Essays in Financial Econometrics Term Structure Modeling and Forecasting

著者	Wali Ullah
学位授与機関	Tohoku University
学位授与番号	11301甲第15970号
URL	http://hdl.handle.net/10097/58519

Essays in Financial Econometrics

Term Structure Modeling and Forecasting

WALI ULLAH

Dissertation submitted to the Graduate School of Economics and Management of Tohoku University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ECONOMICS

GRADUATE SCHOOL OF ECONOMICS AND MANAGEMENT
TOHOKU UNIVERSITY JAPAN
May, 2013

Essays in Financial Econometrics

Term Structure Modeling and Forecasting

A Dissertation

by

WALI ULLAH

Approved as to style and content by:

Prof. Yasumasa MATSUDA
Dissertation Supervisor

May, 2013

Abstract

Pricing assets in the fixed-income market is an important field in financial econometrics. The most basic asset in the fixed-income market is a zero-coupon bond. The complete set of zero-coupon bonds of all maturities results in the term structure of interest rates that forms the basis of the fixed-income market. The term structure of interest rates is also an important element in macroeconomics and finance. At a certain point of time, the yield curve can have different shapes. These shapes, representing a time-varying relationship of the interest rate and maturity, are of great significance for various economic and financial decisions. In this dissertation, we conduct a formal econometric analysis of dynamic term structure models for the Japanese government bond yields. We present some fundamental concepts of government bonds yield to provide a basis for yield curve modeling and forecasting in the first chapter.

Chapter two compares the in-sample fit and out-of-sample forecast accuracy of the Cox-Ingersoll-Ross (CIR) and Nelson-Siegel models. For the in-sample fit, there is a significant lack of information on the short-term CIR model. The CIR model should also be considered too poor to describe the term structure in the simulation based context. It generates a downward slope average yield curve. Contrary to CIR model, Nelson-Siegel model is not only compatible to fit attractively the yield curve but also accurately forecast the future yield for various maturities. Furthermore, the non-linear version of the Nelson-Siegel model outperforms the linearized one. In the simulation based context, the Nelson-Siegel model is capable to replicate most of the stylized facts of the Japanese market yield curve.

Chapter three empirically examines the role of macroeconomic and stock market variables in the dynamic Nelson-Siegel framework with the purpose of fitting and forecasting the term structure of interest rate. The Nelson-Siegel type models in state-space framework considerably outperform the benchmark simple time series forecast models such as an AR(1) and random walk. The yields-macro model incorporating macroeconomic factors leads to a better in-sample fit of the term structure than the yields-only model. The out-of-sample predictability of the former for the short horizon forecasts is superior to the latter for all maturities, and for the longer horizons the former is still compatible to the latter. Inclusion of macroeconomic factors dramatically reduces the autocorrelation of forecasts errors which has been a common phenomenon of statistical analysis in the previous term structure models.

The monetary policy targets the short rates, however, during the zero interest rate policy (ZIRP) the short end of yield curve will not be any more a policy instrument. Relying on the joint yields-macro latent factors model, chapter four examines the effect of monetary policy stances on term structure and the possible feed-back effect on the real sector using the Japanese experience of ZIRP. The analysis indicates that it is the entire term structure that transmits the policy shocks to the real economy rather than the yields spread only. The monetary policy signals passes through the yield curve level and slope factors to stimulate the economic activity. The curvature

factor, besides reflecting the cyclical fluctuations of the economy, acts as a leading indicator for future inflation. In addition, policy influence tends to be low, as the short end becomes segmented toward medium/long-term of the yield curve. Furthermore, volatility in bond markets is found to be asymmetrically affected by positive and negative shocks and the long end tends to be less sensitive to stochastic shocks than the short maturities. The analysis indicates that the traditional expectation hypothesis (with time invariant term premia) does not hold during the ZIRP period.

The results in this dissertation have several implications for policy. The analysis of the dynamic Nelson and Siegel (1987) yield curve model is relevant for how central banks and financial institutions analyze the term structure. If financial institutions and central banks are looking for a model to study the evolution of the yield curve in Japanese market, the Nelson-Siegel family of models could be a good candidate. Furthermore, the models presented in chapter three and four provides a framework to understand important aspects of the recent intertwined financial crisis, economic recessions and monetary policy regimes. It highlights the importance of yield curve factors for policy analysis that can serve as leading pro-cyclical or counter-cyclical indicators.

Dedicated to my Parents

and loving Noor Jehan.

Acknowledgements

First and foremost I would like to thank Allah Almighty for all the mercies and blessings He has shown toward me. The completion of this research is one of His many blessings.

I would like to thank my both dissertation supervisors Prof. Yasumasa Matsuda and Prof. Yoshihiko Tsukuda for their valuable suggestions, guidance, comments and discussions. I extend my special thanks to Prof. Yasumasa Matsuda for his active involvement in my research and special efforts, care and craft. In particular, I am thankful to Prof. Yoshihiko Tsukuda for showing me the first step in conducting research is to decide on how a problem should be approached from econometric point of view and a general guide toward the selection and specification of a model. Even after becoming the vice president at the Tohoku University, he was always there to teach me Financial and Time Series Econometrics. I enjoyed his intensive and stimulating classes.

The author would like to thank his colleague class fellows Jiasha Fu and Liao Pinchieh for their valuable suggestions in the seminar classes. Particularly, I am grateful to my friend and colleague Dr. Jarallah, Shaif for his valuable comments and suggestions for improving the organization, style and clarity of the text. My special acknowledgement is intended for Dr. Shahzadah Nayyar Jehan and Dr. Shin-Ichi Nishiyama for their valuable suggestions and generous help. I am also grateful to several of my colleagues at Graduate School of Economics and Management, Tohoku University, Mohammad Mohsin, Khalil Ullah M., Xinli Wang, David P. V. Saille and Puyu Hu for their support.

Behind the scenes there is a constant support and encouragement, I received from my father. He kept me challenged, made many helpful comments and lengthy discussions concerning my research. Furthermore, I am grateful for the help and support of my mother and wife. Their advice and assistance in more practical matters proved invaluable, particularly in stressful times.
In spite of all my efforts, some errors might have crept into the text. The sole responsibility for any error rests with me.

Tuesday, May 7, 2013. Waliullah

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The bond market is the largest financial market around the world. The total outstanding volume is more than $69,938 billion, compared to $63,102 billion outstanding equity market volume. The Japan accounts for $15,139 billion (21.646%) of the global bond market value, with $12,967 billion (85.653%) coming from government issued bonds.[1]An interesting aspect of looking at Japanese government treasuries and bonds (JGBs), besides with the largest size of the market (comparing in terms of the government bonds outstanding volume), is that they serve as a yardstick for many other traded securities in the Japanese market.[2] As they are issued by the government, the risk of default is negligible. Therefore, return on government bonds is commonly referred to as the risk free rate.[3] The yield earned on these securities serves as a benchmark for all other securities that are subjected to default risk, such as those issued by corporations and financial institutions.

Treasuries issued by the Japanese government are traded with various maturities, typically classified in four groups. Treasuries expiring within a year are commonly referred to as short-term bills, treasuries between one and five years maturity are medium-term notes, treasuries ranging from five to ten years are long-term bonds while treasuries with maturities beyond ten years are called super long-term bonds.[4]

The short-term JGBs are all discount bonds, meaning that they are issued at the price lower than the face value. No interest payments are made, but at maturity the principal amounts are redeemed at face value. These short-term bills are also known as zero-coupon bonds. On the other hand, all medium, long and super long-term bonds are treasuries with coupon rate.[5]

The yield (interest rate) for the zero-coupon bonds can directly be observed as they are

---

[1] The bond market volume is taken from the Bank for International Settlements (BIS) Quarterly Review of March 2012. The data used is the most recent complete available data (September 2011) from the statistical appendix. The equity market volume comes from the World Federation of Exchanges, and is also for August 2011.

[2] The US accounts for $26,176 billion (37.427%) of the world market, the largest one, of which $12,550 billion (47.945%) comes from government treasuries.

[3] This does not mean that government bonds are completely free of any risk. Due to inflation, the real amount of the coupons and face value change over time, and there is risk due to changes in interest rates if the bond is held for a time shorter than the remaining maturity. Strictly speaking, only government bonds expiring in very short-term can be considered risk free.

[4] However, throughout this dissertation, we often use the label “bonds” to refer to treasuries in general.

[5] The Japanese government issues two types of coupon-bearing bond, JGBs with fixed coupon rate and JGBs with floating rates. The JGBs with floating rate are issued for 10 and 15 years maturities. The data used in this study does not include the floating rate bonds.

traded at discounted face value. However, for bonds with maturities of a horizon longer than one year, the zero-coupon yields are not directly observable. The reason behind this is that yield to maturity on coupon-bearing bonds suffer from the coupon effect which implies that two bonds which are identical in every respect except for bearing different coupon rates can have a different yield to maturity. Therefore, longer maturities zero-coupon yields need to be derived from coupon-bearing treasury notes and bonds.

In the market, there are many zero-coupon rates referring to various maturities. The relationship between these different rates and maturities is known as the term structure of interest rate. This relationship is also referred to as yield curve. At certain point of time, the yield curve can have different shapes. These shapes, representing a time-varying relationship of the interest rate and maturity, are of great significance for various economic and financial decisions. Furthermore, cross section of yields is closely tied to various policy issues and macro economy. Therefore, researchers try to summarize the information in the yield curve with a few common factors driving all these rates. To figure out the exact driving force of bond market, various models have been developed in the course of time, yet several aspects of bond yields are still absent from the main stream literature.

Understanding what moves bond yields and why researchers should look at the yield curve is important for various reasons. First, due to the relation between yields on bonds with long and short maturities, the yield curve contains information about the future path of the economy. Yields on long maturity bonds are expected values of average future short yields after an adjustment for risk. This means that the current yield curve contains information about the future path of the economy. Yield spreads have indeed been useful for forecasting not only future short yields (Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005; Fama and Bliss, 1987) but also real activity (Ang *et al.* 2006; Estrella and Hardouvelis, 1991; Hamilton and Kim, 2002; Harvey, 1988) and inflation (Fama, 1990; Mishkin, 1990), even though these forecasting relationships may be unstable (Stock and Watson, 2003). These forecasts provide a basis for investment decisions of firms, savings decisions of consumers, and policy decisions.

Secondly, the shape of the yield curve matters for monetary policy. The central bank seems to be able to move the short end but it is the long end of curve that affects the aggregate demand and real investment in the economy. Therefore, for a given state of the economy, a model of the yield curve helps to understand how movements at the short end translate into longer term yields. This involves understanding both how the central bank conducts policy and how the signals transmit from short to long end and then to the real sector (transmission mechanism of monetary policy).

Debt policy constitutes a third reason. When issuing new debt, governments need to decide about the maturity of the new bonds. It is necessary to find out how the entire yield curve responds to an increase in supply of bonds of certain maturity for optimal debt policy management. The outcome of certain policies such as selling short maturity debt and buying long maturity notes may flatten or invert the yield curve that will push the economy in recession.

Fourth, the yield curve matters for derivative pricing and hedging. The price of many derivatives depends on the entire yield curve (for example futures and interest rate options). Furthermore, banks also need to hedge their interest rate risk exposure due to short-term interest payments and long-term interest collections.

This chapter reports some fundamental concepts of government bond yields to provide a basis for yield curve modeling and forecasting. First, we discuss the research objectives and the problems associated with interest rate modeling and forecasting. Section 1.2 discusses the three rates, i.e., discount, forward and zero-coupon rates, while section 1.3 deals with the Fama-Bliss forward and spot rates calculation methodology. In section 1.4, we present the dataset and its descriptive features, used in this dissertation, followed by some relevant literature about modeling the yield curve in section 1.5. Section 1.6 links the yield curve factors to the fundamentals of an economy. Lastly, some basic definitions and notations are introduced that are recurrently used in this dissertation in 1.7.

1.1. Research Aims and Objectives

An important area of financial econometrics is the modeling and forecasting of interest rates. Interest rate models share important commonalities with models of other financial markets but the bond market displays some distinct features that present unique challenges to interest rate modeling and forecasting. The bond market contains a wide variety of different but closely related assets, in particular bonds with different maturities. This gives rise to the notion of term structure of interest rate. The literature on term structure has evolved mostly in continuous time, where stochastic calculus reigns and partial differential equations (PDEs) spit fire. Those who are dealing with this strand of literature are faced with considerable obstacles either from theoretical or empirical perspectives. The first problem faced in term structure modeling is how to summarize the price information at any point in time for the large number of bonds that are traded in the market. The models so far developed are trying to define the basis which spans the set of investment opportunities and derive the entire term structure from this basis. These bases are also referred to as factors. More precisely, the models try to summarize the multi-dimensional information pertaining to a large number of contingent claims in few factors to construct the entire term structure. Since, all the methods are based on approximations and estimations, the amount of inaccuracy accumulates, which make it often difficult to comprehend which class of the models comes with encouraging results.

Secondly, these models are developed keeping in view various different objectives/goals. The earlier term structure models are derived from economic theory, usually under the assumption of absence of arbitrage. However, the appealing characteristics of no-arbitrage and a sound economic foundation often come at the cost of poor fit and this class of models is, therefore, empirically found to be unable to beat a naive random walk forecast of interest rates. The other class, on the other hand, is based merely on statistical grounds and is known for its relatively good empirical fit. This strand, constituting the statistical class of models, typically focuses on the consistency of asset prices across markets with little regard for the underlying economic fundamentals. Being at the two brinks, it is necessary to come up with a model that, besides having sound theoretical foundation and describing the market trends (optimally fit and precisely forecast), can also serve for policy analysis in order to understand important aspects of the recent intertwined financial crisis, economic recessions and policy regimes.

Taking into consideration these issues, in chapter 2, we introduce the Cox-Ingersoll-Ross (1985) and dynamic Nelson-Siegel (1987) models and compare the in-sample fit as well as out-of-sample forecast performance of the two classes of models. Based on the representative agent general equilibrium macroeconomic approach, Cox-Ingersoll-Ross (CIR) model presents an explicit analytical expression for the equilibrium interest rate dynamics and bond prices in order to fit the observed yield curve. On the other hand, the Nelson-Siegel model is based on the stylized facts that can be inferred from empirical analysis and belongs to the statistical class of models. The motivation for choosing the CIR short rate model is due to its two key features: mean reversion and non-constant (time-varying) volatility. In particular, it allows for the short rate to revert to a long run mean so that if the current rate is above (below) the long run mean, it is expected to decrease (increase) towards the long-term mean in the future. The volatility of interest rates is often made to be dependent on the level of interest rates so that when rates are higher, they are more volatile. Whereas, the motivation for the Nelson-Siegel model comes from its relative simplicity, ease of estimation and to the fact that there is some underlying economic interpretation in the three factors it is based on, which represent level, slope and curvature of the yield curve.

However, the downside of the statistical class of models is that they often lack theoretical support and do not assume absence of arbitrage. In chapter 3 and 4, we try to overcome this disadvantage and close the gap at least partially by investigating the interaction of the models with the macroeconomy. In chapter 3, we focus on an area of macro-finance research that examines the relationship between the term structure of interest rates and the economy in an interdisciplinary fashion. We evaluate the out-sample forecast performance of the dynamic Nelson-Siegel model with and without macroeconomic variables in the state-space representation and provide in-depth treatment of aspects of the interplay between the yield curve and the macroeconomy. Regarding the policy issues, we formulate a framework that integrates monetary policy as well as real economy factors in the term structure model in chapter 4. The objective is to figure out the missing string between monetary policy and real activity. More specifically the chapter tries to find out the transmission mechanism of monetary policy and the dynamic interaction between yield curve factors and macroeconomy.

1.2. Discount, Spot and Forward Rates and Rates Conversion

Although many theoretical models in financial economics hinge on an abstract interest rate, in reality there are many different interest rates. For example, the rate of a three month treasury bill is different from that of a six month treasury bill. The relationship between these different rates and maturities is known as the term structure of interest rates. The term structure of interest rates can be described in terms of spot rates, discount rates or forward rates.

The discount function gives the present value of unit face value bond which is repaid in m periods. The corresponding continuously compounded yield to maturity $R_t(m)$ of the investment is given as:

$$P_t(m) = \exp[-R_t(m)m] \tag{1.1}$$

where $P_t(m)$ is the current price of zero-coupon bond, t denotes a moment in time, m is the time to maturity and $R_t(m)$ is the corresponding continuously compounded yield to maturity, i.e., zero-coupon rate. Obviously, the discount function as in (1.1) is an exponentially decaying function of the maturity. The continuously compounded yields are also referred to as spot rates. Rewriting (1.1) in other way to describe the yield to maturity as a function of the value of bond:

$$R_t(m) = -\frac{1}{m}\log[P_t(m)] \tag{1.2}$$

An analysis based on spot rates, for different maturities, gives information about the term structure of interest rates. When we plot the spot rates $R_t(m)$ against maturities m, we get the spot rates curve.

From the discount function (1.1), we obtain the instantaneous forward rate, which is the interest rate contracted now and to be paid for a future investment. The forward rate as a function of maturity is referred to as implied forward rate curve. Assuming continuous compounding, we observe the following relationship between the value of bond and forward rates.

$$f_t(m) = -\frac{P_t^{'}(m)}{P_t(m)} \tag{1.3}$$

where $P_t^{'}(m) = \partial P_t(m)/\partial m$. Taking the periods infinitesimally closer, we obtain the instantaneous implied forward rate $f_t(m)$. It is essential to differentiate the implied forward rate from the market forward rate. The implied forward rates are derived theoretically from spot rates. Market forward rates, on the other hand, are the actual rates that are realized in a forward or future contracts in the future market.

An important result is that the spot rate is the weighted average of the implied forward rates. Assuming continuous compounding, the relationship between the yield to maturity and the implied forward rates is:

$$R_t(m) = \frac{1}{m}\int_0^m f_t(u)du \tag{1.4}$$

which implies that the zero-coupon yield is an equally-weighted average of the forward rates. Given the yield curve or the forward curve, we can price any coupon bond as the sum of the present values of future coupon and principal payments. This important relationship between the zero-coupon and the instantaneous forward rates is a critical component of the Nelson-Siegel model of the term structure. The expression in (1.3) and (1.4) are derived and explained in appendix A. The previous three relations (1.1, 1.3 and 1.4) show that knowledge of any one of these functions is sufficient to solve for the other two.

1.3. Fama and Bliss Forward and Spot Rates Method

The zero-coupon rate is the return (yield), on a bond corresponding to a single cash payment at a particular time in the future. This would represent the return on an investment in a zero-coupon bond with a particular time to maturity. The prices of zero-coupon bonds can be directly used to construct the term structure, however, the lack of market liquidity and the limited available maturity spectrum necessitates the estimation based on observed coupon-bearing bond prices. Therefore, in practice, yield curves are not observed. Instead, they must be estimated from observed bond prices on zero-coupon and coupon-bearing bonds.

There are several approaches to construct zero-coupon rates; one can proceed by fitting a smooth discount curve and then convert to yields at the relevant maturities using (1.2). The first discount curve approach to yield curve construction is due to McCulloch (1971, 1975), who suggested to model the discount curve using polynomial splines. As the discount rate is a monotonically decreasing function of maturity and, therefore, price of bonds can be expressed as a linear combination of discount rates. However, the fitted discount curve of the functional form, used by McCulloch, diverges at long maturities due to the polynomial structure.[6] Furthermore, the McCulloch (1971, 1975) methods operate on linear combinations of discount functions; the implied forward rate curve usually has some undesirable features. In addition, the choice of knot points for polynomial splines is also rather ad hoc. To avoid these problems with polynomial spline methods, Fisher *et al.* (1995) proposed to use smoothing splines for interpolating the term structure of interest rates. Another improved discount curve approach to yield curve construction is due to Vasicek and Fong (1982), who model the discount curve using exponential splines. They use negative transformation of maturity rather than maturity itself, ensures that forward rates and zero-coupon yields converge to a fixed limit as maturity increases. Their approach is more attractive for fitting yield curves at long ends.

Notwithstanding, the progress of polynomial, smoothing, and exponential spline based discount curve approaches remain potentially problematic. Besides, having some undesirable features at the long end of the term structure, these approaches usually do not generate good out-of-sample forecasts.

Another popular approach is the Bootstrapping for constructing a (zero-coupon) fixed-income yield curve from the prices of a set of coupon-bearing bonds. In first stage, the zero-coupon bonds are used to derive spot rates for the available maturities, and in the second stage the entire term structure is recovered from the bond yields by recursively forward

---

[6] The corresponding yield curve flattens out with maturity, as emphasized in Shea (1984).

substitution. This iterative process is called the Bootstrap method. Given that, in general, we lack data points in a yield curve (there are only a fixed number of bonds in the market) and more importantly these have varying coupon frequencies, it makes sense to construct a curve of zero-coupon instruments from which we can price any yield, whether forward or spot, without the need of more external information.

An alternative and popular approach to yield curve construction is due to Fama and Bliss (1987), who construct spot rates not from an estimated discount curve, but rather from estimated forward rates at the observed maturities. Their method sequentially constructs the forward rates necessary to price successively longer maturity bonds. Those forward rates are often called "unsmoothed Fama-Bliss" forward rates, and they are transformed to unsmoothed Fama-Bliss yields by appropriate averaging, using (1.4). The unsmoothed Fama-Bliss yields exactly price the included bonds. Unsmoothed Fama-Bliss yields are often the "raw" yields to which researchers fit empirical yield curves, such as members of the Nelson-Siegel family, about which we will have much to say throughout this dissertation. Such fitting effectively smoothes the unsmoothed Fama-Bliss yields.[7]

To apply the Fama-Bliss method, we assume that we have a sequence of bonds with possibly irregularly spaced maturity dates. Furthermore, we also assume that the implied forward rate between two successive maturity dates is constant.

Consider the following simple example to clarify the concept of Fama-Bliss unsmooth implied forward and spot rates. Suppose that we have: (i) the i^{th} bond matures at time t_i with market price P_i per unit face value, which equals the redemption value, for $i = 1,2,3,4$ and $0 < t_1 < t_2 < t_3 < t_4$, (ii) bonds 1 and 2 have no coupons, while bond 3 has two coupons, at time t_{31}^* with $t_1 < t_{31}^* < t_2$ and at maturity t_3 and bond 4 has three coupons, with coupon dates t_{41}^*, t_{42}^* and t_4, where $t_2 < t_{41}^* < t_3$ and $t_3 < t_{42}^* < t_4$, and (iii) the force of forward rate $f_t(m)$ follows a step function, taking constant values between successive maturities, i.e., $f_t(m) = f_{i,t}(m)$ for $t_{i-1} < t \leq t_i$, $i = 1,2,3,4$ with $t_0 = 0$ (details are given in table 1.1).

Table 1.1: The Bond Price and Coupon Rate Data

Bond	Coupon rate	Coupon date	Maturity	Bond price
1	0	–	t_1	P_1
2	0	–	t_2	P_2
3	C_3	$t_{31}^*, t_1 < t_{31}^* < t_2$	t_3	P_3
	C_3	t_3		
4	C_4	$t_{41}^*, t_2 < t_{41}^* < t_3$	t_4	P_4
	C_4	$t_{42}^*, t_3 < t_{42}^* < t_4$		
	C_4	t_4		

The table presents the data on both zero-coupon and coupon bearing bonds, their maturity time; coupon rate and coupon payments dates in typical standard notations to derive the analytical expressions for the Fama-Bliss unsmoothed implied forward rates and spot rates in standard financial notations.

[7] In the empirical part of this dissertation, the analysis is carried out using the unsmoothed Fama-Bliss spot rates.

As bond 1 is a zero-coupon, its value equation is:

$$P_1 = \exp[-t_1 R_{t_1}] = \exp\left[-\int_0^{t_1} f(u)du\right] = \exp[-f_1 t_1]$$

$$f_1 = -\frac{1}{t_1}\log P_1 \tag{1.5}$$

which is forward rate for the interval $(0, t_1]$.

For bond 2 (which is also a zero-coupon bond), the valuation equation take the form of:

$$P_2 = \exp[-t_2 R_{t_2}] = \exp\left[-\int_0^{t_2} f(u)du\right]$$

$$= \exp[-f_1 t_1 - f_2(t_2 - t_1)] \tag{1.6}$$

Substituting (1.5) in (1.6) and solving for f_2, we obtain:

$$f_2 = -\frac{1}{t_2 - t_1}[\log P_2 + f_1 t_1]$$

$$f_2 = -\frac{1}{t_2 - t_1}\log\left[\frac{P_2}{P_1}\right] \tag{1.7}$$

which is the forward rate for an interval $(t_1, t_2]$.

For bond 3, there is a coupon payment of amount C_3 at time t_{31}^*, with $t_1 < t_{31}^* < t_2$. Since, the equation of value for bond 3 can be:

$$P_3 = (1 + C_3)\exp\left[-\int_0^{t_3} f(u)du\right] + C_3\exp\left[-\int_0^{t_{31}^*} f(u)du\right]$$

$$P_3 = (1 + C_3)\exp[-f_1 t_1 - f_2(t_2 - t_1) - f_3(t_3 - t_2)] + C_3\exp[-f_1 t_1 - f_2(t_{31}^* - t_1)] \tag{1.8}$$

Solving (1.8) for f_3 implies:

$$f_3 = -\frac{1}{t_3 - t_2}\log\left[\frac{P_3 - C_3\exp[-f_1 t_1 - f_2(t_{31}^* - t_1)]}{(1 + C_3)P_2}\right] \tag{1.9}$$

By substituting (1.5) and (1.7) in (1.9), we obtain:

$$f_3 = -\frac{1}{t_3 - t_2}\log\left[\frac{P_3 - P_1 C_3\exp[-f_2(t_{31}^* - t_1)]}{(1 + C_3)P_2}\right] \tag{1.10}$$

which is the forward rate for an interval $(t_2, t_3]$.

