0\} \\
&\quad \times \exp\{(C_{n-1}' + B_{n-1}'(\phi_1 - \Sigma_X\lambda_1) - \delta_1')X_t + B_{n-1}'\phi_2 X_{t-1}\}.
\end{aligned}$$

The difference equations are then derived by equating the above obtained results to

$$p_t^n = \exp\{A_n + B_n'X_t + C_n'X_{t-1}\},$$

where the parameters are given by the equations:

$$\begin{aligned}
A_n &= B_{n-1}'(k - \Sigma_X\lambda_0) + \frac{1}{2}B_{n-1}'\Sigma_X'\Sigma_X B_{n-1} + A_{n-1} - \delta_0, \\
B_n' &= C_{n-1}' + B_{n-1}'(\phi_1 - \Sigma_X\lambda_1) - \delta_1', \\
C_n' &= B_{n-1}'\phi_2.
\end{aligned}$$

3.5 Non-Gaussian Affine Term-Structure Models

This chapter introduces a new estimation procedure for non-Gaussian affine term structure models that uses linear regression to construct a concentrated log-likelihood function, see Creal and Wu (2013). The origin of the model approach encompasses from the work of Duffie and Kan (1996) to Aït-Sahalia and Kimmel (2010) to name few, where much of the literature has been conducted in continuous time. The discrete time version of multivariate non-Gaussian model was then extended by Dai *et al.* (2010), continuing from the work of Gouriéroux and Jasiak (2006).

The main focus is on the development of the family of discrete time, non-Gaussian affine term structure models. Although from recently introduced literatures of Ang and Piazzesi (2003) and the recent Christensen *et al.* (2011), Gaussian affine term structure models have become the basic workhorse in macroeconomics and finance for purpose of incorporating the theoretical foundation framework to understanding the relationship between yields of different maturities. The approach provides some useful insights, but cannot capture conditional heteroskedasticity, that is, the level of non-constant volatility into bond prices that cannot be predicted over any period of time.

3.5.1 Dynamics of the State Vector

The vector of non-Gaussian state variables h_{t+1}, that captures the volatility, its stochastic process under physical measure $\mathbb{P}$ is obtained by taking on affine functions under discrete-time settings. This is an equivalence of a multivariate Cox *et al.* (1985) process, given by

$$\begin{aligned} h_{t+1} &= \mu_h + \Sigma_h \omega_{t+1} && (3.5.1)\\ \omega_{t+1} &\sim \Gamma(\nu_{h,i} + z_{i,t+1}), \quad \text{i=1,...,H} && (3.5.2)\\ z_{i,t+1} &\sim \text{Poisson}(e_i' \Sigma_h^{-1} \Phi_h \Sigma_h \omega_t), \quad \text{i=1,...,H} \ . && (3.5.3) \end{aligned}$$

The conditional mean of h_{t+1} in matrix form,

$$\mathbb{E}(h_{t+1}|h_t) = (I_H - \Phi_H)\mu_h + \Sigma_h \nu_h + \Phi_h h_t.$$

The conditional variance h_{t+1} is also an affine function of h_t

$$\mathbb{V}(h_{t+1}|h_t) = \Sigma_{h,t}\Sigma_{h,t}' = \Sigma_h \text{diag}(\nu_h - 2\Sigma_h^{-1}\Phi_h\mu_h)\Sigma_h' + \Sigma_h \text{diag}(2\Sigma_h^{-1}\Phi_h h_t)\Sigma_h',$$

where h_t gives the current information set at time t and the parameter $h_{t+1} = H \times 1$ is a vector of non-Gaussian state variable. The parameter $\nu_h = (\nu_{h,1}, ..., \nu_{h,H})$ gives the shape, Φ_h is the matrix that controls the autocorrelation of h_{t+1}, Σ_h is a scale matrix, and μ_h is a vector determining the lower bound of h_{t+1}. Lastly, e_i denotes the $i-$th column of I_h.

Gouriéroux and Jasiak (2006) built the univariate version of this model and Dai *et al.* (2010) extended it from Equation (3.5.1) to Equation (3.5.3) with $\mu = 0$ and Σ diagonal. In this model for h_{t+1}, shocks are allowed to be correlated through the off-diagonal elements of Σ_h and the process may have a negatively correlated drift through the off-diagonal elements of Φ_h. The vector of conditionally Gaussian state variables g_{t+1} follows vector autoregressive with conditional heteroskedasticity,

$$\begin{aligned} g_{t+1} &= \mu_g + \Phi_g g_t + \Phi_{gh} h_t + \Sigma_{gh}\varepsilon_{h,t+1} + \varepsilon_{g,t+1}, \\ \Sigma_{g,t}\Sigma'_{g,t} &= \Sigma_{0,g}\Sigma'_{0,g} + \sum_{i=1}^{H} \Sigma_{i,g}\Sigma'_{i,g} h_{it} \\ \varepsilon_{h,t+1} &= h_{t+1} - \mathbb{E}(h_{t+1}|h_t), \end{aligned} \tag{3.5.4}$$

where $\varepsilon_{g,t+1} \sim \mathrm{N}(0, \Sigma_{g,t}\Sigma'_{g,t})$ and $\Sigma_{i,g}$ are lower triangles for $i = 0, ..., H$. The Gaussian factors are functions of the non-Gaussian state variables through both the autoregressive term $\Phi_{gh}h_t$ and the covariance term $\Sigma_{gh}\varepsilon_{h,t+1}$. The conditional variance g_{t+1} is also a function of non-Gaussian factors, which introduces conditional heteroskedasticity into bond prices. A useful property of the model stated from Equation (3.5.1) to Equation (3.5.4) for the vector $x_{t+1} = (h'_{t+1}, g'_{t+1})'$ for pricing bonds is that any admissible affine transformation remains within the same family of distributions. Now the state vector has the dynamics.

Proposition 3.2. *Suppose that the state vector $x_t = (h'_t, g'_t)'$ follows the process of Equations (3.5.1), (3.5.3), (3.5.4) and (3.5.2) with parameter θ. Consider an admissible affine transformation of the form:*

$$\begin{pmatrix} h_t^{\mathbb{Q}} \\ g_t^{\mathbb{Q}} \end{pmatrix} = \begin{pmatrix} c_h \\ c_g \end{pmatrix} + \begin{pmatrix} c_{hh} & c_{hg} \\ c_{gh} & c_{gg} \end{pmatrix} \begin{pmatrix} h_t \\ g_t \end{pmatrix}.$$

The process $x_{t+1}^{\mathbb{Q}} = (h_{t+1}^{\mathbb{Q}'}, g_{t+1}^{\mathbb{Q}'})'$ remains in the same family of distributions under updated parameter $\theta^{\mathbb{Q}}$. The parameters ν_h and $\Sigma_h^{-1}\Phi_h\Sigma_h$ are invariant to rotation. The admissibility restrictions and the relationship between the new and old parameterizations are proved below.

3.5.2 Proof of Proposition 3.2

Proof. The necessary admissibility restrictions to keep the non-Gaussian factors positive are

1. $C_{hg} = 0$,

2. C_{hh} is restricted such that all elements $C_{hh}\Sigma_h$ are non-negative,

3. c_h is restricted such that all elements $c_h + C_{hh}\mu_h$ are non-negative,

4. C_{hh} and C_{gg} are full rank.

For some values of θ, these restrictions may allow c_h and C_{hh} to be negative. Under these restrictions, the new process $x_{t+1}^{\mathbb{Q}} = (h_{t+1}^{\mathbb{Q}'}, g_{t+1}^{\mathbb{Q}'})'$ remains in the same family of distributions under updated parameters $\theta^{\mathbb{Q}}$. The proof of Proposition (3.2) is immediate by comparing the Laplace transform of $x_{t+1}^{\mathbb{Q}}$ to the Laplace transform of x_{t+1}. The mapping between the new parameters $\theta^{\mathbb{Q}}$ and the original parameters θ is given by

$$
\begin{aligned}
\mu_h^{\mathbb{Q}} &= c_h + C_{hh}\mu_h \\
\Phi_h^{\mathbb{Q}} &= C_{hh}\Phi_h C_{hh}^{-1} \\
\Sigma_h^{\mathbb{Q}} &= C_{hh}\Sigma_h \\
\mu_g^{\mathbb{Q}} &= c_g + C_{gg}\mu_g - C_{gg}\Phi_g C_{gg}^{-1} + C_{gh}([I_H - \Phi_h]\mu_h + \Sigma_h \nu_h) \\
&- (C_{gh}\Phi_h - C_{gg}\Phi_g C_{gg}^{-1} C_{gh} + C_{gg}\Phi_{gh}) C_{hh}^{-1} c_h \\
\Phi_g^{\mathbb{Q}} &= C_{gg}\Phi_g C_{gg}^{-1} \\
\Phi_{gh}^{\mathbb{Q}} &= (C_{gh}\Phi_h - C_{gg}\Phi_g C_{gg}^{-1} C_{gh} + C_{gg}\Phi_{gh}) C_{hh}^{-1} \\
\Sigma_{gh}^{\mathbb{Q}} &= (C_{gh} + C_{gg}\Sigma_{gh}) C_{hh}^{-1} \\
\Sigma_{0,g}^{\mathbb{Q}}\Sigma_{0,g}^{\mathbb{Q}'} &= C_{gg}\Sigma_{0,g}\Sigma_{0,g}^{'} C_{gg}^{'} - \sum_{i=1}^{H} C_{gg}\Sigma_{i,g}\Sigma_{i,g}^{'} C_{gg}^{'} e_i^{'} C_{hh}^{-1} c_h \\
\Sigma_{i,g}^{\mathbb{Q}}\Sigma_{i,g}^{\mathbb{Q}'} &= \sum_{j=1}^{H} C_{gg}\Sigma_{j,g}\Sigma_{j,g}^{'} C_{gg}^{'} e_j^{'} C_{hh}^{-1} e_i.
\end{aligned}
$$

□

Proposition (3.2) helps to understand the identification of the model. The admissibility constraints ensure that the non-Gaussian state variables always remain positive after applying a transformation from x_t to $x_t^{\mathbb{Q}}$ and that there exists another admissible rotation from $x_t^{\mathbb{Q}}$ back to x_t. Analogous to the popular class of Gaussian ATSMs, we specify $x_t = (g_{t+1}^{\prime}, h_{t+1}^{\prime})'$ to have the same dynamics under both $\mathbb{P}$ and $\mathbb{Q}$ measure. This is done by allowing the parameters controlling the conditional mean to be different under the two probability measures and set the scale parameter Σ_{gh}, Σ_h and $\Sigma_{i,g}$ for $i = 0, ..., H$ to be the same. The location parameter μ must also be the same under both measures to ensure no-arbitrage.

3.5.3 Stochastic Discount Factor

The stochastic discount factor demonstrates the compensated risk exposure of an investor holding a zero-coupon bond under stochastic volatility. The log stochastic discount factor is expressed into the risk free rate and three components describing risk compensation,

$$m_{t+1} = -r_t - \frac{1}{2}\lambda'_{gt}\lambda_{gt} - \lambda'_{gt}\epsilon_{g,t+1} - \lambda'_{wt}\epsilon_{w,t+1} - \lambda'_{zt}\epsilon_{z,t+1},$$

where $\epsilon_{i,t+1}$ are standardized shocks with mean zero and identity covariance matrix. Moreover, in addition to the risk-free rate r_t, an agent gets compensated for being exposed to the Gaussian shock $\epsilon_{g,t+1}$, the gamma shock $\epsilon_{w,t+1}$ and a Poisson shock $\epsilon_{z,t+1}$. The parameter λ_{it} is the price of risk i for each of the three types of shocks in the model, which are defined as

$$\begin{aligned}
\lambda_{gt} &= \mathbb{V}(g_{t+1}|h_t, h_{t+1}, z_{t+1})^{-\frac{1}{2}}[\mathbb{E}(g_{t+1}|h_t, h_{t+1}, z_{t+1}) - \mathbb{E}^{\mathbb{Q}}(g_{t+1}|h_t, h_{t+1}, z_{t+1})], \\
\lambda_{wt} &= \mathbb{V}(w_{t+1}|h_t, z_{t+1})^{-\frac{1}{2}}[\mathbb{E}(w_{t+1}|h_t, z_{t+1}) - \mathbb{E}^{\mathbb{Q}}(w_{t+1}|h_t, z_{t+1})], \\
\lambda_{zt} &= \mathbb{V}(z_{t+1}|h_t)^{-\frac{1}{2}}[\mathbb{E}(z_{t+1}|h_t) - \mathbb{E}^{\mathbb{Q}}(z_t|h_t)].
\end{aligned}$$

The market price of risk has an intuitive form as the sharp-ratio measuring per unit risk compensation. Specially, price of risks are the difference in the conditional means of each shock under measure $\mathbb{P}$ and measure $\mathbb{Q}$, standardized by a conditional standard deviation. These time varying quantities of risk are a feature of non-Gaussian models that are not available in Gaussian models.

3.5.4 Bond Prices

The n^{th} period price of a zero-coupon bond at time t is the expected $(t+1)$-price, discounted by the short rate r_t under risk neutral measure, is given by,

$$P_t^n = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t)P_{t+1}^{n-1}].$$

The short rate as a linear function of the state variables

$$r_t = \delta_0 + \delta'_h h_t + \delta'_g g_t.$$

The exponentially affine function of bond prices leads to the following expression

$$P_t^n = \exp[A_n + B'_{n,h}h_t + B'_{n,g}g_t]. \tag{3.5.5}$$

The bond loadings A_n, $B_{n,h}$, $B_{n,g}$ can be recursively derived to form a matrix form

$$\begin{aligned}
A_n &= -\delta_0 + A_{n-1} + \mu_g^{\mathbb{Q}'} B_{n-1,g} + [\mu_h - \Phi_h^{\mathbb{Q}}\mu_h + \Sigma_h \nu_h^{\mathbb{Q}}]' B_{n-1,h} \\
&\quad + \frac{1}{2} B'_{n-1,g}\Sigma_{0,g}\Sigma'_{0,g} B_{n-1,g} + \mu'_h \Phi_h^{\mathbb{Q}'} \Sigma_h^{-1'} I_H \Sigma'_h B_{n-1,gh} \\
&\quad - \mu'_h \Phi_h^{\mathbb{Q}'} \Sigma_h^{-1'} [\mathrm{diag}(\imath_H - \Sigma'_h B_{n-1,gh})]^{-1} \Sigma'_h B_{n-1,gh} \\
&\quad - \nu_h^{\mathbb{Q}'} [\ln(\imath_H - \Sigma'_h B_{n-1,gh}) + \Sigma'_h B_{n-1,gh}], \\
B_{n,h} &= -\delta_h + \Phi_{gh}^{\mathbb{Q}'} B_{n-1,g} + \Phi_h^{\mathbb{Q}'} B_{n-1,h} + \frac{1}{2}(I_H \otimes B'_{n-1,g})\Sigma_g \Sigma'_g (\imath_H \otimes B_{n-1,g}) \\
&\quad - \Phi_h^{\mathbb{Q}'} \Sigma_h^{-1'} (I_H - [\mathrm{diag}(\imath_H - \Sigma'_h B_{n-1,gh})]^{-1}) \Sigma'_h B_{n-1,gh}, \\
B_{n,g} &= -\Delta_g + \Phi_g^{\mathbb{Q}'} B_{n-1,g}
\end{aligned}$$

where $\Sigma_g \Sigma'_g$ is a $GH \times GH$ block diagonal matrix with diagonal elements $\Sigma_{i,g}\Sigma'_{i,g}$ for $i =, ..., H$ and $B_{n-1,gh} = \Sigma'_{gh} B_{n-1,g} + B_{n-1,h}$. The initial loadings at maturity are $A_1 = -\delta_0$, $B_{1,g} = -\delta_{1,g}$ and $B_{1,h} = -\delta_{1,h}$. The bond yields in linear form are given by

$$\begin{aligned}
y_t^n &\equiv -\frac{1}{n}\ln(P_t^n) \\
&= \bar{A}_n + \bar{B}'_{n,h} h_t + \bar{B}'_{n,g} g_t,
\end{aligned}$$

with $\bar{A}_n = -\frac{1}{n}A_n$, $\bar{B}_{n,h} = -\frac{1}{n}B_{n,h}$ and $\bar{B}_{n,g} = -\frac{1}{n}B_{n,g}$. Stacking y_t^n in order for N different maturities $n_1, n_2, ..., n_N$ gives,

$$Y_t = A + BX_t,$$

where $A = (\bar{A}_{n1}, ..., \bar{A}_{nN})'$, $B = (\bar{B}_{n1}, ..., \bar{B}_{nN})'$. If more yields are observed than the number of factors, i.e. $(N > G + H)$, not all yields can be priced exactly. Standard assumptions are made in Creal and Wu (2013), concerning the ATSM literature that $N_1 = G + H$ linear combinations of the yields Y_t^1, which are priced without error. The remaining $N_2 = N - N_1$ linear combination Y_t^2 are observed with Gaussian measurement errors. The observation equations are:

$$Y_t^1 = A_1 + B_1 X_t, \text{ and } Y_t^2 = A_2 + B_2 X_t + \xi_t \quad \xi_t \sim N(0, \Sigma'\Sigma), \tag{3.5.6}$$

where $A_1 \equiv \Theta_{Y_1} A$, $A_2 \equiv \Theta_{Y_2} A$ and $B_1 \equiv \Theta_{Y_1} B$, $B_2 \equiv \Theta_{Y_2} B$. The model develops a new estimation procedure for the class of discrete-time non-Gaussian affine term structure models that allows for any admissible rotation of the non-Gaussian state variables that can also be estimated. The estimation challenge that are problematic in some Gaussian macro-finance models are improved in this approach by reducing the number of parameter needed to be maximized numerically. Both Gaussian and non-Gaussian models give the best fit on the cross-section of yields. Any affine macro-finance model that allows for restrictions on key parameters of interest rate can also benefit from this non-Gaussian method.

3.6 Diebold, Rudebusch and Aruoba (2006) Approach

The Diebold *et al.* (2006) (DRA) approach introduces a simple way of summarising the dynamic interactions between macroeconomic fundamentals and the yield curve by incorporating the yield curve factors shown to be popular in factor models in the finance literature with observable macroeconomic variables. The characterization of the three factor term structure model introduced in this framework is based on the classical Nelson and Siegel (1987) representation which is parsimonious and suitable for estimating and forecasting the entire yield curve, period by period. This model is quite flexible to capture the yield curve dynamics into a three dimensional parameter that evolves dynamically, as in Diebold and Li (2006), which made it to be crucial for relating yield curve evolution over time to movements in macroeconomic variables. The state variable vector in the DRA model is given by

$$X_t = (l_t, s_t, c_t, g_t, \pi_t, r_t)',$$

which consists of three Nelson-Siegel factors and macroeconomic variables, namely, real activity, inflation, federal funds rate that represents the monetary policy instrument. The joint dynamics of macroeconomic factors to the Nelson and Siegel representation can be extended to give the measurement equation of the form:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ y_t(\tau_3) \\ y_t(\tau_4) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} & \phi_{14} & \phi_{15} & \phi_{16} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} & \phi_{24} & \phi_{25} & \phi_{26} \\ 1 & \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} & \frac{1-e^{-\lambda\tau_3}}{\lambda\tau_3} - e^{-\lambda\tau_3} & \phi_{34} & \phi_{35} & \phi_{36} \\ 1 & \frac{1-e^{-\lambda\tau_4}}{\lambda\tau_4} & \frac{1-e^{-\lambda\tau_4}}{\lambda\tau_4} - e^{-\lambda\tau_4} & \phi_{44} & \phi_{45} & \phi_{46} \end{pmatrix} \times \begin{pmatrix} l_t \\ s_t \\ c_t \\ g_t \\ \pi_t \\ r_t \end{pmatrix} + \begin{pmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \epsilon_t(\tau_3) \\ \epsilon_t(\tau_4) \end{pmatrix}, \tag{3.6.1}$$

with the measurement equation taking the affine form,

$$y_t(t+n) = -\frac{1}{n}(A_n + B_n X_t) + \epsilon_t, \qquad \epsilon_t \sim N(0, \sigma_n^2)$$

where the corresponding coefficients for A_n and B_n are given by,

$$A_n = 0, \tag{3.6.2}$$

$$B_n = [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}, 0, 0, 0]'. \tag{3.6.3}$$

The system of the model follows a non-structural VAR representation of the macroeconomy and the factor dynamics are assumed to follow a VAR(1) process, which has a transition equations of the form:

$$\begin{pmatrix} l_t \\ s_t \\ c_t \\ g_t \\ \pi_t \\ r_t \end{pmatrix} = \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} & \omega_{14} & \omega_{15} & \omega_{16} \\ \omega_{21} & \omega_{22} & \omega_{23} & \omega_{24} & \omega_{25} & \omega_{26} \\ \omega_{31} & \omega_{32} & \omega_{33} & \omega_{34} & \omega_{35} & \omega_{36} \\ \omega_{41} & \omega_{42} & \omega_{43} & \omega_{44} & \omega_{45} & \omega_{46} \\ \omega_{51} & \omega_{52} & \omega_{53} & \omega_{54} & \omega_{55} & \omega_{56} \\ \omega_{61} & \omega_{62} & \omega_{63} & \omega_{64} & \omega_{65} & \omega_{66} \end{pmatrix} \begin{pmatrix} l_{t-1} \\ s_{t-1} \\ c_{t-1} \\ g_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \varepsilon_{3,t} \\ \varepsilon_{4,t} \\ \varepsilon_{5,t} \\ \varepsilon_{6,t} \end{pmatrix}. \tag{3.6.4}$$

The compact form for transition equations is

$$f_t = \Omega f_{t-1} + \varepsilon_t. \tag{3.6.5}$$

The model offers a bidirectional causality of effects between both the macroeconomy and the yield curve, unlike in the Ang and Piazzesi (2003) approach. It is also observed that the expectation hypothesis in relation to the yield curve modelling approach, seem to hold reasonably well in some certain periods under macroeconomic perspective. On the other hand, even though the model benefits from parsimony of Nelson-Siegel and the flexibility of jointly capturing both the dynamics of the yield curve and macroeconomic variables, however, it does not consider no-arbitrage restrictions, which are taken as being crucial in finance perspective.