As for the bond 4, using the similar argument, the equation of value for bond 4 can be written as:

$$P_4 = (1 + C_4)P_3^*\exp[-f_4(t_4 - t_3)] + P_2 C_4\exp[-f_3(t_{41}^* - t_2)]$$
$$+ P_3^* C_4\exp[-f_4(t_{42}^* - t_3)] \tag{1.11}$$

where $P_3^* = \exp[-f_1 t_1 - f_2(t_2 - t_1) - f_3(t_3 - t_2)]$; from (1.11), we obtain:

$$\frac{P_4 - P_2 C_4 \exp[-f_3(t_{41}^* - t_2)]}{P_3^*} = (1 + C_4)\exp[-f_4(t_4 - t_3)] + C_4 \exp[-f_4(t_{42}^* - t_3)] \quad (1.12)$$

The left-hand side of (1.12) can be computed using the data on P_4 and (1.5, 1.7 and 1.10), while the right-hand side contains the unknown quantity f_4. The equation can be solved numerically for the implied forward rate f_4 in the period $(t_3, t_4]$. For the continuously compounded spot rates (zero-coupon rates) using (1.4), we obtain:

$$R_1 = f_1 \qquad \text{for } 0 < t \le t_1$$

$$R_2 = \frac{f_1 t_1 + f_2(t - t_1)}{t} \qquad \text{for } t_1 < t \le t_2$$

$$R_3 = \frac{f_1 t_1 + f_2(t_2 - t_1) + f_3(t - t_2)}{t} \qquad \text{for } t_2 < t \le t_3$$

$$R_4 = \frac{f_1 t_1 + f_2(t_2 - t_1) + f_3(t_3 - t_2) + f_4(t - t_3)}{t} \qquad \text{for } t_3 < t \le t_4$$

1.4. Yield Curve Data in the Japanese Market

Various different yields may be observed at a certain point of time, corresponding to different bond maturities. But, yield curves evolve dynamically; hence, they have not only a cross-sectional, but also a temporal dimension. Here, we address the obvious descriptive question: how do yields tend to behave across different maturities and over time?

The data we use are monthly spot rates for zero-coupon and coupon-bearing bonds, generated using the pricing data of Japanese bonds and treasury bills. We use the end-of-month price quotes (bid-ask average) for Japanese government bonds, from January 2000 to December 2011, taken from the Japan Securities Dealers Association (JSDA) bonds files.[8] In total, there are 144 months in the dataset. Following Fama and Bliss (1987) method, in the first stage, each month a term structure of continuously compounded forward rates is calculated from the available maturities. Bills are used for maturities to a year. To extend beyond a year, the pricing assumption is that the forward rate for the interval between successive maturities is the relevant discount rate in the interval. Suppose forward rates for the month t are calculated for maturities to T and the next bond matures at $t + k$.

Coupons on the bond to be received prior to T are priced with the forward rates from to each payment date. Coupons and the principal to be received after T are priced with the daily forward rates from t to T and with the (solved for) forward rate to T and to $T + k$ that equates the price of the bond at t to the value of all payments. These calculations generate a step function term structure in which forward rates are the same between successive maturities.

[8] Some data for Japanese government bonds (JGB) in 2000 and 2001 is taken from bonds data files published by Tokyo Stock Exchange.

In second stage, forward rates are averaged to generate end of month term structure of yields for all the available maturities. Furthermore, we pool the data into fixed maturities. Because not every month has the same maturities available, we linearly interpolate nearby maturities to pool into fixed quarterly maturities of 3, 6, 9, 12, 15, 18, 21, 24, …, 300 months (100 maturities).

In table 1.2, we show descriptive statistics for yields at various maturities. Several well-known and important yield curve facts emerge. First, averaged yields (the average yield curve) increase with maturity; that is, term premia appears to exist, perhaps due to risk aversion, liquidity preferences, or preferred habitats. Second, the unconditional volatility of yield decreases with maturity, presumably because long rates involve averages of expected future short rates. It shows that long end of curve is less sharp and less volatile. Third, it also seems that the skewness has the downward trend with the maturity. Moreover, kurtosis of the short rates is lower than those of the long rates. Lastly, yields are highly persistent, as evidenced not only by the very large one-month autocorrelations but also by the sizable twelve-month autocorrelations. However, it seems that long rates are less persistent than the short rates.

Table 1.2: Descriptive Statistics of Yield Curve Data

Maturity	Mean	S. Dev.	Max.	Min.	Skewness	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
3	0.167	0.348	0.692	0.002	1.346	3.259	0.892	0.530	0.077
6	0.164	0.345	0.733	0.004	1.367	3.469	0.877	0.548	0.081
9	0.176	0.339	0.770	0.003	1.348	3.412	0.874	0.555	0.092
12	0.224	0.327	0.812	0.004	1.003	2.600	0.878	0.450	-0.001
15	0.250	0.327	0.855	0.003	0.956	2.487	0.870	0.459	0.021
18	0.276	0.304	0.990	0.013	0.974	2.589	0.873	0.455	0.018
21	0.303	0.303	0.990	0.027	0.932	2.475	0.877	0.451	0.022
24	0.327	0.292	1.027	0.019	0.896	2.382	0.875	0.434	0.025
30	0.387	0.284	1.117	0.027	0.871	2.368	0.865	0.403	0.026
36	0.446	0.281	1.186	0.078	0.815	2.315	0.862	0.383	0.035
48	0.594	0.280	1.368	0.121	0.653	2.133	0.855	0.326	0.027
60	0.730	0.273	1.517	0.161	0.509	2.079	0.856	0.269	0.027
72	0.864	0.265	1.627	0.216	0.365	2.137	0.849	0.210	0.025
84	1.011	0.262	1.759	0.285	0.214	2.234	0.842	0.129	0.035
96	1.165	0.260	1.878	0.382	-0.009	2.418	0.830	0.066	0.051
108	1.302	0.246	1.951	0.474	-0.224	2.784	0.832	0.056	0.091
120	1.424	0.231	1.998	0.549	-0.535	3.457	0.830	0.042	0.102
180	1.801	0.217	2.24	0.758	-1.388	6.203	0.841	-0.009	0.183
240	2.061	0.209	2.525	0.934	-1.934	8.291	0.850	-0.018	0.152
300	2.267	0.207	2.860	1.070	-1.774	7.983	0.874	-0.114	-0.045
Level	2.267	0.207	2.860	1.070	-1.774	7.983	0.874	-0.114	-0.045
Slope	2.099	0.311	2.842	1.031	-0.571	4.081	0.874	-0.043	-0.292
Curvature	-1.781	0.382	-0.993	-2.489	0.293	1.972	0.867	-0.017	-0.037

Note: The table shows descriptive statistics for monthly yields at different maturities and for the yield curve level, slope and curvature, where we define the level as the 25-year yield, the slope as the difference between the 25-year and 3-month yields, and the curvature as the twice the 2-year yield minus the sum of the 3-month and 25-year yields. The last three columns contain sample autocorrelations at displacements of 1, 12 and 24 months. The sample period is 2000:01–2011:12. The number of observations is 144.

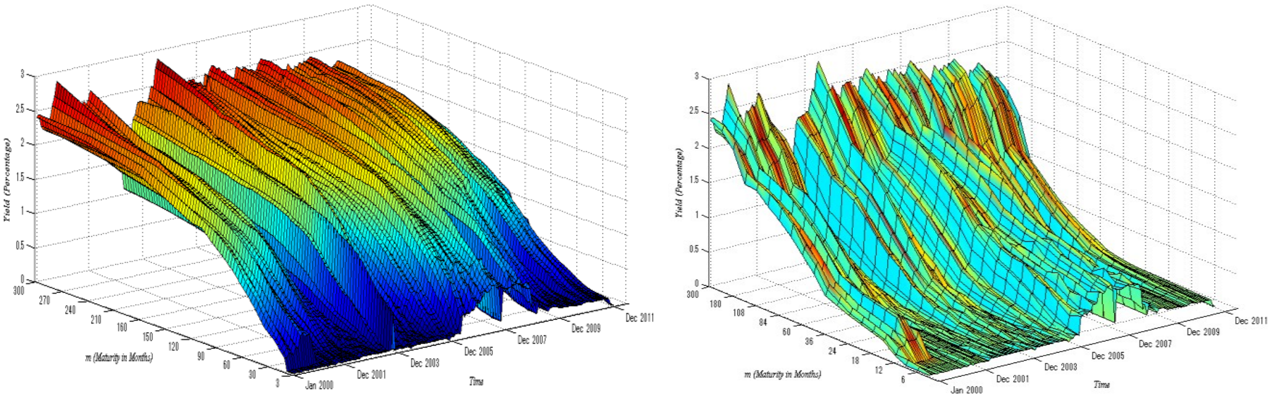


Figure 1.1: Yield Curves, 2000:01–2011:12.
The sample consists of monthly yield data 2000:01–2011:12 (144 months). The figure in the left pane shows the yield curves at fixed quarterly maturities of 3, 6, 9, 12, 15, 18, 21, 24 …300 months (100 maturities), while in the right pane depicts the term structure for some selected maturities such as 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240 and 300 months (20 maturities).

In figure 1.1, we show the three-dimensional surface for the Japanese government bonds market, with yields shown as a function of maturity, over time. The figure reveals that the yield curves move a lot, shifting among different shapes. The first noticeable fact is that yields vary significantly over time from which various common dynamics across all yields can be deduced. Especially, in the years 2000 to 2006, the short rates are nearly zero and on ward from 2006 there is an increasing trend in the yield for all the maturities. Moreover, in our data set, on average, we observe the upward sloping yield curve.

Furthermore, in chapter 3 and 4, we focus an area of macro-finance research that examines the relationship between the term structure of interest rates and the economy. We use monthly

Table 1.3: Descriptive Statistics of Macroeconomic and Stock Market Variables Data

	IP_t	EX_t	INF_t	MS_t	SI_t
Mean	0.705	-2.316	-0.225	2.139	-0.451
Std. Dev.	7.359	9.503	0.801	0.730	4.964
Maximum	16.506	21.233	2.098	3.645	12.011
Minimum	-18.476	-21.189	-2.532	0.439	-20.258
Skewness	-0.706	0.406	0.246	-0.168	-0.375
Kurtosis	3.583	3.077	3.979	2.485	4.065
$\hat{\rho}$ (1)	-0.276	-0.022	0.113	0.945	0.258
$\hat{\rho}$ (6)	0.243	-0.102	-0.229	0.651	-0.131
$\hat{\rho}$ (12)	0.795	-0.101	0.448	0.212	0.063
ADF-Statistic	-2.647	-12.045	-10.558	-2.496	-9.084
	(0.086)	(0.000)	(0.000)	(0.124)	(0.000)

Note: The table presents summary statistics for macroeconomic variables and capital market indicator data 2000:01–2011:12. All the four variables are measured as the last 12 months percentage growth rate. The IP_t is the annual growth rate in industrial production, EX_t is the (¥/$) annual growth of the real exchange rate, INF_t is the 12-month percent change in the consumer price index, MS_t is the growth of M2 money supply, and SI_t is 12 months growth rate of Tokyo Stock Exchange Index (TOPIX). $\hat{\rho}$ (i) denotes the sample autocorrelations at displacements of 1, 6 and 12 months. The last row contains augmented Dickey–Fuller (ADF) unit root test-statistic and its P-value (in parenthesis) .

data from January 2000 to December 2011, for industrial production, real exchange rate, consumer price index, money supply and Tokyo Stock Exchange share prices index (TOPIX) for the Japanese economy. The data for the former four variables is obtained from the International Financial Statistics (IFS) published by International Monetary Fund (IMF) while, for TOPIX is taken from annual reports of Tokyo Stock Exchange for various years. All the five variables are measured as the last 12 months percentage growth rate for two main reasons. First, for the stationarity consideration, as the time series of the variables in their level form are following *I(1)* process. Secondly, for the consistency purpose with the interest rate data, as our yields for all maturities are measured in annual percent format. The IP_t is the growth rate in industrial production, EX_t is the growth in real exchange rate (¥/$), INF_t is the inflation rate measured as 12-months percent change in the consumer price index, MS_t is the annualized growth rate of $M2$ money supply, and SI_t is the last 12-months growth rate of TOPIX. The descriptive statistics of the macroeconomic variables and the capital market indicator are depicted in table 1.3.

1.5. Yield Curve Modeling

The term structure models can be divided into three broad categories. The first stream relies on the optimizing behaviour of economic agents, using the dynamic stochastic general equilibrium (DSGE) framework. These models are all economic in nature. A model that forms the basis for this class of term structure models is the Vasicek (1977) model. The innovative feature of the Vasicek (1977) is that it models the interest rate as a mean reversion process. Other early contributions to the literature of equilibrium pricing include Cox *et al.* (1985), Dunn and Singleton (1986), Campbell (1986, 1993, 1996 and 1999) and more recently, Piazzesi and Schneider (2006). Other influential contributions in this class are Duffie and Kan (1996) and Dai and Singleton (2002). Duffie and Kan (1996) generalize this literature and Dai and Singleton (2002) characterize the set of admissible and identifiable models. However, based on the sound economic foundation, this approach still delivers unsatisfactory results in terms of poor in-sample fit as well as out-of-sample forecasts (Duffee, 2002). Furthermore, estimation of these models is repeatedly found to be challenging, requiring additional restrictions that are often not well motivated statistically or theoretically (Duffee, 2011).

A second stream of the literature adopts only the very basic structure of the DSGE approach to study the term structure of interest rates, usually imposing no-arbitrage condition, when estimating a Vector Autoregressive (VAR) model of yields. Significant development came in this stream of literature in 1990s. These models focus on fitting the term structure at a given point in time to ensure that no-arbitrage opportunities exist, see Hull and White (1990) and Heath *et al.* (1990). These type of models also include some recent studies such as Ang and Piazzesi (2003), Piazzesi (2005), Bikbov and Chernov (2010). The main (empirical) lesson from this stream of literature is that one needs a combination of observed and state variables to explain the dynamics of the term structure of interest rates.

The third relevant literature include statistical models without a structural interpretation; that is, models which synthesize data patterns without necessarily representing the theoretical models that fit under equilibrium and free-arbitrage conditions. This approach constitutes the statistical class of term structure models. This strand includes the methodology of principal components (Litterman & Scheinkman, 1991), curve interpolation models such as splines (McCulloch, 1971), smoothing splines (Shea, 1984), kernel regression (Linton *et al.* 2001), and the factors models for curve fitting such as Nelson and Siegel (1987) and Svensson (1994).

Diebold and Rudebusch, (2012) have emphasized three reasons that why the dynamic factor approach is appealing in modeling the term structure of interest rate. First, the functional form of factor imposed on the spot rate curve provides a highly accurate empirical description of yield curve data and have clear economic interpretations. It provides a framework that all bond price information can be summarized with just a few constructed variables or factors. Second, factors models prove tremendously appealing for statistical reasons. They provide a valuable compression of information, effectively collapsing an intractable high-dimensional modeling situation into a tractable low-dimensional situation.[9] Third, financial economic theory suggests, and routinely invokes, factor structure. There are thousands of financial assets in the markets, but for a variety of reasons, the risk premiums that separate their expected returns as driven by a much smaller number of components, or risk factors. In the equity sphere, for example, the celebrated capital asset pricing model (CAPM) is a single-factor model. Various extensions (e.g., Fama and French, 1993; 1992) invoke a few additional factors but remain intentionally very low-dimensional, almost always with less than five factors. Yield curve factors models are a natural bond market parallel.

One of the most popular subclasses within the statistical class of models is based on the Nelson and Siegel (1987) model. Originally intended to describe cross sectional aspects of yield curves, the Nelson-Siegel model imposes a parsimonious three-factor structure on the link between yields and different maturities, where the factors can be interpreted as level, slope and curvature. Diebold and Li (2006) find that a dynamic reformulation of this model provides forecasts that outperform the random walk and various alternative forecasting approaches. The Nelson-Siegel model thanks its popularity for a large part to its relative simplicity; ease of estimation and to the fact that there is some underlying economic interpretation in the three factors it is based on (Diebold and Li, 2006; De Pooter, 2007 and Almeida *et al.* 2007). Furthermore, the functional form in Nelson and Siegel (1987) imposes more smoothness on the shapes of the curves, as desirable by macroeconomists (Gürkaynak *et al.* 2007).

1.6. Yield Curve and Macroeconomic Factors

In the last decade, the use of the term structure of interest rates has been one of the most important topics of research in macroeconomics and finance. In the early literature of empirically oriented term structure models, the focus was on the optimal fit of the yield curve without taking

[9] As the data actually have factor structure, we need low-dimensional factor structure for statistical tractability.

into consideration the economic fundamentals. Since, finance models typically have no macroeconomic content, but instead focus on the consistency of asset prices across markets.[10] Similarly, in macro models, the entire financial sector is often represented by a single interest rate (short rate) with no yields spread and curvature for credit or liquidity risk and no role for financial intermediation or financial frictions. In order to understand the important aspects of the recent financial crisis and the economic recession, a joint macro-finance perspective is likely necessary. Therefore, in this section, we focus on an area of macro-finance research that has examined the relationship between the term structure of interest rates and the economy.

From a finance perspective, the short rate is a fundamental building block for rates of other maturities because long yields are risk-adjusted averages of expected future short rates. From a macro perspective, the short rate is a key monetary policy instrument, which is adjusted by the central bank in accordance to achieve economic stabilization goals. Taken together, a joint macro-finance perspective would suggest that understanding the way central banks move the policy rate in response to fundamental macroeconomic shocks should explain movements in the short end of the yield curve; furthermore, with the consistency between long and short rates enforced by the no-arbitrage assumption, expected future macroeconomic variation should account for movements farther out in the yield curve as well.

The expected real rate may be associated with expectations of future monetary policy and of future real growth. Moreover, because inflation tends to be positively related to activity, the expected inflation component may also be informative about future growth.

Although the yield curve has clear advantages as a predictor of future economic events, several other variables have also been widely used to forecast the path of the economy. Among financial variables, stock prices have received much attention. Finance theory suggests that stock prices are determined by expectations about future dividend streams, which in turn are related to the future state of the economy. Among macroeconomic variables, in a monetary policy context, forward rates are potentially useful as indicators of market expectations of future interest rates, inflation rates and exchange rates as discussed by Svensson (1994) and Sodelind and Svensson (1997), and the yield curve carries information about future GDP growth as shown by Estrella and Mishkin (1996, 1998). This macroeconomic explanation of term structure of interest rate has also been supported by empirical results, such as the model of Hördahl *et al.* (2006), in which forecast ability outperforms the random walk hypothesis. Similarly, Diebold *et al.* (2006) characterize the relationship among Level, Slope and Curvature factors of yield curve and the macro economy. They found strong evidence of macroeconomic effects on future yield curve and somewhat weaker evidence that yield curve affects future macroeconomic variables.

In addition to these studies, Redebusch and Wu (2003) and Ang *et al.* (2006) construct joint models and they both found that there is bidirectional link between yield curve and macroeconomic variables. Movements in macroeconomic fundamentals will have an impact on the shape of the curve and the level of interest rates. Although, economic theory would suggest

[10] The word "finance models" is used interchangeable for statistical class of models.

that the impact of the macro economy on yields is stronger than vice versa.

1.7. Definitions and Notations

In this section, we discuss the definitions and the notions that occur resiliently in studying term structure models. For the sake of readability, we use the same notations for parameters and variables in the different models throughout the entire study unless stated otherwise.

1.7.1. Short Rates and Long Rates

The short rate is the annualized interest rate (yield) for an infinitesimally short period of time. In practice, however, the three-month rate is considered a better approximation of the short rate because, for example, overnight loans are affected by factors that term structure models do not aim to cover (Schumacher, 2009). In this study, the short rate r_t is defined as:

$$r_t = R_t(0) = \lim_{m \to 0} R_t(m) \tag{1.15}$$

The long rate is the annualized spot rate (yield) for long horizon maturity, theoretically as maturity approaches to infinity. In practice, however, the infinite maturity bonds do not exist and the yield on bonds with maturity of ten years or more is considered a better approximation of the long rate. The long rate l_t is defined as:

$$l_t = R_t(\infty) = \lim_{m \to \infty} R_t(m) \tag{1.16}$$

1.7.2. Wiener Process

The Wiener process (also referred to as Brownian motion) is a continuous time stochastic process with small, independent increments. In finance, the Wiener process is used to describe changes in prices of options or in interest rates.

A continuous time process (W_t) is said to be a Wiener process if it satisfies the following three properties (Schumacher, 2009):

1. $W_0 = 0$.
2. Any two increments are independents, such as if $t_1 < t_2 < t_3 < t_4$, then the increments Δ_2 and Δ_4 are independent, where $\Delta_2 = (W_{t2} - W_{t1})$ and $\Delta_4 = (W_{t4} - W_{t3})$.
3. For any given t_1 and t_2 with $t_2 > t_1$, the distribution of the increment $\Delta_2 = (W_{t2} - W_{t1})$ is Gaussian with mean 0 and variance $\Delta t = (t_2 - t_1)$, i.e., $\Delta_2 \sim N\left(0, \sqrt{\Delta t}\right)$.

1.7.3. Mean Reversion

A process is said to be mean reverting if the process tends to fall (rise) after hitting a maximum (minimum) towards its central location. Interest rates are known to be subject to mean reversion over a longer horizon. Therefore, some models of term structure use an Ornstein-Uhlenbeck process for the dynamic evolution of interest rates.

The Ornstein-Uhlenbeck process is given by the following stochastic differential equation (SDE):

$$dy_t = \kappa(\mu - y_t)dt + \sigma dW_t \tag{1.17}$$

In (1.17), μ denotes the mean to which the process will revert; κ indicates the speed at which the process y_t reverts to μ and W_t is a Wiener process. As with the mean reverting process, the parameters κ, μ and σ are strictly positive. Furthermore, the amount of randomness is indicated by parameter σ. Also note that whenever y_t is high (i.e., larger than μ), the process is likely (rather than surely because of the randomness involved) to move downwards. The opposite also holds true.

1.8. Onward

In the chapters that follow, we address the issues and questions raised in this chapter, and many others. We discuss and compare the Cox-Ingersoll-Ross (CIR) model and Nelson and Siegel (1987) exponential components framework to distill the entire term structure of zero-coupon yield curve in chapter 2. The CIR model is calibrated using maximum likelihood estimation method and the dynamic Nelson and Siegel (1987) model using cross-sectional data by the non-linear least squares procedure. For comparison, we also estimate the linearized version of Nelson-Siegel model with ordinary least squares (OLS). The objective is to compare two classes of models in terms of in-sample fit and out-of-sample forecasts of the term structure of interest rate.

In chapter 3, we evaluate the out-of-sample forecast performance of the dynamic Nelson-Siegel model with and without macroeconomic variables in the state-space representation. We explicitly incorporate the three macroeconomic variables, i.e., the level of economic activity, exchange rate and inflation rate and one stock market activity indictor (TOPIX) in the state-space representation of yield curve model to analyze its impact in the in-sample fit and subsequently the efficiency gain in forecasting the yields for various maturities.

Chapter 4 deals with the term structure of interest rates and monetary policy. We use a three-factor term structure model, based on the classic contribution of Nelson and Siegel (1987). We incorporate three macroeconomic variables, i.e., the level of economic activity, money supply and inflation rate in the state-space representation along with stochastic volatility component in the yield curve model to examine the effect of monetary policy stances on term structure and the possible feed-back effect on the real sector using the Japanese experience of zero interest rate policy (ZIRP).

Lastly, in chapter 5, we provide the main conclusion from this study; highlight aspects of the current frontier and attempting to point out the way toward additional progress.

Chapter 2

Term Structure Modeling and Forecasting of Government Bond Yields

Does a Good In-Sample Fit Imply Reasonable Out-of-Sample Forecasts?

2.1. Introduction

Nothing in economy is watched much closer on a minute by minute basis than the yield curve. The central banks around the world try to manage it and everyone tries to forecast it. Its shape is a key to the profitability of many businesses and investment strategies. Equally important is the ability of the model to forecast the future term structure as it can be interpreted as a predictor of the future state of economy.[11] Therefore, accurate modeling, estimation and precise forecasting of the term structure of interest rate are of crucial importance in many areas of finance and macroeconomics.

Although the prices of zero-coupon bonds can be directly used to construct the term structure, however, due to the limited available maturity spectrum and lack of market liquidity of the zero-coupon bonds, it is essential to estimate the yields based on the observed coupon-bearing bond prices. Therefore, several term structure models have been developed in the course of time to plot the yield curve. A model that forms the basis of many other term structure models is the Vasicek (1977) model. The innovative feature of the Vasicek (1977) is that it models the interest rate as a mean reversion process. A famous extension to the Vasicek model is the Cox-Ingersoll-Ross (1985) model, which aims to cope with some of the drawbacks of the Vasicek model. The Cox *et al.* (1985) model describes the evolution of the short rates and distills the entire term structure by only one stochastic variable. Other famous extensions are the Vasicek and Fong (1982), Hull and White (1990) and Black *et al.* (1990) models.

However, more positive results have emerged recently based on the framework of Nelson and Siegel (1987). Originally intended to describe the cross sectional aspects of the yield curves, the Nelson-Siegel model imposes a parsimonious three-factor structure on the link between yields of different maturities, where the factors can be interpreted as level, slope and curvature. Though statistical in nature, the standard Nelson-Siegel model is still widely used due to its good fit of the

[11] These forecasts are used by companies in their investment decisions and discounting future cash flows, consumers in their saving decisions, and economists in the policy decisions.

observed term structure.[12]

This chapter discusses the Cox-Ingersoll-Ross (CIR) model and the Nelson-Siegel exponential components framework to distill the entire term structure of zero-coupon yields. Being derived from dynamic stochastic general equilibrium (DSGE) specification, the CIR model was characterized for theoretical purposes, whereas, the motivation for the Nelson-Siegel model comes from the stylized facts that can be inferred from empirical analysis. The CIR model is compared with the Nelson-Siegel model to find out which of the two classes is appropriate for forecasting purposes. The comparison between the Nelson-Siegel and the CIR models will help to find out which of the two can appropriately represent the true characteristics of the market. We also compare the in-sample fit of Nelson-Siegel model for the linear and non-linear estimation methods.

Furthermore, we simulate the CIR and the Nelson-Siegel models to find out whether simulation results match the larger trends and statistics (i.e., stylized facts) of the actual interest rate data. In this context, we aim to understand that:

- Which of the two classes of models well explain the entire term structure of interest rates?
- Does non-linear estimation of Nelson-Siegel model lead to a better in-sample fit than the linear estimation process?
- Does better fit imply reasonable simulation results?

The motivation to simulate interest rates may be to examine the out-of-sample performance of the two classes of term structure models. An interesting reading on this topic for the Nelson-Siegel model is in Diebold and Li (2006), which indicates that the model produces term structure forecasts at both short and long horizons with encouraging results.

The chapter contributes to the existing literature in two ways. In calibrating the multi-factor Nelson-Siegel model, we estimate the dynamic version of the model by employing the non-linear least squares estimation procedure and allow all the four parameters to vary over time.[13] We show that how the non-linearized version of the model (assuming the time-varying τ) leads to a better in-sample fit as compared to the linearized one. Secondly, we model the four time-varying factors of Nelson-Siegel model to simulate the yield curve, contrary to the previous studies in which parameter τ_t is fixed to a pre-specified value and they model three factors to forecast the term structure. Lastly, in estimating the CIR and Nelson-Siegel models, some new empirical facts

[12] For instance, De Pooter (2007) states that nine out of thirteen central banks that report their curve estimation methods to the Bank of International Settlements (BIS) use either the Nelson-Siegel model or its variation. Furthermore, Diebold and Li (2006) find that the dynamic reformulation of this model provides forecasts that outperform the random walk and various alternative forecasting approaches.

[13] In the earlier studies, the parameter τ is pre-specified to a fixed value without estimation. For example, Diebold and Li (2006) argue that τ is to be taken as a constant with little degradation of fit, but it greatly simplifies the estimation procedure. They fix τ to 30 months that maximizes the loading of the curvature factor. Similarly, Fabozzi *et al.* (2005) set the shape parameter τ to 3 leaving the hump located at 5.38 years, arguing for the computational efficiency (no iterations through τ need to be performed). However, in some studies τ is considered as time invariant unknown parameter (does not pre-specify). Such as Diebold, *et al.* (2006) estimate τ to be 23.3 months ($\lambda = 0.077$). In Ullah *et al.* (2013), the estimated τ is 71.420 implying that the loading on the curvature factor is maximized at a maturity of about 6 years.

will emerge from the Japanese market data. Of particular importance, short-term yields such as the three and six-month yields were essentially stuck at zero during most of the period from 2000 to 2006. It will also be interesting to figure out that how the short rate CIR model fits the very low short-term interest rate to compute the entire term structure.

The chapter is organized as follows. In section 2.2, the Cox-Ingersoll-Ross (1985) model and the dynamic multi-factor Nelson-Siegel (1987) model are discussed. Section 2.3 describes the Japanese interest rate data and estimates the parameters of the models. We evaluate the forecasting performance of the two competing term structure models in section 2.4, while section 2.5 concludes the chapter.

2.2. Term Structure Models

The term structure of interest rates describes the relationship between interest rates and time to maturity. At a certain point of time for various maturities, the term structure can have different shapes. The curves that encounter in reality can be upward, downward sloping, flat or humped shape. These typical shapes can be generated by a class of functions associated with the solutions of differential or difference equations. Cox *et al.* (1985) developed a general equilibrium model with explicit analytical expression for the equilibrium interest rate dynamics and bond prices using the first order stochastic differential equation (SDE). Being a general equilibrium model, it contains all the elements of the traditional expectation hypothesis. On the other hand, Nelson and Siegel (1987) introduced a model for term structure which explains 96% of the variation of the yield curve across maturities with the help of second order differential equation.

Based on the definitions and notations in section 1.2 and 1.7, in the next two subsections we present the models.

2.2.1. Cox-Ingersoll-Ross (CIR) Model

Vasicek (1977) developed a one-factor model of the term structure which depends on only one uncertainty factor, i.e., the short rate. Vasicek defines the short rate process as:

$$dr_t = \kappa(\mu - r_t)dt + \sigma dW_t \tag{2.1}$$

As with the mean reverting process, the three parameters κ, μ and σ are strictly positive and W_t is a Wiener process. A major drawback of the Vasicek model is that the model can produce negative interest rates.[14] Cox *et al.* (1985) adopt a general equilibrium approach to endogenously determine the risk-free rate. They reformulated the Vasicek model, in order to prevent the short rate from becoming negative, as:

$$dr_t = \kappa(\mu - r_t)dt + \sigma\sqrt{r_t}dW_t \tag{2.2}$$

The $\kappa(\mu - r_t)dt$ is a drift term which represents the mean reversion and is similar to the drift

[14] If real interest rates are to be modeled, this does not necessarily have to be a big problem as real interest rates can be negative in reality. Nominal rates, on the contrary, will never be negative in practice.

term in the Vasicek model. The difference between the two models is the square root in the second (volatility) term, which prevents the short rate from becoming negative.[15]

Furthermore, the short rate r_t as in (2.2) follows a non-central chi-square distribution with $(2q+2)$ degrees of freedom, and the parameter of non-centrality $2u$ is proportional to the current spot rate. The probability density of the interest rate r_{t_i} at time t_i conditional on $r_{t_{i-1}}$ at t_{i-1} is given as:

$$f_{\mathrm{CIR}}\left(r_{t_i} | r_{t_{i-1}}; \xi, \Delta t\right) = c[\exp(-u-v)]\left(\frac{v}{u}\right)^{\frac{q}{2}} I_q\left(2\sqrt{uv}\right) \tag{2.3}$$

where $\xi = (\kappa, \mu, \sigma)'$ is the parameters vector,

$$c = \frac{2\kappa}{\sigma^2[1-\exp(-\kappa\Delta t)]}$$
$$u = cr_{t_{i-1}} \exp(-\kappa\Delta t)$$
$$v = cr_{t_i}$$
$$q = \frac{2\kappa\mu}{\sigma^2} - 1$$

and $I_q\left(2\sqrt{uv}\right)$ is a modified Bessel function of the first kind of order q.

Valuing the zero-coupon bond, Cox *et al.* (1985) show that the pricing function in the CIR model can be expressed as:

$$P_t(m) = A_t(m)\exp[-B_t(m)]\, r_t \tag{2.4}$$

where

$$A_t(m) = \left[\frac{2\theta.\exp\left\{\frac{m}{2}(\theta+\kappa)\right\}}{2\theta + (\kappa+\theta)[\exp(m\theta)-1]}\right]^{2\kappa\mu/\sigma^2} \tag{2.5}$$

$$B_t(m) = \frac{2[\exp(m\theta)-1]}{2\theta + (\kappa+\theta)[\exp(m\theta)-1]} \tag{2.6}$$

$$\theta = \sqrt{\kappa^2 + 2\sigma^2} \tag{2.7}$$

The bond price in (2.4) is a decreasing concave function of maturity m and decreasing convex function of the short-term interest rate r_t and mean interest rate level μ. Furthermore, $P_t(m)$ is an increasing concave (decreasing convex) function of κ (the speed of adjustment parameter), if the short-term interest rate r_t is greater (less) than θ. The bond price is also an increasing

---

[15] When r_t approaches zero, the volatility term $\sigma\sqrt{r_t}$ approaches zero. In this case, the short rate will only be affected by the drift term, resulting the short rate to revert to the mean again. Cox *et al.* (1985) show that whenever $2\kappa\mu > \sigma^2$, the interest rate is strictly larger than zero. Furthermore, there is empirical evidence that whenever interest rates are high, the volatility is likely to be high as well, which justifies the volatility term in the CIR model.

concave function of the interest rate variance σ^2.[16]

Rewriting the expression for $P_t(m)$ in (2.4) and substituting it in (1.2), implies a function to compute the term structure of interest rate in the CIR model as:

$$R_t(m) = \frac{1}{m}\left[\left(B_t(m)r_t\right) - \log\left(A_t(m)\right)\right] \quad (2.8)$$

with $A_t(m)$, $B_t(m)$ and θ are as in (2.5), (2.6) and (2.7) respectively.

On a time grid $0 = t_0, t_1, t_2, \ldots\ldots$ with time step $\Delta t = t_i - t_{i-1}$, the discretized version of the CIR model is defined as:

$$r_{t+\Delta t} = r_t + \kappa(\mu - r_t)\Delta t + \sigma\sqrt{\Delta t}\sqrt{r_t}\varepsilon_t \quad (2.9)$$

with $\varepsilon_t \sim N(0,1)$. Various different shapes of the term structure can be computed by the CIR model by changing the parameters values in (2.8).

2.2.2. Nelson-Siegel Model

Motivation for Nelson-Siegel model comes from the expectation hypothesis. According to the expectation hypothesis, forward rates will behave in such a way that there is no arbitrage opportunity in the market. In other words, the theory suggests that implied forward rates are the rationally expected spot rates of the future periods. Nelson and Siegel (1987) propose that if spot rates are generated by a differential equation, then implied forward rates will be the solutions to this equation. Assuming a second-order differential equation, to describe the movements of the yield curve, with the assumption of real and equal roots, the solution will be the instantaneous implied forward rate function as:

$$f_t(m) = \beta_{1t} + \beta_{2t}\exp\left(\frac{-m}{\tau_t}\right) + \beta_{3t}\left[\left(\frac{m}{\tau_t}\right)\exp\left(\frac{-m}{\tau_t}\right)\right] \quad (2.10)$$

for $t = 1,2,\ldots,T$ and time-varying parameters vector $\psi_t = (\beta_{1t}, \beta_{2t}, \beta_{3t}, \tau_t)'$.

The model may be viewed as a constant plus a Laguerre function, that is, a polynomial times an exponential decay term, which belongs to a mathematical class of approximating functions. The solution for the yield as a function of maturity, using (1.4), is:

$$R_t(m) = \beta_{1t} + \beta_{2t}\left[\frac{1 - \exp(-m/\tau_t)}{m/\tau_t}\right] + \beta_{3t}\left[\frac{1 - \exp(-m/\tau_t)}{m/\tau_t} - \exp\left(\frac{-m}{\tau_t}\right)\right] \quad (2.11)$$

The Nelson-Siegel specification of yield in (2.11) can generate several shapes of the yield curve including upward sloping, downward sloping and (inverse) humped shape with no more than one maxima or minima. The functional form imposed on the forward interest rates as in (2.10) leads

[16] It is due to that larger σ^2 value indicates more uncertainty about future real production opportunities, and thus more uncertainty about future consumption. In such a world, risk-averse investors would value the guaranteed claim in a bond more highly.

to a flexible, smooth parametric function of the term structure that is capable of capturing many of the typically observed shapes that the yield curve assumes over time and captures most of the properties of the term structure.

The limiting path of $R_t(m)$, as m increases, is its asymptote β_{1t}; and, when m is small, the limit is $(\beta_{1t} + \beta_{2t})$. β_{1t} is the asymptotic value of the spot rate function, which can be seen as the long-term interest rate and is assumed (required) to be positive ($\beta_{1t} > 0$). Furthermore, the loading of β_{1t} equals one (constant and independent of m) and, therefore, the term structure at different maturities is affected by β_{1t} equally, which justifies the interpretation of β_{1t} as a level factor. The instantaneous short rate is given by $\beta_{1t} + \beta_{2t}$, which is constrained to be greater than zero. Furthermore, β_{2t} determines the rate of convergence with which the spot rate function approaches its long-term trend. The slope will be negative if $\beta_{2t} > 0$ and vice versa. The loading of β_{2t} approaches to one as $m \to 0$ and to zero as $m \to \infty$. Therefore, the yield curve is primarily affected by β_{2t} in the shorter run, so a change in β_{2t} implies a change in the slope of the term structure. Therefore, it is legitimate to interpret β_{2t} as the slope factor. The loading that comes with β_{3t} starts at 0, increases, and then decays to zero. Since, β_{3t} has the greater impact on medium-term yields and can be termed as the curvature factor, because it affects the curvature of the term structure. Furthermore, the parameter β_{3t} determines the size and the form of the hump, i.e., $\beta_{3t} > 0$ results in a hump, whereas $\beta_{3t} < 0$ produces a U-shape.

Finally, the parameter τ_t determines the maturity time at which the loading of the β_{3t} is optimal. It also specifies the location of the hump or the U-shape on the yield curve. Therefore, the range of shapes the curve can take is dependent on τ_t, it can be interpreted as a shape factor. The small values of τ_t, which have rapid decay in regressors, tend to fit low maturities interest rates quite well and larger values of τ_t lead to more appropriate fit of longer maturities spot rates. It has an interesting rule and economic interpretation as it shows a point of maturity m that separates the short rate from the medium/long-term rates.

2.3. Parameter Calibration and Estimation

Taking into account three dimensions–yield, maturity and time–of the data, different estimation methods can be used. To estimate the CIR model, one could choose to do a cross-sectional or time series estimation. For the Nelson-Siegel as the factors are time-dependent, one can proceed with cross-sectional or multivariate time series estimation. The differences between the estimates should be small if the employed model of the term structure is true. In this study, we estimate the CIR model using the time series data and the Nelson-Siegel model via the cross-sectional data for each observed month in the dataset.

2.3.1. Data

We use the monthly spot rates of Japanese government bond with fixed quarterly maturities of 3, 6, 9, 12, 15, 18, 21, 24,…,300 months (100 maturities). The spot rates are derived from the bid/ask average price quotes, from January 2000 through December 2011, using the Fama and

Bliss (1987) methodology. The data for price quotes for Japanese government bonds is taken from the Japan Securities Dealers Association (JSDA) bonds files. The descriptive statistics and three dimensional plot of the data is given in section 1.4 (table 1.2 and figure 1.1).

2.3.2. Calibration of the Cox-Ingersoll-Ross Model

The parameters vector of the Cox-Ingersoll-Ross model $\xi = (\kappa, \mu, \sigma)'$, as introduced in (2.2), is estimated using the time series data. To estimate the parameters vector ξ by maximum likelihood method, we use the CIR density given in (2.3). For T be the number of observations, e.g., the number of months the interest rate is observed, the likelihood function is given by:

$$L(\xi) = \prod_{i=1}^{T} f_{CIR}\left(r_{t_i} \middle| r_{t_{i-1}}; \xi, \Delta t\right) \tag{2.12}$$

for $i = 1,2, \dots \dots, T$. Moreover, maximizing the log-likelihood function is often easier than maximizing the likelihood function itself, we take natural logarithm on both sides in (2.12), resulting in:

$$\log[L(\xi)] = T\log(c) + \sum_{i=1}^{T} \left[-u - v + \frac{q}{2}\log\left(\frac{v}{u}\right) + \log\left(I_q\left(2\sqrt{uv}\right)\right)\right] \tag{2.13}$$

Maximizing (2.13) over its parameter space yields maximum likelihood estimates $\hat{\xi}$.[17] Matlab built-in function `fminsearch` is used to minimize the negative log-likelihood function to obtain $\hat{\xi}$. However, direct implementation of the Bessel function $I_q\left(2\sqrt{uv}\right)$ into Matlab causes the program to crash. A failure occurs because the Bessel function diverges to plus infinity on a high pace. To cope with this problem, scaled Bessel function [denoted by $I_q^{scaled}\left(2\sqrt{uv}\right)$], defined as $I_q\left(2\sqrt{uv}\right)\left[\exp\left(-2\sqrt{uv}\right)\right]$, is used. To take the scaled Bessel function into account, the log-likelihood function in (2.13) is adjusted as:[18]

$$\log[L(\xi)] = T\log(c) + \sum_{i=0}^{T} \left[-u - v + \frac{q}{2}\log\left(\frac{v}{u}\right) + \log\left(I_q^{scaled}\left(2\sqrt{uv}\right)\right) + 2\sqrt{uv}\right] \tag{2.14}$$

We use the OLS estimators as the start values of the discrete version of the CIR model (2.9) for the optimization problem defined in (2.14). To estimate the parameter vector ξ, using (2.14), one can use the time series data of three months, six months, one year or two years maturity yields. Obviously, taking different yield data implies different parameter estimates. We choose to

[17] Note that, as the logarithmic function is a monotonically increasing function, maximization of the likelihood function also maximizes the log-likelihood function. That is, the location of the maximum does not change.

[18] In (2.14) the term $2\sqrt{uv}$ appears because $\left[\exp\left(-2\sqrt{uv}\right)\right]$ in the scaled Bessel function should be canceled out to keep the log-likelihood function the same.

calibrate the model on the two years maturity yields, for two reasons. On the one hand, the CIR model is a short rate model, so the time to maturity should not be too large. On the other hand, taking a short maturity time, say three or six months, might yield strange estimates because of the extremely low interest rates and high volatilities in the initial years of the data from 2000 to the end of 2006.[19] Moreover, the data is on a monthly interval, the time step is set equal to 1/12.

The results of initial estimates of OLS along with the global optimal estimates, using the maximum likelihood method, are depicted in the first panel of table 2.1. Given the initial estimates, the maximum likelihood estimates (MLE) in panel 1 of table 2.1 shows that the fitted yield curve is upward sloping.

Figure 2.1 (upper pane) plots the average observed and the estimated yield curve. It is clearly visible that the CIR model plots an upward sloping yield curve like the observed one. In the perfect case, the two curves would match exactly. However, we observe that estimated yield is closer to the actual yield curve up to two years maturity and the discrepancy between the two is the increasing function of maturity beyond two years, as the residuals curve is upward sloping. It may be largely due to the low interest rates from 2000 to 2006. In order to get deeper insight of the behavior of the yield curve during the prolonged period of zero policy rate, we also estimated the CIR model for the two sub-periods, i.e., sub-period 1 (January 2000 to December 2006) and sub-period 2 (January 2007 to December 2011). In the second panel of table 2.1, we provide the initial and MLE estimates for the two subsets of data, i.e., the zero interest rate period (2000 to 2006) and the non-zero interest rate period (2007 to 2011). Furthermore, the estimated yield

Table 2.1: Results of the MLE Estimation of the CIR Model

	$\hat{\kappa}$	$\hat{\mu}$	$\hat{\sigma}$	log L
Panel 1. Full Period Sample Results (2000:01–2011:12)				
Initial (OLS)	0.9287	0.0030	0.0809	
MLE	1.6149	0.0031	0.0775	6702.800
Panel 2. Results for Two Sub-Periods Samples				
Sub-Period I (2000:01–2006:12)				
Initial (OLS)	0.6185	0.0035	0.0760	
MLE	1.4591	0.0030	0.0738	3881.000
Sub-Period II (2007:01–2011:12)				
Initial (OLS)	1.6540	0.0033	0.0879	
MLE	2.1960	0.0035	0.0838	2935.000

Note: The table presents the initial OLS and MLE estimated results of $\hat{\xi}$ vector using the time series data of two years maturity. log L denotes the log likelihood value of the MLE estimation. Panel 1 consists the results of the full sample period, 2000:01–2011:12 (144 observations), while panel 2 presents the results for two sub-periods, i.e., sub-period 1 (2000:01–2006:12) and sub-period 2 (2007:01–2011:12). The number of observations for the first sub-period and second sub-period is 84 and 60 respectively.

[19] We also tried the 3 months, 6 months, one year and 18 months short rates and the results are almost same with the 24 months short rates results. However, the 24 months yield data fits the estimated yield curve a slightly better than the 3, 6, 12, and 18 months at short maturity. Estimated results are reported in appendix B.

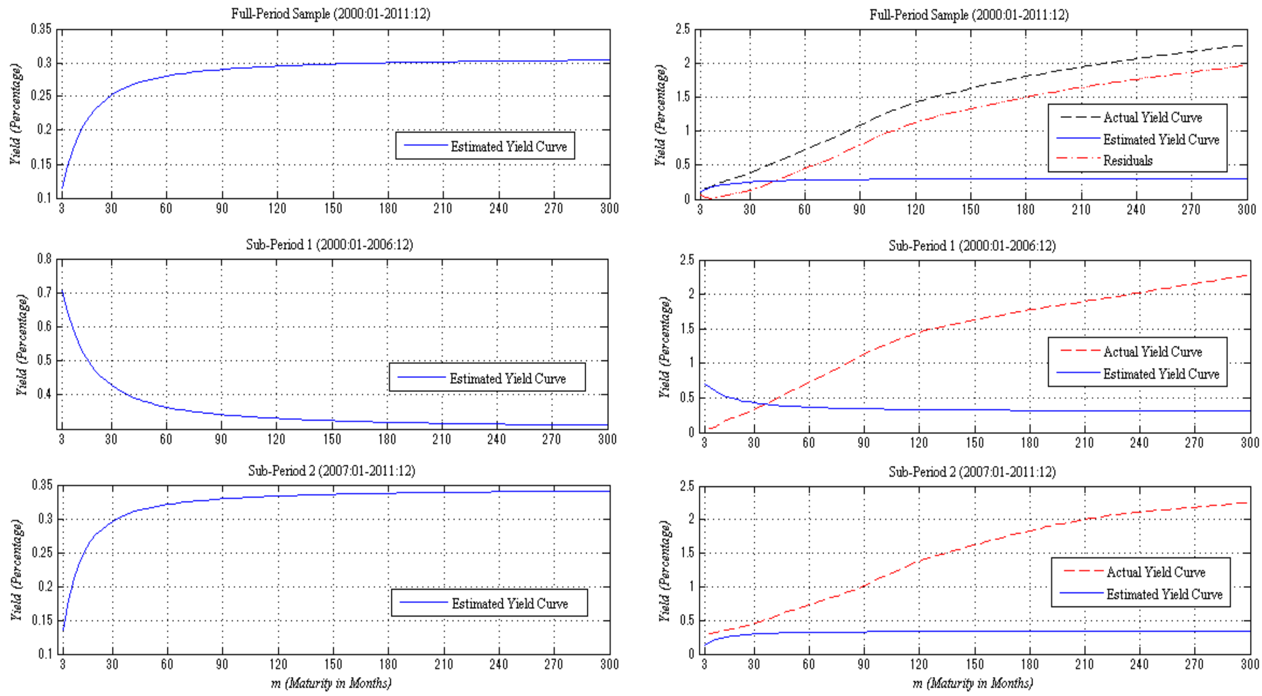


Figure 2.1: Fitted Yield Curve with the CIR Model
Actual average (data-based) and fitted (model-based) yield curve along the residuals for the entire sample (2000:01–2011:12) and two sub-periods, i.e., sub-period 1 (2000:01–2006:12) and sub-period 2 (2007:01–2011:12) are plotted. The fitted yield curves are obtained by evaluating the CIR function at the MLE estimated $\hat{\kappa}, \hat{\mu}$ and $\hat{\sigma}$ from the table 2.1.

curves for both sub-periods are depicted in the lower two panes of figure 2.1.

The maximum likelihood estimates for the first sub-period show that the fitted yield curve is negatively sloped, however for the second sub-period the estimated yield curve has an upward slope. Furthermore, the plots of estimated yield curve for both the sub-periods in figure 2.1 also support this view.

2.3.3. Estimation of the Nelson-Siegel Model

The Nelson-Siegel model in (2.11) forms the basis for our estimation procedure. For estimating the parameters of the model, we consider the functional form as:

$$R_t(m) = \Lambda(\tau_t)\beta_t + \varepsilon_t \tag{2.15}$$

where $R_t(m)$ denotes the (N×1) vector of yield rates at time t for N distinct maturities, $\Lambda(\tau_t)$ is (N×3) matrix of loadings and $\varepsilon_t \sim N(0, \sigma^2 I_N)$ is the error term, which accounts for whatever is not captured in the function $R_t(m)$ about how bonds are priced. The $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the vector of unknown parameters. Furthermore, τ_t is also unknown parameter.

Contrary to the prior studies, we do not fix τ_t to a pre-specified value, but allow it to vary over time and can be optimally determined in the estimation process in order to obtain a better in-sample fit. As the dynamic Nelson-Siegel function of spot rates results in a non-linear model, we employ the non-linear least squares method to estimate the model parameters $\psi_t = (\beta_{1t}, \beta_{2t}, \beta_{3t}, \tau_t)'$ for each month t. To minimize the sum of squared zero-yield errors, the objective function $F(\beta_t, \tau_t)$ is given by:

$$F(\beta_t, \tau_t) = [R_t(m) - \Lambda(\tau_t)\beta_t]^2 \tag{2.16}$$

We derive the analytical gradient $\nabla F(\beta_t, \tau_t)$ for the objective function in (2.16) and solve numerically for the optimal $\hat{\psi}_t$.[20] The analytical gradient of $F(\beta_t, \tau_t)$ is reported in appendix C. Moreover, following Nelson and Siegel (1987), we set τ_t to the median value estimated in the non-linearized version of Nelson-Siegel model (in previous stage) and estimate it by the ordinary least squares (OLS) in order to make a comparison between the linearized and the non-linearized versions of the model.