Chapter 4

Dynamic Nelson-Siegel Models

This chapter discusses a second approach of modelling the term structure of interest rates, by considering first the original formulation of the extremely popular, Nelson and Siegel (1987) interpolation and several Nelson-Siegel type models that imposes theoretical restrictions that considers no-arbitrage assumption. A reformulation and extension version is proposed by Diebold and Li (2006), which is the dynamic Nelson-Siegel model. Furthermore, we touch upon a derived discrete time arbitrage-free Nelson-Siegel class that establishes the uniqueness and exact solution of the arbitrage-free Nelson-Siegel model, (see Linlin and Gengming (2012)).

4.1 Classical Nelson-Siegel Curve Fitting

Nelson and Siegel (1987) formulation steps aside from the solid theoretical foundation to rule out riskless arbitrage opportunities. The model aims to introduce "A simple parsimonious model that is quite flexible enough to represent the range of shapes associated with yield curves". Following from the extensions made by Siegel and Nelson (1988), to fit a smooth yield curve to unsmooth yields, Nelson and Siegel (1987) begin to give a forward rate of the form,

$$f(\tau) = b_1 + b_2 e^{-\lambda\tau} + b_3 \lambda\tau e^{-\lambda\tau}. \tag{4.1.1}$$

The Nelson-Siegel forward rate curve can be viewed as a constant plus a Laguerre function, which is a polynomial times an exponential decay term and is a popular mathematical approximating function. The corresponding yield curve as a function of maturity can be easily derived from Equation (2.1.6)

$$\begin{aligned} y(\tau) &= \frac{1}{\tau}\int_0^{\tau}(b_1 + b_2 e^{-\lambda s} + b_3\lambda s e^{-\lambda s})ds \\ &= b_1 + b_2\frac{1-e^{-\lambda\tau}}{\lambda\tau} + b_3\frac{1}{\tau}\int_0^{\tau}\lambda s e^{-\lambda s}ds \\ &= b_1 + b_2\frac{1-e^{-\lambda\tau}}{\lambda\tau} + b_3(-e^{-\lambda\tau} + \frac{1-e^{-\lambda\tau}}{\lambda\tau}) \\ &= b_1 + b_2[\frac{1-e^{-\lambda\tau}}{\lambda\tau}] + b_3[\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}]. \end{aligned} \tag{4.1.2}$$

This is the cross-sectional linear projection of $y(\tau)$ for fixed t on variables $(1, [\frac{1-e^{-\lambda\tau}}{\lambda\tau}], [\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}])$ with the parameters b_1, b_2, b_3 and a decay rate λ. The Nelson-Siegel yield curve also corresponds to a discount curve that begins at one at zero maturity and approaches zero at infinite maturity, as appropriate. Even though there are some historically developed models used that showed some success, (see Fisher *et al.* (1995)), but the Nelson-Siegel model turns out to have some very appealing features. Firstly, Nelson-Siegel form enforces some basic constraints from financial economic theory, that is, the corresponding discount curve satisfies that $P(0) = 1$, and $\lim_{\tau\to\infty} P(\tau) = 0$ as appropriate. In addition, the Nelson-Siegel zero-coupon curve satisfies that

$$\lim_{\tau\to 0} y(\tau) = f(0) = r,$$

the instantaneous short rate. Secondly, the Nelson-Siegel form provides a parsimonious approximation. Parsimony is essential as it promotes smoothness i.e yields tend to be very smooth functions of maturity. Moreover, the model also protects against in-sample over fitting, for producing good forecast and also promotes empirically tractable and trustworthy estimation. Thirdly, Nelson-Siegel form not only provides parsimonious approximation but also a flexible one as the yield curve assumes a variety of shapes at different times, depending on the values of four parameters. Fourthly, from mathematical approximation-theoretic viewpoint, Nelson-Siegel form is far from arbitrary. The forward rate curve Equation (4.1.1), as noted by Nelson-Siegel, corresponds to the yield curve Equation (4.1.2), is viewed as a constant plus a Laguerre function on the domain $[0, \infty)$, which matches the domain for the term structure. Moreover, the desirable approximation-theoretic properties of Nelson-Siegel go well beyond their Laguerre structure. For all of the four reasons, Nelson-Siegel has become very popular for static curve fitting in practice, particularly among financial market practitioners, (see Gürkaynak *et al.* (2007)), and central banks, use the Nelson-Siegel model, or an extension thereof, to model the term structure of interest rates (Pooter (2007)). To estimate the model, suppose we have N yield rates $y(\tau)$ for different maturities $\tau_1, ..., \tau_N$. Using the model in Equation (4.1.2), we seek the optimal parameters b_1, b_2, b_3 and λ in terms of best fitting the given yield rates $y(\tau)$. That is,

$$\begin{pmatrix} y(\tau_1) \\ y(\tau_2) \\ \vdots \\ y(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}.$$

The system can be written into matrix form:

$$y(\tau) = X\lambda b.$$

where $y(\tau)$ is an N dimensional vector, $X\lambda$ is the $M \times 3$ matrix and b is the $3-$dimensional vector. This model has the appealing feature that the yields follow the functional form of Nelson and Siegel (1987) that fits yield curves quite well on the observed term structure. The noted drawbacks regarding the Nelson-Siegel model is that the model admits arbitrage because it has a deterministic foundation and therefore fails to account for the convexity arising from Jensen's inequality effects when yield factors are stochastic. Moreover, Nelson-Siegel model cannot enforce no-arbitrage because the parameters that determine the dynamic evolution of yield curve factors are not linked to the parameters that determine the shape and location of the yield curve. The elements of the state vector are just Nelson and Siegel's level, slope, and curvature measures. However, this model showed to forecast poorly in an out of sample forecast because in practice b_1, b_2, b_3 and λ tend to vary over time. Diebold and Li (2006) suggest to overcome this caveat by dynamically extending the model.

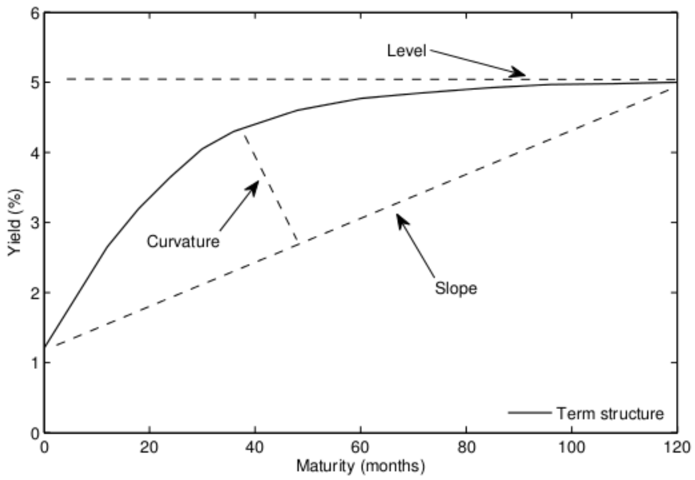
4.2 Introducing Dynamics

Diebold and Li (2006) extended the classical Nelson and Siegel (1987) framework, by interpreting and extracting latent dynamic factors as level, slope and curvature. Moreover, Diebold and Li (2006) suggest a rearrangement on factors of the model to give the factors a clear and distinct economic interpretation. Allowing for time dependency on parameters, Diebold and Li (2006) suggest to match the original Nelson and Siegel (1987) $b_{1t} = \beta_{1t}$, $b_{2t} = \beta_{2t} + \beta_{3t}$, $b_{3t} = \beta_{3t}$, such that parameters, which are now variables from a time-series perspective, change at each time t. In so doing, this produces a dynamic Nelson-Siegel (DNS) model. The measurement equation for fixed τ becomes

$$\begin{aligned} y_t(\tau) &= \beta_{1t} + (\beta_{2t} + \beta_{3t})\frac{1-e^{-\lambda\tau}}{\lambda\tau} - \beta_{3t}e^{-\lambda\tau} \\ &= \beta_{1t} + \beta_{2t}[\frac{1-e^{-\lambda\tau}}{\lambda\tau}] + \beta_{3t}[\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}]. \end{aligned} \tag{4.2.1}$$

Then, this is a time-series linear projection of $y_t(\tau)$ on variables β_{1t}, β_{2t} and β_{3t} which are interpreted as level, slope and curvature factor, respectively, with parameters $(1, [\frac{1-e^{-\lambda\tau}}{\lambda\tau}], [\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}])$. The dynamic latent factors in the DNS model determine entirely the dynamics of y_t for any τ, and the factor loadings are functions of maturity τ and a shape parameter λ, determine entirely the cross section of $y(\tau)$ for any t. The DNS model forms part of the popular dynamic factor models[1] where many yields across maturities are driven by just three latent factors. Figure 4.1 shows how a given term structure can be broken down into these three latent factors.

Figure 4.1: Term structure representation broken down into three factors



The interpretation of the latent factors in respect to their respective loadings in the Dynamic Nelson-Siegel model are as follows: the parameter λ governs the exponential decay rate. The linear combination of these variables, $[\beta_{1t}, \beta_{2t}, \beta_{3t}]$ have the associated loadings $\left[1, [\frac{1-e^{-\lambda\tau}}{\lambda\tau}], [\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}]\right]$ where β_{1t} has a constant one that does not approach zero for all yields, as the yield maturity varies. That is,

$$\lim_{\tau\to\infty} y_t(\tau) = \beta_{1t}.$$

Hence β_{1t} may be viewed as a long-term factor that governs the level of the yield curve. The loading on β_{2t} is closely related to the slope, and the corresponding loading is a monotone function, $\frac{1-e^{-\lambda\tau}}{\lambda\tau}$, that starts at one when maturity τ turns to zero and quickly to zero as τ approaches infinity; hence β_{2t}, which often called a short term factor, mostly affecting short-term yields. The loading on β_{3t} has a humped-shaped, $\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}$, which starts at zero (and is thus

[1]Dynamic factor structure is very popular to the theory and practice of financial asset pricing and very convenient statistically for it provides tractability in governing dimensional situations.

not short-term), increases and then decays to zero (and thus is not long-term), with the speed of decay and location of maximal hump determined by the parameter λ; hence β_{3t} indicates the contribution of the medium-term factor.

An important insight is that the three factors, which following the literature can thus far be called long-term, short-term and medium-term, in reference to the yield maturities at which they have maximal respective relative impact. Moreover, factors may also be interpreted in terms of their effect on the overall yield curve as level, slope and curvature. That is, an increase in β_{1t} increases all yields equally, as the loading is identical at all maturities, thereby changing the level of the yield curve. Furthermore, an increase in β_{2t} increases short yields more than long yields, because the short rates load on β_{2t} more heavily, thereby changing the slope of the yield curve. It is interesting to note, moreover, that the instantaneous yield depends on both the level and slope factors, because the second limit can be calculated as:

$$\begin{aligned}
\lim_{\tau \to 0^+} y_t(\tau) &= \lim_{\tau \to 0^+} \beta_{1t} + (\beta_{2t} + \beta_{3t}) \frac{1 - e^{-\lambda\tau}}{\lambda\tau} - \beta_{3t} e^{-\lambda\tau} \\
&= \beta_{1t} - \beta_{3t} + (\beta_{2t} + \beta_{3t}) \lim_{\tau \to 0^+} \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \\
&= \beta_{1t} - \beta_{3t} + (\beta_{2t} + \beta_{3t}) \lim_{\tau \to 0^+} \frac{\lambda e^{-\lambda\tau}}{\lambda} \\
&= \beta_{1t} - \beta_{3t} + (\beta_{2t} + \beta_{3t}) \\
&= \beta_{1t} + \beta_{2t}.
\end{aligned}$$

That is, the instantaneous rate is determined by $\beta_{1t} + \beta_{2t}$. Finally, an increase in β_{3t} will have little effect on very short or very long yields, which load minimally on it, but will increase medium-term yields, which load more heavily on it, thereby increasing yield curve curvature. Figure 4.2 shows the loading on the three latent factors as a function of maturity.

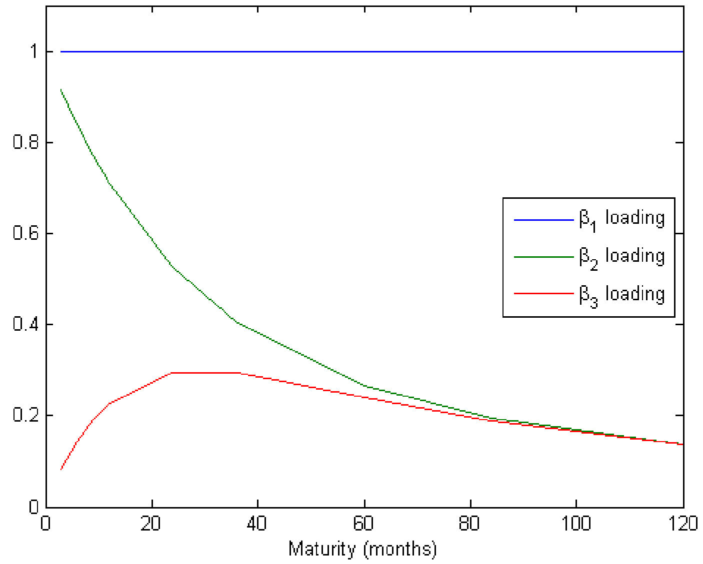


Figure 4.2: Dynamic Nelson-Siegel Factor Loadings, for $\lambda = 0.0609$

The parameter λ governs the exponential decay rate. Small values of λ produce slow decay, resulting in a better fit of the yield curve at long maturities, while large values of λ produce fast decay and can better fit the yield curve at short maturities. Moreover, the parameter λ also governs where the loading on β_{3t} achieves its maximum. As stated in Diebold and Li (2006), λ might be taken as a fixed constant at a perspective value. The time series estimates of β_{1t}, β_{2t} and β_{3t} are simplified because the measurement Equation (4.2.1) becomes a linear regression.

4.3 State Space Representation

In the original Nelson and Siegel (1987) factorization, the loadings $b_{2t} = \frac{1-e^{-\lambda\tau}}{\lambda\tau}$ and $b_{3t} = e^{-\lambda\tau}$ give similar monotonically decreasing shape, which gives similar interpretation to the underlying factors. In Diebold and Li (2006), the three factors are given distinctly different interpretations denoted as L_t, S_t, and C_t, such that, $X_t = [L_t, S_t, C_t]'$, is modeled as an unrestricted VAR(1). The measurement equation is then specified as

$$y_t(\tau) = L_t + S_t[\frac{1-e^{-\lambda\tau}}{\lambda\tau}] + C_t[\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}]. \tag{4.3.1}$$

Adding stochastic error term $\epsilon_t(\tau)$ to the deterministic DNS curve, stochastically relates any set of N yields of various maturities, that matures at time τ_N, to the three latent yield factors. The measurement equation of yields, is expressed as

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} & \frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1} \\ 1 & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} & \frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} & \frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N} \end{pmatrix} \begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} + \begin{pmatrix} \epsilon_t(\tau_1) \\ \epsilon_t(\tau_2) \\ \vdots \\ \epsilon_t(\tau_N) \end{pmatrix}. \quad (4.3.2)$$

The state space system of the model, given in matrix notation as,

$$y_t = \Theta f_t + \epsilon_t. \quad (4.3.3)$$

for $t = 1, ..., T$, where y_t representing the $N \times 1$ vector of yields, $N \times 3$ factor loading matrix Θ, 3×1 latent factor vector f_t and $N \times 1$ measurement errors of the yields . The stochastic errors $\epsilon_t(\tau)$, are interpreted as maturity specific factors, where each yield $y_t(\tau)$ is driven in part by three latent factors and in part by its stochastic error $\epsilon_t(\tau)$. The common factor dynamics of the three latent factors under VAR(1) give the state space system with the following transition equation which governs the time series dynamics of the latent factors,

$$\begin{pmatrix} L_t \\ S_t \\ C_t \end{pmatrix} = \begin{pmatrix} \mu_L \\ \mu_S \\ \mu_C \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix} \begin{pmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{pmatrix} + \begin{pmatrix} \nu_t(L) \\ \nu_t(S) \\ \nu_t(C) \end{pmatrix}. \quad (4.3.4)$$

The state space equation is then given by,

$$f_t = \mu + F(f_{t-1} - \mu) + \nu_t, \quad (4.3.5)$$

where μ is the 3×1 factor mean, and the 3×3 coefficient matrix F governs the factor dynamics. To specify the covariance structure of the measurement and transition disturbances, it is assumed that the white noise transition and measurement disturbances ϵ_t and ν_t are orthogonal to each other and to the initial state vector:

$$\begin{pmatrix} \epsilon_t \\ \nu_t \end{pmatrix} \sim WN \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right], \quad (4.3.6)$$

$$\begin{aligned} \mathbb{E}(F_0 \nu_t') &= 0 \\ \mathbb{E}(F_0 \epsilon_t') &= 0. \end{aligned} \quad (4.3.7)$$

Throughout this work, it is assumed that H and Q matrices are diagonal and non-diagonal, respectively. Assumptions of diagonal matrix H, imply that yields are uncorrelated for different maturities. Assumptions regarding the non-diagonal matrix Q allow shocks to the three term structure factors to be correlated. Continuing from the interpretation of factors given in Section (4.1), the empirical counterpart of the factors is that the level of loading on

L_t is the long term yield of the yield curve. The factor loading S_t correlates negatively with the spread between long and short yields, i.e. the slope of the yield curve. Loading on C_t is close to the empirical measure of curvature of the yield curve. DNS model introduced by Diebold and Li (2006) seem to be theoretically satisfactory as it enjoys the advantage of parsimony, goodness of fit, and superior forecast performance. Theoretically, DNS imposes certain economically desirable properties, such as requiring the discount function to approach zero with maturity. That is,

$$\lim_{\tau \to 0} P(\tau) = 0.$$

However, despite its good empirical performance and a certain amount of theoretical appeal, DNS fails on an important theoretical restrictions necessary to rule out opportunities for risk free arbitrage.

4.4 Compatibility of DNS Models and ATSMs

The type of affine term structure models with closed form, in discrete time, with no-arbitrage restrictions, have gained much popularity in macro-finance, following Ang and Piazzesi (2003). The model considers the combination of observable and unobservable variables to explain the entire yield curve. However, even though it considers theoretical restrictions for consistency, estimation and identification of price of risk is quite difficult, which leads to poor forecasting performance of future yields. This has been pointed out by Duffee (2002) that the canonical arbitrage free models exhibits poor empirical time series performance, especially when forecasting the future yields. On the other hand, the DNS model is shown to be extremely popular among practitioners and central bankers, due to its simplicity and effectiveness. However, the DNS does consider arbitrage opportunities, in that way, it lacks the traditional financial theory when applied to an efficient market. In this section, we make a comparison to identify how to adapt the DNS model to the no-arbitrage restrictions in discrete time. Both models are featuring a constant stochastic volatility to capture the bulk of variations in the yield curve. Recalling the solutions of the essentially affine arbitrage-free term structure model, following the set-up in Ang and Piazzesi (2003), the DNS model is shown to be incompatible with the arbitrage-free conditions, because the measurement Equation (4.3.1) corresponds to the representation of the affine function (3.3.21), with the coefficients taking the following form:

$$\begin{aligned} A_n &= 0; && (4.4.1) \\ B_n &= [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]', && (4.4.2) \end{aligned}$$

for all n, which does not satisfy the pricing Equation (3.3.18). In order for DNS model to be compatible with arbitrage-free assumptions, it must be proved that for any parameter A_n that follow the difference equation under measure $\mathbb{Q}$ or $\mathbb{P}$, finding the parameter B_n in the difference equation

$$B'_{n+1} \quad = \quad B'_n(\Phi_X^{\mathbb{P}} - \Sigma_X \Lambda) + B'_1,$$

such that the Nelson-Siegel factor loadings shown in Equation (4.4.2) for B_n holds.