Applying the non-linear least squares to the yield data for each month gives us a time series of estimated parameters vector $\hat{\psi}_t$ and the corresponding panel of fitted yields $\hat{R}_t(m)$ and residuals $\hat{\varepsilon}_t$ (pricing errors). The first panel of table 2.2 shows the descriptive statistics of the estimates of the Nelson-Siegel model of the non-linear least squares method.

The estimated vector of parameters $\hat{\psi}_t$ is highly statistically significant.[21] From the autocorrelations in the table 2.2 (panel 1) of the four factors, we can see that the $\hat{\beta}_{3t}$ and $\hat{\beta}_{1t}$ are the most persistent, and that the second factor is a bit less persistent than the first. It suggests that long rates are slightly more persistent than short rates. Although the lag autocorrelation is reasonably high, the Augmented Dickey–Fuller (1979) test for unit root suggests that all the estimated factors $\hat{\beta}_{1t}$,$\hat{\beta}_{2t}$,$\hat{\beta}_{3t}$ and $\hat{\tau}_t$ are *I(0)* process and stationary at level.[22] However, $\hat{\beta}_{1t}$ solely determines the long run limiting behavior of the Nelson-Siegel model. The results also indicate that the residuals autocorrelation is low, justifying the use of non-linear least squares method. The average R^2 and residuals indicate that the average yield curve is fitted very well.

Furthermore, the time series plot of the $\hat{\tau}_t$ in figure 2.2 shows that the optimal point of β_{3t} loading ranges from 1.6 to 10 years. It indicates that there is a large degree of variability in the $\hat{\tau}_t$ over the period selected. Testing the sample with the median value of $\hat{\tau}_t$ leads to a small loss of accuracy of the fitted curve but there is a large variation in the $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$.[23] The descriptive statistic results of estimated $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$ for the fixed value of τ (median value of $\tau = 38.068$), estimated by OLS, are presented in the second panel of table 2.2.

The degree of loss of fit ranges from 1.4% to 5.7%. Comparing the results in panel 1 and 2 of table 2.2, there is significant difference in the estimated factors of Nelson-Siegel model for the

[20] The optimization problem stated in (2.16) is non-convex and may have multiple local optima, which increases the dependence of the numeric solution on the starting values. Arbitrarily choosing the start parameters possibly may not reach to a global optimum. This phenomena is avoided by applying the one-dimensional grid search to the system to estimate τ_t (denoted as $\tilde{\tau}_t$) and substituted in the (2.11) to linearize the dynamic model. Subsequently, OLS is employed to estimate the parameter vector β_t (denoted as $\tilde{\beta}_t$). The grid search $\tilde{\tau}_t$ and the OLS estimated $\tilde{\beta}_{1t}$,$\tilde{\beta}_{2t}$ and $\tilde{\beta}_{3t}$ are used as the initial values to find the numeric solution of optimization problem defined in (2.16).

[21] The p-value of individual t-statistic (not reported) is less than 0.03 in almost every period for of all the four factors.

[22] Based on the SIC criteria, optimal lag 3 has been selected for all the four variables in employing the augmented Dickey–Fuller unit-root test. The MacKinnon critical values for rejection of hypothesis of a unit root are -4.023at the one percent level, -3.441 at the five percent level and -3.145 at the ten percent level.

[23] The median value of $\hat{\tau}_t$ is 38.068.

two estimation processes. The linearized version of model either under-estimate or over-estimate the actual yield curve, whereas the non-linear estimation application leads to a reasonable fit of the yield curve. It suggests that standardizing the parameter τ_t to a prespecified value, not only reduces the degree of fit but also leads to a significant biased in the estimated parameters $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$.

Table 2.2: Descriptive Statistics of the Nelson-Siegel Estimated Factors

	$\hat{\beta}_{1t}$	$\hat{\beta}_{2t}$	$\hat{\beta}_{3t}$	$\hat{\tau}_t$	$\hat{\varepsilon}_t$	R^2
Panel 1. Non-Linearized Version of the Model (Time-varying τ_t)						
Mean	2.940	-2.759	-2.426	46.876	0.000	0.996
Std. Dev.	0.417	0.391	1.925	6.156	0.000	0.002
Maximum	3.805	-1.374	5.201	119.999	0.000	0.999
Minimum	1.219	-3.671	-4.676	19.348	0.000	0.987
Skewness	-1.566	0.891	1.420	1.530	0.068	-1.355
Kurtosis	6.690	4.943	5.156	4.996	2.796	6.116
$\hat{\rho}$ (1)	0.802	0.784	0.840	0.688	0.015	0.497
$\hat{\rho}$ (12)	0.055	-0.027	0.112	0.127	-0.067	-0.070
$\hat{\rho}$ (24)	-0.118	-0.355	-0.208	-0.128	-0.066	-0.040
ADF Stat.	-4.255	-4.147	-3.163	-5.129	-11.789	-
Panel 2. Linearized Version of the Model ($\tau = 38.068$)						
Mean	2.055	-2.989	-2.831	-	0.000	0.956
Std. Dev.	0.118	0.177	0.775	-	0.015	0.012
Maximum	3.094	-1.266	3.036	-	0.071	0.988
Minimum	1.124	-3.275	-3.395	-	-0.083	0.931
Skewness	0.614	0.388	-0.722	-	0.205	-0.893
Kurtosis	2.133	1.789	2.544	-	2.371	3.940
$\hat{\rho}$ (1)	0.866	0.857	0.860	-	0.215	0.436
$\hat{\rho}$ (12)	0.275	0.399	0.439	-	-0.178	0.014
$\hat{\rho}$ (24)	-0.168	-0.091	0.053	-	--0.125	-0.134
ADF Stat.	-3.355	-3.324	-3.297	-	-7.756	-

Note: The table presents descriptive statistics for Nelson-Siegel estimated factors, R^2 and $\hat{\varepsilon}$ averaged over the different maturity times using monthly yield data 2000:01–2011:12. Panel 1 presents the features of the results obtained from non-linearized version of the Nelson-Siegel model by applying non-linear least squares method, while panel 2 shows the features of the results estimated by ordinary least squares (OLS) methods for pre-specified value (median value obtained from non-linear estimation) of the shape parameter ($\hat{\tau} = 38.068$). $\hat{\rho}\ (i)$ denotes the sample autocorrelations at displacements of 1, 12, and 24 months. The last row contains augmented Dickey–Fuller (ADF) unit root test statistics. The number of observations is 144.

Furthermore, to empirically test whether the factors β_{1t}, β_{2t} and β_{3t} are legitimately called a level, slope and curvature factors respectively, as suggested in Diebold and Li (2006), we construct a level, slope and curvature from the observed zero-coupon yields data and compare them with $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$ (estimated with time-varying τ_t) respectively. The level of the yield curve (L_t) is defined as the 25-year yield. We compute the slope (S_t) as the difference between the 25-year and three-month yield and the curvature (C_t) is worked out as two times the two-year yield minus the sum of the 25-year and three month zero-coupon yields. The pairwise correlation of empirically defined factors and estimated (model based) factors is $\rho(L_t, \hat{\beta}_{1t}) = 0.694$, $\rho(S_t, \hat{\beta}_{2t}) = -0.741$ and $\rho(C_t, \hat{\beta}_{3t}) = 0.660$. Pairwise correlations between the estimated factors and the empirically defined level, slope and curvature is almost smaller by 0.28 points than the results of earlier empirical studies, particularly for the US and Canadian markets.[24] Furthermore, to be precise, the estimated correlation and the time series plot in figure 2.2 show that $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$ may truly be called level, slope and curvature factors respectively, as the estimated factors and their empirical proxies seem to follow the same pattern.

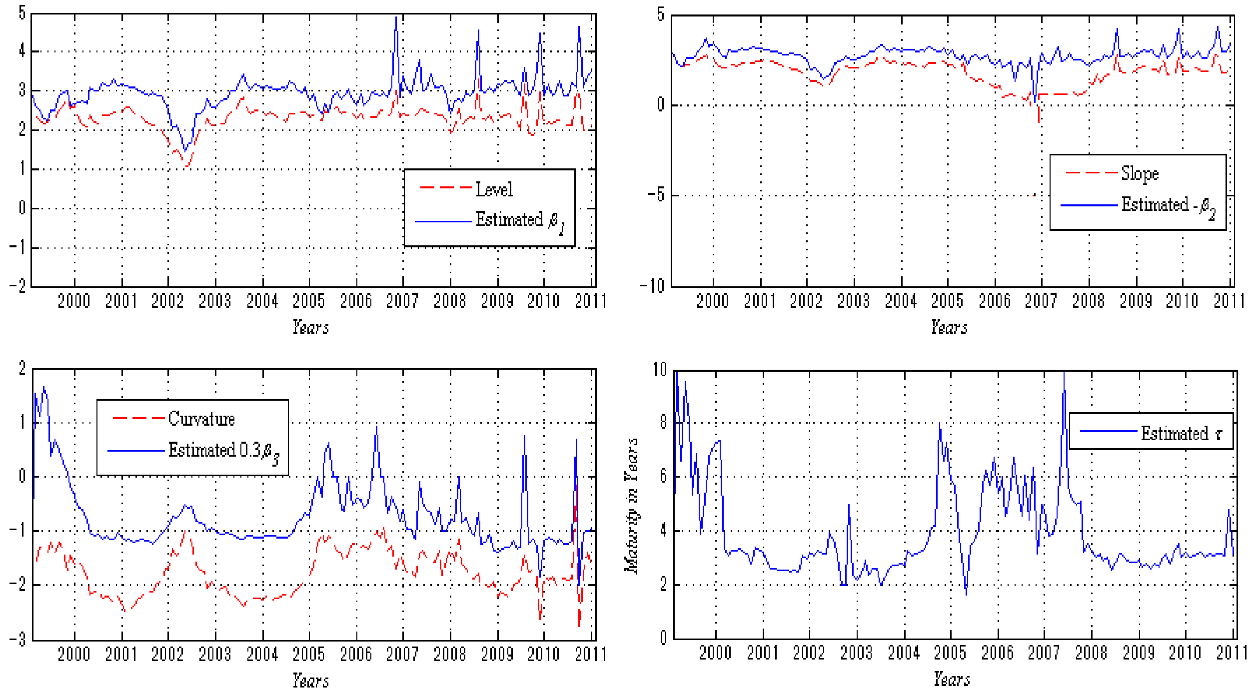


Figure 2.2: Time Series Plot of Nelson-Siegel Estimated Factors and Empirical Level, Slope and Curvature
Model-based level, slope and curvature (i.e., estimated factors vector $\hat{\beta}_t$) for time-varying τ_t vs. data-based level, slope and curvature are plotted, where level is defined as the 25-year yield, slope as the difference between the 25-year and 3-month yields and curvature as two times the 2-year yield minus the sum of the 25-years and 3- month zero-coupon yields. Rescaling of estimated factors is based on Diebold and Li (2006).

Using the estimates of Nelson-Siegel model for both time-varying and fixed τ, in figure 2.3, we plot the implied average fitted yield curves, the actual yield curve and the residuals. It seems that the curve fits pretty well and the two vary quite closely for time-varying τ. It does, however, have difficulties at some dates, especially when yields are dispersed, with multiple interior minima and maxima. For the fixed τ the discrepancy between the actual and estimated average yield curve is clearly visible. It under-estimates the actual yield up to 30 months maturity and over-estimates beyond 30 months. Similarly, the average residuals plot in the right panel of figure 2.3 also supports this argument.

[24] Diebold and Li (2006) perform a similar exercise based on zero-coupon yields generated using end-of-month price quotes for U.S. treasuries, from 1985:01 through 2000:12. Their estimated correlations are $\rho(L_t, \hat{\beta}_{1t}) = 0.97$, $\rho(S_t, \hat{\beta}_{2t}) = -0.99$ and $\rho(C_t, \hat{\beta}_{3t}) = 0.99$. Similarly, Elen van (2010) used the monthly Canadian zero-coupon yields from 1986:01 to 2009:012 and has reported the correlations as $\rho(L_t, \hat{\beta}_{1t}) = 0.943$, $\rho(S_t, \hat{\beta}_{2t}) = -0.929$ and $\rho(C_t, \hat{\beta}_{3t}) = 0.784$.

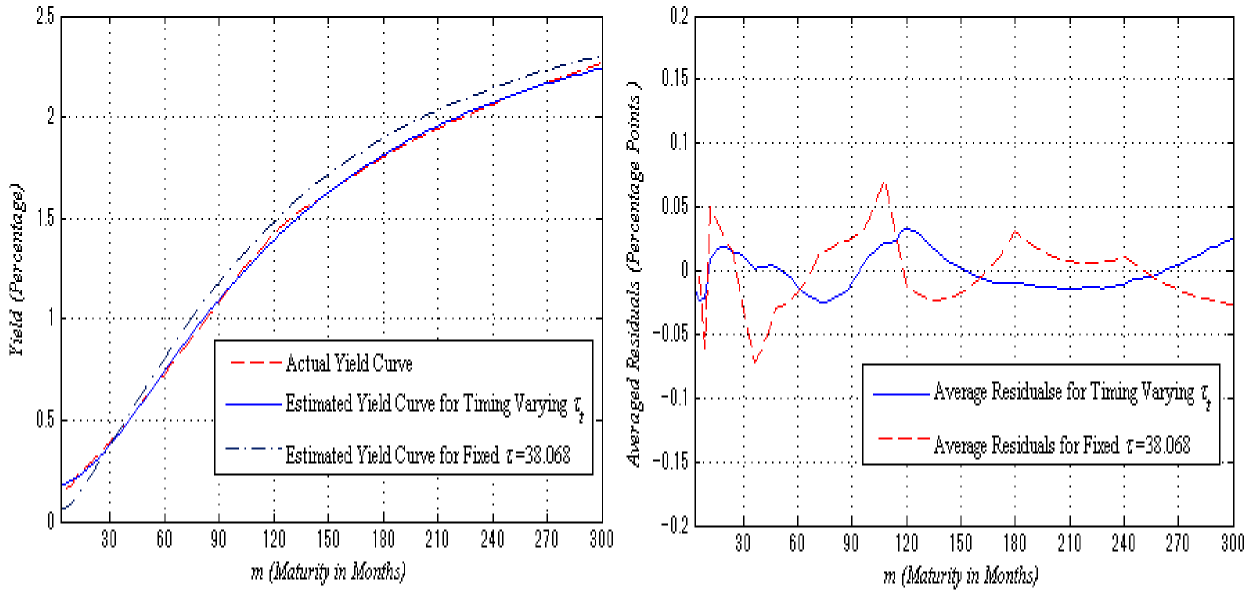


Figure 2.3: Average Fitted Yield Curve and Residuals of the Nelson–Siegel Model
Actual (data-based) and estimated (model-based) average yield curves and average residuals for both time-varying $\hat{\tau}_t$ and fixed $\hat{\tau} = 38.068$ are plotted. The fitted yield curves are obtained by taking average of the estimated yield of the Nelson-Siegel model over 144 months. Similarly, the residuals are also averaged over 144 months for the various maturities.

Table 2.3: Descriptive Statistic of the Nelson-Siegel Yield Curve Residuals for Time-varying τ

Maturity	Mean	S. Dev.	MAE	RMSE	Skewness	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
3	-0.004	0.023	0.022	0.023	0.278	1.370	0.740	0.229	0.034
6	-0.011	0.019	0.020	0.022	0.874	2.381	0.579	0.038	0.061
9	-0.009	0.017	0.017	0.020	0.769	2.508	0.671	-0.036	-0.119
12	0.002	0.018	0.015	0.018	-0.174	1.833	0.587	0.300	0.111
15	0.008	0.018	0.017	0.019	-0.569	1.983	0.610	0.332	0.155
18	0.010	0.018	0.019	0.021	-0.753	2.243	0.572	0.211	0.122
21	0.011	0.018	0.019	0.021	-0.799	2.298	0.650	0.079	0.200
24	0.008	0.018	0.017	0.020	-0.551	2.044	0.623	0.119	0.167
30	0.007	0.019	0.018	0.020	-0.395	1.786	0.694	0.227	-0.090
36	0.001	0.022	0.020	0.022	-0.034	1.379	0.806	0.373	-0.096
48	0.008	0.020	0.018	0.020	-0.096	1.463	0.775	0.300	0.017
60	-0.008	0.020	0.019	0.021	0.568	1.826	0.743	0.308	0.111
72	-0.014	0.018	0.020	0.022	1.088	2.841	0.756	0.156	-0.072
84	-0.006	0.021	0.020	0.022	0.462	1.643	0.855	0.152	0.077
96	0.004	0.024	0.023	0.024	-0.255	1.283	0.891	0.420	0.072
108	0.008	0.019	0.018	0.020	-0.492	1.814	0.758	0.519	0.161
120	0.016	0.016	0.021	0.023	-1.437	3.975	0.425	0.188	0.022
180	-0.003	0.021	0.019	0.021	0.272	1.543	0.869	0.435	0.174
240	-0.005	0.017	0.016	0.018	0.353	1.844	0.795	0.334	-0.004
300	0.005	0.025	0.024	0.025	-0.385	1.265	0.856	0.403	0.029

Note: The table presents summary statistics of the residuals $\hat{\varepsilon}$ for different maturity times of the Nelson–Siegel model using monthly yield data 2000:01–2011:12 for time-varying τ. MAE is mean absolute errors, RMSE is the root mean squared errors and $\hat{\rho}(i)$ denotes the sample autocorrelations at displacements of 1,12, and 24 months. The number of observations is 144.

Furthermore, table 2.3 and figure 2.4 present the descriptive statistics and the three dimensional plot of the residuals of Nelson-Siegel model estimation by non-linear least squares (for the time-varying τ_t) respectively. It turns out that the fit is more appealing in most cases. Some months, however, especially those with multiple maxima and/or minima are not fitted very well. Multiple maxima and/or minima occur in the term structure of months in the mid-2005 and onward, which becomes apparent by the large residuals in these months.

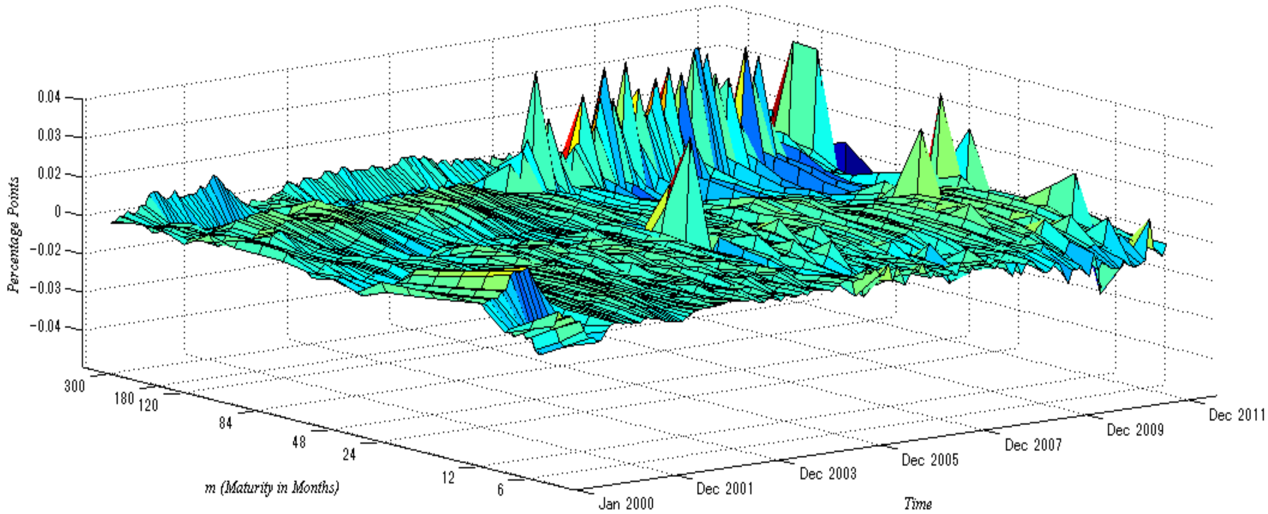


Figure 2.4: Nelson–Siegel Model based Yield Curves Residuals, 2000:01–2011:12 for Time-varying τ. The sample consists of monthly residuals, obtained from the non-linear least squares estimation of the Nelson-Siegel model using the data 2000:01–2011:12 (144 months), at fixed quarterly maturities of 3, 6, 9, 12, 15, 18,…300 months.

In summary, there is a significant lack of information on the short-term CIR model to fit the term structure of interest rate. It is not capable to fit the yield curve as the discrepancy between the two curves is significantly large. Contrarily, the Nelson-Siegel model provides an evolution of the term structure closer to reality. It distills the term structure of interest rate quite well and can describe the evolution and the trends of the market. Fixing the τ to the median value leads to fit the yield curve better than the CIR model but not than the time-varying τ estimation process (non-linear least squares) of the Nelson-Siegel model.

2.4. Term Structure Forecasting

A good approximation to yield-curve dynamics should not only fit well in-sample, but also produces satisfactorily out-of-sample forecasts. In this section, we simulate the interest rates to find out whether the simulated yields for various maturities based on the CIR and Nelson-Siegel models can replicate the stylized facts of the actual observed yields data. The stylized facts derived from the actual yields data for Japanese bonds are:

1. The average yield curve is upward sloping and concave.
2. Short rates are more volatile than long rates.
3. Long rates are less persistent than short rates.
4. Skewness has the downward trend with the maturity.
5. Kurtosis of the short rates are lower than those of the long rates.

2.4.1. Forecasting with the Cox-Ingersoll-Ross Model

Using the parameters in panel 1 of table 2.1, we simulate the short rates using the discrete version of CIR model as in (2.9) for 10,000 times. The starting point of the short rates simulation process is the two-year yield at December 2011, being 0.071. Using the simulated short rates, the entire term structure of yield is computed by using equation (2.8), that is, we compute 10,000 matrices of (144×100), containing yields for all maturity times and for all months.

Table 2.4 displays descriptive statistics that are of interest (e.g., mean, variance and autocorrelations) of the simulated yields for various maturities. This table may be compared with the statistical properties of actual yields in table 1.2 (section 1.4).

Summary statistics in table 2.4 indicate that the CIR model is not capable of replicating the interest rates' general trends. The CIR model generates the same skewness coefficients, the same kurtosis and the same autocorrelations for all maturity times. The short rates seem more volatile than the long rates, although the volatility is underestimated for all maturity times compared to the actual yields data. Moreover, the mean has a downward trend with increasing maturity. Figure 2.5 shows a plot of the downward shaped average yield curve (averaged over simulation times), implying that the simulated yield curve is not in line with the first stylized fact. The figure also shows that the CIR model is capable to produce term structure's other shapes.

Table 2.4: Descriptive Statistics of the Simulated Yields Using the CIR Model

Maturity	Mean	S. Dev.	Max	Min	Skewness	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
3	0.314	0.204	1.835	0.016	1.455	6.221	-0.008	-0.005	0.017
6	0.314	0.170	1.582	0.037	1.455	6.221	-0.008	-0.005	0.017
12	0.313	0.123	1.229	0.059	1.455	6.221	-0.008	-0.005	0.017
18	0.312	0.093	1.007	0.120	1.455	6.221	-0.008	-0.005	0.017
24	0.312	0.074	0.861	0.160	1.455	6.221	-0.008	-0.005	0.017
36	0.311	0.051	0.689	0.206	1.455	6.221	-0.008	-0.005	0.017
60	0.311	0.031	0.539	0.247	1.455	6.221	-0.008	-0.005	0.017
120	0.310	0.015	0.425	0.279	1.455	6.221	-0.008	-0.005	0.017
180	0.310	0.010	0.386	0.289	1.455	6.221	-0.008	-0.005	0.017
240	0.310	0.008	0.367	0.294	1.455	6.221	-0.008	-0.005	0.017
300	0.310	0.006	0.356	0.297	1.455	6.221	-0.008	-0.005	0.017

Note: The table shows descriptive statistics for simulated yields at different maturities for the CIR model. The entire term structure of yield is computed by the CIR yield curve model using the simulated short rates. The simulation exercise is done 10,000 times for 144 months. The last three columns contain the first, 12^{th} and 24^{th} order sample autocorrelation coefficients. The number of observations is 10,000.

One may conclude that the CIR model performs unsatisfactorily and seems not useful in the simulation based context. As opposed to the CIR model, the Nelson-Siegel model does not fall within the standard class of affine term structure models. Therefore, yields forecasts and their stylized facts simulated with the Nelson-Siegel model will likely be significantly different from the yields produced by the CIR model.

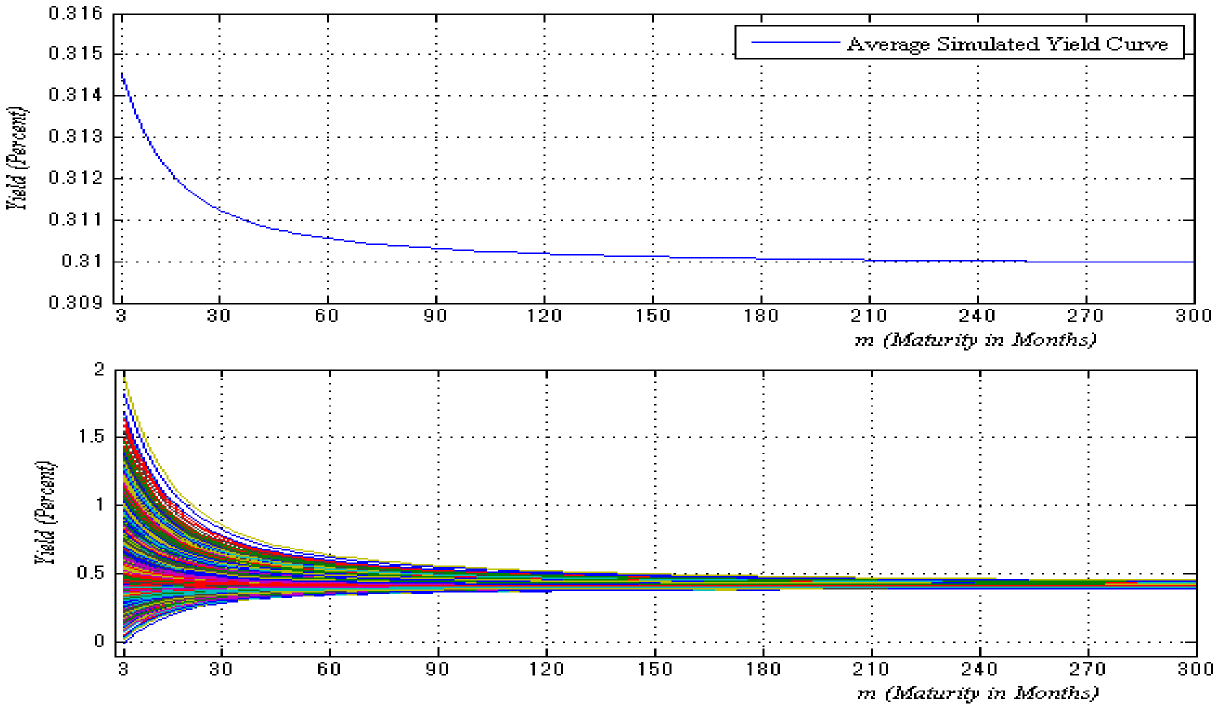


Figure 2.5:Average and All Simulated Yield Curves with the CIR Model
The entire term structure of yield is computed by the CIR yield curve model using the simulated short rates. The simulation exercise is done 10,000 times for 144 months. The 10000 simulated yield curves along with average simulated yield curve are plotted at fixed quarterly maturities of 3, 6, 9, 12, 15, 18, 21, 24 …300 months (100 maturities).

2.4.2. Forecasting with the Nelson-Siegel Model

Since the four parameters of the Nelson-Siegel model give a full description of the term structure of interest rate, one can model them and can use various methodologies to make out-of-sample forecast of the yield curve.[25] Here, the four time-varying estimated Nelson–Siegel factors are modeled as univariate AR(1) processes to simulate the term structure of interest rate.[26] The yield forecasts based on underlying univariate AR(1) factor specifications are:

$$\hat{R}_t(m) = \Lambda(\hat{\tau}_t)\hat{\beta}_t \tag{2.17}$$

$$\hat{\psi}_t = A_0 + A_1\hat{\psi}_{t-1} + \varepsilon_t \tag{2.18}$$

where A_0 is (4×1) vector of constants, A_1 is (4×4) diagonal matrix, $\hat{\psi}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\tau}_t)'$ and $\varepsilon_t \backsim N(0, \Sigma)$ is (4×1) error vector. For comparison, we also include the VAR(1) forecasts of yield because the pairwise correlation between estimated factors is reasonably high. This might produce out-of-sample forecasts with greater accuracy. The multivariate VAR(1) model specification is the same as in (2.18), but we modify A_1 to be (4×4) full matrix rather than a diagonal matrix.

[25] It is concluded in the previous section that the non-linear estimation (with time-varying τ) leads to a better fit of the yield curve; therefore, non-linear least squares estimated parameters are modeled to carry out the simulation exercise.

[26] Following Diebold and Li (2006), we also computed out-of-sample forecasts for one month, 6 months and 1 year. The summary results are given in appendix D for reference.

Estimation of AR(1) and VAR(1) models specified in (2.18) is straight forward. We estimate the parameters vector A_0 and matrix A_1 of both AR(1) and VAR(1) using the time series of $\hat{\psi}_t$ that we obtained from the non-linear least squares regression on (2.15) by employing the maximum likelihood method, assuming the normal density for ε_t. We use a forecasting period of ten years with a time step of one month. That is, we simulate 120 months, starting with the January 2012 until December 2021. Using the AR(1) and VAR (1) estimated parameters, we simulate the time series of size 120 months for 10,000 times.

Table 2.5 displays summary statistics of the four simulated factors for both AR(1) and VAR (1) specifications, averaged over number of simulations. This table may be compared with the actual estimated factors from table 2.2 (panel 1).

Comparing the simulated Nelson-Siegel factors of AR(1) and VAR(1) models with the estimated factors in panel 1 of table 2.2, the results show that in terms of most descriptive statistical properties, particularly the mean, skewness and kurtosis, the VAR(1) simulated factors and estimated factors are close alternatives. However, relatively the estimated factors are less persistent than the simulated factors. In terms of lag autocorrelation, the estimated factors are almost similar to the AR(1) results but regarding the mean and other descriptive features the AR(1) overestimates $\hat{\beta}_2$ and $\hat{\beta}_3$ and accurately estimates the $\hat{\beta}_1$ and $\hat{\tau}$.

Table 2.5: Descriptive Statistics of the Simulated Nelson-Siegel Factors

	AR(1)				VAR(1)			
	$\hat{\beta}_{1t}$	$\hat{\beta}_{2t}$	$\hat{\beta}_{3t}$	$\hat{\tau}_t$	$\hat{\beta}_{1t}$	$\hat{\beta}_{2t}$	$\hat{\beta}_{3t}$	$\hat{\tau}_t$
Mean	2.959	-2.799	-2.932	3.552	2.939	-2.753	-2.621	3.792
Std. Dev.	0.004	0.004	0.016	0.018	0.005	0.005	0.019	0.015
Maximum	2.966	-2.787	-2.892	3.596	2.953	-2.742	-2.583	3.824
Minimum	2.948	-2.807	-2.967	3.504	2.928	-2.764	-2.662	3.760
Skewness	-0.494	0.592	-0.101	-0.159	0.343	-0.178	-0.025	-0.015
Kurtosis	3.227	3.024	2.290	2.929	2.941	2.368	2.254	2.432
$\hat{\rho}$ (1)	0.786	0.359	0.115	-0.129	0.830	0.395	0.118	0.013
$\hat{\rho}$ (12)	0.881	0.453	0.123	-0.125	0.866	0.582	0.212	-0.292
$\hat{\rho}$ (24)	0.796	0.481	0.140	-0.070	0.622	0.333	0.016	-0.147

Note: The table presents descriptive statistics of the simulated Nelson-Siegel factors averaged over number of simulations for both AR(1) and VAR(1) specifications. The four factors of the Nelson-Siegel specification are modeled as first order AR and VAR to forecast the yield curve for 120 months, 2012:01–2021:12, for 10,000 times. The last three rows contain their first, 12th and 24th order sample autocorrelation coefficients. The computation of descriptive statistics is based on 120 observations.

Averaged over the number of simulations and the different months, both the simulated yield curves are upward sloping (figure 2.6). Comparing the simulated yield curves with the actual in figure 2.6, one notices the curves to be much alike. This may be attributed to the fact that the standard Nelson-Siegel is not only capable to generate a better in-sample fit but also performs satisfactorily in out-of-sample forecasts. At lower maturities, the VAR(1) simulated average yield curve is a bit nearer to the actual yield curve but at longer maturities both the VAR(1) and AR(1) are identical. Overall the results show that VAR(1) can replicate the properties of the estimated yield features better than the AR(1) specification.

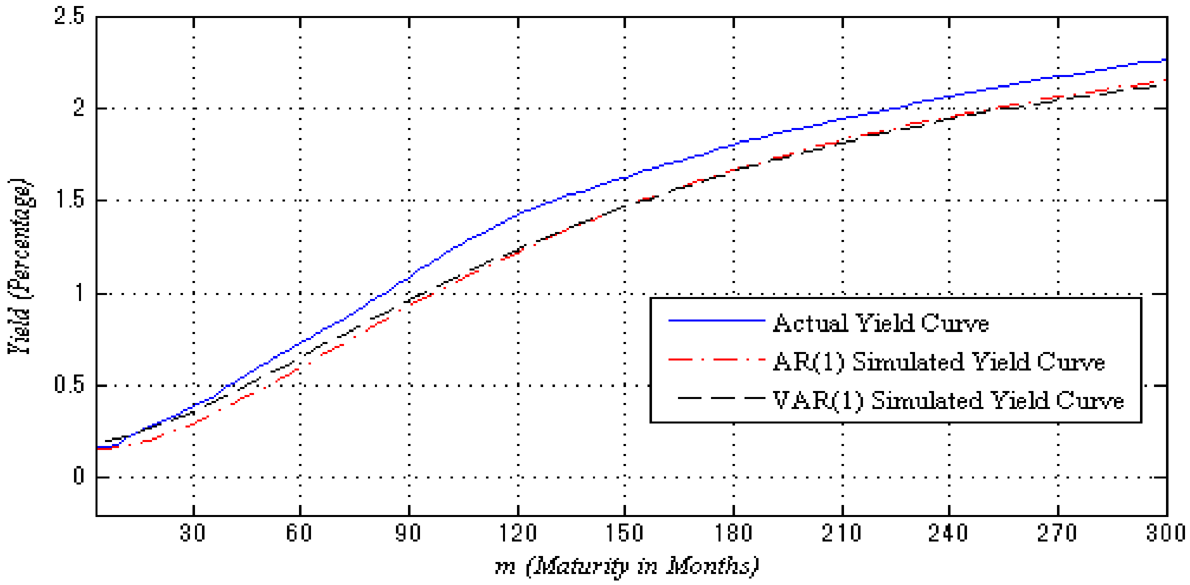


Figure 2.6: Simulated Average Yield Curves with the Nelson-Siegel Model
The four factors of the Nelson-Siegel specification are modeled as first order AR and VAR to forecast the yield curve for 120 months, 2012:01–2021:12, for 10,000 times. The average simulated yield curves for both AR(1) and VAR(1) specifications are obtained by averaging the simulated yields over different months as well as number of simulations. Actual (data-based) average yield curve is also plotted for comparison. All three yield curve are plotted at fixed quarterly maturities of 3, 6, 9, 12, 15, 18, 21, 24 …300 months (100 maturities).

To check for the other stylized facts, we compute yields for all maturities, by substituting the simulated vector $\hat{\psi}_t$ (at each simulation) in (2.17), for all 120 different months. Accordingly, we compute 10,000 matrices – one for each scenario (simulation) – of dimensions (120×100), containing the yields on every month for all maturity times. Then, we compute the statistical properties that are of interest (e.g., variances and autocorrelations) of the simulated yields for all maturities. Table 2.6 shows the descriptive statistics of the simulated yields for maturities of 3, 6, 12, 18, 24, 36, 60, 120, 180, 240, and 300 months for AR(1) and VAR(1) specifications, that can be compared to the actual yield statistical properties in table 1.2 (section 1.4).

Here, it can be seen that the simulated short rates of both AR(1) and VAR(1) indeed are more volatile than the long rates. It also seems that in simulation the skewness catches the downward trend with maturity in both cases. Moreover, kurtosis of the simulated short rates are lower than those of the simulated long rates, as can also be found in the observed nominal yields.

The numeric values of the average yield of actual yield data for various maturities resembles with the VAR (1) simulated yields. The volatilities of both AR(1) and VAR(1) are much smaller than the actual yield. The actual volatilities vary within the range of 0.207 and 0.348, whereas the simulated yields unconditional volatility of VAR(1) model vary between 0.002 and 0.004 and between 0.002 and 0.003 for AR(1) specification. Numeric values for the skewness coefficients and kurtosis, however, deviate from the observed yields. The Japanese data shows skewness coefficients between -1.934 and 1.360, the simulation shows values somewhere between -0.201 and -0.045 for AR(1) and between -0.500 and -0.086 for the VAR(1) model. Furthermore, the kurtosis ranging from approximately 2.079 and 8.291, while the simulation produces kurtosis ranging from roughly 1.970 up to 3.984 for both AR(1) and VAR (1) specifications. One may also deduce from table 2.6 that the simulated yield short rates are more persistent than the long rates as we observe in the nominal data.

Table 2.6: Descriptive Statistics of Simulated Yields Using the Nelson-Siegel Model

Maturity	Mean	S. Dev.	Max	Min	Skewness	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
Simulated Yields Descriptive Statistics for AR (1) Specification									
3	0.192	0.003	0.196	0.189	-0.045	1.979	0.866	-0.195	-0.175
6	0.202	0.003	0.205	0.199	0.012	2.042	0.851	-0.254	-0.130
12	0.229	0.003	0.233	0.225	-0.105	2.338	0.837	-0.235	-0.116
18	0.265	0.003	0.269	0.260	-0.232	2.441	0.837	-0.159	-0.147
24	0.310	0.003	0.314	0.304	-0.290	2.427	0.840	-0.100	-0.178
36	0.414	0.003	0.418	0.408	-0.307	2.361	0.833	-0.037	-0.211
60	0.655	0.003	0.660	0.649	-0.198	2.353	0.745	-0.020	-0.187
120	1.235	0.003	1.241	1.228	-0.183	2.843	0.448	-0.033	0.037
180	1.657	0.002	1.665	1.648	-0.339	3.059	0.558	0.048	0.017
240	1.939	0.002	1.949	1.930	-0.291	3.029	0.652	0.089	-0.006
300	2.129	0.002	2.140	2.119	-0.204	3.986	0.708	0.107	-0.013
Simulated Yields Descriptive Statistics for VAR(1) Specification									
3	0.158	0.004	0.161	0.154	-0.086	2.214	0.865	-0.041	-0.175
6	0.160	0.004	0.164	0.156	-0.407	2.475	0.875	0.008	-0.199
12	0.176	0.003	0.181	0.171	-0.557	2.659	0.891	0.075	-0.188
18	0.206	0.002	0.211	0.199	-0.502	2.567	0.900	0.108	-0.160
24	0.246	0.002	0.252	0.239	-0.425	2.429	0.905	0.125	-0.134
36	0.349	0.002	0.354	0.342	-0.292	2.182	0.903	0.137	-0.091
60	0.601	0.002	0.606	0.595	-0.115	2.027	0.849	0.121	-0.017
120	1.221	0.002	1.227	1.215	0.017	2.349	0.632	0.019	-0.038
180	1.664	0.002	1.671	1.658	0.025	2.312	0.629	0.001	-0.131
240	1.955	0.002	1.962	1.947	-0.095	2.430	0.657	0.001	-0.180
300	2.148	0.002	2.155	2.140	-0.252	2.628	0.684	0.010	-0.202

Note: The table shows descriptive statistics for monthly simulated yields at different maturities for both AR(1) and VAR(1) specifications of the four factors vector $\hat{\psi}_t$of the Nelson-Siegel Model. The four simulated factors are substituted in (2.17) to compute the simulated yields for various maturities for 120 months, 2012:01–2021:12, for 10,000 times. The last three columns contain the first, 12th and 24th order sample autocorrelation coefficients. The computation of descriptive statistics is based on 120 observations.

In summary, we conclude that the CIR model cannot replicate the interest rates' general trends and should be considered weak to describe the term structure in the simulation based context. On the other hand, the out-of-sample forecast results of the Nelson-Siegel seem reasonably well. In a simulation based context, the Nelson-Siegel model is capable to replicate most of the stylized facts of the Japanese market yield curve and the VAR(1) based specification of factors is able to replicate the properties of the estimated factors as well as actual yield data better than the AR (1) model of the factors.

2.5. Conclusion

The term structure of interest rates is the most important factor in the capital markets and probably the economy. It is widely used for pricing contingent claims, determining the cost of capital and managing financial risk. In this study, we implement the CIR and the Nelson-Siegel models and compare the in-sample fit as well as the out-of-sample forecast performance using monthly Japanese government bonds zero-coupon data (yield to maturity) from January 2000 until December 2011.

For the in-sample fit, the results show that there is a significant lack of information on the short-term CIR model. The CIR model plots upward sloping yield curve, however, the discrepancy between the actual and the estimated is an increasing function of maturity beyond two years maturity. Contrary to CIR model, the Nelson-Siegel model provides an evolution of the term structure closer to reality. The Nelson-Siegel model is capable to distill the term structure of interest rate quite well and describe the evolution and the trends of the market. Furthermore, fixing the shape parameter τ to the median value leads to a better yield curve fit than the CIR model but not as striking as the time-varying τ estimation process (non-linear least squares) does.

Regarding the term structure forecast, we conclude that the CIR model cannot accomplish to replicate the interest rates' general trends. The CIR model generates the same skewness, kurtosis and autocorrelations for all maturity times. The volatility is underestimated for all maturity times and more importantly, it produces a downward slope average yield curve, implying that CIR model should be considered too poor to describe the term structure evolution in the simulation based context. On the other hand, the out-of-sample forecast results of the Nelson-Siegel model seem reasonably well. The Nelson-Siegel model is capable to replicate most of the stylized facts of the Japanese market yield curve. Between the AR(1) and VAR(1) specification of factors, the descriptive features of the actual yield data and estimated factors are more closely in line with the VAR(1) simulated yields features.

Summarizing, it turns out that the model proposed by Nelson and Siegel (1987) is compatible to fit attractively the yield curve (in-sample fit) and accurately forecast the future yields for various maturities. These successes account for the continued popularity of statistical class of models and its use by central banks around the world. Furthermore, the Nelson-Siegel model (non-linear version) could be a good candidate to study the evolution of the yield curve in Japanese market.

Chapter 3

Term Structure Forecasting of Government Bond Yields with Latent and Macroeconomic Factors

Do Macroeconomic Factors Imply Better Out-of-Sample Forecasts?

3.1. Introduction

The trade-off between the in-sample fit of yield curve, that is obtained by employing statistical models without a reference to economic theory, and the lack of fit by economic models that do provide a basis for the underlying economic theory is one of the key features of the term structure of interest rate literature. Therefore, estimation and forecasting time series of the cross-section of yields have proven to be a challenging task.

The initial work on the fitting of yield curve has a strong theoretical foundation. It relies on the optimization behaviour of economic agent, using the dynamic stochastic general equilibrium (DSGE) framework. A model that forms the basis for this class of term structure models is the Vasicek (1977) model. The innovative feature of the Vasicek (1977) is that it models the interest rate as a mean reversion process. Other early contributions to the literature of equilibrium pricing include Cox *et al.* (1985), Dunn and Singleton (1986), Campbell (1996 and 1999) and more recently, Piazzesi and Schneider (2006). However, based on the underline economic theory, this approach delivers unsatisfactory results and suffers from the so called equity premium puzzle, lack of yield curve fitting and is incapable to accurately forecast the future interest rate term structure (Ullah, 2012).

Motivation for statistical models comes from the stylized facts that can be inferred from empirical analysis. Watching a film that shows the random evolution of the yield curves and forward curves over the past several decades reveals that this class of curves can be generated either by solution to differential equation or difference equation. Within the class of statistical models, more positive results have emerged recently based on the framework of Nelson and Siegel (1987). Originally intended to describe the cross sectional aspects of yield curve, Nelson and Siegel (1987) impose a parsimonious three-factor structure on the link between yields and different maturities. Diebold and Li (2006) find that the dynamic reformulation of this model provides forecasts that outperform the random walk and various alternative forecasting approaches. Whilst being statistical in nature it has the advantage that the components carry a clear economic interpretation. Various recent publications such as Diebold and Li (2006) and Diebold *et al.* (2006) have strengthened the importance of the Nelson-Siegel model and more importantly, Christensen *et al.* (2011) have derived the Nelson-Siegel framework in a standard affine term structure model.

The yield curve models that have theoretical foundation are developed mainly by macroeconomists, which focus on the role of expectations of inflation and future real economic activity in the determination of yield. On the other hand, the statistical yield curve models mainly focus on the shape and better fit of the yield curve and eschew any explicit role for such determinants. Many recent papers have also modeled the yield curve, and they can be categorized by the extent and nature of the linkages permitted between yield and macroeconomic variables. In this regard, the more related studies include Ang and Piazzesi (2003), Hördahl *et al*. (2006), Wu (2002) and Diebold *et al.* (2006), who explicitly incorporate macroeconomic determinants into multi-factors yield curve models. In these studies two hypotheses of yield and macroeconomic factors interaction— yields-to-macro and macro-to-yields links — are testable one. These studies focus on the existence of either unidirectional or bidirectional causality of yield curve and macroeconomy. However, the literature lacks the role of macroeconomic and financial market factors in the yield curve forecasting. This study takes a step toward bridging this gap by formulating and forecasting the yield curve that integrates macroeconomic and financial market factors in the yield curve model.

We examine the role of macroeconomic and stock market variables in the dynamic Nelson-Siegel framework for fitting and forecasting the term structure of interest rate on the Japanese government bond market. To assess the role of macroeconomic variables, we use a three-factor term structure model based on the classic contribution of Nelson and Siegel (1987), interpreted as a model of level, slope, and curvature. We explicitly incorporate three macroeconomic variables, i.e., the level of economic activity, exchange rate, and inflation rate and one stock market activity indicator (Stock Market Index) in the state-space representation of yield curve model to analyze their impacts in the in-sample fit and subsequently the efficiency gain in forecasting the yields for various maturities. It will be to get a clue about the role of macroeconomic variables in the yield curve dynamics and forecasting. The motivation to analyze the importance of macroeconomic and stock market indicators in forecasting the interest rates may be to examine the out-of-sample forecasts errors persistency.[27] The empirical results on the Japanese market indicate that inclusion of the two new factors of exchange rate and stock market indicator into the Nelson-Siegel model improves the out-of-sample forecast of the yield curve and reduces (or eliminates) auto-correlations in the forecast errors.

[27] Although, the studies that focus on the forecast performance of statistical class of models come with encouraging results, particularly in terms of lower RMSE than various standard benchmark forecasts, but these errors are highly persistent for most maturities and at various horizons (Bliss,1997; de Jong, 2000; Diebold and Li, 2006 and Ullah, 2012).

The remainder of the chapter is organized as follows. In the next section, we present the dynamic Nelson-Siegel model with and without macroeconomic factors (we call the former yields-macro model and the latter yields-only model) and explain the estimation method. The section 3.3 deals with the data structure and compare the in-sample fit results for the two competing models. In section 3.4, we present the out-of-sample forecast performance and the results of various tests to compare the forecast errors evolution over time and maturities. Finally, the section 3.5 presents the conclusion.

3.2. Term Structure Models and Estimation Method

At a certain point of time, the yield curve is the paired set of yields of zero-coupon treasury securities and maturity. In practice, the central banks issue a limited number of securities with different maturities and coupons; therefore, obtaining the yield curve at each moment requires estimation, i.e., inferring what the zero-coupon yields would be across the whole maturity spectrum. Yield curve estimation requires the assumption of some model for the shape of the yield curve, so that the gaps may be filled in by analogy with the yields seen in the observed maturities. Once a model is selected, estimates of its coefficients are chosen, so that the weighted sum of the squared deviations between the actual prices of treasury securities and their predicted prices is minimized.

At calendar time t, for a zero-coupon bond with unit face value maturing in m periods with the current price $P_t(m)$, the continuously compounded yield $R_t(m)$ is $P_t(m) = \exp[-R_t(m)m]$. The instantaneous forward rate $f_t(m)$, which is the interest rate contracted now and to be paid for a future investment, is given by $f_t(m) = -[P_t'(m)/P_t(m)]$ or correspondingly the zero-coupon yield is $R_t(m) = m^{-1} \int_0^m f_t(u)du$, which implies that the zero-coupon yield is an equally-weighted average of instantaneous forward rates.

In the next two subsections, we present yields-only spot rate model and yields-macro model (extended model) that incorporates macroeconomic as well as stock market variables in the standard yield curve model in state-space framework.

3.2.1. Yields-Only Factors Model (Yield Curve Model without Macroeconomic Factors)

The class of curves first proposed by Nelson-Siegel (1987) does well in capturing the overall shape of the yield curve and is being popular among practitioners and central banks alike. They modeled the forward rates with the three-component exponential approximation to the cross-section of yields as a function of maturity m at any moment in time t as:[28]

$$f_t(m_i) = \beta_{1t} + \beta_{2t}\exp\left(\frac{-m_i}{\tau}\right) + \beta_{3t}\left[\left(\frac{m_i}{\tau}\right)\exp\left(\frac{-m_i}{\tau}\right)\right] \tag{3.1}$$

with the time-varying parameter vector $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ and time invariant parameter τ.

[28] These types of exponential functions to fit and forecast the observed yield curve become popular as they reconcile the following characteristics: (i) sufficiently flexible to reflect the important and typical patterns of the observed market data (ii) relatively robust against disturbances from individual observations (iii) applicable with only a few observations and (iv) results in more stable yield curves.