4.5 Arbitrage-Free Dynamic Nelson-Siegel Model

The arbitrage-free Nelson-Siegel (AFNS) class model imposes a merged Nelson-Siegel functional coefficient structure on the canonical representation of arbitrage free affine dynamics derived in continuous time framework by Christensen *et al.* (2011). This combination of AFNS model addresses the theoretical weakness of the most popular Nelson-Siegel model and to overcome the empirical implementation problem concerning the canonical arbitrage-free ATSMs, by incorporating DNS elements which retains its empirical tractability. The model is derived under risk neutral dynamics to establish the sufficient conditions that guarantees the existence of a solution. Since it has been noted that the loadings in Duffie and Kan (1996) are exactly equal to those of DNS, Christensen *et al.* (2011) work tries to make DNS arbitrage free by finding the best approximation to DNS within the Duffie-Kan arbitrage free class, with factor loadings closest to the DNS. Christensen *et al.* (2011) considers the full three factor DNS model and also explicitly characterized the yield adjustment term, showing how DNS can be adjusted to fall within the Duffie-Kan workhorse model. The yield adjustment term is the key wedge between AFNS and DNS under $\mathbb{P}$-measure, and the constant depends only on maturity, not on time. Furthermore, AFNS supplies the missing piece through the yield adjustment term which links the transition and measurement equation in a way that ensures that discounted bond prices are semi-martingales and therefore arbitrage free.

Looking at the relationship between AFNS and DNS, the Christensen *et al.* (2011) AFNS model blends two important and successful approaches to yield curve modelling: the DNS empirically based one and the no-arbitrage theoretically based one. Models in the DNS tradition fit and forecast well, but they lack theoretical rigor as they may admit arbitrage possibilities, while models in the arbitrage free tradition are theoretically rigorous as they enforce absence of arbitrage, but they may fit and forecast poorly. AFNS bridges with a DNS model that enforces absence of arbitrage. Approach from the arbitrage free side, AFNS maintains the arbitrage free theoretical restrictions of the canonical affine models but can be easily estimated, because the dynamic Nelson-Siegel structure helps identify the latent yield curve factors and delivers analytical solution for the zero-coupon bond prices. Approach from the DNS side, AFNS maintains the simplicity and empirical tractability of the popular DNS model, while simultaneously enforcing the theoretically desired property of absence of riskless arbitrage. Another insightful information is that AFNS as compared to DNS, is very poweful in separating risk premia and expectations on future short rates.

4.5.1 Discrete-time AFNS (Yield Only) Model

Linlin and Gengming (2012) and Li *et al.* (2012) depict the model in discrete-time settings to establish a necessary condition and a uniqueness of a solution for a yield curve with dynamic Nelson-Siegel factor loadings to be arbitrage free. The main important assumption in AFNS model is that the mean-reversion coefficient matrix of the state dynamics, under risk-neutral measure, takes a specific form

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

The suitable $\Phi_X^{\mathbb{Q}}$, equates the factor loadings of the three yield factors to the dynamic Nelson-Siegel model, such that,

$$B_n' = B_1' \sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i = [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]. \tag{4.5.1}$$

That is, the parameters should be a function of λ in the Nelson-Siegel factor loadings. Equation (4.5.1) is key in understanding the solution and the proof to the arbitrage-free Nelson-Siegel class of models. Furthermore, in order to exclude arbitrage opportunities, the measurement equations of yields have to be adjusted by a constant term, which differs from the original Nelson-Siegel model, where $A_n = 0$. Hence, A_n gives the following equation under measure $\mathbb{Q}$, for $n > 1$. That is,

$$A_n = A_{n-1} + B_{n-1}' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_{n-1}' \Sigma_X \Sigma_X' B_{n-1} + A_1. \tag{4.5.2}$$

In order to prove the arbitrage-free Nelson-Siegel class of models and find the exact solution in discrete-time, the solution must be derived in discrete-time setting. The merged affine arbitrage free assumption and the Nelson-Siegel interpolation, under discrete time setting is shown to prove the existence and euniqueness of the solution for the AFNS model.

Proposition 4.1. *The sufficient condition for the DNS model to be arbitrage-free implies the existence of the solution for the AFNS model.*

Assumption 1. *Assume that the one-period risk-free rate is an affine function on a vector of all state variables X_t, which comprises of three latent Nelson-Siegel factors. The one-period risk-free rate is then defined by*

$$r_t = \delta_0 + \delta_1' X_t,$$

where

$$\delta_1 = [1, \frac{1-e^{-\lambda}}{\lambda}, \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda}]'.$$

Assumption 2. *The vector of state variables X_t follows a first-order Gaussian VAR process under the risk-neutral measure $\mathbb{Q}$,*

$$X_t = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}},$$

with the transition matrix taking the specific form

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

Using Assumption (1) and Assumption (2), the following implications can be deduced. That is, the zero-coupon bond prices, are exponentially affine on vector X_t, given by,

$$P_t^n = \exp(A_n + B_n' X_t),$$

where B_n is the Nelson-Siegel factor loading,

$$B_n = [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]'$$

and A_n satisfies the following difference equation for $n > 1$:

$$A_n = A_{n-1} + B_{n-1}' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_{n-1}' \Sigma_X \Sigma_X' B_{n-1} + A_1. \tag{4.5.3}$$

4.5.2 Proof of Proposition 4.1

Proof. If $n = 1$, we have the following

$$B_1 = -\delta_1 = -[1, \frac{1-e^{-\lambda}}{\lambda}, \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda}]' \quad \text{and} \tag{4.5.4}$$

$$A_1 \neq 0. \tag{4.5.5}$$

The parameter B_n in the difference equation (3.3.36) under the risk-neutral measure in discrete-time ATSM, gives the following expression:

$$B_{n+1}' = B_n' \Phi_X^{\mathbb{Q}} + B_1'.$$

Iteratively, this can be written as

$$B'_n = -\delta_1 \sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i \quad \forall n > 1,$$

where the i^{th} transition matrix

$$\begin{aligned}
\left(\Phi_X^{\mathbb{Q}}\right)^i &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}^i = e^{-i\lambda} \begin{bmatrix} e^{\lambda} & 0 & 0 \\ 0 & 1 & \lambda \\ 0 & 0 & 1 \end{bmatrix}^i \\
&= e^{-i\lambda} \begin{bmatrix} e^{i\lambda} & 0 & 0 \\ 0 & 1 & i\lambda \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-i\lambda} & i\lambda e^{-i\lambda} \\ 0 & 0 & e^{-i\lambda} \end{bmatrix}, \qquad (4.5.6) \\
\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i &= \begin{bmatrix} n & 0 & 0 \\ 0 & \sum_{i=0}^{n-1} e^{-i\lambda} & \sum_{i=0}^{n-1} i\lambda e^{-i\lambda} \\ 0 & 0 & \sum_{i=0}^{n-1} e^{-i\lambda} \end{bmatrix} \\
&= \begin{bmatrix} n & 0 & 0 \\ 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} & \lambda \left[\frac{e^{-\lambda}-e^{-n\lambda}}{(1-e^{-\lambda})^2} - \frac{(n-1)e^{-n\lambda}}{1-e^{-\lambda}} \right] \\ 0 & 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} \end{bmatrix}.
\end{aligned}$$

It follows that

$$\begin{aligned}
B_n &= -\left[\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i \right]' \delta_1 \\
&= -\begin{bmatrix} n & 0 & 0 \\ 0 & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} & 0 \\ 0 & \lambda \left[\frac{e^{-\lambda}-e^{-n\lambda}}{(1-e^{-\lambda})^2} - \frac{(n-1)e^{-n\lambda}}{1-e^{-\lambda}} \right] & \frac{1-e^{-n\lambda}}{1-e^{-\lambda}} \end{bmatrix} \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{bmatrix} \\
&= -\begin{bmatrix} n \\ \frac{1-e^{-n\lambda}}{\lambda} \\ \frac{e^{-\lambda}-e^{-n\lambda}}{1-e^{-\lambda}} - (n-1)e^{-n\lambda} + \frac{1-e^{-n\lambda}}{\lambda} - \frac{e^{-\lambda}-e^{-(n+1)\lambda}}{1-e^{-\lambda}} \end{bmatrix} \\
&= \left[-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda} \right]'.
\end{aligned}$$

Hence, the constant adjustment term which can be derived recursively, becomes

$$A_{n+1} = A_n + B'_n \mu_X^{\mathbb{Q}} + \frac{1}{2} B'_n \Sigma_X \Sigma'_X B_n + A_1. \qquad (4.5.7)$$

Therefore, existence of the solution is satisfied. □

Proposition 4.2. *The necessary conditions for the DNS model implies a uniqueness of the solution for the AFNS model.*

To show that AFNS has a unique solution;

Assumption 3. *We assume that the DNS model with a constant adjustment term on measurement equation of yields, a_n, fits the yield curve completely. This gives,*

$$y_t(n) = a_n + L_t + \left(\frac{1-e^{-\lambda n}}{\lambda n}\right) S_t + \left(\frac{1-e^{-\lambda n}}{\lambda n} - e^{-\lambda n}\right) C_t. \tag{4.5.8}$$

Assumption 4. *The estimated Nelson-Siegel yield curve factors follow a $VAR(1)$ process under the risk-neutral measure $\mathbb{Q}$,*

$$X_t = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}. \tag{4.5.9}$$

The underlying Assumption (3), Assumption (4) of yield curve $y_t(t+n)$, implies that the instantaneous risk-free rate needs to follow the affine process

$$r_t = \delta_0 + \delta_1' X_t, \tag{4.5.10}$$

where

$$\delta_1 = [1, \frac{1-e^{-\lambda}}{\lambda}, \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda}]' \qquad \text{and} \qquad \delta_0 = a_1;$$

Furthermore, the risk-neutral dynamic coefficient matrix $\Phi_X^{\mathbb{Q}}$ of the Nelson-Siegel three latent factors needs to satisfy

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

When $n = 1$, the one period short-rate in (4.5.10) is easily satisfied by the measurement equation in Assumption (3). Since the state dynamics, X_t and the short rate r_t are assumed to be affine functions of the term structure model, this implies that the function of yields is affine

$$y_t^{t+n} = -\frac{1}{n}\left(A_n + B_n' X_t\right), \tag{4.5.11}$$

where the difference equations

$$A_{n+1} = A_n + B_n' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_n' \Sigma_X \Sigma_X' B_n + A_1. \tag{4.5.12}$$

$$B_{n+1}' = B_n' \Phi_X^{\mathbb{Q}} + B_1'. \tag{4.5.13}$$

4.5.3 Proof of Proposition 4.2

Proof. By comparing the yield curve Equation (4.5.11) with Equation (4.5.8), we have the following coefficients

$$\begin{aligned} B_n &= [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]'. \\ A_n &= -na_n. \end{aligned} \tag{4.5.14}$$

Solving the three latent factor matrix, let

$$\Theta_X^{\mathbb{Q}} = \begin{bmatrix} \phi_{11} & \phi_{21} & \phi_{31} \\ \phi_{12} & \phi_{22} & \phi_{32} \\ \phi_{13} & \phi_{23} & \phi_{33} \end{bmatrix}.$$

We then insert Equation (4.5.14) in both sides of Equation (4.5.13) and take the transpose, giving

$$\begin{bmatrix} n+1 \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} \\ \frac{1-e^{-\lambda(n+1)}}{\lambda} - (n+1)e^{-\lambda(n+1)} \end{bmatrix} = \Theta_X^{\mathbb{Q}} \begin{bmatrix} n \\ \frac{1-e^{-\lambda n}}{\lambda} \\ \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n} \end{bmatrix} + \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda}}{\lambda} \\ \frac{1-e^{-\lambda}}{\lambda} - e^{-\lambda} \end{bmatrix}.$$

We can rearrange it such that

$$\begin{bmatrix} n \\ \frac{e^{-\lambda}-e^{-\lambda(n+1)}}{\lambda} \\ \frac{e^{-\lambda}-e^{-\lambda(n+1)}}{\lambda} - (n+1)e^{-\lambda(n+1)} + e^{-\lambda} \end{bmatrix} = \Theta_X^{\mathbb{Q}} \begin{bmatrix} n \\ \frac{1-e^{-\lambda n}}{\lambda} \\ \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n} \end{bmatrix}.$$

Solving the elements of $\Theta_X^{\mathbb{Q}}$, coefficients need to be compared, equation by equation. The first equation implies that

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

The second equation implies that

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

where the second line requires the following to hold:

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

The third equation implies that

$$\phi_{13} = 0;$$

and

$$\frac{e^{-\lambda}}{\lambda} + e^{-\lambda} - e^{-\lambda}(1+\lambda)\frac{e^{-\lambda n}}{\lambda} - e^{-\lambda}ne^{-\lambda n} = \phi_{23}\frac{1}{\lambda} + \phi_{33}\frac{1}{\lambda} - (\phi_{23}+\phi_{33})\frac{e^{-\lambda n}}{\lambda} \\ - \phi_{33}ne^{-\lambda n},$$

where the second line requires

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

To summarise, we have

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

The risk-neutral dynamic coefficient matrix $\Phi_X^{\mathbb{Q}}$ of the three latent factors is satisfied, and AFNS has a unique solution. □

4.5.4 Discrete-Time AFGNS Model

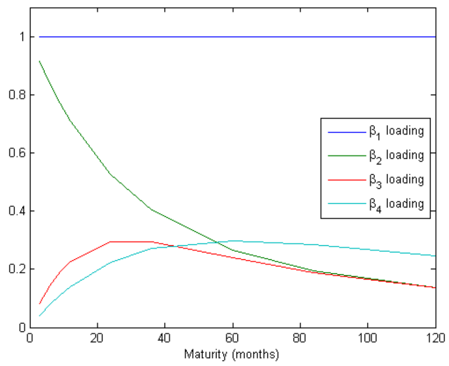
To estimate the yield curve central banks use different models. The table 4.1 provides a review of the use of certain types of models in some countries. It can be seen that in the most foreign countries dominate Svensson (1995) and Nelson and Siegel (1987) models.

Table 4.1: Central Banks and Yield Curve Models

Central Bank	**Model**
Belgium	Nelson-Siegel, Svensson
Canada	Svensson
USA	Fisher-Nychka-Zervos (Spline)
Finland	Nelson-Siegel
France	Nelson-Siegel, Svensson
Germany	Svensson
Italy	Nelson-Siegel
Japan	Fisher-Nychka-Zervos (Spline)
Norway	Svensson
Spain	Svensson
UK	Anderson and Sleath (Spline) (Until 2001 Svensson)
Sweden	Fisher-Nychka-Zervos (Spline), before Svensson
Switzerland	Svensson
EU	Svensson

Source: Pereda, J. (2009), "Estimación de la Curva de Rendimiento Cupón Cero para el Perú", Banco Central de reserva del Perú, Revista Estudios Económicos N° 17

Nelson-Siegel model is extremely popular in the practice; both individual investors and the central banks use this model. Svensson model is an extension of the Nelson-Siegel model, which adds a second curvature term to the DNS model and provides sufficient precision adjustment and smooth curve shape of periodic rents. It has become very popular in the mid 90's. Since then, it has been used by substantial number of central banks worldwide to assess the bond structure with a single coupon and future interest rates (forward rates). Amongst some of the recent work in the literature, Christensen *et al.* (2009) make use of both these popular models, by proposing an arbitrage-free generalized Nelson-Siegel term structure model. The AFGNS model adds an additional slope factor to the dynamic Nelson-Siegel-Svensson (DNSS) model to see whether risk-neutral restrictions can be found that make the five-factor model arbitrage free, and also to allow for a better fit along maturities.



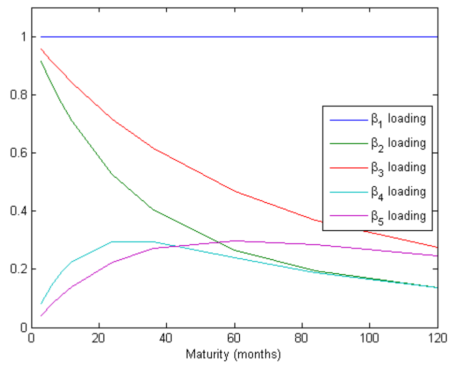


Figure 4.3: Factor loadings of DNSS and DGNS Models. The figure on the left shows the factor loading of the four state variables in the function of DNSS model with $\lambda_1 = 0.0609$ and $\lambda_2 = 0.0295$. The figure on the Right shows the factor loading of the five state variables in the function of DGNS model with $\lambda_1 = 0.0609$ and $\lambda_2 = 0.0295$.

The Nelson-Siegel in-sample problem, regarding fitting the long maturities is that, the slope and the curvature factor loadings decay rapidly and approach zero with maturity. Thus, only the level factor available to fit the yields with long maturities. Empirically, the problem results into a lack of fit of the long term maturities as noted by Christensen *et al.* (2009). The Svensson model seems to overcome this caveat of fitting the cross section of yields. The dynamic factor of the Nelson-Siegel-Svensson curve (DNSS) is represented by the measurement equation of the form

$$\begin{aligned} y_t(n) &= L_t + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n}\right) S_t + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n} - e^{-\lambda_1 n}\right) C_{1,t} \\ &\quad + \left(\frac{1-e^{-\lambda_2 n}}{\lambda_2 n} - e^{-\lambda_2 n}\right) C_{2,t}, \end{aligned}$$

where the four factor loadings are functions of maturity as shown in Figure 4.3. The four factor DNSS has similar critiques on riskless arbitrage noted in DNS model. To derive an arbitrage-free Gaussian version of DNSS that satisfies the mechanics of Proposition AFNS, Christensen *et al.* (2009), give a generalization of factor loadings to match the second curvature factor in DNSS. The five factor arbitrage-free generalised Nelson-Siegel model (AFGNS) yield function, gives the form

$$\begin{aligned} y_t(n) &= a_n + L_t + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n}\right) S_{1,t} + \left(\frac{1-e^{-\lambda_2 n}}{\lambda_2 n}\right) S_{2,t} \\ &\quad + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n} - e^{-\lambda_1 n}\right) C_{1,t} + \left(\frac{1-e^{-\lambda_2 n}}{\lambda_2 n} - e^{-\lambda_2 n}\right) C_{2,t}. \end{aligned}$$

The following proof derives the existence and uniqueness of the solution for the AFGNS measurement equation of the yield curve in discrete time.

Proposition 4.3. *The sufficient condition for the DNS model implies the existence of a solution for the AFGNS.*

Assumption 5. *Assume that the instantaneous risk-free rate is defined by*

$$r_t = \delta_0 + \delta_1' X_t,$$

where $\delta_0 = \left[1, \frac{1-e^{-\lambda_1}}{\lambda_1}, \frac{1-e^{-\lambda_2}}{\lambda_2}, \frac{1-e^{-\lambda_1}}{\lambda_1} - e^{-\lambda_1}, \frac{1-e^{-\lambda_2}}{\lambda_1} - e^{-\lambda_2}\right]'$.

Assumption 6. *The state variables in* $X_t = (X_t^1, X_t^2, X_t^3, X_t^4, X_t^5)$ *follow a VAR process under the risk-neutral measure* $\mathbb{Q}$*, such that*

$$X_t = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}.$$

Moreover, the crucial element of the solution is that the dynamic coefficient matrix, $\Phi_X^{\mathbb{Q}}$*, takes the following form*

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & e^{-\lambda_1} & 0 & \lambda_1 e^{-\lambda_1} & 0 \\ 0 & 0 & e^{-\lambda_2} & 0 & \lambda_2 e^{-\lambda_2} \\ 0 & 0 & 0 & e^{-\lambda_1} & 0 \\ 0 & 0 & 0 & 0 & e^{-\lambda_2} \end{bmatrix}.$$

From the above state dynamics Assumption (5) and Assumption (6), imply that the zero-coupon bond prices, are exponentially affine on vector X_t,

$$P_t^n = \exp(A_n + B_n' X_t).$$

The parameter B_n are Nelson-Siegel factor loadings with a second set of slope and curvature loadings determined by a different shape parameter. That is, the unique solution for the system of ODE (4.5.13) is

$$B_n = \left[-n, -\tfrac{1-e^{-\lambda_1 n}}{\lambda_1}, -\tfrac{1-e^{-\lambda_2 n}}{\lambda_2}, ne^{-\lambda_1 n} - \tfrac{1-e^{-\lambda_1 n}}{\lambda_1}, ne^{-\lambda_2 n} - \tfrac{1-e^{-\lambda_2 n}}{\lambda_2}\right]',$$

and A_n satisfies the difference equation of the form

$$A_n = A_{n-1} + B_{n-1}' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_{n-1}' \Sigma_X \Sigma_X' B_{n-1}, \text{ for } n > 1.$$

4.5.5 Proof of Proposition 4.3

Proof. Recall that the ordinary difference equation for B_n under the risk-neutral measure in discrete-time ATSM, is given by

$$B_{n+1} \quad = \quad \Phi_X^{\mathbb{Q}'} B_n - \delta_1. \tag{4.5.15}$$

Simplifying Equation (4.5.15), gives

$$\begin{bmatrix} B_{n+1}^1 \\ B_{n+1}^2 \\ B_{n+1}^3 \\ B_{n+1}^4 \\ B_{n+1}^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & e^{-\lambda_2} & 0 & 0 & 0 \\ 0 & 0 & e^{-\lambda_2} & 0 & 0 \\ 0 & \lambda_1 e^{-\lambda_1} & 0 & e^{-\lambda_1} & 0 \\ 0 & 0 & \lambda_2 e^{-\lambda_2} & 0 & e^{-\lambda_2} \end{bmatrix} \begin{bmatrix} B_n^1 \\ B_n^2 \\ B_n^3 \\ B_n^4 \\ B_n^5 \end{bmatrix} - \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda_1}}{\lambda_1} \\ \frac{1-e^{-\lambda_2}}{\lambda_2} \\ \frac{1-e^{-\lambda_1}}{\lambda_1} - e^{-\lambda_1} \\ \frac{1-e^{-\lambda_2}}{\lambda_2} - e^{-\lambda_2} \end{bmatrix},$$

which can be divided into two sets of independent difference equations:

$$\begin{bmatrix} B_{n+1}^1 \\ B_{n+1}^2 \\ B_{n+1}^4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda_1} & 0 \\ 0 & \lambda_1 e^{-\lambda_1} & e^{-\lambda_1} \end{bmatrix} \begin{bmatrix} B_n^1 \\ B_n^2 \\ B_n^4 \end{bmatrix} - \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda_1}}{\lambda_1} \\ \frac{1-e^{-\lambda_1}}{\lambda_1} - e^{-\lambda_1} \end{bmatrix},$$

and

$$\begin{bmatrix} B_{n+1}^1 \\ B_{n+1}^3 \\ B_{n+1}^5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda_2} & 0 \\ 0 & \lambda_2 e^{-\lambda_2} & e^{-\lambda_2} \end{bmatrix} \begin{bmatrix} B_n^1 \\ B_n^3 \\ B_n^5 \end{bmatrix} - \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda_2}}{\lambda_2} \\ \frac{1-e^{-\lambda_2}}{\lambda_2} - e^{-\lambda_2} \end{bmatrix}.$$

Then, following the similar manner as shown in proof of Proposition (4.1) for the AFNS model, this simply gives,

$$
\begin{aligned}
\left(\Phi_X^{\mathbb{Q}}\right)^i &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-i\lambda_1} & 0 \\ 0 & i\lambda_1 e^{-i\lambda_1} & e^{-i\lambda_1} \end{bmatrix}, \\
\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i &= \begin{bmatrix} n & 0 & 0 \\ 0 & \frac{1-e^{-n\lambda_1}}{1-e^{-\lambda_1}} & 0 \\ 0 & \lambda_1 \left[\frac{e^{-\lambda_1}-e^{-n\lambda_1}}{(1-e^{-\lambda_1})^2} - \frac{(n-1)e^{-n\lambda_1}}{1-e^{-\lambda_1}} \right] & \frac{1-e^{-n\lambda_1}}{1-e^{-\lambda_1}} \end{bmatrix}.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
B_n &= -\left[\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}})^i \right]' \Delta \\
&= \left[-n, -\frac{1-e^{-\lambda_1 n}}{\lambda_1}, ne^{-\lambda_1 n} - \frac{1-e^{-\lambda_1 n}}{\lambda_1} \right]'.
\end{aligned}
$$

Following the same pattern for the second set, we get that

$$
B_n = \left[-n, -\tfrac{1-e^{-\lambda_1 n}}{\lambda_1}, -\tfrac{1-e^{-\lambda_2 n}}{\lambda_2}, ne^{-\lambda_1 n} - \tfrac{1-e^{-\lambda_1 n}}{\lambda_1}, ne^{-\lambda_2 n} - \tfrac{1-e^{-\lambda_2 n}}{\lambda_2} \right].
$$

Moreover, it can be shown that the constant adjustment term

$$
A_n = A_{n-1} + B'_{n-1}\mu_X^{\mathbb{Q}} + \frac{1}{2} B'_{n-1}\Sigma_X \Sigma'_X B_{n-1},
$$

can be computed recursively from the measurement equation. □

Proposition 4.4. *The necessary condition for the DNS model implies the uniqueness of the solution for the AFGNS model.*

Assumption 7. *Assume that the following dynamic generalized Nelson-Siegel model with constant term can fit the yield curve completely, where $\lambda_1 \neq \lambda_2$. That is,*

$$
\begin{aligned}
y_t(n) &= a_n + L_t + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n} \right) S_{1,t} + \left(\frac{1-e^{-\lambda_2 n}}{\lambda_2 n} \right) S_{2,t} \\
&\quad + \left(\frac{1-e^{-\lambda_1 n}}{\lambda_1 n} - e^{-\lambda_1 n} \right) C_{1,t} + \left(\frac{1-e^{-\lambda_2 n}}{\lambda_2 n} - e^{-\lambda_2 n} \right) C_{2,t}.
\end{aligned} \tag{4.5.16}
$$

Assumption 8. *The general Nelson-Siegel latent factors follow a VAR(1) process under the risk-neutral measure $\mathbb{Q}$, such that*

$$
X_t = \mu_X^{\mathbb{Q}} + \Phi_X^{\mathbb{Q}} X_{t-1} + \Sigma_X \epsilon_t^{\mathbb{Q}}.
$$

According to the yield curve Assumption (7) and state dynamics Assumption (8), this implies that the risk-free rate follows the affine process,

$$r_t = \delta_0 + \delta_1' X_t,$$

where

$$\delta_1 = [1, \frac{1-e^{-\lambda_1}}{\lambda_1}, \frac{1-e^{-\lambda_2}}{\lambda_2}, \frac{1-e^{-\lambda_1}}{\lambda_1} - e^{-\lambda_1}, \frac{1-e^{-\lambda_2}}{\lambda_1} - e^{-\lambda_2}]', \text{ and } \delta = a_1.$$

Furthermore, risk neutral dynamics for the latent factors satisfy

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & e^{-\lambda_1} & 0 & \lambda_1 e^{-\lambda_1} & 0 \\ 0 & 0 & e^{-\lambda_2} & 0 & \lambda_2 e^{-\lambda_2} \\ 0 & 0 & 0 & e^{-\lambda_1} & 0 \\ 0 & 0 & 0 & 0 & e^{-\lambda_2} \end{bmatrix}.$$

When $n = 1$ the one period short rate r_t, is guaranteed by the measurement Equation (4.5.16), which is affine. Following Assumption (8) and the implication that the short rate takes on the affine form, shows that the model is also affine, such that the yields are affine functions of X_t, given by

$$y_t(t+n) = -\frac{1}{n}(A_n + B_n' X_t),$$

where

$$A_{n+1} = A_n + B_n' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_n' \Sigma_X \Sigma_X' B_n + A_1. \tag{4.5.17}$$

$$B_{n+1}' = B_n' \Phi_X^{\mathbb{Q}} + B_1'. \tag{4.5.18}$$

4.5.6 Proof of Proposition 4.4

Proof. By comparing the difference Equation (4.5.18) and the measurement Equation (4.5.16), we have

$$B_n = \left[-n, -\frac{1-e^{-\lambda_1 n}}{\lambda_1}, -\frac{1-e^{-\lambda_2 n}}{\lambda_2}, ne^{-\lambda_1 n} - \frac{1-e^{-\lambda_1 n}}{\lambda_1}, ne^{-\lambda_2 n} - \frac{1-e^{-\lambda_2 n}}{\lambda_2} \right].$$

Substitute the factor loading B_n into both sides of Equation (4.5.18) and take the transpose, we then have

$$\begin{bmatrix} B_{n+1}^1 \\ B_{n+1}^2 \\ B_{n+1}^3 \\ B_{n+1}^4 \\ B_{n+1}^5 \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} & \phi_{51} \\ \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} & \phi_{52} \\ \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} & \phi_{53} \\ \phi_{14} & \phi_{24} & \phi_{34} & \phi_{44} & \phi_{54} \\ \phi_{15} & \phi_{25} & \phi_{35} & \phi_{45} & \phi_{55} \end{bmatrix} \begin{bmatrix} B_n^1 \\ B_n^2 \\ B_n^3 \\ B_n^4 \\ B_n^5 \end{bmatrix} - \begin{bmatrix} 1 \\ \frac{1-e^{-\lambda_1}}{\lambda_1} \\ \frac{1-e^{-\lambda_2}}{\lambda_2} \\ \frac{1-e^{-\lambda_1}}{\lambda_1} - \lambda_1 e^{-\lambda_1} \\ \frac{1-e^{-\lambda_2}}{\lambda_2} - \lambda_2 e^{-\lambda_2} \end{bmatrix}.$$

After we rearrange, let

$$\Pi_X^{\mathbb{Q}} = \begin{bmatrix} \phi_{11} & \phi_{21} & \phi_{31} & \phi_{41} & \phi_{51} \\ \phi_{12} & \phi_{22} & \phi_{32} & \phi_{42} & \phi_{52} \\ \phi_{13} & \phi_{23} & \phi_{33} & \phi_{43} & \phi_{53} \\ \phi_{14} & \phi_{24} & \phi_{34} & \phi_{44} & \phi_{54} \\ \phi_{15} & \phi_{25} & \phi_{35} & \phi_{45} & \phi_{55} \end{bmatrix},$$

then,

$$\begin{bmatrix} n \\ \frac{e^{-\lambda_1}-e^{-\lambda_1(n+1)}}{\lambda_1} \\ \frac{e^{-\lambda_2}-e^{-\lambda_2(n+1)}}{\lambda_2} \\ \frac{e^{-\lambda_1}-e^{-\lambda_1(n+1)}}{\lambda_1} - (n+1)e^{-\lambda_1(n+1)} + e^{-\lambda_1} \\ \frac{e^{-\lambda_2}-e^{-\lambda_2(n+1)}}{\lambda_2} - (n+1)e^{-\lambda_2(n+1)} + e^{-\lambda_2} \end{bmatrix} = \Pi_X^{\mathbb{Q}} \begin{bmatrix} n \\ \frac{1-e^{-\lambda_1 n}}{\lambda_1} \\ \frac{1-e^{-\lambda_2 n}}{\lambda_2} \\ \frac{1-e^{-\lambda_1 n}}{\lambda_1} - ne^{-\lambda_1 n} \\ \frac{1-e^{-\lambda_2 n}}{\lambda_2} - ne^{-\lambda_2 n} \end{bmatrix}.$$

Next, we solve for the elements of $\Phi_X^{\mathbb{Q}}$ by comparing the coefficients in $\Pi_X^{\mathbb{Q}}$, equation by equation. The first equation implies that

$$\begin{aligned} \phi_{11} &= 1, \\ \text{and} \quad \phi_{21} &= \phi_{31} = \phi_{41} = \phi_{51} = 0. \end{aligned}$$

The second and third equations imply that

$$\begin{aligned} \phi_{12} &= \phi_{32} = \phi_{42} = \phi_{52} = 0, \\ \phi_{22} &= e^{-\lambda_1}, \\ \phi_{13} &= \phi_{33} = \phi_{43} = \phi_{53} = 0, \\ \text{and} \quad \phi_{33} &= e^{-\lambda_2}. \end{aligned}$$

The fourth equation implies that

$$\phi_{14} = \phi_{34} = \phi_{54} = 0,$$

and

$$\begin{aligned} \frac{e^{-\lambda_1}}{\lambda_1} + e^{-\lambda_1} - e^{-\lambda_1}(1+\lambda_1)\frac{e^{-\lambda_1 n}}{\lambda_1} - e^{-\lambda_1} ne^{-\lambda_1 n} = \phi_{24}\frac{1}{\lambda_1} + \phi_{44}\frac{1}{\lambda_1} - (\phi_{24} + \phi_{44}) \\ \times \frac{e^{-\lambda_1 n}}{\lambda_1} - \phi_{44} ne^{-\lambda_1 n}, \end{aligned}$$

where the second line requires

$$\begin{aligned} \phi_{44} &= e^{-\lambda_1}, \\ \text{and} \quad \phi_{24} &= \lambda_1 e^{-\lambda_1}. \end{aligned}$$

Similarly, for the fifth equation, we have

$$\begin{aligned} \phi_{15} &= \phi_{25} = \phi_{45} = 0, \\ \phi_{55} &= e^{-\lambda_2}, \\ \text{and} \quad \phi_{35} &= \lambda_2 e^{-\lambda_2}. \end{aligned}$$

In summary, we have

$$\Phi_X^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & e^{-\lambda_1} & 0 & \lambda_1 e^{-\lambda_1} & 0 \\ 0 & 0 & e^{-\lambda_2} & 0 & \lambda_2 e^{-\lambda_2} \\ 0 & 0 & 0 & e^{-\lambda_1} & 0 \\ 0 & 0 & 0 & 0 & e^{-\lambda_2} \end{bmatrix}.$$

The risk-neutral dynamic coefficient matrix 1 of four latent factors is satisfied, and AFGNS has a unique solution. □

The arbitrage-free Nelson-Siegel (AFNS) model incorporates the parsimony of the traditional Nelson-Siegel model and the theoretical rigor of affine no-arbitrage term structure model. However, it is found to be significantly restrictive, (see Li *et al.* (2012)), since its consistency is within the presence of exactly three state variables and does not consider the inclusion of observable macro variables. Some of the recent work that uses four[2] to five factor model, obtained some improvement in fitting and estimating the curve, such that the models, as noted by Almeida *et al.* (2009), outperforms the three factor Diebold and Li (2006) model in forecasting. Nevertheless, the important point is that, restricting state variables to a three dimensional factor structure is typically sufficient.

4.6 Specification Analysis

As discussed in Dai and Singleton (2000), there often exists an over-identification problem in affine term structure models. For example, some parallel transformation to the drift, $\mu_X^{\mathbb{Q}}$, in the state dynamics and intercept, δ, in the short rate equation may results in the same distribution of the yields. Hence, we need to perform specification analysis to define the canonical representation for identification.

In the AFNS model, we let $\mu_X^{\mathbb{Q}} = [\mu_L^{\mathbb{Q}}, 0, 0]'$ and $\delta = 0$. Equivalently, we can restrict $\mu_X^{\mathbb{Q}} = 0$ but free up δ. In our analysis we choose the former, with which

[2] Cochrane and Piazzesi (2005) use a fourth factor which they argue has strong predictive content.

the difference equations of coefficients in the measurement equation becomes

$$A_{n+1} = A_n + n\mu_L^{\mathbb{Q}} + \frac{1}{2}B_n'\Sigma_X\Sigma_X'B_n, \quad \text{with} \quad A_1 = 0. \tag{4.6.1}$$

$$B_n = [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]'. \tag{4.6.2}$$

Likewise, for the AFGNS model, similar restrictions on specification can be imposed. From the specification analysis, it is clear that the AFNS model only has one additional parameter, $\mu_L^{\mathbb{Q}}$, compared to the DNS model. This guarantees the maximal parsimony of the AFNS model. Christensen *et al.* (2011) considered the full three factor DNS model and also explicitly characterized the yield adjustment term, showing how DNS can be adjusted to fall within the Duffie and Kan (1996) affine class of arbitrage models. According to Christensen *et al.* (2011), DNS requires only a time invariant yield adjustment term, which accounts for Jensen's inequality effects, to make it completely arbitrage free.

4.7 Analysis of Yield only Representation

ATSM, DNS and AFNS, all rely on yield curve information to extract factors that determine the dynamics and span the space of the yield curve. However, we know that yields at different maturities do not stand alone from the macroeconomy and its dynamics. Fisher equation (Fisher (1930)) states that in equilibrium nominal interest rate and inflation should have one-to-one relationship.

In modern economies, central banks influence the short term interest rate through open market operations by reacting to economic conditions such as unemployment rate and inflation, (Taylor (1993)), and even to a large amount of macroeconomic indicators, see (Bernanke and Boivin (2003), Bernanke *et al.* (2005)). Since long-term interest rates are weighted expected short term interest rates adjusted for risks, the state and dynamics of the economy and the monetary policy transmission mechanism are useful in understanding the yield curve movements at proper frequencies.

On the other hand, due to the limited information available for macroeconomic data, in order to investigate the dynamics and structure of the macroeconomy, it would be very helpful to study the joint dynamics of the yield curve and the macroeconomy in an integrated framework, because rich data existing in the bond market can provide valuable information about market expectations of the underlying macroeconomy.

4.8 Main Findings for Discrete time AFNS

The discrete time version of Christensen *et al.* (2011) derives the A FNS class of ATSMs with an exact solution and proof of uniqueness. In this model by Li *et al.* (2012) both sufficient and necessary conditions for NS factor loadings to be arbitrage-free are important. The sufficient conditions guarantee the existence of a solution. The necessary conditions are crucial firstly for providing the uniqueness of the solution that is characterised by a specific form of transition matrix in risk neutral state dynamics. Empirically, necessary conditions are also important for estimation and identification, as this rules out alternative solutions which might lead to observationally equivalent yields. To further improve the application of the model, Li *et al.* (2012) designed a simple, fast and reliable estimation procedure to embedded multi-step regression. The empirical results for their estimated discrete time AFNS model with US yield curve data for the past three decades and the recent financial crisis showed some interesting implications as compared to DNS. That is, the model performs better with in-sample forecast. In addition, AFNS credits the US bond yield conundrum to unusually low risk premia and notifications with sharply rising risk premia before the financial crisis.

Chapter 5

Generalised AFNS with Macro-factors

This chapter introduces a generalised arbitrage-free macro-finance term structure model with both Nelson-Siegel latent factors and observable macro factors, to be noted as arbitrage-free Nelson-Siegel macro-finance (AFNSMA). The model is an enlarged framework of the affine arbitrage class of Nelson-Siegel term structure models as shown in Christensen *et al.* (2009) derived under discrete time by Linlin and Gengming (2012). Two sub-classes for the model derivation are introduced: spanned versus unspanned model, depending on whether the macro factors span the yield curve. The unspanned class is where the macro variables do not span the yield curve, such that yields are still determined only by the three Nelson-Siegel yield factors and a constant adjustment term, under no-arbitrage assumptions. The spanned class, is where macro factors have non-zero factor loadings on the yield curve. In both model-classes, the yield factors and macro factors interact with each other in the dynamic state equation under the physical measure $\mathbb{P}$. To overcome the shortcomings of the macro-finance affine term structure model and the reduced-form Nelson-Siegel model with macro factors, Li *et al.* (2012) propose a unified modelling framework to integrate the Nelson-Siegel yield factors and macro factors under the premise of absence of arbitrage.

5.1 Unspanned Model

This part of the unspanned model class derives the existence of the solution for AFNSMA under the notion that macro variables do not span the cross-section of yields.

Proposition 5.1. *The sufficient condition for the DNS model to be arbitrage-free implies the existence of the solution for the AFNSMA model.*

Assumption 9. *Assume that the instantaneous risk-free rate is affine on a vector of variables X_t, which contains three latent Nelson-Siegel factors X_t^u and observable variables X_t^o,*

$$r_t = \delta_0 + \delta_1^{u'} X_t + \delta_1^{o'} X_t,$$

where

$$\delta_1^u = [1, \frac{1-e^{-\lambda}}{\lambda}, \frac{1-e^{-\lambda}}{\lambda} - \lambda e^{-\lambda}]'$$

and $X_t^u = (X_t^1, X_t^2, X_t^3)'$. Define $A_1 = \delta_0$ and $B_1 = \delta_1$. The state variables are described by the following VAR under the risk-neutral $\mathbb{Q}$,

$$\begin{pmatrix} X_t^u \\ X_t^o \end{pmatrix} = \mu_X^{\mathbb{Q}} + \begin{pmatrix} \Phi_u^{\mathbb{Q}} & 0 \\ \Phi_{uo}^{\mathbb{Q}} & \Phi_o^{\mathbb{Q}} \end{pmatrix} \begin{pmatrix} X_{t-1}^u \\ X_{t-1}^o \end{pmatrix} + \epsilon_t^{\mathbb{Q}},$$

where,

$$\Phi_u^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

Then the zero-coupon bond price is given by,

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

with B_n^u the Nelson-Siegel factor loading

$$\begin{aligned} B_n^u &= -\delta_1 = [n, \frac{1-e^{-\lambda n}}{\lambda}, \frac{1-e^{-\lambda n}}{\lambda} - ne^{-\lambda n}]' \\ B_n^o &= 0, \end{aligned}$$

and A_n satisfies the difference equations for $n > 1$:

$$A_{n+1} = A_n + B_n' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_n^{u'} \Omega_X^u B_n^u + A_1, \tag{5.1.1}$$

where $\Omega_X = \Sigma_X \Sigma_X'$ gives the upper left block of the variance-covariance matrix[1] corresponding to innovations to Nelson-Siegel factors in the state dynamics.