The forward rate representation chosen by Nelson-Siegel belongs to a class of Laguerre functions. These functions are characterized by a polynomial times a decaying exponential term. The use of Laguerre functions is a well-known approximation procedure. The solution for the yield as a function of maturity m can be found by integrating (3.1), resulting in:

$$R_t(m_i) = \beta_{1t} + \beta_{2t}\left[\frac{1-\exp(-m_i/\tau)}{m_i/\tau}\right] + \beta_{3t}\left[\frac{1-\exp(-m_i/\tau)}{m_i/\tau} - \exp\left(\frac{-m_i}{\tau}\right)\right] + \varepsilon_t \tag{3.2}$$

for $i = 1,2,\dots,N$ and $t = 1,2,\dots,T$.

The Nelson-Siegel specification in (3.2) can generate several shapes of the yield curve including upward sloping, downward sloping and (inverse) hump shaped with no more than one maxima or minima.

In Nelson-Siegel framework as in (3.2), β_{1t} may be interpreted as the overall level of the yield curve, as its loading is constant for all maturities; β_{2t} has a maximum loading (equal to 1) at the shortest maturity, which then monotonically decays through zero as maturity increases; β_{3t} has a loading that is null at the shortest maturity, increases until an intermediate maturity and then falls back to zero in the limit. Hence, β_{2t} and β_{3t} may be interpreted as the short end and medium-term latent components of the yield curve respectively, because shocks in β_{2t} predominantly affect only short end of yield curve and thus induce variations in yield spreads and shocks in β_{3t} dominantly affect the yield curve's curvature. The parameter τ is ruling the rate of decay of the loading towards the short-term factor and specifies the maturity where the medium-term factor has maximum loading. It also identifies the location of the hump or the U-shape on the yield curve. Therefore, the range of shapes the curve can take is dependent on τ, it can be interpreted as the shape parameter.

Here, we assume that the three time-varying latent factors β_{1t}, β_{2t} and β_{3t} follow a vector autoregressive process of first order, which allows us to formulate the yield curve latent factors model in state-space form and to use the Kalman filter for obtaining maximum-likelihood estimates of the hyper-parameters and the implied estimates of the time-varying latent factors vector β_t.

The state-space form comprises the measurement system, relating a set of observed zero-coupon yields of N distinct maturities to the three latent factors as:

$$\begin{bmatrix} R_t(m_1) \\ R_t(m_2) \\ \vdots \\ R_t(m_N) \end{bmatrix} = \begin{bmatrix} 1 & \frac{1-e^{-m_1/\tau}}{m_1/\tau} & \frac{1-e^{-m_1/\tau}}{m_1/\tau} - e^{-m_1/\tau} \\ 1 & \frac{1-e^{-m_2/\tau}}{m_2/\tau} & \frac{1-e^{-m_2/\tau}}{m_2/\tau} - e^{-m_2/\tau} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-m_N/\tau}}{m_N/\tau} & \frac{1-e^{-m_N/\tau}}{m_N/\tau} - e^{-m_N/\tau} \end{bmatrix} \begin{bmatrix} \beta_{1t} \\ \beta_{2t} \\ \beta_{3t} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix} \tag{3.3}$$

where $t = 1,2,\dots,T$, and ε_t is (N×1) vector of measurement errors, i.e., deviations of the observed yields in period t for each maturity m from the implied yields defined by the shape of the fitted yield curve.

If one is interested in fitting the term structure then the measurement equations are sufficient. However, in order to construct term structure forecasts we also need a model for the factors dynamics. We follow the dynamic framework of Diebold *et al.* (2006) by specifying first-order vector autoregressive processes for the factors. The state-space form of the model comprises the state system as:

$$\begin{bmatrix}\beta_{1t}-\mu_1\\ \beta_{2t}-\mu_2\\ \beta_{3t}-\mu_3\end{bmatrix}=\begin{bmatrix}A_{11} & A_{12} & A_{13}\\ A_{21} & A_{22} & A_{23}\\ A_{31} & A_{32} & A_{33}\end{bmatrix}\begin{bmatrix}\beta_{1,t-1}-\mu_1\\ \beta_{2,t-1}-\mu_2\\ \beta_{3,t-1}-\mu_3\end{bmatrix}+\begin{bmatrix}\eta_{1t}\\ \eta_{2t}\\ \eta_{3t}\end{bmatrix} \tag{3.4}$$

where μ_1, μ_2 and μ_3 are the mean values of the three latent factors, and η_{1t}, η_{2t} and η_{3t} are innovations to the autoregressive processes of the latent factors. In order to simplify the mathematical computation and notations, the state-space form of the model can be written as:

$$R_t(m)=\Lambda(\tau)\beta_t+\varepsilon_t \tag{3.5}$$

$$\xi_t=A\xi_{t-1}+\eta_t \tag{3.6}$$

The measurement equation in (3.5) specify the vector of yields, which contains N different maturities, $R_t(m)=[R_t(m_1)\dots\dots R_t(m_N)]'$, as the sum of a Nelson-Siegel spot rate curve $\Lambda(\tau)$ plus a vector of yield errors which are assumed to be independent across maturities but with different variance terms $\sigma^2(m_i)$. Furthermore, $\xi_t=(\beta_{1t}-\mu_1,\beta_{2t}-\mu_2,\beta_{3t}-\mu_3)'$ being the (3×1) vector of factors, matrix A is (3×3) and $\Lambda(\tau)$ is the (N×3) matrix of factors loadings which are potentially time-varying if the shape parameter τ is estimated alongside the factors.

For the Kalman filter to be the optimal linear filter, it is assumed that the innovations of both observation and state vectors are orthogonal to initial state: $E(\xi_0\eta_t')=0$ and $E(\xi_0\varepsilon_t')=0$. Lastly, we assume that the innovations of the measurement and of the transition systems are white noise, mutually uncorrelated and have Gaussian distribution.

$$\begin{bmatrix}\varepsilon_t\\ \eta_t\end{bmatrix}\sim N\left(\begin{bmatrix}0\\0\end{bmatrix},\begin{bmatrix}\Omega & 0\\ 0 & \Sigma\end{bmatrix}\right) \tag{3.7}$$

where Σ is (3×3), the covariance matrix of innovations of the transition system and is assumed to be unrestricted, while the covariance matrix Ω of the innovations to the measurement system of (N×N) dimension is assumed to be diagonal. The latter assumption means that the deviations of the observed yields from those implied by the fitted yield curve are uncorrelated across maturities and time. Given the large number of observed yields used, the diagonality assumption of covariance matrix of the measurement errors is necessary for computational tractability. Moreover, it is also a quite standard assumption, as for example, *iid* errors are typically added to observed yields in estimating no-arbitrage term structure models. The assumption of an unrestricted Σ matrix, which is potentially non-diagonal, allows the shocks to the three term structure factors to be correlated.

3.2.2. Yields-Macro Factors Model (Yield Curve Model with Macroeconomic Factors)

Given the ability of the estimated factors of the Nelson-Siegel model to provide a good representation of the yield curve for the Japanese market data (Ullah, 2012), it is of immense interest to relate the Nelson-Siegel factors to macroeconomic and equity market variables and analyze the dynamic interaction among them and the efficiency gain in forecasting the yields for various maturities.

The link between the level of the yield curve and inflationary expectations, as suggested by the Fisher equation, is a common theme in the recent macro-finance literature, including Kozicki and Tinsley (2001), Dewachter and Lyrio (2002), Hördahl *et al.* (2006) and Rudebusch and Wu (2003). According to Fisher's theory, the nominal rate has a one-to-one relationship with the expected inflation. Therefore, the term structure could be a predictor for future inflation. An increase in the long-term interest rate will be interpreted as a rise in inflation expectations and vice versa. The central banks around the world that have an implicit inflation target, can affect the short end of yield curve (rising short rates) in order to lower inflation expectations of the market and to indirectly influence the long end of the yield curve (Schich, 1999).

Regarding the economic growth, the yield curve is also widely used for understanding investors' collective sentiments about the future condition of the economy. The relation between the term spread (slope of yield curve) and economic activity may be that the term spread reflects the stance of monetary policy. If the policy makers raise short-term interest rates, long-term rates are usually not increasing one-to-one with them but slightly less. Hence, the spread tightens and even might become negative. Higher interest rates slow down overall spending and economic growth will stagnate. Therefore, a small or negative slope of the yield curve will be an indication for slower growing economy in the future.

The uncovered interest rate parity relationship forms the basis of the interaction between exchange rate and yield curve and to describe the relation between short and long-term interest rates. The effect of monetary policy actions on the exchange rate mainly depends on how the long-term rate reacts to this change. If the central bank raises interest rate and the long end shifts upwards as well, the domestic currency appreciates. In case that the long-term rate moves sideways, the higher short-term rate will cause the domestic currency to depreciate (Inci and Lu, 2004; Clostermann and Schnatz, 2000 and Byeon and Ogaki, 1999).

Furthermore, if the yield curve can predict the economy, it should be of some use in gauging the overall risk/reward potential of the stock market as well. That is because both corporate profits and stock prices depend heavily on the strength of the economy. So, if the economy is likely to improve, so shall corporate profits and stock prices. However, there is no guarantee that stocks will do well during periods when the yield curve has a normal positive slope, but recent research does suggest that the risk/reward trade-off for stocks is much better during periods when the yield curve is positively sloped.

Thus, the term structure includes significant amount of information about the market's expectations of future inflation, exchange rate, economic growth and state of equity market as suggested by the recent macro-finance literature mentioned above, it will be interesting to analyze its role in the in-sample fit and out-of-sample forecast performance of the yield curve. In line with the arguments of the studies that show the dynamic interaction of yield and macroeconomic factors, we expect that yield curve level factor has strong correlation with the exchange rate and inflation level, while the spread and curvature factors are related to the overall economic activity measures and risk premium of stocks. However, Diebold *et al.* (2006) report negligible responses of macroeconomic variables to shocks in the curvature factor, but conversely, Monch (2006) argues that flattening of the yield curve is associated with the changes in the curvature factor and can be linked to an economic slowdown.

To assess the role of macroeconomic and financial variables in the yield curve dynamics and forecasting, it can be done readily in an expanded version of the state-space framework of yields-only model. Regarding the macroeconomic variable, we include three key variables: the annual growth rate of industrial production (IP_t), real exchange rate (EX_t) (¥/$) and annual price inflation (INF_t). These variables represent, respectively, the level of real economic activity, foreign market competitiveness and the inflation rate, which are widely considered to be the minimum set of fundamentals needed to capture basic macroeconomic dynamics. As far as the stock market is concerned, the annual growth rate of stock market aggregate index (SI_t) is considered in the model as an indicator of the capital market performance. Though, the stock market aggregate index is an equity market indicator, in this study we call all the four variables $(IP_t, EX_t, INF_t, SI_t)$ as macroeconomic variable for the ease of interpretation and writing.

A straightforward extension of the yields-only model adds the four macroeconomic factors to the set of state equations, which leads to the following system of equations.

$$\begin{bmatrix} R_t(m) \\ Z_t \end{bmatrix} = \begin{bmatrix} \Lambda(\tau) & 0 \\ 0 & I_4 \end{bmatrix} \begin{bmatrix} \beta_t \\ \tilde{Z}_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} \tag{3.8}$$

$$\xi_t = A\xi_{t-1} + \eta_t \tag{3.9}$$

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma \end{bmatrix} \right) \tag{3.10}$$

where $\xi_t = \left(\beta_{1t} - \mu_1, \beta_{2t} - \mu_2, \beta_{3t} - \mu_3, \widetilde{IP}_t - \mu_4, \widetilde{EX}_t - \mu_5, \widetilde{INF}_t - \mu_6, \widetilde{SI}_t - \mu_7\right)'$ is the (7×1) vector of yield curve and macroeconomic factors, $Z_t = (IP_t, EX_t, INF_t, SI_t)'$ is the (4×1) vector of macroeconomic factors, A and Σ are (7×7) matrices, μ is (7×1) mean vector of factors and I_4 is (4×4) identity matrix. The dimension of $R_t(m)$, $\Lambda(\tau)$ and Ω are same as in yields-only model.

This system forms our yields-macro model, to which we will compare our earlier yields-only model. Our baseline yields-macro model continues to assume a non-diagonal Σ matrix and a diagonal Ω matrix. It is worth noting that the signal equation of yield curve [represented by the first equation in (3.8)] implies no change from the previous version of the model, recognizing the fact that the yield curve is fully described by the three latent factors (level, slope, and curvature) and to ensure that $\hat{R}_t(m)$ is positive semi-definite. The inclusion of macroeconomic variables in the signal equation does not guarantee that $\hat{R}_t(m) \geq 0$ and may imply negative yield at short maturities.

Furthermore, in (3.7) and in (3.10), we assume that the innovations of both, the measurement equation ε_t as well as the transition system η_t, are normally distributed. While real data are never exactly multivariate normal, the normal density is often a useful approximation to the true population distribution. Additionally, the multivariate normal density is mathematically tractable and nice results can be obtained. Moreover, the distribution of many multivariate statistics is approximately normal, regardless of the form of the parent population because of the central limit theorem.

3.2.3. Estimation Method

There are several approaches to estimate the latent factors and parameters in the Nelson-Siegel model. These approaches depend on whether the measurement and state equations are estimated separately or simultaneously and on the assumptions regarding the shape parameter.

The most straightforward approach is the two-step procedure as used by Fabozzi *et al.* (2005) and Diebold and Li (2006). In the first step, the measurement equations are treated as a cross-sectional model and Least Squares method is used to estimate the parameters for every period separately. In the second step, time series models are specified and fitted for the factors. The alternative to the two-step approach is to estimate all parameters simultaneously. This approach uses the Kalman filter to estimate the factors.

We consider the dynamic Nelson-Siegel model in (3.5–3.6) and (3.8–3.9) as linear Gaussian state-space models. The state vector of unobserved factors ξ_t can be estimated conditional on the past and current observations $R_1, R_2, \dots, R_t$ via the Kalman filter. Defining $\hat{\xi}_{t|s}$ as the minimum mean square linear estimator (MMSLE) of ξ_t given $R_1, R_2, \dots, R_s$ with mean square error (MSE) matrix $W_{t|s}$, for $s = t - 1$. For given values of $\hat{\xi}_{t|t-1}$ and $W_{t|t-1}$, the Kalman filter first computes $\hat{\xi}_{t|t}$ and $W_{t|t}$, when observation R_t becomes available, using the filtering step as:

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + W_{t|t-1}\Lambda(\tau)' F_{t|t-1}^{-1} v_{t|t-1} \quad (3.11)$$

$$W_{t|t} = W_{t|t-1} - W_{t|t-1}\Lambda(\tau)' F_{t|t-1}^{-1} \Lambda(\tau) W_{t|t-1} \quad (3.12)$$

with

$$v_{t|t-1} = R_t - \Lambda(\tau)\hat{\beta}_{t|t-1} \quad (3.13)$$

$$F_{t|t-1} = \Lambda(\tau) W_{t|t-1} \Lambda(\tau)' + \Omega \quad (3.14)$$

where $v_{t|t-1}$ is the prediction error vector and $F_{t|t-1}$ is the prediction error covariance matrix. The MMSLE of the state vector for the next period $t + 1$, conditional on $R_1, R_2, \dots, R_t$, is given by the prediction step as:

$$\hat{\xi}_{t+1|t} = A\hat{\xi}_{t|t} \tag{3.15}$$
$$W_{t+1|t} = AW_{t|t}A' + \Sigma \tag{3.16}$$

For a given time series of $R_1, R_2, \ldots, R_t$, the Kalman filter computations are carried out recursively for $t = 1,2, \ldots, T$ with initializations $\hat{\xi}_{1|0} = \mu$ (the unconditional mean) and $W_{1|0} = \Pi$, where Π is the covariance matrix of ξ_t as we assume that $\xi_t \sim N(\mu, \Pi)$.

An attractive feature of models in state-space form is that they can allow obtaining smooth optimal extractions of the latent level, slope and curvature factors. The smoothing algorithm associated with the Kalman filter produces the smoothed estimates of the latent vector for all periods and is based on the all available observations in the dataset. The estimation procedure itself does not change depending on data availability. Moreover, the smoothed estimates of the factors do also generate smoothed estimates of the interest rates and corresponding residuals for all maturities. This property ranks it among the most popular term structure estimation methods.

The smoothed estimates of state vector can be calculated as follows. First, we run the data through the Kalman filter, storing the sequences $W_{t|t}$ and $W_{t+1|t}$ as calculated in (3.12) and (3.16) and storing $\hat{\xi}_{t|t}$ and $\hat{\xi}_{t+1|t}$ as obtained in (3.11) and (3.15) respectively for $t = 1,2, \ldots, T$. The terminal value for $\hat{\xi}_{t|t}$ then gives the smoothed estimates for the last date in the sample $\hat{\xi}_{T|T}$ and $W_{T|\mathrm{T}}$ is its covariance matrix. The sequence of smoothed estimates $\hat{\xi}_{t|T}$ is then calculated in reversed order by iterating on:

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\left(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t}\right) \tag{3.17}$$

for $t = T-1, T-2, \ldots, 1$. The corresponding covariance matrix is similarly found by iterating on:

$$W_{t|T} = W_{t|t} - \left(W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\right)\left(W_{t+1|T} - W_{t+1|t}\right)\left(W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\right)' \tag{3.18}$$

in reverse order for $t = T-1, T-2, \ldots, 1$.

The parameters in the VAR(1), the constants vector μ, coefficients matrix A and both the covariance matrices (Ω and Σ) along with the shape parameter τ are treated as unknown coefficients, which are collected in the parameter vector θ. Estimation of θ is based on the numerical maximization of the log-likelihood function that is constructed via the prediction error decomposition and given by:

$$\log L(\theta) = -\frac{NT}{2}\log(2\pi) - \frac{1}{2}\sum_t \log\left[\left|F_{t|t-1}(\theta)\right|\right] - \frac{1}{2}\sum_t \left[v_{t|t-1}'\left[F_{t|t-1}(\theta)\right]^{-1} v_{t|t-1}\right] \tag{3.19}$$

The specification in (3.19) is a function of the parameter set $\theta = (\tau, \mu, A, \Omega, \Sigma)$. The likelihood is comprised of the (N×1) yield prediction error vector; $v_{t|t-1} = R_t - \hat{R}_{t|t-1}$, where $\hat{R}_{t|t-1}$ is the vector of in-sample yield forecasts given information up to time $t-1$, and of the (N×N) conditional covariance matrix of the prediction errors $F_{t|t-1}$.[29] The shape parameter τ is assumed to be constant over time. The log likelihood $\log L(\theta)$ in (3.19) can be evaluated by the Kalman filter for a given value of θ. Marquardt non-linear optimization algorithm is employed for numerical maximization of the log-likelihood function in (3.19).

3.3. Empirical Results

Taking into account the three dimensions of data – yield, time to maturity and calendar time –in this study, we follow the one-step procedure to estimate and forecast the yield curve dynamics. In general, state-space representation provides a powerful framework for analysis and estimation of dynamic models. In addition, the one-step Kalman filter approach is preferable to the two-step approach because the simultaneous estimation of all parameters produces correct inference via standard theory. This innovative feature grades it among the widely held term structure estimation methods. The two-step procedure, in contrast, suffers from the fact that the parameters estimation and signal extraction uncertainty associated with the first step is not acknowledged in the second step.

3.3.1. Data

The data we use are monthly spot rates for zero-coupon and coupon bearing bonds, generated using pricing data of Japanese government bonds and treasury bills. We use end-of-month price quotes (bid-ask average) for Japanese government bonds, from January 2000 to December 2011, taken from the Japan Securities Dealers Association (JSDA) bonds files. In total, there are 144 months in the dataset. Following Fama and Bliss (1987) method, we calculate the daily continuously compounded forward rates and subsequently convert them to spot rates using (1.4) for the fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240 and 300 months (20 maturities).

Concerning the macroeconomic variables, we use monthly data from January 2000 to December 2011, for industrial production, real exchange rate, consumer price index and Tokyo Stock Exchange share prices index (TOPIX) for the Japanese economy. The IP_t is growth rate in industrial production, EX_t is the growth in real exchange rate (¥/$), INF_t is the inflation rate and is measured as 12-months percent change in the consumer price index, and SI_t is last 12-months growth rate of TOPIX. Detailed description of both, yield data and macroeconomic variables along with the three dimensional plot of the monthly spot rates for various maturities is given in section 1.4.

3.3.2. Estimation of the Models

We apply the Kalman filter to the state-space representation for yields-only model (3.5–3.7) and yields-macro model (3.8–3.10) to compute optimal yields predictions and the corresponding

[29] see Kim and Nelson (1999) for further details.

prediction errors, after which we proceed to evaluate the Gaussian likelihood function using the prediction-error decomposition of the likelihood. The Kalman filter is initialized using the unconditional mean (zero) and unconditional covariance matrix of the state vector, which are derived from the Gaussian distribution and assuming that the innovations of both signal and state equations are normally distributed.[30] The log-likelihood function as specified in (3.19) is maximized by iterating the Marquardt algorithm, using numerical derivatives. The non-negativity condition is imposed on all estimated variances (diagonal elements of all covariance matrices) by estimating log-variances and subsequently converted to variances by exponentiating and then asymptotic standard errors are computed using the delta method. As the Kalman filter algorithm is sensitive to the initializing values of parameters, we use the two-step method of Diebold and Li (2006). In the first step, we use the non-linear least squares method to estimate the measurement equation and obtain time series of β_t and τ_t and subsequently use the estimated β_t vector to compute startup parameter values (initial transition equation matrix). Furthermore, we initialize all variances at 1.0 and τ at 38.068 (the median value) given in Ullah (2012).

We present estimation results of vector μ and matrix A for the yields-macro model in the first panel of table 3.1, while in second panel for the yields-only model. The results show that the estimated vector μ is highly statistically significant for both the models as the estimated standard errors are sufficiently small, compared to the estimated coefficients.[31] The estimate of the matrix A indicates highly persistent own dynamics of β_{1t}, β_{2t} and β_{3t} with estimated own-lag coefficients of 0.911, 0.928 and 0.904 for the yields-macro model, whereas 0.903, 0.901 and 0.866 for yields-only model, respectively. Cross factor dynamics of yield factors appear unimportant, with the exception of a minor but statistically significant effect of β_{2t-1} on β_{1t} in both models. Furthermore, for the yields-macro model the estimates of the effect of macro-factors on yield curve factors are small in magnitude as compared to the effect of yield curve factors on macroeconomic variables, but statistically significant and consistent with the yield-macroeconomic dynamics literature. The results in first panel show that industrial production and exchange rate are positively while the inflation rate is negatively related to the overall yield level. The most important result is that of statistically significant relationship of overall economic activity (represented by growth rate of industrial production) and the extent of stock market activity with the yield curve slope factor. This suggests that yield curve spread has a consistent predictive power of the future state of overall economic activity and stock market performance. Furthermore, this negative relationship is consistent with the idea that during recessions, premia on long-term bonds tend to be high and yields on short bonds tend to be low. Hence, during recessions, upward sloping yield curves not only indicate bad times today, but better times tomorrow. Moreover, the exchange rate has a positive statistically significant effect on the yield curve curvature.

[30] For detail of initializing the Kalman filter, see Hamilton (1994).

[31] The p-value for all intercept terms are smaller than 0.020.

Table 3.1: Latent Factors VAR(1) Model Parameters Estimates

	μ	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	IP_{t-1}	EX_{t-1}	INF_{t-1}	SI_{t-1}
Panel 1: Yields-Macro Model								
β_{1t}	**2.997**	**0.911**	**0.015**	0.018	**0.012**	**0.007**	**-0.011**	-0.008
	(0.157)	(0.012)	(0.005)	(0.012)	(0.001)	(0.002)	(0.005)	(0.016)
β_{2t}	**-2.855**	0.139	**0.928**	-0.001	**-0.015**	-0.036	-0.009	**-0.013**
	(0.234)	(0.189)	(0.155)	(0.029)	(0.004)	(0.125)	(0.024)	(0.003)
β_{3t}	**-2.866**	-0.260	-0.196	**0.904**	-0.055	**0.010**	0.006	-0.010
	(0.437)	(0.428)	(0.257)	(0.063)	(0.246)	(0.002)	(0.025)	(0.056)
IP_t	**1.061**	-0.412	**-0.511**	0.474	**0.413**	0.041	**1.761**	**0.345**
	(0.414)	(0.844)	(0.201)	(0.766)	(0.075)	(0.046)	(0.728)	(0.032)
EX_t	**3.766**	**-0.006**	0.233	-0.487	-0.033	**0.579**	-0.633	0.017
	(1.301)	(0.002)	(0.531)	(0.717)	(0.033)	(0.020)	(0.558)	(0.014)
INF_t	**-0.006**	**-0.488**	-0.123	**-0.286**	-0.001	0.201	**0.687**	0.041
	(0.003)	(0.051)	(0.819)	(0.024)	(0.211)	(0.182)	(0.085)	(0.105)
SI_t	**3.674**	0.014	-0.041	0.081	**0.217**	0.027	**0.966**	**0.643**
	(1.025)	(0.756)	(0.849)	(0.118)	(0.052)	(0.031)	(0.022)	(0.024)
Panel 2: Yields-Only Model								
β_{1t}	**2.977**	**0.903**	**0.021**	0.011				
	(0.173)	(0.120)	(0.015)	(0.023)				
β_{2t}	**-2.819**	0.041	**0.901**	-0.001				
	(0.305)	(0.210)	(0.155)	(0.033)				
β_{3t}	**-2.723**	-0.560	-0.371	**0.866**				
	(0.538)	(0.502)	(0.312)	(0.085)				

Panel 3:Test for Joint-Significance of the Individually Insignificant Coefficients

	Yields-Macro Model			**Yields-Only Model**		
Wald Test Statistic	Value	df	P-value	Value	df	P-value
Chi-square	41.557	27	0.000	29.557	5	0.000

Note: The table reports the estimates for the parameters of the transition equation for both, yields-macro and yields-only, models. The panel 1 and 2 present the estimates for the vector μ and matrix A for the yields-macro and yields-only models respectively. The standard errors are in parenthesis. Bold entries denote parameter estimates significant at the 5 percent level. The panel 3 presents the results of the Wald-test for the joint significance of individually insignificant coefficients in matrix A. The null hypothesis is that insignificant coefficients are simultaneously equal to zero. The test statistic is Chi-square with their respective degrees of freedom (df). P-value is the probability value of the test statistic.