[1]From Ang and Piazzesi (2003), notice that $\Sigma_X \Sigma_X' = \Omega$ and Σ_X is identified as the $K \times K$ matrix. But in AFNSMA model only Ω that has been considered, which result in risk price parameters λ_0 and λ_1 to have different meaning and scale compared to Ang-Piazzesi model. Moreover, the pricing kernel is defined as $M_{t+1} = \exp(-\frac{1}{2}\lambda_t' \Omega \lambda_t - \delta_0 - \delta_1' X_t - \lambda_t' \nu_{t+1}^{\mathbb{P}})$, where $\nu_{t+1} \sim N(0, \Omega)$. In Ang and Piazzesi (2003) $\nu_{t+1} = \Sigma_X' \epsilon_{t+1}$, with $\epsilon_{t+1} \sim N(0, I)$.

5.1.1 Proof of Proposition 5.1

Proof. Recall the ordinary difference equation of B_n as shown in Equation (3.3.36), we have

$$B_{n+1} = B_n' \Phi_X^{\mathbb{Q}} - \delta_1,$$

that is

$$\begin{bmatrix} B_{n+1}^u \\ B_{n+1}^o \end{bmatrix} = \begin{bmatrix} \Phi_u^{\mathbb{Q}'} & \Phi_{uo}^{\mathbb{Q}'} \\ 0 & \Phi_o^{\mathbb{Q}'} \end{bmatrix} \begin{bmatrix} B_n^u \\ B_n^o \end{bmatrix} - \begin{bmatrix} \delta_1^u \\ \delta_1^o \end{bmatrix}.$$

Since $B_1^o = -\delta_1^o = 0$, this gives,

$$\begin{aligned} B_{n+1}^o &= B_n^o \Phi_o^{\mathbb{Q}'} = 0, \quad \text{for all} \quad n > 0, \\ B_{n+1}^u &= B_n^u \Phi_u^{\mathbb{Q}'} - \delta_1^u = 0, \quad \text{for all} \quad n > 0. \end{aligned}$$

Iteratively, it can be shown that,

$$B_n^{u'} = B_1' \sum_{i=0}^{n-1} (\Phi_u^{\mathbb{Q}})^i \quad \text{for all} \quad n > 1,$$

where $(\Phi_u^{\mathbb{Q}})^i = (\Phi_X^{\mathbb{Q}})^i$ as shown in Equation (4.5.6). It follows that

$$\begin{aligned} B_n^u &= -\left[\sum_{i=0}^{n-1} (\Phi_X^{\mathbb{Q}u})^i \right]' \delta_1 \\ &= \left[-n, -\frac{1 - e^{-\lambda n}}{\lambda}, n e^{-\lambda n} - \frac{1 - e^{-\lambda n}}{\lambda} \right]'. \end{aligned}$$

Since $B_n^o = 0$, we have

$$B_n' \Omega B_n = B_n^{u'} \Omega^u B_n^u,$$

where Ω^u is the upper-left 3×3 block of the variance-covariance matrix Ω of the state innovation in the VAR. Hence, the constant adjustment term becomes

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

□

5.2 Spanned Model

The spanned model class proves the existence of the solution for the AFNSMA where macro variables have non-zero factor loadings on the yield curve.

Proposition 5.2. *The sufficient condition for the DNS model to be arbitrage-free implies the existence of the solution for the AFNSMA model.*

Assumption 10. *Assume that the one period risk-free rate is affine on a vector of variables X_t, which contains three latent Nelson-Siegel factors X_t^u and observable variables X_t^o,*

$$r_t = \delta_0 + \delta_1^{u'} X_t + \delta_1^{o'} X_t,$$

where

$$\delta_1^u = [1, \frac{1-e^{-\lambda}}{\lambda}, \frac{1-e^{-\lambda}}{\lambda} - \lambda e^{-\lambda}]'$$

and $X_t^u = (X_t^1, X_t^2, X_t^3)'$. The state variables are described by the following VAR under the risk-neutral $\mathbb{Q}$

$$\begin{pmatrix} X_t^u \\ X_t^o \end{pmatrix} = \tilde{\mu}_X + \begin{pmatrix} \Phi_u^{\mathbb{Q}} & \Phi_{uo}^{\mathbb{Q}} \\ 0 & \Phi_o^{\mathbb{Q}} \end{pmatrix} \begin{pmatrix} X_{t-1}^u \\ X_{t-1}^o \end{pmatrix} + \epsilon_t^{\mathbb{Q}},$$

where,

$$\Phi_u^{\mathbb{Q}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & e^{-\lambda} & \lambda e^{-\lambda} \\ 0 & 0 & e^{-\lambda} \end{bmatrix}.$$

Then the zero-coupon bond price is given by,

$$P_t^n = \exp(A_n + B_n^{u'} X_t^u + B_n^{o'} X_t^0),$$

with B_n^u the Nelson-Siegel factor loading

$$B_n^u = [-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}]'$$

and A_n satisfies the difference equations for $n > 1$:

$$A_{n+1} = A_n + B_n' \mu_X^{\mathbb{Q}} + \frac{1}{2} B_n' \Omega_X B_n + A_1, \quad \text{for} \quad n > 0$$

with $A_1 = -\delta$.

5.2.1 Proof of Proposition 5.2

Proof. Recall the ordinary difference equation of B_n as shown in Equation (3.3.36), we have

$$B_{n+1} = B_n' \Phi_X^{\mathbb{Q}} - \delta_1,$$

that is

$$\begin{bmatrix} B_{n+1}^u \\ B_{n+1}^o \end{bmatrix} = \begin{bmatrix} \Phi_u^{\mathbb{Q}'} & 0 \\ \Phi_{uo}^{\mathbb{Q}'} & \Phi_o^{\mathbb{Q}'} \end{bmatrix} \begin{bmatrix} B_n^u \\ B_n^o \end{bmatrix} - \delta.$$

Then,

$$B^u_{n+1} = \Phi^{\mathbb{Q}'}_u B^u_n - \delta^u_1.$$

Hence, by Proposition (5.2), we have that

$$\begin{aligned} B_n &= \left[-n, -\frac{1-e^{-\lambda n}}{\lambda}, ne^{-\lambda n} - \frac{1-e^{-\lambda n}}{\lambda}\right]', \\ B^o_{n+1} &= \Phi^{\mathbb{Q}'}_{uo} B^u_n + \Phi^{\mathbb{Q}'}_o B^o_n - \delta^o_1, \end{aligned}$$

and

$$A_{n+1} = A_n + B'_n \mu^{\mathbb{Q}}_X + \frac{1}{2} B'_n \Omega_X B_n + A_1, \quad \text{for} \quad n > 0.$$

The constant adjustment term A_{n+1} on both sub-classes, which takes on no-arbitrage assumptions for consistency, satisfies the difference Equation (5.1.1). □

5.3 Risk Premia

There is limited focus based on risk price on both AFNS and AFDNS models that imposes explicit assumptions on the form of risk price. However, this is claimed not to affect the identification of other parameters such that inference on risk premia and excess returns can be easily made. If the component for risk compensation is defined by risk premium of bond prices, then this can be derived as the difference between theoretical yield and the expected yield with zero risk compensation when assuming the risk parameters $\lambda_0 = 0$ and $\lambda_1 = 0$. The risk parameters can be identified from this relationship

$$\begin{aligned} \mu^{\mathbb{Q}}_X &= \mu^{\mathbb{P}}_X - \Omega_X \lambda_0, \\ \Phi^{\mathbb{Q}}_X &= \Phi^{\mathbb{P}}_X - \Omega_X \lambda_1. \end{aligned}$$

such that

$$\begin{aligned} \lambda_0 &= \Omega^{-1}_X (\mu^{\mathbb{P}}_X - \mu^{\mathbb{Q}}_X), \\ \lambda_1 &= \Omega^{-1}_X (\Phi^{\mathbb{P}}_X - \Phi^{\mathbb{Q}}_X). \end{aligned}$$

The risk premium is then defined as

$$RP^n_t = y^n_t - \tilde{y}^n_t.$$

The expected yield with zero risk premium is defined by

$$\tilde{y}^n_t = -\frac{1}{n}(\tilde{A}_n + \tilde{B}'_n),$$

where

$$\begin{aligned}
\tilde{B}'_{n+1} &= \tilde{B}'_n \Phi_X^{\mathbb{P}} - \delta'_1 \\
\tilde{A}_{n+1} &= \tilde{A}_n + \tilde{B}'_n \mu + \frac{1}{2} \tilde{B}'_n \Omega \tilde{B}_n - \delta_0.
\end{aligned}$$

The risk premium can be calculated as

$$\begin{aligned}
RP_t^n &= y_t^n - \tilde{y}_t^n \\
&= -\frac{1}{n}(A_n + B'_n X_t) + \frac{1}{n}(\tilde{A}_n + \tilde{B}'_n X_t) \\
&= \frac{1}{n}\left[(\tilde{A}_n - A_n) + (\tilde{B}'_n - B'_n)X_t\right].
\end{aligned}$$

Since there are no imposed restrictions in the dynamic coefficient $\Phi_X^{\mathbb{P}}$ so that in general the dynamics of the vector $\tilde{B}_n^o \neq 0$ and then the macro variables in a non-zero vector X_t^o that do not span the yield curve will affect the variation of risk premia.

5.3.1 Excess Return

The excess return of holding a bond of maturity n for k periods is given by

$$\begin{aligned}
xr_{t+k}^{n,k} &= -(n-k) y_{t+k}^{n-k} + n y_t^n - k r_{t+k} \\
&= \ln(P_{t+k}^{n-h}) - \ln(P_t^n) + \ln(P_{t+n}^1) \\
&= (A_{n-k} - A_n + kA_1) + (B_{n-k} + kB_1)' X_{t+k} - (B_n)' X_t.
\end{aligned}$$

This shows that both the current and future macro state variables under unspanned model, do affect the excess return. As for the unspanned model for $B_n^o = 0$, the excess return becomes

$$xr_{t+k}^{n,k} = (A_{n-k} - A_n + kA_1) + (B_{n-k}^u + kB_1^u)' X_{t+k}^u - (B_n^u)' X_t^u.$$

Although the macro variables at time t do not have direct impact from the last term $(B_n^u)' X_t^u$, it affects excess return through its impact on future X_{t+k}^u as

$$X_{t+k}^u = \sum_{i=0}^{k-1} \Phi^i \mu + \Phi^k X_t.$$

Since $\Phi_X^{\mathbb{P}}$ is not restricted under the physical measure, X_t^o has impact on X_{t+k}^u if the upper right block of Φ_X^k is non-zero. Similarly, the macro variables will influence the expected excess returns through its expectated impact on X_{t+k}^u,

$$\mathbb{E}_t(xr_{t+k}^{n,k}) = (A_{n-k} - A_n + kA_1) + (B_{n-k}^u + kB_1^u)' X_{t+k}^u - (B_n^u)' X_t^u.$$

5.4 Main Findings of AFGNS

The work of Li *et al.* (2012) introduce an arbitrage-free macro finance modelling framework where the state vector contains the popular Nelson-Siegel yield factors and observable macro factors, which is regarded as an extension of Christensen *et al.* (2009). The two subclasses that the model presents, the unspanned specification is the no-arbitrage counterpart of the macro-finance model of Diebold *et al.* (2006), where the three Nelson-Siegel yield factors are combined with three macroeconomic variables in a reduced form state-space framework. Unspanned model is more parsimonious compared to spanned model and forecasts better with less problems in identification of risk-neutral dynamic coefficients. Moreover, the model provides more rigor and structure without increasing the dimension of parameters, a problem often suffered by affine term structure models with macro factors which involve significant number of risk price parameters to estimate and identify. Li *et al.* (2012) analysis risk premia of yields and excess returns within popular interpolation of level slope and curvature.

5.5 No-arbitrage Taylor Rule Model

Taylor (1993) introduced a rule for guiding and assessing monetary policy performance. This section introduces an estimate of Taylor (1993) rule and identify monetary policy shocks using no-arbitrage pricing techniques. The no-arbitrage framework also accommodates backward-looking and forward-looking Taylor rules.

5.5.1 Specifications of No-arbitrage Taylor Rule

No-arbitrage Taylor rule model continues from the growing literature on linking the dynamics of the term structure with macro factors, (see Ang and Piazzesi (2003)), by estimating backward Taylor rules and forward Taylor rules in an affine term structure model with time varying risk-premia. This model structure describes expectations of future short rates and a vector autoregressive for macroeconomic variable under the imposition of arbitrage-free restrictions. The system uses four legs of macroeconomic factors, and the general set-up of the model gives the following state variable, the output gap g_t, at quarter t; the continuously compounded inflation rate π_t, from quarter $t-4$ to t and a latent term structure variable. The dynamics of the states variable takes the form:

$$
\begin{aligned}
f_t^o &= \mu_1 + \begin{pmatrix}\Phi_{11} & \Phi_{12}\end{pmatrix}\begin{pmatrix}f_{t-1}^o \\ f_{t-1}^u\end{pmatrix} + \begin{pmatrix}\Phi_{13} & \Phi_{14} & \Phi_{15}\end{pmatrix}\begin{pmatrix}f_{t-2}^o \\ f_{t-3}^o \\ f_{t-4}^o\end{pmatrix} + \nu_t^o, \\
f_t^u &= \mu_2 + \begin{pmatrix}\Phi_{21} & \Phi_{22}\end{pmatrix}\begin{pmatrix}f_{t-1}^o \\ f_{t-1}^u\end{pmatrix} + \nu_t^u, \qquad (5.5.1)
\end{aligned}
$$

where $f_t^o = [g_t \quad \pi_t]'$, is the vector of observable macro variables, f_t^u is the latent factor. Lastly, an i.i.d vector

$$
\nu_t = \begin{pmatrix}\nu_t^o \\ \nu_t^u\end{pmatrix} \sim N(0, \Sigma_\nu \Sigma_\nu'). \qquad (5.5.2)
$$

The state vector comprises of four lags of all the state variables in a vector of $K = 12$ elements, given by

$$
X_t = [g_t \quad \pi_t \quad f_t^u \quad \dots \quad g_{t-3} \quad \pi_{t-3} \quad f_{t-3}^u]',
$$

which gives the state equation in a compact form as,

$$
X_t = \mu + \Phi X_{t-1} + \Sigma \varepsilon_t, \qquad (5.5.3)
$$

where

$$
\varepsilon_t = \begin{pmatrix}\nu_t \\ 0_{9\times 1}\end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix}\Sigma_\nu & 0_{3\times 9} \\ 0_{9\times 3} & 0_{9\times 9}\end{pmatrix}.
$$

The main focus of the model is to show that the latent factor f_t^u, is related to monetary policy and how no-arbitrage conditions can identify various policy rules. Following from the work of Ang and Piazzesi (2003), the state dynamics, such as the short rate, the stochastic discount factor, the price of risk, and zero-coupon bond prices, are quite the same equations for this model. The one period excess holding period return, by using the equivalence relationship in Equations (3.3.21), is given by:

$$
\begin{aligned}
rx_{t+1}^n &= \ln\left(\frac{P_{t+1}^{n-1}}{P_t^n}\right) - r_t \\
&= ny_t^n - (n-1)y_{t+1}^{n-1} - r_t. \qquad (5.5.4)
\end{aligned}
$$

The conditional expected excess holding period return can be computed using:

$$
\begin{aligned}
\mathbb{E}_t[rx_{t+1}^n] &= -\frac{1}{2}B_{n-1}'\Sigma_X\Sigma_X' B_{n-1} + B_{n-1}'\Sigma_X\lambda_0 + B_{n-1}'\Sigma_X\lambda_1 X_t \\
&= A_n^x + B_n^{x'}, \qquad (5.5.5)
\end{aligned}
$$

which also takes an affine form.

5.5.2 The Benchmark Taylor Rule

This section shows how Taylor rules can be identified with no-arbitrage conditions by jointly estimate standard backward-looking and forward-looking Taylor rules, in a single consistent framework. Following Taylor's original specification, we define the benchmark Taylor rule to be:

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi} \pi_t + \varepsilon_t^{MP,T}, \tag{5.5.6}$$

where the short rate, r_t is set by the Federal Reserve in response to current output, g_t and inflation, π_t. The basic Taylor rule in Equation (5.5.6) can be interpreted as the short rate Equation (3.3.4), (see Ang and Piazzesi (2003)), in a standard affine term structure model, where the unobserved monetary policy shock $\varepsilon_t^{MP,T}$ corresponds to a latent term structure factor, $\varepsilon_t^{MP,T} = \gamma_{1,u} f_t^u$. This corresponds to the short rate Equation (3.3.4) in the term structure model with

$$\delta_1 \equiv \begin{pmatrix} \delta_{1,g} \\ \delta_{1,\pi} \\ \delta_{1,u} \\ 0_{9\times 1} \end{pmatrix} = \begin{pmatrix} \gamma_{1,g} \\ \gamma_{1,\pi} \\ \gamma_{1,u} \\ 0_{9\times 1} \end{pmatrix},$$

which has zeros for all coefficients on lagged g and π. The benchmark Taylor rule in Equation (5.5.6) is estimated consistently using ordinary least squares (OLS), under the assumption that the policy shock, $\varepsilon_t^{MP,T}$ or f_t^u , is contemporaneously not correlated with the output gap and inflation. This assumption is satisfied if macro indication only react slowly to policy shocks. However, it is more advantageous estimating the policy coefficients, $\gamma_{1,g}$ and $\gamma_{1,\pi}$, and extracting the monetary policy shock, $\varepsilon_t^{MP,T}$, using no-arbitrage restrictions in several ways. Firstly, no-arbitrage restrictions can free up the contemporaneous correlation between the macro and latent factors. Secondly, no-arbitrage assumptions produce more consistent and efficient estimates. Thirdly, the term structure model provides estimates of the effect of a policy or macro shock on any stage of the yield curve. Finally, the model allows one to trace the predictability of risk premia in bond yields to macroeconomic or monetary policy. The Taylor rule in Equation (5.5.6) is independent of the past level of the short rate and its movements are not well explained by the expression on the right-hand side variables in Equation (5.5.6), which statistically proves to be highly persistent and this leads to the implied shock to inherit the dynamics of the highly persistent short rate.

5.5.3 Backward-Looking Taylor Rules

The backward-looking Taylor rule pioneered by Clarida *et al.* (2000), considers a modified Taylor rule that consists of current as well as multiple lagged values of macro variables and the previous short rate, to influence the actions of the

Federal funds rate. This is given by,

$$r_t = \gamma_0 + \gamma_{1,g} g_t + \gamma_{1,\pi}\pi_t + \gamma_{2,g} g_{t-1} + \gamma_{2,\pi}\pi_{t-1} + \gamma_{2,r} r_{t-1} + \varepsilon_t^{MP,B}, \qquad (5.5.7)$$

where $\varepsilon_t^{MP,B}$ is the implied monetary policy shock from the backward-looking Taylor rule. Such type of a rule is advantageous to the central bank as its objective is to smooth interest rates. The short rate Equation (3.3.5) can be modified to take the same form as Equation (5.5.7). Using the notation f_t^o and f_t^u to refer to the observable macro and latent factors, respectively, the short rate dynamics takes the form

$$r_t = \delta_0 + \delta_{1,o}' f_t^o + \delta_{1,u} f_t^u, \qquad (5.5.8)$$

where,

$$\delta_1 \equiv \begin{pmatrix} \delta_{1,o} \\ \delta_{1,u} \\ 0_{9\times 1} \end{pmatrix}.$$

Using Equation (5.5.1), we can substitute for the dynamics of f_t^u in Equation (5.5.8) to obtain

$$\begin{aligned} r_t &= \delta_0 + \delta_{1,o}' f_t^o + \delta_{1,u}\mu_2 + \delta_{1,u}\Phi_{21} f_{t-1}^o + \Phi_{22}(r_{t-1} - \delta_0 - \delta_{1,o}' f_{t-1}^o) + \delta_{1,u}\nu_t^u \\ &= (1-\Phi_{22})\delta_0 + \delta_{1,u}\mu_2 + \delta_{1,o}' f_t^o + (\delta_{1,u}\Phi_{21}' - \delta_{1,o}\Phi_{22}')' f_{t-1}^o \\ &+ \Phi_{22} r_{t-1} + \varepsilon_t^{MP,B}, \end{aligned} \qquad (5.5.9)$$

where the backward-looking monetary policy shock is defined to be $\varepsilon_t^{MP,B} \equiv \delta_{1,u}\nu_t^u$. The short rate in the expression is a function of current and lagged macro factors, f_t^o and f_{t-1}^u, the lagged short rate, r_{t-1}, and a monetary policy shock, $\varepsilon_t^{MP,B}$. Equating the coefficients with Equation (5.5.7) to identify the structural coefficients as:

$$\begin{aligned} \gamma_0 &= (1-\Phi_{22})\delta_0 + \delta_{1,u}\mu_2 \\ \begin{pmatrix} \gamma_{1,g} \\ \gamma_{1,\pi} \end{pmatrix} &= \delta_{1,o} \\ \begin{pmatrix} \gamma_{2,g} \\ \gamma_{2,\pi} \end{pmatrix} &= (\delta_{1,u}\Phi_{21}' - \delta_{1,o}\Phi_{22}^{'}) \\ \gamma_{2,r} &= \Phi_{22}. \end{aligned} \qquad (5.5.10)$$

The backward Taylor rules in Equation (5.5.9) and Equation (5.5.6), seem to give observational equivalence reduced-form dynamics for interest rates and macro variables, since f_t^o is identical to the benchmark Taylor rule.

5.5.4 Forward-Looking Taylor Rules

Another type of policy rule that has been proposed by Clarida *et al.* (2000), a forward-looking Taylor rule, which is another version of Taylor rule. According to this rule, the central banks reacts to expected inflation and expected output-gap. A forward-looking Taylor rule using expected inflation and expected output-gap over the next period takes the form:

$$r_t = \gamma_0 + \gamma_{1,g}\mathbb{E}_t(g_{t+1}) + \gamma_{1,\pi}\mathbb{E}_t(\pi_{t+1}) + \varepsilon_t^{MP,F}, \quad (5.5.11)$$

where $\varepsilon_t^{MP,F}$ is defined to be the forward-looking Taylor rule monetary policy shock. The mapping of the forward-looking Taylor rule in Equation (5.5.11) into the framework of an affine term structure models gives the conditional expectations of future output gap and inflation to be functions of the current state variable:

$$\mathbb{E}_t(X_{t+1}) = \mu + \Phi X_t.$$

Denoting e_i as a vector of zeroes with a one in the ith position, we can write Equation (5.5.11) as:

$$r_t = \gamma_0 + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)'\mu + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)'\Phi X_t + \varepsilon_t^{MP,F}, \quad (5.5.12)$$

as g_π and π_t are ordered as the first and second elements in X_t. Equation (5.5.12) is an affine short rate equation, where the short rate coefficients are a function of parameters of the dynamics of X_t :

$$r_t = \bar{\delta}_0 + \bar{\delta}_1{}' X_t, \quad (5.5.13)$$

where

$$\begin{aligned} \bar{\delta}_0 &= \gamma_0 + (\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2)'\mu \\ \bar{\delta}_1{}' &= \Phi'(\gamma_{1,g}e_1 + \gamma_{1,\pi}e_2) + \gamma_{1,u}e_3', \end{aligned}$$

and $\varepsilon_t^{MP,F} \equiv \gamma_{1,u}f_t^u$ with $\gamma_{1,u} = \delta_{1,u}$. A forward Taylor rule is defined by making use of $\bar{\delta}_0$ and $\bar{\delta}_1$ coefficients instead of the normal bond price recursions in Equation (3.3.18) and (3.3.20). This gives a complete set-up of the term structure model, except the use of the basic short rate Equation (3.3.5). The relation in Equation (5.5.9), explicitly show that the forward-looking Taylor rule structural coefficients $(\gamma_0, \gamma_{1,g}, \gamma_{1,\pi})$, impose restrictions on the parameters of an affine term structure model.

The new no-arbitrage bond restriction using the restricted coefficients $\bar{\delta}_0$ and $\bar{\delta}_1$, reflect conditional expectations of output gap and inflation that enter in the short rate Equation (5.5.9). Furthermore, the conditional expectations, $\mathbb{E}_t(g_{t+1})$ and $\mathbb{E}_t(\pi_{t+1})$ are those implied by the underlying dynamics of g_t and π_t in the VAR(2) process. A more general forward-looking Taylor rule based on expected output gap and inflation over the next $k-$period ahead:

$$r_t = \gamma_0 + \gamma_{1,g}\mathbb{E}_t(g_{t+k,k}) + \gamma_{1,\pi}\mathbb{E}_t(\pi_{t+k,k}) + \varepsilon_t^{MP,F}, \quad (5.5.14)$$

where $g_{t+k,k}$ and $\pi_{t+k,k}$ represents output gap and inflation over the next k periods:

$$g_{t+k,k} + \frac{1}{k}\sum_{i=1}^{k} g_{t+i} \quad \text{and} \quad \pi_{t+k,k} + \frac{1}{k}\sum_{i=1}^{k} \pi_{t+i}.$$

As noted in Clarida *et al.* (2000), the general case in Equation (5.5.14) also nests the benchmark Taylor rule Equation (5.5.6) as a special case, by letting $k = 0$.

Alternatively, to fix infinite future forecasting for the next period k, Fed can be viewed as discounting both expected future output gap and expected future inflation at the same discount rate, β. Under this stated assumption, the forward-looking Taylor rule takes the form:

$$r_t = \gamma_0 + \gamma_{1,\hat{g}}\hat{g}_t + \gamma_{1,\hat{\pi}}\hat{\pi}_t + \varepsilon_t^{MP,F}, \tag{5.5.15}$$

where $\hat{g}_t$ and $\hat{\pi}_t$ are infinite sums of expected future output gap and inflation, respectively, and both discounted at rate β per period. We can estimate the discount rate β as part of a standard term structure model by using the dynamics of X_t in Equation (5.5.3) to write $\hat{g}_t$ as:

$$\begin{aligned}
\hat{g}_t &= \sum_{i=0}^{\infty} \beta^i e_1' \mathbb{E}_t(X_{t+i}) \\
&= e_1'(X_t + \beta\mu + \beta\Phi X_t + \beta^2(I+\Phi)\mu + \beta^2\Phi^2 X_t + ...) \\
&= e_1'(\mu\beta + (I+\Phi)\mu\beta^2 + ...) + e_1'(I + \Phi\beta + \Phi^2\beta^2 + ...)X_t \\
&= \frac{\beta}{(1-\beta)} e_1'(I - \Phi\beta)^{-1}\mu + e_1'(I-\Phi\beta)^{-1}X_t.
\end{aligned}$$

Similarly, we can also write discounted future inflation π_t in a similar fashion as:

$$\hat{\pi}_t \equiv \sum_{i=0}^{\infty} \beta^i e_2' \mathbb{E}_t(X_{t+i}) = \frac{\beta}{(1-\beta)} e_2'(I-\Phi\beta)^{-1}\mu + e_2'(I-\Phi\beta)^{-1}X_t.$$

To place the forward-looking rule with discounting in a term structure model, we re-write the short rate Equation (3.3.5) as:

$$r_t = \hat{\delta}_0 + \hat{\delta}_1' X_t, \tag{5.5.16}$$

where

$$\begin{aligned}
\hat{\delta}_0 &= \gamma_0 + [\gamma_{1,\hat{g}} e_1 \quad \gamma_{1,\hat{\pi}} e_2]' \left(\frac{\beta}{(1-\beta)} (I - \Phi\beta)^{-1}\mu \right), \\
\hat{\delta}_1' &= [\gamma_{1,\hat{g}} e_1 \quad \gamma_{1,\hat{\pi}} e_2]'(I - \Phi\beta)^{-1} + \gamma_{1,u} e_3'.
\end{aligned}$$

Similarly, we modify the bond price recursions for the standard affine model in equation (3.3.19) by using the new $\hat{\delta}_0$ and $\hat{\delta}_1$ coefficients that embody restrictions on $\beta, \gamma_0, \gamma_{1,\hat{g}}, \gamma_{1,\hat{\pi}}, \mu$ and Φ.

5.5.5 Forward and Backward-looking Taylor Rules

To compute the monetary policy rule, a combined forward-looking and backward-looking Taylor rules that takes into account the forward-looking expectations of macro variable is illustrated. The standard forward-looking Taylor rule in Equation (5.5.11), of the observable macro variables, gives

$$r_t = \gamma_0 + \gamma_{1,o}' \mathbb{E}_t(f_{t+1}^o) + \varepsilon_t^{MP,F},$$

where the vector $\mathbb{E}_t(f_{t+1}^o) = [\mathbb{E}_t(g_{t+1}) \quad \mathbb{E}_t(\pi_{t+1})]'$ and $\varepsilon_t^{MP,F} = \gamma_{1,u} f_t^u$. We substitute for f_t^u using the implied short rate Equation (5.5.12) that is implied by the forward-looking Taylor rule (5.5.11):

$$\begin{aligned} r_t &= \gamma_0 + \gamma_{1,u}\mu_2 - \frac{\gamma_{1,u}\Phi_{22}\bar{\delta}_0}{\bar{\delta}_{1,u}} + \gamma_{1,o}'\mathbb{E}_t(f_{t+1}^o) + \frac{\gamma_{1,u}\Phi_{22}\bar{\delta}_0}{\bar{\delta}_{1,u}} r_{t-1} \\ &\quad + \gamma_{1,u}\Phi_{21}' f_{t-1}^o + \frac{\gamma_{1,u}\Phi_{22}\bar{\delta}_0}{\bar{\delta}_{1,u}} (\bar{\delta}_1' X_{t-1} - \bar{\delta}_{1,u} f_{t-1}^u) + \varepsilon_t^{MP,FB}, \quad (5.5.17) \end{aligned}$$

where $\bar{\delta}_{1,u}$ is a coefficient on f_t^u in $\bar{\delta}_1$. Equation (5.5.17) expresses the short rate as a function of both expected future macro factors and lagged macro factors, the lagged short rate, r_{t-1}, and a forward- and backward-looking monetary policy shock, $\varepsilon_t^{MP,FB} = \gamma_{1,u}\nu_t^u$. This combined Taylor rule is an equivalent expression of forward-looking Taylor rule (5.5.11). Similarly, the coefficients on macro variables in the backward-looking Taylor rule (5.5.9), are identical to the coefficients in the benchmark Taylor rule (5.5.6).

Chapter 6

Conclusion

The first part of the thesis reviews the traditional model literature about the term structure of interest rates stemming from the affine arbitrage-free work by Vasicek (1977), Cox *et al.* (1985) and Duffie and Kan (1996) models and the macro finance approach from the seminal work by Ang and Piazzesi (2003). The second part focuses on the alternative for modelling the term structure that considers reduced form statistical approach, the three factor interpolation Nelson and Siegel (1987). Based on this approach, further extensions have been discussed, starting from dynamic Nelson-Siegel up until the generalised arbitrage-free macro finance term structure model with both Nelson-Siegel latent factors and observable macro factors.

Amongst macro finance models that considers no-arbitrage conditions, there are proven important advantages that makes macro finance models to out perform most term structure models. However, regardless of models success, macro finance models suffers from some limitations. Recall from the important work of Ang and Piazzesi (2003) model which is attempting to find a connection between macroeconomics and the financial description of the term structure of interest rates to improve model performance, the model does not show success when it comes to forecasting purposes. This results from tremendous numerical difficulties encountered in estimating the necessary parameters. The identification of risk price is also problematic.

As for the reduced form models, theoretically they give satisfactory results because of the parsimony of the models. But as noticed in DNS, models of this nature fail on ruling out opportunities for risk-free arbitrage. On the other hand, Li *et al.* (2012) offer a generalization of the dynamic Nelson-Siegel model. This model contributes towards both the no-arbitrage approach and reduced form approach. The main advantage as opposed to other macro finance models is that there is no effect on the identification of parameters and this leads to predictions based on risk premia be easily made. Imposition of arbitrage-free assumptions is a powerful tool for analysing risk price and risk premia, when these estimates are successfuly made, the model is highly likey to provide a good analysis about the yield curve movements. Both Linlin and Gengming (2012) and Li *et al.* (2012) models, that introduced an arbitrage-free macro finance modelling framework where the state vector contains the popular Nelson-Siegel yield factors and observable macro factors seem to be the leading ones at this stage, as they address theoretical weakness of the most popular Nelson-Siegel model and overcome the empirical implementation problem concerning the canonical arbitrage-free affine term structure models in macro-finance literature.

Appendices

Appendix A

Term Structure Model Specification

A.1 Nominal Pricing Kernel

No arbitrage opportunity between bonds with different maturities guarantees the existence of a unique and positive discount factor m, which is stochastic, linking the current price of maturity n with the yield of maturity $n-1$ in the next period. The stochastic discount factor is used to price all nominal assets in the economy. This implies that asset prices obey in such a way that their returns multiplied with the stochastic discount factor equal one. Under discrete time approach the first order condition of the utility maximization problem is given by

$$\max \mathbb{E}_t^{\mathbb{P}}\{\sum_{j=0}^{\infty} \delta^j U(C_{t+j})\},$$

where δ is a discount factor, $U(\cdot)$ is a utility function and C_t denotes the investor's consumption, a function that expresses consumer spending. The first order conditions of the investors optimal consumption can be formed into a useful identity using consumption Euler equation

$$\begin{aligned} U'(C_t) &= \delta \mathbb{E}_t^{\mathbb{P}}[(1+R_{t+1}^n)U'(C_{t+1})] \\ 1 &= \mathbb{E}_t^{\mathbb{P}}[(1+R_{t+1}^n)\frac{\delta U'(C_{t+1})}{U'(C_t)}] \\ 1 &= \mathbb{E}_t^{\mathbb{P}}[(1+R_{t+1}^n)m_{t+1}], \end{aligned} \tag{A.1.1}$$

where R_{t+1}^n is the real rate of return on an asset with n periods left to maturity of the underlying asset from time t to $t+1$. The stochastic discount factor $m_{t+1} = \frac{\delta U'(C_{t+1})}{U'(C_t)}$ which is also called the pricing kernel in this case, on a probability space $(\Omega, \mathcal{F}_t, \mathbb{P})$, where $\mathcal{F}_t$ is a natural filtration generated by m_{t+1}. The pricing kernel is interpreted as the investor's preferences regarding substitution of consumption and investment through time, can also be called the intertemporal marginal rate of substitution, the rate at which a person will trade present consumption for a small increase in future consumption while maintaining the level of utility. Since the cash flows are deterministic on fixed income settings, the covariance of the cash flows with the pricing of kernel depends only on time. The holding-period return on a bond, which is the total return on an asset or portfolio over the period during which it was held, is given by

$$1 + R_{t+1}^{n} = \frac{P_{t+1}^{n-1}}{P_t^n}, \tag{A.1.2}$$

where P denotes the value of a pure discount bond, n is the number of periods left to maturity at time t. Substituting Equation (A.1.2) into Equation (A.1.1), we get

$$P_t^n = \mathbb{E}_t^{\mathbb{P}}[P_{t+1}^{n-1} m_{t+1}]. \tag{A.1.3}$$

The current price of a zero-coupon bond given as the conditional expectation of the product of its future price and a statistical discount factor. The price for today can be computed as follows,

$$\begin{aligned} P_t^n &= \exp(A_n + B_n' X_t) \text{ for all } t \text{ and } n \geq 0, \\ P_t^0 &= \exp(A_0 + B_0' X_t) = 1. \end{aligned}$$

Initial conditions $A_0 = B_0 = 0$, both of the parameters are identically zero. Therefore, for one period bond we have,

$$\begin{aligned} P_t^1 &= \mathbb{E}_t^{\mathbb{P}}[P_{t+1}^0 m_{t+1}] \\ &= \mathbb{E}_t^{\mathbb{P}}[m_{t+1}]. \end{aligned}$$

Equation (A.1.3) is extremely important in modelling of the pure-discount bond price and the term structure of interest rates, which results in describing the dynamics of the pricing kernel. If we assume that the bond price and the stochastic discount factor are jointly lognormal, then we have,

$$\begin{aligned} \ln P_t^n &= \ln\left(\mathbb{E}_t\left[P_{t+1}^{n-1} m_{t+1}\right]\right) \\ &= \mathbb{E}_t\left[\ln P_{t+1}^{n-1} + \ln m_{t+1}\right] + \frac{1}{2}\mathrm{Var}_t\left[\ln P_{t+1}^{n-1} + \ln m_{t+1}\right]. \end{aligned}$$

A.2 Affine Recursive Form of the Discount Bond Price

Given the coefficients

$$A_{n+1} = A_n - \delta_0 + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X \lambda_0) + \frac{1}{2} B_n' \Sigma_X \Sigma_X' B_n, \tag{A.2.1}$$

$$B_{n+1}' = B_n'(\Phi_X^{\mathbb{P}} - \Sigma_X \lambda_1) - \delta_1'. \tag{A.2.2}$$

The recursive relations of A_n and B_n, for $n \geq 0$ is as follows:

$$
\begin{aligned}
A_{n+1} &= (n+1)A_1 + \sum_{i=0}^{n} B^{(i)}, \text{ where } B^{(i)} = B_i'\left(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0\right) + \frac{1}{2}B_i'\Sigma_X\Sigma_X' B_i \\
A_n &= nA_1 + \sum_{i=0}^{n-1}\left[B_i'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_i'\Sigma_X\Sigma_X'B_i\right] \qquad \text{(A.2.3)} \\
&\vdots \\
A_3 &= 3A_1 + \sum_{i=0}^{2}\left[B_i'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_i'\Sigma_X\Sigma_X'B_i\right] \\
&= -3\delta_0 + B_1'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_1'\Sigma_X\Sigma_X'B_1 \\
&\quad + B_2'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_2'\Sigma_X\Sigma_X'B_2 \\
&= -3\delta - \delta_1'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}\delta_1'\Sigma_X\Sigma_X'\delta_1 \\
&\quad + \sum_{i=1}^{n-2}\left[B_{i+1}'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}B_{i+1}'\Sigma_X\Sigma_X'B_{i+1}\right] \\
A_2 &= -2\delta_0 - \delta_1'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda_0) + \frac{1}{2}\delta_1'\Sigma_X\Sigma_X'\delta_1 \\
A_1 &= A_1 = -\delta_0. \qquad \text{(A.2.4)}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
A_{n+1} &= (n+1)A_1 + \sum_{i=0}^{n}\left[B_i'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda) + \frac{1}{2}B_i'\Sigma_X\Sigma_X'B_i\right] \\
&= A_1 + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda) + \frac{1}{2}B_n'\Sigma_X\Sigma_X'B_n \\
&\quad \underbrace{+\, nA_1 + \sum_{i=0}^{n-1}\left[B_i'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda) + \frac{1}{2}B_i'\Sigma_X\Sigma_X'B_i\right]}_{A_n} \\
A_{n+1} &= A_n + B_n'(\mu_X^{\mathbb{P}} - \Sigma_X\lambda) + \frac{1}{2}B_n'\Sigma_X\Sigma'B_n + A_1. \qquad \text{(A.2.5)}
\end{aligned}
$$

$$
\begin{aligned}
B_{n+1} &= \sum_{i=0}^{n} (\Phi'_X - \lambda'_1 \Sigma_X)^i B_1 && \text{(A.2.6)} \\
B_n &= \sum_{i=0}^{n-1} (\Phi'_X - \lambda'_1 \Sigma_X)^i B_1 && \text{(A.2.7)} \\
&\vdots \\
B_3 &= \sum_{i=0}^{2} (\Phi'_X - \lambda'_1 \Sigma_X)^i B_1 \\
&= B_1(\Phi'_X - \lambda'_1 \Sigma_X)^{i-2} + B_1(\Phi'_X - \lambda'_1 \Sigma_X)^{i-1} + B_1(\Phi'_X - \lambda'_1 \Sigma_X)^i \\
B_2 &= B_1(\Phi'_X - \lambda'_1 \Sigma_X)^{i-1} + B_1(\Phi'_X - \lambda'_1 \Sigma_X)^i \\
&= -\delta_1 - \delta_1(\Phi'_X - \lambda'_1 \Sigma_X) \\
B_1 &= B_1 = -\delta_1. && \text{(A.2.8)}
\end{aligned}
$$

Then,

$$
\begin{aligned}
B_{n+1} &= B_1 + \sum_{i=1}^{n} (\Phi'_X - \lambda'_1 \Sigma_X)^i B_1 \\
&= B_1 + (\Phi'_X - \lambda'_1 \Sigma_X)^1 \underbrace{\sum_{i=1}^{n} (\Phi'_X - \lambda'_1 \Sigma_X)^{i-1} B_1}_{B_n} \\
&= B_1 + (\Phi'_X - \lambda'_1 \Sigma_X) B_n,
\end{aligned}
$$

therefore,

$$
\begin{aligned}
B'_{n+1} &= [B_1 + (\Phi'_X - \lambda'_1 \Sigma_X) B_n]' \\
&= B'_1 + B'_n(\Phi^{\mathbb{P}}_X - \Sigma_X \lambda_1), && \text{(A.2.9)}
\end{aligned}
$$

Both A_{n+1} and B_{n+1} satisfies the coefficients of the measuremnt equation under no-arbitrage assumptions.