Regarding the impact of yield curve factors on macroeconomic variables, the results show that exchange rate and inflation rate are negatively related to the level of interest rate. It suggests that the long end of yield curve contains important information about the future inflation. The negative significant impact of long rates on exchange rate indicates that domestic currency appreciates because of capital inflow due to the attractiveness of domestic bonds. Furthermore, as the long end of yield curve goes down, inflationary expectations become stronger as a consequence of rise in aggregate demand. The spread term β_{2t} is negatively related to the level of economic activity, while the impact on stock market performance is statistically insignificant. Thus, a decrease in the slope of yield curve (becoming flat or negatively sloped) can be considered as a signal of economic slowdown. The macroeconomic variables have negligible responses to shocks in the curvature factor except the inflation rate. The inflation rate is negatively related to the curvature factor.[32]

As many of the coefficients in matrix A for both the models are statistically insignificant, Wald-test for their joint significance is conducted and the results are presented in the third panel of table 3.1. The test statistic rejects the null-hypothesis of joint insignificance of the 27 and 5 individually insignificant coefficients in the yields-macro and yields-only model respectively. This suggests that inclusion of macroeconomic factors in the Nelson-Siegel specification of yield curve improves the model's overall fit and prediction power.

Table 3.2: Estimates of Covariance Matrix Σ

	Yields-Macro Model							**Yields-Only Model**		
	$\hat{\eta}_{1t}$	$\hat{\eta}_{2t}$	$\hat{\eta}_{3t}$	$\hat{\eta}_{4t}$	$\hat{\eta}_{5t}$	$\hat{\eta}_{6t}$	$\hat{\eta}_{7t}$	$\hat{\eta}_{1t}$	$\hat{\eta}_{2t}$	$\hat{\eta}_{3t}$
$\hat{\eta}_{1t}$	**3.571**	**-0.035**	-0.039	**0.205**	-0.115	0.016	-0.005	**3.523**	**-0.032**	-0.039
	(0.214)	(0.007)	(0.114)	(0.012)	(0.418)	(0.422)	(0.057)	(0.226)	(0.010)	(0.022)
$\hat{\eta}_{2t}$		**2.987**	**0.061**	0.074	0.189	-0.003	**-0.042**		**2.991**	**0.054**
		(0.185)	(0.021)	(0.067)	(0.499)	(0.615)	(0.008)		(0.263)	(0.026)
$\hat{\eta}_{3t}$			**1.358**	0.345	0.047	**-0.084**	0.095			**1.248**
			(0.234)	(0.339)	(0.865)	(0.024)	(0.752)			(0.183)
$\hat{\eta}_{4t}$				**3.834**	-0.029	-0.086	**3.071**			
				(0.073)	(0.120)	(0.074)	(1.059)			
$\hat{\eta}_{5t}$					**1.607**	**0.648**	-0.959			
					(0.341)	(0.321)	(0.628)			
$\hat{\eta}_{6t}$						**2.133**	-0.241			
						(0.269)	(0.459)			
$\hat{\eta}_{7t}$							**1.386**			
							(0.228)			

Test for Diagonality of Covariance Matrix Σ

Wald Test Statistic	**Yields-Macro Model**			**Yields-Only Model**		
	Value	df	P-value	Value	df	P-value
Chi-square	31.409	21	0.000	20.136	3	0.000

Note: The upper panel of table reports the estimates of covariance matrix of innovations of the transition equation for both the models (yields-macro and yields-only models). The standard errors are in parenthesis. The lower panel presents the results of the Wald-test for the null hypothesis that covariance matrix **Σ** is diagonal. The test statistic is Chi-square with their respective degrees of freedom (df). P-value is the probability value of the test statistic. Bold entries denote parameter estimates significant at the 5 percent level.

The estimates of covariance matrix of the state innovations as depicted by Σ in (3.7 and 3.10) and Wald-test of its diagonality in both models are shown in table 3.2. There is only one individually insignificant covariance term (between $\hat{\eta}_{1t}$ and $\hat{\eta}_{3t}$) for the yields-only model. However, for yield-macro model only 7 out of 21 covariance terms are statistically significant at

[32] We also estimated the state-space model, considering the macroeconomic variable as exogenous in the transitional equation (3.9) that can be expressed as:

$$\xi_t = \mathrm{A}\xi_{t-1} + \Gamma Z_{t-1} + \eta_t \quad , \qquad \eta_t \sim N(0, \Sigma)$$

where $\xi_t = (\beta_t - \mu)'$ is (3×1) vector of yield curve factors, $Z_t = (IP_t\,, EX_t\,, INF_t\,,\ SI_t)'$ is the (4×1) vector of macroeconomic variables, μ, A and Γ are (3×1) vector, (3×3) and (3×4) matrix of unknown parameters respectively. Σ is (3×3) covariance matrix of the error term (η_t) of state equation. We estimate and forecast with observation equation (3.5) and above mentioned state equation and the results of estimated state vector ξ_t and its forecast values are almost similar to our earlier representation of yields-macro model.

5% level of significance. We also perform the Wald-test for the joint significance of the off-diagonal elements of the matrix and the test statistic clearly rejects the null-hypothesis of the diagonality of the Σ matrix for yields-macro model as well as yields-only model. The result is consistent with our prior expectation that the innovations of transition system are cross correlated.

Using cross-sectional as well as information concerning the evolution of yields over time, we employ the Kalman smoother algorithm to obtain optimal extractions of the latent level, slope and curvature factors and corresponding covariance matrix using (3.17) and (3.18) respectively. Table 3.3 shows the descriptive statistics of the three time-varying Kalman filter smooth estimates of factors along with averaged smoothed residuals for both models, i.e., the yields-macro and yields-only model.

Comparing the mean, standard deviation and other descriptive features of the estimated factors across models show that both the models give rather similar estimates for the level, slope and curvature factors in magnitude. From the autocorrelations in the table 3.3 of the estimated factors, we can see that β_{1t} is more persistent than the rest of two factors for both the models. The results suggest the high persistency and low volatility of long rates. The results also show that the lag autocorrelation of the residuals is low, justifying the reliability of standard errors of the estimated factors. The residuals descriptive features indicate that the average yield curve is fitted very well. Finally, the estimated τ for both models, i.e., yields-macro models is 71.293 and yields-only model is 71.420, implies that the loading on the curvature factor is maximized at a maturity of about 6 years.

Table 3.3: Descriptive Statistics of the Nelson-Siegel Factors Estimates

	Yields-Macro Model				**Yields-Only Model**			
Factors	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\varepsilon}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\varepsilon}$
Mean	2.951	-2.780	-2.655	0.001	2.994	-2.813	-2.722	-0.003
Std. Dev.	0.381	0.483	1.202	0.014	0.367	0.462	1.180	0.014
Maximum	3.789	-1.392	0.681	0.034	3.803	-1.432	0.478	0.027
Minimum	1.453	-3.892	-4.273	-0.059	1.500	-3.900	-4.341	-0.055
Skewness	-1.165	0.383	0.627	-1.255	-1.346	0.423	0.636	-0.670
Kurtosis	5.888	2.766	2.432	6.000	6.617	2.960	2.457	4.254
$\hat{\rho}$ (1)	0.904	0.882	0.889	0.464	0.903	0.881	0.885	0.426
$\hat{\rho}$ (6)	0.518	0.454	0.531	0.301	0.610	0.440	0.513	0.353
$\hat{\rho}$ (12)	0.289	-0.048	0.136	0.164	0.301	-0.071	0.116	0.126
$\hat{\tau}$		71.293	(0.025)			71.420	(0.028)	

Note: The table shows descriptive statistics for smoothed estimates of β_t vector and averaged smoothed residuals $\hat{\varepsilon}$ (averaged over the different maturities) of the yields-macro as well as yields-only model using monthly data 2000:01–2011:12. $\hat{\rho}\ (i)$ denotes the sample autocorrelations at displacements of 1, 6 and 12 months. $\hat{\tau}$ is the optimal estimate of the shape parameter and its standard errors are in parenthesis. The number of observations is 144.

Furthermore, the time-series of the factors' smoothed estimates as well as the series of their empirical proxies and potentially related macroeconomic variables are plotted in figures 3.1 and 3.2. The level of the yield curve (L_t) is defined as the 25-year yield. We compute the slope (S_t)

as the difference between the 25-year and three-month yield and finally, the curvature (C_t) is defined as two times the two-year yield minus the sum of the 25-year and three month zero-coupon yields. Comparing the factors' estimates for both the models give rather close similar approximations for the level, slope and curvature factors. The pairwise correlation of empirically defined factors and estimated factors of the yields-macro model is $\rho(L_t, \hat{\beta}_{1t}) = 0.768$, $\rho(S_t, \hat{\beta}_{2t}) = -0.902$ and $\rho(C_t, \hat{\beta}_{3t}) = 0.837$ and for the yields-only model is $\rho(L_t, \hat{\beta}_{1t}) = 0.739$, $\rho(S_t, \hat{\beta}_{2t}) = -0.897$ and $\rho(C_t, \hat{\beta}_{3t}) = 0.831$. To be precise, the estimated factors and their empirical proxies seem to follow the same pattern and hence, may truly be called level, slope and curvature factors, respectively.

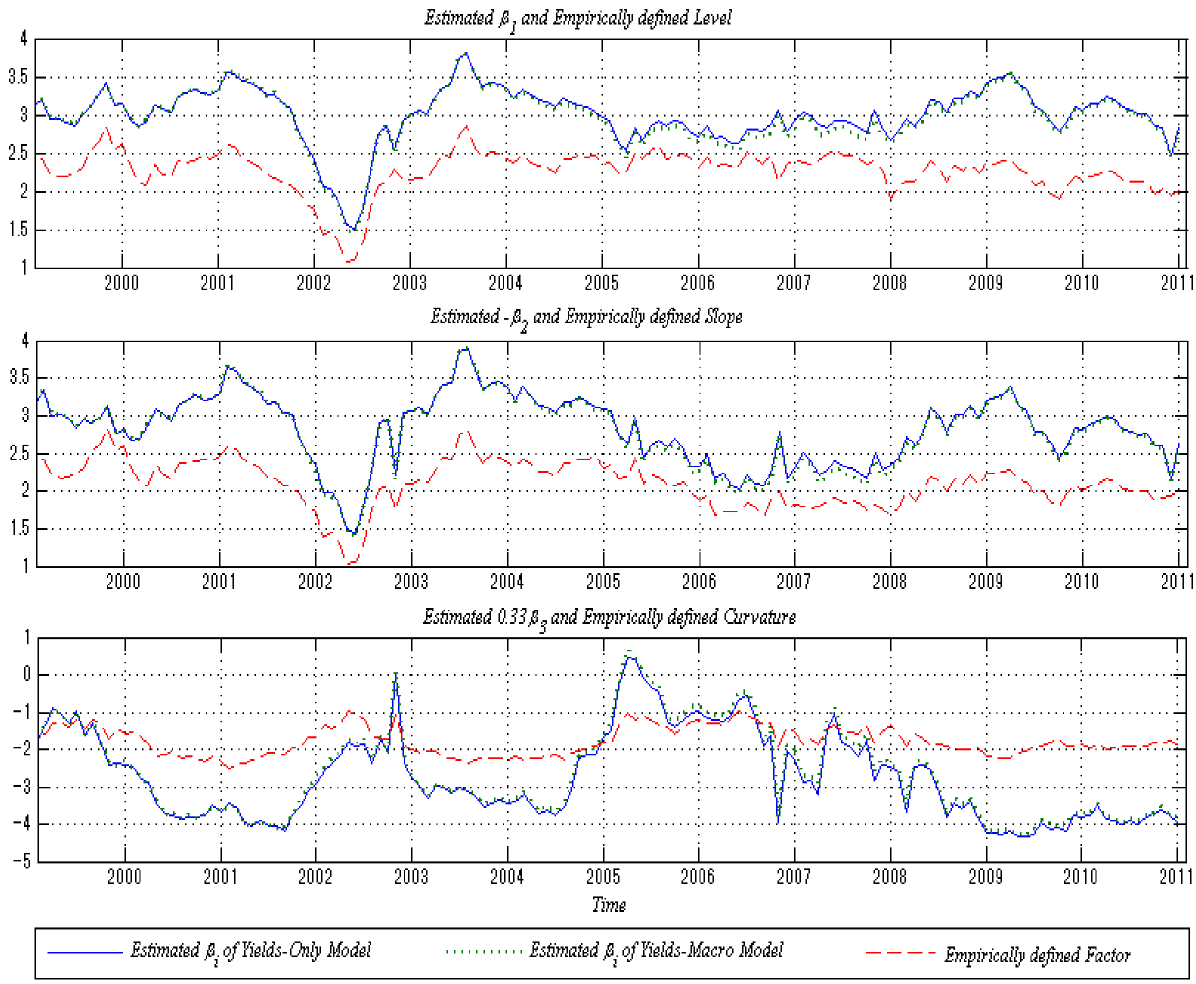


Figure 3.1: Time Series Plot of Nelson-Siegel Estimated Factors and Empirical Level, Slope and Curvature
Model-based level, slope and curvature (i.e., estimated factors) vs. data-based level, slope and curvature (i.e., empirical proxies), where level is defined as the 25-year yield, slope as the difference between the 25-year and 3-month yields and curvature as two times the 2-year yield minus the sum of the 25-years and 3- month zero coupon yields. Rescaling of estimated factors is based on Diebold and Li (2006).

The level factor is closely related to annual growth of money supply as depicted in figure 3.2.[33] The correlation between $\hat{\beta}_{1t}$ and annual growth of money supply is -0.352, consistent with inflationary expectations hypothesis. It confers that shocks to monetary policy are important

[33] Growth rate of money supply is defined as $MS_t = 100 \times [(M_t - M_{t-12})/M_{t-12}]$, where M is level of seasonally adjusted money supply M2 and its data has been retrieved from International Financial Statistics, IMF.

sources of variation in long end of the yield curve and pricing the long-term maturity bonds. Thus, monetary policy surprises and acts as to drive up short rates and alter expectations about future interest rates by shifting the long end of yield curve up in a persistent way (consistent with the costly price adjustment hypothesis of monetary policy) and thereby stimulates the economy. It suggests that the shift of long end and hence, the shape of yield curve has important information of the state of economy. It indicates that monetary policy is an important source of variation even in the zero interest rate policy regime. Moreover, the variation in inflation is closely explained by the curvature factor of the yield curve. The correlation between $\hat{\beta}_{3t}$ and INF_t is 0.391. The CPI based inflation rate closely follows the pattern of curvature factor of yield curve as depicted in the lower panel of figure 3.2.

Moreover, one can observe that $\hat{\beta}_{1t}$ and $-\hat{\beta}_{2t}$ follow almost the same pattern in figure 3.1. There is a sharp decline in $\hat{\beta}_{1t}$ as well as the slope factor $\hat{\beta}_{2t}$ in early 2001 till mid-2002 and is followed by the gradual recovery process. This behavior of the level and slope factors is closely related to the monetary policy regimes of the Japanese economy during the decade. In early 1998 in the Japanese economy, the demand was falling and the economy was heading into a recession and financial instability.[34] In order to avoid the severe recession, the so-called zero interest rate policy (ZIRP) was introduced and an easy monetary policy was adopted.[35] The economy did not respond quickly, however, it started to show some signs of recovery in the spring of 2000 and as a consequence, the ZIRP was lifted in August 2000. Almost as soon as the interest rate was raised, the Japanese economy entered into another recession and many urged changes in monetary policy and return to ZIRP.[36] In February 2001, the Bank of Japan (BOJ) introduced the Lombard lending facility as well as cut the official discount rate from 0.5% to 0.35%.[37] However, these measures did not show any significant impact and further steps to easing in monetary policy are taken. The target inter-bank rate was lowered immediately to 0.15 percent, and would go down to zero, as conditions warranted. The official discount rate was sharply cut to 0.1 percent. During this regime, we observe that the long rates as well as the slope of yield curve have a downward trend. This confers that fall in long rates was larger than the decline in short rates. Hence, the slope was falling along with alteration in the shape of yield curve.

---

[34] The effects of Asian financial crisis were heading towards the Japanese economy and financial instability became prominent as one large bank and one small bank, a large securities firm and a medium-size securities firm all failed and credit lines between western financial institutions and Japanese financial institutions became severely limited in November 1997 (Ito and Mishkin, 2004).

[35] The overnight call rate was radically reduced to 0.25% in September 1998 and to 0.15% in early 1999 from 0.5%.

[36] First, the ICT bubble ended and stock prices in the Japanese stock market were heading down, suggesting investment and consumption would be adversely affected in the near future. Second, the US economy was beginning to show weakness and Japanese exports to the United States were expected to decline in the future. Third, the inflation rate was still negative, and there was no sign of an end to deflation. It was not known at the time, but the official date for the peak of the business cycle turned out to be October 2000. The growth rate of 2000:III turned negative, which was offset to some extent by a brief recovery in 2000:IV.

[37] The Lombard lending facility was to lend automatically to banks with collateral at the official discount rate, so that the interest rate would be capped at 0.35%. However, the market rate was at around 0.2 – 0.25%, so there was little real impact from the introduction of the Lombard facility.

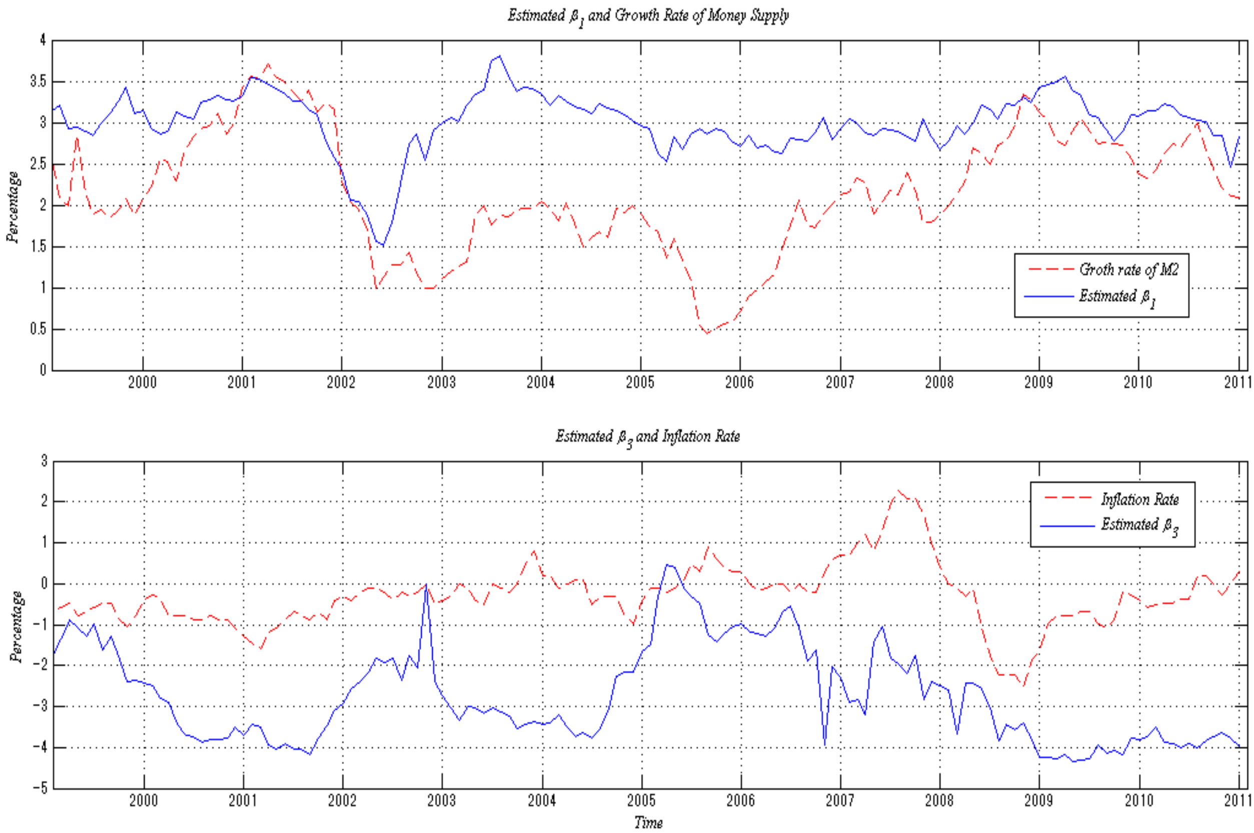


Figure 3.2: Time Series Plot of Nelson-Siegel Estimated Factors with Macroeconomic Variables
The estimated level and curvature factors ($\beta_{1t}$and β_{3t}) are plotted vs. annual growth of the M2 (Money Supply) and Inflation rate respectively. Inflation rate is the 12-month percent change in the consumer price index.

During the last quarter of 2002 the regime switched as in September 2002, the Bank started to purchase equities that the commercial banks held. The action was justified by the BOJ on the ground that it would reduce the risk of commercial banks, and it was made clear that it is not intended as monetary policy tool, but rather as financial market stabilization policy. However, it was not explained why the resulting risk to the BOJ balance sheet due to financial stabilization policy was not a big concern, while it was for monetary policy (Ito and Mishkin, 2004). Furthermore, the Bank made it explicit that it would continue ZIRP until deflationary concerns subside and the inflation rate is clearly above zero. The new policy was a big improvement over the last regime. Despite the good performance in the GDP growth rate in 2003:IV, the financial and capital market participants were expecting that ZIRP will continue for a long time. Thus, during the recovery regime the long end is gradually rotating and hence the slope is on increasing trend. The process completes around late 2004.

Thus, during the initial period of ZIRP and severe recession, we observe a sharp decline in the yields of long-term bonds and the slope of yield curve; and during the period of recovery, the yield curve long end as well as the slope is on the increasing trend. This suggests that the state of economy was clearly depicted by the behavior of level and slope factors of the yield curve and yield curve depicts stances of monetary policy and is an important leading indicator of the business condition and the state of economy.

Table 3.4: Descriptive Statistic of the Yield Curve Residuals

Maturity	Mean	Std. Dev.	RMSE	MAE	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
Yields-Macro Model							
3	-0.014	0.094	0.094	0.068	0.489	0.260	0.157
6	-0.028	0.083	0.087	0.057	0.436	0.238	0.056
9	-0.032	0.081	0.087	0.05	0.545	0.338	0.014
12	-0.003	0.029	0.029	0.022	0.368	0.296	0.186
15	0.001	0.015	0.015	0.012	0.375	0.279	0.262
18	0.004	0.013	0.011	0.008	0.156	-0.022	0.125
21	0.004	0.012	0.013	0.009	0.037	0.044	0.087
24	-0.001	0.015	0.015	0.011	0.139	-0.003	0.118
30	-0.001	0.018	0.018	0.014	0.260	0.102	0.057
36	-0.008	0.022	0.022	0.017	0.371	0.358	0.033
48	0.001	0.014	0.014	0.011	0.212	0.077	-0.013
60	-0.009	0.023	0.024	0.02	0.447	0.332	-0.017
72	-0.016	0.040	0.043	0.034	0.491	0.360	0.094
84	-0.006	0.060	0.060	0.047	0.621	0.486	0.227
96	0.020	0.067	0.070	0.054	0.632	0.501	0.293
108	0.038	0.051	0.063	0.051	0.632	0.525	0.338
120	0.050	0.051	0.070	0.056	0.455	0.350	0.214
180	0.004	0.074	0.063	0.055	0.689	0.604	0.411
240	-0.003	0.097	0.077	0.063	0.688	0.529	0.364
300	0.030	0.081	0.064	0.063	0.484	0.253	0.171
Yields-Only Model							
3	-0.019	0.093	0.095	0.069	0.591	0.326	0.141
6	-0.032	0.083	0.089	0.058	0.594	0.255	0.047
9	-0.035	0.081	0.088	0.052	0.684	0.350	0.007
12	-0.005	0.029	0.030	0.023	0.344	0.288	0.203
15	0.001	0.015	0.015	0.012	0.423	0.246	0.282
18	0.003	0.012	0.012	0.008	-0.043	0.033	0.082
21	0.004	0.013	0.013	0.009	-0.057	-0.026	0.135
24	0.001	0.015	0.015	0.011	0.161	-0.022	0.132
30	-0.001	0.019	0.019	0.014	0.333	0.078	0.069
36	-0.007	0.022	0.023	0.017	0.519	0.344	0.045
48	0.001	0.014	0.014	0.011	0.333	0.088	-0.043
60	-0.008	0.023	0.024	0.02	0.475	0.339	-0.023
72	-0.017	0.042	0.045	0.035	0.567	0.353	0.105
84	-0.008	0.063	0.064	0.050	0.682	0.476	0.242
96	0.016	0.072	0.074	0.057	0.678	0.485	0.308
108	0.033	0.057	0.065	0.051	0.684	0.491	0.362
120	0.043	0.056	0.070	0.057	0.558	0.275	0.262
180	-0.012	0.066	0.066	0.056	0.742	0.609	0.407
240	-0.024	0.079	0.082	0.066	0.735	0.563	0.333
300	0.005	0.083	0.083	0.068	0.516	0.328	0.117

Note: The table presents summary statistic of the residuals $\hat{\varepsilon}$ for different maturity times of the measurement equation of both models using monthly data 2000:01–2011:12. RMSE and MAE are the root mean squared errors and mean absolute error respectively. $\hat{\rho}\,(i)$ denotes the sample autocorrelations at displacements of 1, 6, and 12 months. The number of observations is 144.