A.3 Forward-Looking Taylor Rules

In this section of the appendix, we describe how to compute $\bar{\delta}_0$, $\bar{\delta}_1$ in Equation (5.5.13) of a forward-looking Taylor rule for a k-quarter horizon. From the dynamics of X_t in Equation (5.5.3), the conditional expectation of k-quarter ahead GDP growth and inflation can be written as:

$$
\begin{aligned}
\mathbb{E}_t(g_{t+k,k}) &= \mathbb{E}_t\left(\frac{1}{k}\sum_{i=1}^{k} g_{t+i}\right) = \frac{1}{k} e'_1 \left(\sum_{i=1}^{k} \Phi_i \mu + \Phi_k \Phi X_t\right) \\
\mathbb{E}_t(\pi_{t+k,k}) &= \mathbb{E}_t\left(\frac{1}{k}\sum_{i=1}^{k} \pi_{t+i}\right) = \frac{1}{k} e'_2 \left(\sum_{i=1}^{k} \Phi_i \mu + \Phi_k \Phi X_t\right), && \text{(A.3.1)}
\end{aligned}
$$

where e_i is a vector of zeros with a 1 in the $i-$th position, and Φ_i is given by:

$$\tilde{\Phi}_i = \sum_{j=0}^{i-1} \Phi^j = (I - \Phi)^{-1}(I - \Phi^i). \tag{A.3.2}$$

The bond price recursions for the standard affine model are thus based on the short rate equation $r_t = \bar{\delta}_0 + \bar{\delta}_1' X_t$, where

$$\bar{\delta}_0 = \gamma_0 + \frac{1}{k}[\gamma_{1,g}e_1 \quad \gamma_{1,\pi}e_2]' \left(\sum_{i=1}^{k} \tilde{\Phi}_i \right) \mu, \tag{A.3.3}$$

$$\bar{\delta}_1' = \frac{1}{k}[\gamma_{1,g}e_1 \quad \gamma_{1,\pi}e_2]' \tilde{\Phi}_k \Phi + \gamma_{1,u} e_3'. \tag{A.3.4}$$

As $k \to \infty$, both $\mathbb{E}_t(g_{t+k,k})$ and $\mathbb{E}_t(\pi_{t+k,k})$ approach their unconditional means and there is no state dependence. Hence, the limit of the short rate equation in Equation (4.5.13) as $k \to \infty$ is:

$$r_t = \gamma_0 + [\gamma_{1,g}e_1 \quad \gamma_{1,\pi}e_2]'(I - \Phi)^{-1}\mu + \gamma_{1,u} f_t^u, \tag{A.3.5}$$

which implies that when k is large, the short rate effectively becomes a function only of f_t^u , and g_t and π_t can only indirectly affect the term structure through the feedback in the VAR Equation (5.5.3). In the limiting case, $k = \infty$, the coefficients $\gamma_{1,g}$ and $\gamma_{1,\pi}$ are unidentified because they act exactly like the constant term γ_0 .

Appendix B

Discrete Time Non-Gaussian Models

Appendix B introduces the distributions which are useful for implementing the procedures in practice.

B.1 Gamma and Multivariate Gamma Distributions

A univariate gamma random variable x_{t+1} that is gamma-distributed with shape k and scale θ is denoted $x_{t+1} \sim \Gamma(k, \theta) \equiv \text{Gamma}(k, \theta)$, has a probability density function,

$$p(x_{t+1}|k, \theta) = \frac{1}{\Gamma(k)} x_{t+1}^{k-1} \theta^{-k} \exp(-\frac{x_{t+1}}{\theta}) \quad \text{for } x > 0 \text{ and } k, \theta > 0. \tag{B.1.1}$$

The Laplace transform,

$$\begin{aligned}
\mathbb{E}[\exp(u x_{t+1})] &= \int_0^\infty \exp(u x_{t+1}) p(x_{t+1}|k, \theta) dx_{t+1} \\
&= \int_0^\infty \exp(u x_{t+1}) \frac{1}{\Gamma(k)} x_{t+1}^{k-1} \theta^{-k} \exp(-\frac{x_{t+1}}{\theta}) \\
&= \int_0^\infty \frac{x_{t+1}^{k-1} e^{-(1-u\theta)\frac{x_{t+1}}{\theta}}}{\Gamma(k)\theta^k} \\
&= \left(\frac{1}{1 - \theta u} \right)^k,
\end{aligned}$$

which exists only if $\theta u < 1$. The mean and the variance are,

$$\begin{aligned}
\mathbb{E}(x_{t+1}) &= k\theta \\
\mathbb{V}(x_{t+1}) &= k\theta^2.
\end{aligned}$$

A multivariate gamma random vector $h_{t+1} \sim$ Mult. Gamma(k_h, Σ_h, μ_h) can be obtained by shifting and rotating a vector of uncorrelated gamma random variables, and given by $h_{t+1} = \mu_h + \Sigma_h x_{t+1}$, where x_{t+1} is an $H \times 1$ vector with elements $x_{i,t+1} \sim$ Gamma$(k_{h,i}, 1)$ for $i = 1, ..., H$. The $H \times 1$ vector of non-negative local parameters is μ_h, Σ_h is a full rank $H \times H$ matrix of (non-negative) scale parameters, and $k_h > 0$ is a $H \times 1$ vector of shape parameters. The probability density function of h_{t+1} can be determined by a standard change of variables. That is,

$$p(h_{t+1}|k_h, \Sigma_h, \mu_h) = |\Sigma_h^{-1}| \prod_{i=1}^{H} \frac{\exp(-e_i'\Sigma_h^{-1}[h_{t+1} - \mu_h])}{\Gamma(k_{h,i}) \left(e_i'\Sigma_h^{-1}[h_{t+1} - \mu_h]\right)^{k_{h,i}}},$$

where e_i is an $H \times 1$ unit vector that selects out the $i - th$ element of a vector. The mean and the variance are

$$\begin{aligned} \mathbb{E}[h_{t+1}] &= \mu_h + \Sigma_h k_h, \\ \mathbb{V}[h_{t+1}] &= \Sigma_h \text{diag}(k_h) \Sigma_h'. \end{aligned}$$

The Laplace transform is

$$\begin{aligned} \mathbb{E}[\exp(u'h_{t+1})] &= \int_0^\infty \exp(u'h_{t+1}) p(h_{t+1}|k_h, \Sigma_h, \mu_h) dh_{t+1} \\ &= \exp(u'\mu_h) \int_0^\infty \exp(u'\Sigma_h x_{t+1}) |\Sigma_h^{-1}| \\ &\quad \times \prod_{i=1}^{H} \frac{\exp(-e_i'\Sigma_h^{-1}[h_{t+1} - \mu_h])}{\Gamma(k_{h,i}) \left(e_i'\Sigma_h^{-1}[h_{t+1} - \mu_h]\right)^{k_{h,i}}} dh_{t+1} \\ &= \exp(u'\mu_h) \int_0^\infty \exp(u'\Sigma_h x_{t+1}) \prod_{i=1}^{H} \frac{1}{\Gamma(k_{h,i})} (x_{i,t+1}^{k_{h,i}})^{-1} \\ &\quad \times \exp(-x_{t+1}) dx_{t+1} \\ &= \exp(u'\mu_h) \prod_{i=1}^{H} \left(\frac{1}{1 - e_i'\Sigma_h' u} \right)^{k_{h,i}} \\ &= \exp\left(u'\mu_h - \sum_{i=1}^{H} k_{h,i} \ln[1 - e_i'\Sigma_h' u] \right). \end{aligned}$$

The Laplace transform exists only if $e_i'\Sigma_h' u < 1$ for $i = 1, ..., H$.

B.2 Mixture of Gaussian and Mult-NCG Distributions

Gouriéroux and Jasiak (2006) built the univariate version of the non-Gaussian process shown in Equation (3.5.1) and Equation (3.5.3), to construct one factor term structure models. Dai *et al.* (2010) extended the process to allow for multiple factors. Now, consider a $(G+H)\times 1$ vector $x_{t+1} = (h'_{t+1}, g'_{t+1})'$, where h_{t+1} is an $H \times 1$ vector having a multivariate NCG distribution $p(h_{t+1}|\nu_h, \Phi_h h_t, \Sigma_h, \mu_h)$ and g_{t+1} is a $G \times 1$ vector of conditionally Gaussian random variable $g_{t+1} \sim N(\mu_g + \Sigma_{gh} h_{t+1}, \Sigma_g \Sigma'_g)$. Let $u = (u'_h, u'_g)$ where u_h and u_g are $H \times 1$ and $G \times 1$ vectors, respectively. Using the law of iterated expectations, the Laplace transform is

$$\begin{aligned}
\mathbb{E}[\exp(u'x_{t+1})] &= \mathbb{E}[\exp(u'_g g_{t+1})\exp(u'_h h_{t+1})] = \mathbb{E}_h[\mathbb{E}_{g|h}[\exp(u'_g g_{t+1}\exp(u'_h h_{t+1})] \\
&= \mathbb{E}_h\left[\exp\left((\mu_g + \Sigma_{gh}h_{t+1})'u_g + \frac{1}{2}u'_g\Sigma_g\Sigma'_g u_g\right)\exp(u'_h h_{t+1})\right] \\
&= \exp\left(u'_g\mu_g + \frac{1}{2}u'_g\Sigma_g\Sigma'_g u_g\right)\mathbb{E}_h[\exp([u'_g\Sigma_{gh} + u'_h]h_{t+1})] \\
&= \exp[u'_g\mu_g + \frac{1}{2}u'_g\Sigma_g\Sigma'_g u_g + u'_{gh}\mu_h - \sum_{i=1}^{H}\nu_{h,i}\ln(1 - e'_i\Sigma'_h u_{gh})] \\
&\quad + \exp\left(\sum_{i=1}^{H}\frac{e'_i\Sigma'_h u_{gh}}{1 - e'_i\Sigma'_h u_{gh}}e'_i\Sigma_h^{-1}\Phi_h(h_t - \mu_h)\right),
\end{aligned}$$

where $u_{gh} = \Sigma'_{gh}u_g + u_h$ is an $H \times 1$ vector.

B.3 The Stochastic Discount Factor

We define the stochastic discount factor as:

$$M_{t+1} = \frac{\exp(-r_t)p(g_{t+1}|\mathcal{F}_t, h_{t+1}, z_{t+1};\theta,\mathbb{Q})p(h_{t+1}|\mathcal{F}_t, z_{t+1};\theta,\mathbb{Q})p(z_{t+1}|\mathcal{F}_t;\theta,\mathbb{Q})}{p(g_{t+1}|\mathcal{F}_t, h_{t+1}, z_{t+1};\theta,\mathbb{P})p(h_{t+1}|\mathcal{F}_t, z_{t+1};\theta,\mathbb{P})p(z_{t+1}|\mathcal{F}_t;\theta,\mathbb{P})},$$

where the distributions are conditionally Gaussian, conditionally gamma, and Poisson. This is the exact (non-linear) SDF with no approximations, which we use during estimation. For intuition, consider breaking the log-stochastic discount factor m_{t+1} into three terms; one for each of the shocks that the economic agent faces. Given by,

$$m_{t+1} = -r_t + m_{g,t+1} + m_{h,t+1} + m_{z,t+1},$$

where $m_{i,t+1}$ is the compensation for risk i. Let $\lambda_g = \mu_g - \mu_g^{\mathbb{Q}}$, $\lambda_h = \nu_h - \nu_h^{\mathbb{Q}}$, $\Lambda_g = \Phi_g - \Phi_g^{\mathbb{Q}}$, $\Lambda_h = \Phi_h - \Phi_h^{\mathbb{Q}}$, $\Lambda_{gh} = \Phi_{gh} - \Phi_{gh}^{\mathbb{Q}}$.

B.3.1 Gaussian Risks

Starting with the Gaussian portion, we find

$$m_{g,t+1} = -\frac{1}{2}\lambda'_{gt}\lambda_{gt} - \lambda'_{gt}\epsilon_{g,t+1},$$

where $\epsilon_{g,t+1} = \Sigma_{g,t}^{-1}\xi_{g,t+1}$ is a standard, zero mean Gaussian shock. The price of Gaussian risk is

$$\lambda_{gt} = \Sigma_{g,t}^{-1}((\lambda_g + \Lambda_g g_t + \Lambda_{gh} h_t) - \Sigma_{gh}[\Sigma_h \lambda_h + \Lambda_h(h_t - \mu_h)]).$$

This is a clear generalization of the expression for Gaussian ATSMs. The key difference is a time-varying quantity of risk $\Sigma_{g,t}$.

B.4 Bond Pricing Recursions

The bond pricing recursions can be solved by induction. Assume that the bond prices are

$$P_t^n = \exp(A_n + B'_{n,h} h_t + B'_{n,g} g_t),$$

for some coefficients $A_n, B_{n,h}$ and $B_{n,g}$. At maturity $n = 1$, when the payoff is $P_{t+1}^0 = 1$, we find that

$$P_t^1 = \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t) P_{t+1}^0] = \exp(-\delta_0 - \delta'_h h_t - \delta'_g g_t),$$

such that $A_1 = -\delta_0, B_{1,gh} = -\delta_h$ and $B_{1,g} = -\delta_g$. Consider an $n-$period bond whose price in the next period is P_{t+1}^{n-1}. That is,

$$\begin{aligned} P_t^n &= \mathbb{E}_t^{\mathbb{Q}}[\exp(-r_t) P_{t+1}^{n-1}] = \mathbb{E}_t^{\mathbb{Q}}[\exp(-\delta_0 - \delta'_h h_t - \delta'_g g_t) \\ &\quad \times \exp(A_{n-1} + B'_{n-1,h} h_{t+1} + B'_{n-1,g} g_{t+1})] \\ &= \exp(-\delta_0 - \delta'_h h_t - \delta'_g g_t - A_{n-1}) \mathbb{E}_t^{\mathbb{Q}}[\exp(B'_{n-1,h} h_{t+1} + B'_{n-1,g} g_{t+1})], \end{aligned}$$

where the expectation is taken with respect to the distribution of the random vector $x_{t+1} = (h'_{t+1}, g'_{t+1})'$ under $\mathbb{Q}$ such that

$$\begin{aligned} h_{t+1} &\overset{\mathbb{Q}}{\sim} \text{Mult-NCG}(\nu_{h,i}^{\mathbb{Q}}, \Phi_h^{\mathbb{Q}} h_t, \Sigma_h, \mu_h), \\ g_{t+1} &\overset{\mathbb{Q}}{\sim} N(\mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \Phi_{gh}^{\mathbb{Q}} h_t + \Sigma_{gh}[h_{t+1} - \{(I_H - \Phi_h^{\mathbb{Q}})\mu_h + \Sigma_h \nu_h^{\mathbb{Q}} + \Phi_h^{\mathbb{Q}} h_t\}], \\ &\quad \Sigma_{g,t}\Sigma'_{g,t}). \end{aligned}$$

This expectations has the same form as the Laplace transform provided in Appendix B.1. Using e_i to denote a $H \times 1$ unit vector, we find

$$
\begin{aligned}
P_t^n &= \exp(-\delta_0 - \delta_g g_t - \delta_h h_t + A_{n-1} + \frac{1}{2}\bar{B}'_{n-1,g}\Sigma_{g,t}\Sigma'_{g,t}B_{n-1,g} \\
&\quad + [\mu_g^{\mathbb{Q}} + \Phi_g^{\mathbb{Q}} g_t + \Phi_{gh}^{\mathbb{Q}} h_t - \Sigma_{gh}((I_H - \Phi_h^{\mathbb{Q}})\mu_h + \Sigma_h \nu_h^{\mathbb{Q}} + \Phi_h^{\mathbb{Q}} h_t)]' B_{n-1,g} \\
&\quad + [\Sigma'_{gh} B_{n-1,g} + \bar{B}_{n-1,h}]'\mu_h + \sum_{i=1}^{H} \frac{e'_i \Sigma'_h B_{n-1,gh}}{1 - e'_i \Sigma'_h B_{n-1,gh}} e'_i \Sigma_h^{-1} \Phi_h [h_t - \mu_h] \\
&\quad - \sum_{i=1}^{H} \nu_{h,i}^{\mathbb{Q}} \ln(1 - e'_i \Sigma'_h B_{n-1,gh})) \\
&= \exp(-\delta_0 + A_{n-1} + \mu_g^{\mathbb{Q}'} B_{n-1,g} + [\Sigma'_{gh} B_{n-1,g} + \bar{B}_{n-1,h}]'\mu_h \\
&\quad - ((I_H - \Phi_h^{\mathbb{Q}})\mu_h + \Sigma_h \nu_h^{\mathbb{Q}})' \Sigma'_{gh} B_{n-1,g} + \frac{1}{2}\bar{B}_{n-1,g}\Sigma_{g,t}\Sigma'_{g,t}B_{n-1,g} \\
&\quad - \sum_{i=1}^{H} \nu_{h,i}^{\mathbb{Q}} \ln(1 - e'_i \Sigma'_h B_{n-1,gh}) - \sum_{i=1}^{H} \frac{e'_i \Sigma'_h B_{n-1,gh}}{1 - e'_i \Sigma'_h B_{n-1,gh}} e'_i \Sigma_h^{-1} \Phi_h \mu_h \\
&\quad + [\bar{B}'_{n-1,g}\Phi_g^{\mathbb{Q}} - \delta_g] g_t + \sum_{i=1}^{H} \frac{e'_i \Sigma'_h B_{n-1,gh}}{1 - e'_i \Sigma'_h B_{n-1,gh}} e'_i \Sigma_h^{-1} \Phi_h h_t \\
&\quad + B_{n-1,g}(\Phi_{gh}^{\mathbb{Q}} - \Sigma_{gh}\Phi_h^{\mathbb{Q}}) h_t - \Delta_h h_t) \\
&= \exp(-\delta_0 + A_{n-1} + \mu_g^{\mathbb{Q}'} B_{n-1,g} + \mu'_h [\bar{B}_{n-1,h} + \Phi_h^{\mathbb{Q}'} \Sigma'_{gh} B_{n-1,g}] \\
&\quad - \nu_h^{\mathbb{Q}'} \Sigma'_g \Sigma'_{gh} B_{n-1,g} + \frac{1}{2}\bar{B}'_{n-1,g}\Sigma_{0,g}\Sigma'_{0,g}B_{n-1,g} \\
&\quad - \sum_{i=1}^{H} \nu_{h,i}^{\mathbb{Q}} \ln(1 - e'_i \Sigma'_h B_{n-1,gh}) - \sum_{i=1}^{H} \frac{e'_i \Sigma'_h B_{n-1,gh}}{1 - e'_i \Sigma'_h B_{n-1,gh}} e'_i \Sigma_h^{-1} \Phi_h \mu_h \\
&\quad + [\bar{B}'_{n-1,g}\Phi_g^{\mathbb{Q}} - \delta_g] g_t + [\sum_{i=1} \frac{e'_i \Sigma'_h B_{n-1,gh}}{1 - e'_i \Sigma'_h B_{n-1,gh}} e'_i \Sigma_h^{-1} \Phi_h \\
&\quad + B'_{n-1,g}(\Phi_{gh}^{\mathbb{Q}} - \Sigma_{gh}\Phi_h^{\mathbb{Q}}) - \Delta_h + \frac{1}{2}(I_H \otimes B_{n-1,g})'\Sigma_g \Sigma'_g (\imath_H \otimes B_{n-1,g})] h_t),
\end{aligned}
$$

where $\Sigma_g \Sigma'_g$ is a $GH \times GH$ matrix with diagonal elements $\Sigma_{i,g}\Sigma'_{i,g}$ for $i = 1, ..., H$. The expression

$$
B_{n-1,gh} = \Sigma'_{gh} B_{n-1,g} + B_{n-1,h},
$$

is an $H \times 1$ vector. The Laplace transform exists only if $e'_i B_{n-1,gh} < 1$ for $i = 1, ..., H$.

List of References

Aït-Sahalia, Y. and Kimmel, R.L. (2010). "Estimating affine multifactor term structure models using closed-form likelihood expansions". *Journal of Financial Economics*, vol. 98, no. 1, pp. 113–144.

Almeida, C., Gomes, R., Leite, A., Simonsen, A. and Vicente, J. (2009). Does curvature enhance forecasting? *International Journal of Theoretical and Applied Finance*, vol. 12, no. 08, pp. 1171–1196.

Almeida, C. and Vicente, J. (2008). "The role of no-arbitrage on forecasting: Lessons from a parametric term structure model". *Journal of Banking & Finance*, vol. 32, no. 12, pp. 2695–2705.

Amisano, G. and Giannini, C. (1997). "*Topics in Structural VAR Econometrics*". Springer-Verlang Berlin Heidelberg.

Ang, A., Dong, S. and Piazzesi, M. (2007). "No-arbitrage Taylor rules". working paper series, Columbia University.

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

Ang, A., Piazzesi, M. and Wei, M. (2006). "What does the yield curve tell us about GDP growth?". *Journal of Econometrics*, vol. 131, no. 1, pp. 359–403.

Bauer, M., Rudebusch, G. and Wu, J.C. (2011). "Unbiased estimation of dynamic term structure models". *Chicago Booth Research Paper*, , no. 11-29.

Becker, R.G., Dwolatzky, B., Karakitsos, E. and Rustem, B. (1986). "The simultaneous use of rival models in public optimisation". *Economic Journal*, vol. 96, no. 382, pp. 425–448.

Bekaert, G., Cho, S. and Moreno, A. (2010). "New Keynesian macroeconomics and the term structure". *Journal of Money, Credit and Banking*, vol. 42, no. 1, pp. 33–62.

Bekaert, G. and Hodrick, R.J. (2001). "Expectations hypotheses tests". *The journal of finance*, vol. 56, no. 4, pp. 1357–1394.

Bernanke, B.S. and Boivin, J. (2003). "Monetary policy in a data-rich environment". *Journal of Monetary Economics*, vol. 50, no. 3, pp. 525–546.

Bernanke, B.S., Boivin, J. and Eliasz, P. (2005). "Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach". *The Quarterly Journal of Economics*, vol. 120, no. 1, pp. 387–422.

Bikbov, R. and Chernov, M. (2010). "No-arbitrage macroeconomic determinants of the yield curve". *Journal of Econometrics*, vol. 159, no. 1, pp. 166–182.

Black, F. and Scholes, M. (1973). "The pricing of options and corporate liabilities". *The Journal of Political Economy*, vol. 81, no. 3, pp. 637–654.

Blaskowitz, O.J., Herwartz, H. and Santiago, C. (2005). "Modeling the FIBOR/EURIBOR swap term structure: an empirical approach". SFB 649 discussion paper.

Brandt, M.W. and Yaron, A. (2003). "Time-consistent no-arbitrage models of the term structure". Working papers, National Bureau of Economic Research, Inc.

Campbell, J.Y. (1986). "A defense of traditional hypotheses about the term structure of interest rates". *The Journal of Finance*, vol. 41, no. 1, pp. 183–193.

Campbell, J.Y. and Shiller, R.J. (1991). "Yield spreads and interest rate movements: A bird's eye view". *The Review of Economic Studies*, vol. 58, no. 3, pp. 495–514.

Carriero, A. (2011). "Forecasting the yield curve using priors from no arbitrage affine term structure models". *International Economic Review*, vol. 52, no. 2, pp. 425–459.

Chadha, J.S. and Waters, A. (2011). "Quantitative easing and bond yields: Results from a macro-finance yield curve". Preliminary and incomplete.

Christensen, J.H., Diebold, F.X. and Rudebusch, G.D. (2009). "An arbitrage-free generalized Nelson-Siegel term structure model". *The Econometrics Journal*, vol. 12, no. 3, pp. C33–C64.

Christensen, J.H., Diebold, F.X. and Rudebusch, G.D. (2011). "The affine arbitrage-free class of Nelson-Siegel term structure models". *Journal of Econometrics*, vol. 164, no. 1, pp. 4–20.

Christiano, L.J. and Eichenbaum, M. (1994). "Interest rate smoothing in an equilibrium business cycle model". In: *Federal Reserve Bank of San Francisco Proceedings*.

Clarida, R., Gali, J. and Gertler, M. (2000). "Monetary policy rules and macroeconomic stability: Evidence and some theory". *The Quarterly Journal of Economics*, vol. 115, no. 1, pp. 147–180.

Cochrane, J.H. and Piazzesi, M. (2005). "Bond risk premia". *American Economic Review*, vol. 95, no. 1, pp. 138–160.

Cochrane, J.H. and Piazzesi, M. (2008). "Decomposing the yield curve". *Graduate School of Business, University of Chicago, Working Paper.*

Collin-Dufresne, P., Jones, C.S. and Goldstein, R.S. (2005). "Can interest rate volatility be extracted from the cross section of bond yields? An investigation of unspanned stochastic volatility". Working paper.

Cortés Espada, J.F. and Ramos-Francia, M. (2008). "A macroeconomic model of the term structure of interest rates in mexico". Tech. Rep., Working Papers, Banco de México.

Cox, J.C., Ingersoll, J.E. and Ross, S.A. (1981). "A re-examination of traditional hypotheses about the term structure of interest rates". *The Journal of Finance*, vol. 36, no. 4, pp. 769–799.

Cox, J.C., Ingersoll JR, J.E. and Ross, S.A. (1985). "A theory of the term structure of interest rates". *Journal of Econometric Society*, vol. 53, no. 2, pp. 385–407.

Creal, D.D. and Wu, J.C. (2013). "Estimation of non-Gaussian affine term structure models". *Chicago Booth Research Paper*, , no. 13-72.

Dai, Q. and Singleton, K. (2000). "Specification analysis of the term structure models". *Journal of Finance*, vol. 55, no. 5, pp. 1943–1978.

Dai, Q., Singleton, K.J. and Le, A. (2010). "Discrete-time affine term structure models with generalized market prices of risk". *Review of Financial Studies*, vol. 23, no. 5, pp. 2184–2227.

Dewachter, H. and Iania, L. (2012). "An extended macro-finance model with financial factors". *Journal of Financial and Quantitative Analysis*, vol. 46, pp. 1893–1916.

Dewachter, H. and Lyrio, M. (2008). "Learning, macroeconomic dynamics and the term structure of interest rates". In: *Asset Prices and Monetary Policy*, pp. 191–245. National Bureau of Economic Research, Inc.

Dewachter, H., Lyrio, M. and Maes, K. (2006). "A joint model for the term structure of interest rates and the macroeconomy". *Journal of Applied Econometrics*, vol. 21, no. 4, pp. 439–462.

Diebold, F.X. and Li, C. (2006). "Forecasting the term structure of government bond yields". *Journal of Econometrics*, vol. 130, no. 2, pp. 337–364.

Diebold, F.X., Li, C. and Yue, V.Z. (2008). "Global yield curve dynamics and interactions: a dynamic Nelson-Siegel approach". *Journal of Econometrics*, vol. 146, no. 2, pp. 351–363.

Diebold, F.X., Piazzesi, M. and Rudebusch, G. (2005). "Modeling bond yields in finance and macroeconomics". *American Economic Review*, vol. 95, no. 2, pp. 415–420.

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

Diebold, F.X., Rudebusch, G.D. and Aruoba, S.B. (2003). "The macroeconomy and the yield curve: A nonstructural analysis". CFS working paper series, Center for Financial Studies.

Diebold, F.X., Rudebusch, G.D. and Boragan Aruoba, S. (2006). "The macroeconomy and the yield curve: A dynamic latent factor approach". *Journal of econometrics*, vol. 131, no. 1, pp. 309–338.

Duffee, G. (2002). "Term premia and interest rate forecasts in affine models". *Journal of finance*, vol. 57, no. 1, pp. 405–443.

Duffie, D. (2010). "*Dynamic asset pricing theory*". Princeton University Press.

Duffie, D. and Kan, R. (1996). "A yield-factor model of interest rates". *Mathematical finance*, vol. 6, no. 4, pp. 379–406.

Duffie, D. and Lando, D. (2001). "Term structures of credit spreads with incomplete accounting information". *Econometrica*, vol. 69, no. 3, pp. 633–664.

Duffie, D., Pan, J. and Singleton, K. (2000). "Transform analysis and asset pricing for affine jump-diffusions". *Econometrica*, vol. 68, no. 6, pp. 1343–1376.

Estrella, A. and Mishkin, F. (1996). "The yield curve as a predictor of U.S. recessions". *Current Issues in Economics and Finance*, vol. 2, no. 7.

Evans, C.L. (1998). "Real-time Taylor rules and the federal funds futures market". *Economic Perspectives-Federal Reserve Bank of Chicago*, vol. 22, pp. 44–55.

Evans, C.L. and Marshall, D.A. (2007). "Economic determinants of the nominal treasury yield curve". *Journal of Monetary Economics*, vol. 54, no. 7, pp. 1986–2003.

Fabozzi, F.J. (2002). "*Interest rate, term structure, and valuation modeling*", vol. 125. John Wiley & Sons.

Fama, E.F. (1990). "Stock returns, expected returns, real activity". *Journal of Finance*, vol. 45, no. 4, pp. 1089–1108.

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

Favero, C.A., Niu, L. and Sala, L. (2012). "Term structure forecasting: No-arbitrage restrictions versus large information set". *Journal of Forecasting*, vol. 31, no. 2, pp. 124–156.

Filipozzi, F. (2006). "Macro-finance models of the term structure: a review". *Tallinn University of Technology.*

Fisher, I. (1930). "*The theory of interest*". Macmillan New York.

Fisher, M., Nychka, D.W. and Zervos, D. (1995). Fitting the term structure of interest rates with smoothing splines.

Gomez, A., Maheu, J.M. and Maynard, A. (2008). "Improving forecasts of inflation using the term structure of interest rates". Working papers, University of Toronto, Department of Economics.

Gouriéroux, C. and Jasiak, J. (2006). "Autoregressive gamma processes". *Journal of Forecasting*, vol. 25, no. 2, pp. 129–152.

Guidolin, M. and Timmermann, A. (2009). "Forecasts of U.S. short-term interest rates: A flexible forecast combination approach". *Journal of Econometrics*, vol. 150, no. 2, pp. 297–311.

Gürkaynak, R.S., Sack, B. and Wright, J.H. (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, vol. 54, no. 8, pp. 2291–2304.

Gürkaynak, R.S. and Wright, J.H. (2012). "Macroeconomics and the term structure". *Journal of Economic Literature*, vol. 50, no. 2, pp. 331–367.

Hamilton, J.D. and Jordá, O. (2002). "A model of the Federal Funds Rate Target". *Journal of Political Economy*, vol. 110, pp. 1135–1167.

Hamilton, J.D. and Wu, J.C. (2012). "Identification and estimation of Gaussian affine term structure models". *Journal of Econometrics*, vol. 168, no. 2, pp. 315–331.

Harrison, J.M. and Kreps, D.M. (1979). "Martingales and arbitrage in multiperiod securities markets". *Journal of Economic theory*, vol. 20, no. 3, pp. 381–408.

Harrison, J.M. and Pliska, S.R. (1983). "A stochastic calculus model of continuous trading: complete markets". *Stochastic processes and their applications*, vol. 15, no. 3, pp. 313–316.

Hassanein, M. and Azzam, I. (2010). "Ex post and ex ante returns and risks under different maturities of treasury bonds: evidence from developed and emerging markets". *Macroeconomics and Finance in Emerging Market Economies*, vol. 3, no. 1, pp. 103–118.

Henderson, D. and McKibben, W. (1993). "A comparison of some basic monetary policy regimes for open economies: Implications of different degrees of instrument adjustment and wage persistence". *Journal of Monetary Economics*, vol. 39(1), no. 4, pp. 221–31.

Hicks, J.R. (1946). "*Value and capital*", vol. 2. Clarendon Press Oxford.

Hördahl, P., Tristani, O. and Vestin, D. (2006). "A joint econometric model of macroeconomic and term-structure dynamics". *Journal of Econometrics*, vol. 131, no. 1, pp. 405–444.

Hull, J.C. and White, A.D. (1994). "Numerical procedures for implementing term structure models i: Single-factor models". *The Journal of Derivatives*, vol. 2, no. 1, pp. 7–16.

Jongen, R., Verschoor, W.F. and Wolff, C.C. (2011). "Time-variation in term premia: International survey-based evidence". *Journal of International Money and Finance*, vol. 30, no. 4, pp. 605–622.

"Joslin, S., Le, A. and Singleton, K.J. (2013). Jfec invited paper: Gaussian macro-finance term structure models with lags. *Journal of Financial Econometrics*, vol. 11, no. 4, pp. 581–609.

Kargin, V. and Onatski, A. (2004). "Dynamics of interest rate curve by functional auto-regression". *Journal of Multivariate Analysis*, vol. 99, pp. 2508–2526.

Keynes, J.M. (2006). "*General theory of employment, interest and money*". Atlantic Publishers & Dist.

Kim, D.H. (2009). "Challenges in macro-finance modeling". *Federal Reserve Bank of St. Louis Review*, vol. 91, no. 5, Part 2, pp. 519–544.

Kuttner, K.N. (2001). "Monetary policy surprises and interest rates: Evidence from the fed funds futures market". *Journal of monetary economics*, vol. 47, no. 3, pp. 523–544.

Lemke, W. and Werner, T. (2009). "The term structure of equity premia in an affine arbitrage-free model of bond and stock market dynamics". Working paper series, European Central Bank.

Lengwiler, Y. and Lenz, C. (2010). "Intelligible factors for the yield curve". *Journal of Econometrics*, vol. 157, no. 2, pp. 481–491.

Li, C., Niu, L. and Zeng, G. (2012). "A generalized arbitrage-free Nelson-Siegel term structure model with macroeconomic fundamentals". *Available at SSRN 1787262*.

Li, C. and Wei, M. (2013). "Term structure modelling with supply factors and the Federal Reserve's Large Scale Asset Purchase programs". *International Journal of Central Banking*, vol. 9, no. 1, pp. 3–39.

Lildholdt, P., Panigirtzoglou, N. and Peacock, C. (2007). "*An affine macro-factor model of the UK yield curve*". Bank of England.

Linlin, N. and Gengming, Z. (2012). "The discrete-time framework of the arbitrage-free Nelson-Siegel class of term structure models". *Available at SSRN 2015858*.

Litterman, R. and Scheinkman, J. (1991). "Common factors affecting bond returns". *The Journal of Fixed Income*, vol. 1, no. 1, pp. 54–61.

McCallum, B.T. (1999). "Issues in the design of monetary policy rules". *Handbook of macroeconomics*, vol. 1, pp. 1483–1530.

McCulloch, J.H. (1971). "Measuring the term structure of interest rates". *The Journal of Business*, vol. 44, no. 1, pp. 19–31.

McCulloch, J.H. (1975). "The tax-adjusted yield curve ". *The Journal of Finance*, vol. 30, no. 3, pp. 811–830.

McCulloch, J.H. (1993). "A re-examination of traditional hypotheses about the term structure: a comment". *The Journal of Finance*, vol. 48, no. 2, pp. 779–789.

Merton, R.C. (1973). "Theory of rational option pricing". *The Bell Journal of Economics and Management Science*, pp. 141–183.

Mishkin, F. and Jorion, P. (1991). "A multicountry comparison of term-structure forecasts at long horizons ". *Journal of Financial Economics*, vol. 29, no. 1, pp. 59–80.

Monfort, A. and Pegoraro, F. (2007). "*Multi-lag term structure models with stochastic risk premia*". Banque de France.

Nelson, R. and Siegel, F. (1987). "Parsimonious modeling of yield curves". *Journal of Business*, vol. 60, no. 4, pp. 473–489.

Ochoa, J. (2006). "An interpretation of an affine term structure model of Chile". *Estudios de Economia*, vol. 33, pp. 155–184.

Palomino, F. (2012). "Bond risk premiums and optimal monetary policy". *Review of Economic Dynamics*, vol. 15, no. 1, pp. 19–40.

Pericoli, M. and Taboga, M. (2012). "Bond risk premia, macroeconomic fundamentals and the exchange rate". *International Review of Economics & Finance*, vol. 22, no. 1, pp. 42–65.

Piazzesi, M. (2005). "Bond yields and the federal reserve". *Journal of Political Economy*, vol. 113, no. 2, pp. 311–344.

Pooter, M.D. (2007). "Examining the Nelson-Siegel class of term structure models". Tech. Rep., Tinbergen Institute.

Realdon, M. (2006). "The target rate and term structure of interest rates". Tech. Rep., Department of Economics, University of York.

Reifschneider, D. and Williams, J.C. (2000). "Three lessons for monetary policy in a low-inflation era". *Journal of Money, Credit and Banking*, pp. 936–966.

Ross, S.A. (1976). "The arbitrage theory of capital asset pricing". *Journal of economic theory*, vol. 13, no. 3, pp. 341–360.

Rotemberg, J.J. and Woodford, M. (1997). "An optimization-based econometric framework for the evaluation of monetary policy". *NBER, Macroeconomic Annual*, vol. 2, pp. 297–346.

Rudebusch, G.D. (1995). "Federal reserve interest rate targeting, rational expectations, and the term structure". *Journal of Monetary Economics*, vol. 35, no. 2, pp. 245–274.

Rudebusch, G.D., Swanson, E.T. and Wu, T. (2006). "The bond yield "conundrum" from a macro-finance perspective". *Monetary and Economic Studies*, vol. 24, no. S-1, pp. 83–109.

Rudebusch, G.D. and Wu, T. (2007). "Accounting for a shift in term structure behavior with no-arbitrage and macro-finance models". *Journal of Money, Credit and Banking*, vol. 39, no. 2-3, pp. 395–422.

Sargent, T.J. (1979). "A note on maximum likelihood estimation of the rational expectations model of the term structure". *Journal of Monetary Economics*, vol. 5, pp. 133–143.

Sarno, L., Schneider, P. and Wagner, C. (2012). "Properties of foreign exchange risk premiums". *Journal of Financial Economics*, vol. 105, no. 2, pp. 279–310.

Schuster, P. and Uhrig-Homburg, M. (2012). "The term structure of bond market liquidity conditional on the economic environment: An analysis of government guaranteed bonds". *Working Paper Series in Economics*.

Shea, G.S. (1984). "Pitfalls in smoothing interest rate term structure data: Equilibrium models and spline approximations". *Journal of Financial and Quantitative Analysis*, vol. 19, no. 03, pp. 253–269.

Siegel, A.F. and Nelson, C.R. (1988). "Long-term behavior of yield curves". *Journal of Financial and Quantitative Analysis*, vol. 23, no. 01, pp. 105–110.

Svensson, L.E. (1995). "Estimating forward interest rates with the extended Nelson & Siegel method". *Sveriges Riksbank Quarterly Review*, vol. 3, no. 1, pp. 13–26.

Svensson, L.E. (1997). "Inflation forecast targeting: Implementing and monitoring inflation targets". *European Economic Review*, vol. 41, no. 6, pp. 1111–1146.

Svensson, L.E. (2000). "Open-economy inflation targeting". *Journal of international economics*, vol. 50, no. 1, pp. 155–183.

Taylor, J.B. (1993). "Discretion versus policy rules in practice". *Carnegie-Rochester Conference Series on Public Policy*, vol. 39, no. 1, pp. 195–214.

Taylor, J.B. (1999). "A historical analysis of monetary policy rules". In: *Monetary policy rules*, pp. 319–348. University of Chicago Press.

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

Vasicek, O. and Fong, H. (1982). "Term structure modeling using exponential splines". *Journal of Finance*, vol. 37, no. 2, pp. 339–348.

Wieland, V. (2003). “Monetary policy and uncertainty about the natural unemployment rate”. CFS Working Paper.

Wieland, V. (2006). “Monetary policy and uncertainty about the natural unemployment rate: Brainard-style conservatism versus experimental activism”. *Advances in Macroeconomics*, vol. 6, no. 1, pp. 1–34.

Williams, J.C. (2003). “Simple rules for monetary policy”. *Federal Reserve Bank of San Francisco Economic Review*.

Williams, J.C. and Taylor, J.B. (2009). “A black swan in the money market”. *American Economic Journal: Macroeconomics*, vol. 1, no. 1, pp. 58–83.

Yamamoto, Y. (2012). “Bootstrap inference for impulse response functions in factor-augmented vector autoregressions”. Tech. Rep., Institute of Economic Research, Hitotsubashi University.

Zagaglia, P. (2009). “Forecasting with a DSGE model of the term structure of interest rates: The role of the feedback”. *Research Papers in Economics*, vol. 14.