Furthermore, table 3.4 and figure 3.3 present the descriptive statistics and the three dimensional plot of the smoothed residuals for all the maturities. Both the models fit the yield curve remarkably well. Table 3.4 contain the estimated mean, standard deviation, mean absolute fit error (MAE), root mean squared fit error (RMSE) and autocorrelation at various displacements of the residuals, expressed in basis points, for each of the 20 maturities that we consider. The mean error is negligible at all maturities for both the models. However, comparing with respect to RMSE and MAE, the yields-macro model fits the yield curve slightly pretty than the yields-only model for all maturities. Furthermore, the residuals persistency of yields-macro model is lower than of yields-only model almost for all maturities.

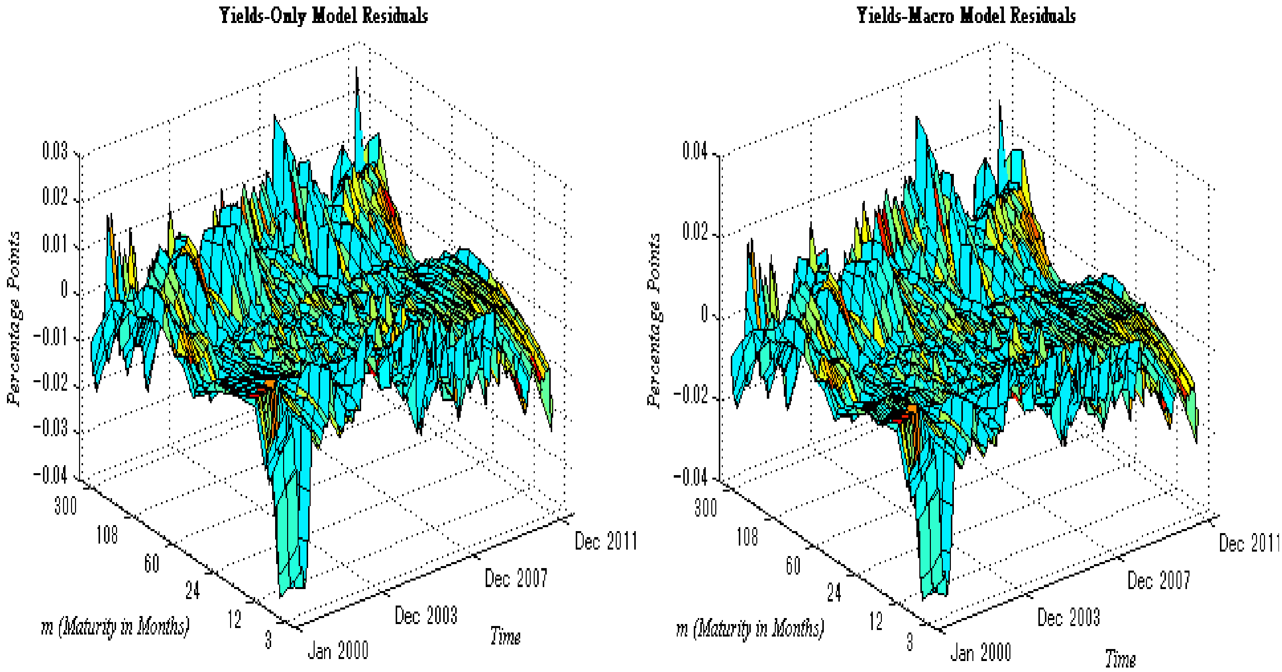


Figure 3.3: Nelson-Siegel Model based Yield Curves Residuals, 2000:01-2011:12
The sample consists of monthly smoothed estimated residuals from January 2000 to December 2011 (144 months) for maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240, and 300 months (20 maturities).

It turns out that the fit is more appealing in most cases. Some months, however, especially those with multiple maxima and/or minima are not fitted very well. It becomes apparent by the large residuals in these months.

Moreover, table 3.5 presents four different criterions to compare the in-sample fit of the yield curve. Table 3.5 contains the estimated Log likelihood ratio, Akaike information criterion, Schwarz information criterion and Hannan-Quinn information criterion for both the models. The Log likelihood ratio of yields-macro model is greater than that of the yield-only model, suggesting that the inclusion of macroeconomic factors leads to estimate the yield curve more accurately. Similarly, the other three criterions AIC, SIC and HQ also support this argument as they are smaller for yields-macro model than of the yields-only model.

Table 3.5: In-sample Fit Diagnostic Statistics of the Nelson-Siegel Model

Models	**Yields-Macro Model**	**Yields-Only Model**
Log likelihood	4384.668	4311.996
AIC criterion	-59.851	-58.347
SIC criterion	-58.133	-57.543
HQ criterion	-58.272	-57.020

Note: The table presents the in-sample fit performance of the yields-macro and yields-only models specified in the state-space representation, using four different criterions. AIC is the Akaike information criterion, SIC is the Schwarz information criterion and HQ is the Hannan-Quinn information criterion.

In summary, we have explained that both the models provide an evolution of the term structure closer to reality. These models in the state-space representation are capable to distill the term structure of interest rate quite well and describe the evolution and the trends of the government bonds market. However, the yields-macro model provides a little better fit of the yield curve than the yields-only model. More importantly, the lag correlation of the signal system innovations in the yields-macro model is lower than of the yields-only model and leads to the reliability of the yields-macro model results. This suggests that the common phenomenon of the high degree of residuals persistency for various maturities in the class of statistical models of yield curve can be avoided by the inclusion of macroeconomic factors in the system of yield curve model. Furthermore, the use of term spread in forecasting future economic activity and stock market seems to have noticeable role and long end of yield curve can explain the exchange rate and inflationary expectations.

3.4. Out-of-Sample Forecasting

A good approximation to yield curve dynamics should not only fit well in-sample, but also produces satisfactory out-of-sample forecasts. For the out-of-sample performance, the similar models are estimated as for the in-sample fit. To assess the forecasting performance of the models, the sample is divided into the initial estimation period January 2000 to December 2007 and the forecasting period January 2008 to December 2011. We estimate and forecast recursively, using data from January 2000 to the time that the forecast is made, beginning in January 2008 and extending through December 2011, i.e., the models are estimated recursively with an expanding data window. Interest rate forecasting is done by constructing factor predictions using the state equations and subsequently substituting these predictions in the measurement equations to obtain the interest rate forecasts. Three forecast horizons, $h = 1, 6$ and 12 months ahead are considered. The h-month ahead factors forecasts, $\hat{\xi}_{t+h}$, are iterated forecasts which follow from forward iteration of the state equations (3.6) for yields-only model and (3.9) for yields-macro model as:

$$\hat{\xi}_{t+h|t} = \hat{A}^h \hat{\xi}_{t|t} \tag{3.20}$$

where $\hat{A}^h$ denotes the matrix $\hat{A}$ multiplied by itself h times and $\hat{\xi}_{t|t}$ is the last available factor estimates. The first three elements of $\hat{\xi}_{t+h|t}$ in (3.20) is subsequently substituted in the observation equations (that are $\hat{\beta}_{t+h|t}$), results in:

$$\hat{R}_{t+h|t}(m) = \Lambda(\hat{\tau})(\hat{\beta}_{t+h|t}) \tag{3.21}$$

where $\hat{\beta}_{t+h|t}$ is the (3×1) vector consists of yield curve three factors. We use the in-sample shape parameter τ estimate to compute the factor loadings in forecasts. Furthermore, we define $\hat{R}_{t+h|t}(m)$ as $\hat{R}_{t,t+h}(m)$ is the forecasted yield in period t for $t+h$ period (for i^{th} maturity).

As a benchmark model for comparing out-of-sample forecast, we use the univariate AR(1) specification of yield:

$$R_{t+h}(m) = \delta_0 + \delta_1 R_t(m) + \varepsilon_{t+h} \tag{3.22}$$

for $h = 1, 6,$ and 12, and $\varepsilon_t \backsim N(0, \sigma^2)$.[38]

3.4.1. Term Structure Forecast Results

In tables 3.6, 3.7 and 3.8, we compute the descriptive statistics of *h-month-ahead* out-of-sample forecasting results of yields-macro, yields-only and AR(1) yield models, for maturities of 3, 6, 12, 24, 60, 120, 180, 240, and 300 months for the forecast horizons of $h = 1, 6,$ and 12 months. We define forecast errors at time t for $t+h$ as $[R_{t+h}(m) - \hat{R}_{t,t+h}(m)]$, where $\hat{R}_{t,t+h}(m)$ is the forecasted yield in period t for $t+h$ period and is not the Nelson–Siegel fitted yield. $R_{t+h}(m)$ is the actual yield in period $t+h$. We examine a number of descriptive statistics for the forecast errors, including mean, standard deviation, mean absolute error (MAE), root mean squared error (RMSE) and autocorrelation at various displacements.

Table 3.6 reports the results of one month ahead forecasts of yields-macro, yields-only as well as AR(1) models of yield. The AR(1) model plays a roll of benchmark for evaluating the empirical performance of the yields-macro and the yields-only models. The one month ahead forecasting results for both the yields-macro and yields-only models considerably outperform those of the AR(1) model in terms of MAE and RMSE for any maturity. The yields-only model appears suboptimal in comparison with the yields-macro as the forecasts errors are serially correlated. However, the lag autocorrelation of the forecasts errors of yields-macro model for all maturities are smaller and negligible as compared to the yields-only and AR(1) models. The mean, MAE and RMSE of forecast errors of yields-macro model are slightly smaller than that of yields-only model for all maturities. Furthermore, MAE, RMSE and lag correlation of the forecast errors of yields-only model are smaller as compared to the AR(1) model of yield. In relative terms, the results indicate that yields-macro model outperform both the yields-only as well as the AR(1) model and yields-only model comes with much accurate forecasts than the AR(1) specification for the one month ahead forecast horizon.

The results of 6 months and one year ahead forecast in table 3.7 and 3.8 respectively reveal that matters worsen radically with longer horizon forecasts. For 6 months ahead forecast the

[38] The random walk model could be a benchmark model. We, in fact, ran the random walk model of yield, but the AR(1) specification of yield outpace the random walk forecasts (results of random walk are not reported).

yields-macro as well as the yields-only models outperform the AR(1) model in term of mean forecast errors, MAE, RMSE and lag autocorrelation for all maturities. Moreover, forecasts of the yields-macro model are much precise than of the yields-only specification for all maturities both in terms of RMSE and error persistency. The 6 month ahead forecasts results seem not good as the one month ahead forecasts in term of lag autocorrelation for all the three models.

Table 3.6: Out-of-Sample 1 Month Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
Forecast Summary for Yields-Macro Model							
3	-0.004	0.128	0.069	0.019	0.522	-0.059	-0.017
6	-0.028	0.120	0.058	0.022	0.419	-0.070	-0.008
12	-0.001	0.119	0.023	0.001	0.319	-0.111	0.077
24	0.004	0.149	0.011	0.001	0.401	0.117	0.185
60	-0.006	0.192	0.020	0.001	0.447	-0.124	-0.010
120	0.044	0.176	0.057	0.005	0.476	-0.163	-0.099
180	-0.012	0.180	0.055	0.005	0.537	-0.028	-0.085
240	-0.024	0.184	0.066	0.009	0.599	0.042	-0.030
300	0.005	0.194	0.068	0.010	0.622	-0.030	-0.086
Forecast Summary for Yields-Only Model							
3	-0.005	0.029	0.084	0.043	0.849	-0.076	-0.181
6	0.043	0.056	0.075	0.044	0.813	-0.115	-0.178
12	0.003	0.012	0.069	0.047	0.601	-0.026	0.038
24	0.005	0.015	0.096	0.062	0.443	-0.045	-0.008
60	-0.019	0.094	0.142	0.077	0.669	-0.092	-0.081
120	0.006	0.083	0.135	0.055	0.673	-0.169	0.040
180	-0.006	0.022	0.136	0.052	0.780	-0.068	-0.019
240	-0.032	0.083	0.146	0.053	0.800	-0.111	-0.042
300	-0.008	0.023	0.150	0.060	0.661	-0.015	0.058
Forecast Summary for AR(1) Model							
3	-0.080	0.126	0.131	0.149	0.884	0.446	-0.004
6	-0.067	0.102	0.104	0.122	0.881	0.447	-0.006
12	-0.029	0.129	0.109	0.133	0.888	0.380	-0.131
24	-0.055	0.203	0.184	0.211	0.846	0.355	-0.061
60	-0.123	0.299	0.278	0.324	0.743	0.286	0.053
120	-0.136	0.212	0.193	0.252	0.567	0.058	0.085
180	0.025	0.185	0.143	0.187	0.494	-0.028	0.037
240	0.035	0.190	0.154	0.193	0.553	-0.161	-0.042
300	-0.138	0.163	0.170	0.214	0.672	-0.186	-0.105

Note: The table reports the results of out-of-sample 1-month-ahead forecasting using state-space specification for the yields-macro and yields-only models along with the AR(1) forecasts of yields for various maturities. We estimate the models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+1$ as $R_{t+1}(m) - \hat{R}_{t,t+1}(m)$, where $\hat{R}_{t,t+1}(m)$ is the $t+1$ month ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors (MAE) and root mean squared errors (RMSE) of the forecast errors, as well as their first, 6th and 12th sample autocorrelation coefficients.

Table 3.7: Out-of-Sample 6 Months Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
Forecast Summary for Yields-Macro Model							
3	-0.005	0.153	0.102	0.039	0.700	-0.237	-0.115
6	-0.018	0.148	0.098	0.037	0.660	-0.254	-0.097
12	0.004	0.169	0.112	0.045	0.671	-0.231	-0.022
24	0.005	0.216	0.147	0.063	0.720	-0.237	0.016
60	-0.008	0.268	0.195	0.087	0.725	-0.188	-0.007
120	0.039	0.230	0.171	0.079	0.726	-0.115	-0.089
180	-0.012	0.240	0.167	0.104	0.766	-0.023	-0.110
240	-0.021	0.251	0.170	0.127	0.790	0.046	-0.047
300	0.011	0.267	0.185	0.137	0.707	0.026	-0.051
Forecast Summary for Yields-Only Model							
3	-0.006	0.147	0.110	0.039	0.745	0.262	0.127
6	-0.019	0.140	0.108	0.039	0.693	0.209	0.144
12	0.004	0.158	0.125	0.046	0.721	0.244	0.149
24	0.006	0.201	0.161	0.072	0.742	0.245	0.174
60	-0.009	0.250	0.210	0.093	0.732	0.178	0.059
120	0.041	0.222	0.176	0.081	0.743	0.190	-0.120
180	-0.014	0.234	0.167	0.118	0.793	0.258	-0.150
240	-0.023	0.245	0.167	0.145	0.822	0.296	-0.098
300	0.012	0.260	0.187	0.160	0.841	0.292	-0.106
Forecast Summary for AR(1) Model							
3	-0.376	0.279	0.411	0.468	0.892	0.551	0.150
6	-0.416	0.286	0.448	0.505	0.832	0.558	0.177
12	-0.216	0.156	0.377	0.195	0.716	0.460	0.074
24	-0.165	0.203	0.242	0.262	0.781	0.415	0.116
60	-0.184	0.262	0.284	0.320	0.742	0.319	0.113
120	-0.148	0.210	0.201	0.256	0.797	0.394	0.091
180	0.109	0.179	0.138	0.179	0.852	0.220	0.046
240	0.144	0.171	0.244	0.176	0.885	0.294	0.133
300	-0.165	0.163	0.289	0.175	0.897	0.310	0.182

Note: The table presents the results of out-of-sample 6-month-ahead forecasting using state-space specification for the yields-macro and yields-only models along with the AR(1) forecasts of yields for various maturities. We estimate the models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+6$ as $R_{t+6}(m) - \hat{R}_{t,t+6}(m)$, where $\hat{R}_{t,t+6}(m)$ is the $t+6$ months ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors (MAE) and root mean squared errors (RMSE) of the forecast errors, as well as their first, 6th and 12th sample autocorrelation coefficients.

For 12 months ahead, both the yield curve models based on Nelson-Siegel specification, i.e., yields-macro and yields-only, perform well than the benchmark AR(1) specification, as the MAE and RMSE of both the former models are lower than of the AR(1) model. However, the autocorrelation of the forecast errors for all the three models is almost same for all the maturities. It is worth noting, that related papers such as Bliss (1997), de Jong (2000) and Diebold and Li (2006) also find serially correlated forecast errors, often with persistence much stronger than ours.

Moreover, MAE and RMSE comparisons at various maturities reveal the preference of yields-macro over yields-only model even at longer horizon forecast.[39]

Table 3.8: Out-of-Sample 12 Months Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
Forecast Summary for Yields-Macro Model							
3	0.006	0.184	0.129	0.046	0.835	-0.040	-0.111
6	-0.005	0.182	0.131	0.044	0.813	-0.052	-0.073
12	0.002	0.211	0.154	0.054	0.832	0.000	-0.044
24	0.003	0.266	0.196	0.084	0.847	-0.018	-0.030
60	-0.019	0.318	0.240	0.117	0.812	-0.046	-0.057
120	0.032	0.259	0.174	0.104	0.801	-0.045	-0.079
180	-0.019	0.260	0.162	0.149	0.814	-0.009	-0.124
240	-0.024	0.269	0.162	0.183	0.818	0.041	-0.077
300	0.008	0.284	0.176	0.202	0.829	0.023	-0.095
Forecast Summary for Yields-Only Model							
3	0.007	0.180	0.133	0.047	0.842	0.010	0.174
6	-0.006	0.178	0.135	0.045	0.821	-0.018	0.167
12	0.003	0.205	0.160	0.055	0.837	0.017	0.182
24	0.003	0.260	0.204	0.084	0.852	-0.024	0.154
60	-0.020	0.309	0.250	0.119	0.829	-0.071	0.065
120	0.034	0.252	0.180	0.108	0.809	-0.072	0.005
180	-0.019	0.256	0.165	0.153	0.833	-0.042	-0.025
240	-0.024	0.266	0.165	0.188	0.854	0.004	0.016
300	0.007	0.281	0.177	0.206	0.869	-0.012	0.024
Forecast Summary for AR(1) Model							
3	-0.080	0.126	0.136	0.149	0.867	0.386	0.252
6	-0.067	0.102	0.144	0.122	0.897	0.352	0.235
12	-0.029	0.129	0.179	0.133	0.856	0.569	0.457
24	-0.055	0.203	0.284	0.211	0.848	0.726	0.601
60	-0.123	0.299	0.278	0.324	0.825	0.698	0.593
120	-0.136	0.212	0.193	0.252	0.921	0.425	0.308
180	0.025	0.185	0.183	0.187	0.858	0.335	0.206
240	0.035	0.190	0.174	0.193	0.840	0.424	0.266
300	-0.138	0.163	0.187	0.214	0.872	0.431	0.300

Note: The table reports the results of out-of-sample 12-month-ahead forecasting using state-space specification for the yields-macro and yields-only models along with the AR(1) forecasts for various maturities. We estimate the models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+12$ as $R_{t+12}(m) - \hat{R}_{t,t+12}(m)$, where $\hat{R}_{t,t+12}(m)$ is the $t+12$ months ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors (MAE) and root mean squared errors (RMSE) of the forecast errors, as well as their first, 6th and 12th sample autocorrelation coefficients.

[39] We also computed the 24 months ahead forecasts for the yields-macro, yields-only and AR(1) models to evaluate the performance for the longer horizon forecasts, though not reported here for the save of space. The Nelson-Siegel specifications perform better than the AR(1) model even for the longer horizon, while the 24 months ahead forecasts are not as good as the short horizon forecasts possibly due to small sample size.

In summary, the out-of-sample forecasts results indicate that the Nelson-Siegel specification of yield curve comes with more accurate and precise forecasts as compared to the benchmark forecasts models such as random walk and AR(1). Furthermore, the forecasts of yields-macro model seem reasonably well in term of lower forecasts errors and lags autocorrelation than the yields-only model. In term of lower RMSE, our results for all the three forecast horizons are preferred than that of related studies, i.e., Bliss (1997), de Jong (2000) and Diebold and Li (2006). The results of yields-macro model suggest that the autocorrelation of forecasts errors could be eliminated/reduced by the inclusion of various yield curve related variables in the model.

3.4.2. Out-of-Sample Forecast Accuracy Comparisons

To assess the overall quality of the out-of-sample forecasts of the models, we use a number of standard forecasts errors evaluation criteria. In the first subsection, we compute the Trace Root Mean Squared Prediction Error (TRMSPE), which summarizes the forecasting performance of each model. TRMSPE does not belong to the class of formal statistical tests but rather is a standard criterion, which is widely used to assess forecast accuracy. To test the statistical accuracy of the forecasts of all the three models, we consider the standard statistical tests in subsections 3.4.2.2 and 3.4.2.3. We employ Diebold and Mariano (1995) test to the loss differential quadratic errors and the standard t-test to the absolute values of forecast errors. Particularly, it is worth noting that the null hypothesis is same and the sign of statistic has a clear meaning in both the tests.[40] Therefore, the results of the two tests can be directly compared in terms of statistical significance and sign.

3.4.2.1. Trace Root Mean Squared Prediction Error

The Trace Root Mean Squared Prediction Error (TRMSPE) combines the forecast errors of all maturities and summarizes the performance of each model, thereby allowing for a direct comparison between models. Given a sample of $(T - T_0)$ out-of-sample forecasts of N distinct maturities with $h-$months ahead forecast horizon, we compute the TRMSPE as follows:

$$TRMSPE = \sqrt{\frac{1}{N(T-T_0)} \sum_{m=1}^{N} \sum_{t=T_0}^{T} \left[R_{t+h}(m) - \hat{R}_{t,t+h}(m)\right]^2} \tag{3.23}$$

where $\hat{R}_{t,t+h}(m)$ is the forecasted yield in period t for $t+h$ period, $[R_{t+h}(m) - \hat{R}_{t,t+h}(m)]$ is the forecast errors at $t+h$ for yield and T_0 is the first month of forecasting period (January 2008).

In table 3.9, we report the TRMSPE for all the three models, i.e., yields-macro, yields-only and AR(1) model for all the three forecasts horizons. The overall performances of yields-macro

[40] One should be aware of the difference that we use squared forecast errors in the Diebold and Mariano (1995) forecast accuracy comparison test and absolute forecast errors in the t-test.

and yields-only models are much superior to that of the AR(1) benchmark model in terms of TRMSPE for all the three forecast horizons. It suggests that both specifications of Nelson-Siegel model (with and without macroeconomic variables) of yield curve outperform the AR(1) forecasts of yield. Table 3.9 shows that the forecasts errors of all the three models are getting larger as, the forecast horizons become longer.

Table 3.9: TRMSPE Results for Out-of-Sample Forecasts Accuracy Comparisons

TRMSE	1 Month Forecasts	6 Months Forecasts	12 Months Forecasts
Yields-Macro Model	0.002	0.041	0.046
Yields-Only Model	0.003	0.044	0.062
AR(1) Model	0.120	0.218	0.248

Note: The table reports the Trace Root Mean Squared Prediction Error (TRMSPE) results of out-of-sample forecasts accuracy comparison for horizons of one, 6, and 12 months for all the three models.

However, the TRMSPE of the yields-macro model for 12 months ahead forecast is much smaller than that of the yields-only model, while the former TRMSPE for the one and six months horizons are marginally better than the latter one. Thus, introducing macroeconomic variables into the yields curve model improves yield curve forecasts as compared to the yields-only model.

Based on the TRMSPE criterion, we can conclude that the yields-macro and yields-only models clearly outperform the AR(1) specification, and the yields-macro model is fairly better than the yields-only model at least for the longer forecast horizon.

3.4.2.2. Diebold-Mariano Test

We employ the Diebold and Mariano (1995) test for the squared forecast errors in order to get a more deep insight. Their test makes a direct comparison between the two competitive models for each maturity and each forecast horizon. We are interested in testing the hypothesis about equality of the squared errors (i) between the yields-only model and AR(1) specification of yield, and (ii) between yields-macro and yields-only models.

The main feature of Diebold and Mariano (DM, 1995) test of forecast accuracy lies in its direct applicability to quadratic loss functions, multi-period forecasts, and forecast errors that are non-Gaussian with non-zero mean and serial and contemporaneous correlation (correlated across maturities as well as over time). Defining the squared forecast errors as:

$$e_t = \left[R_{t+h}(m) - \hat{R}_{t,t+h}(m)\right]^2 \tag{3.24}$$

The basic component of Diebold and Mariano test statistic is the observed difference series of the squared forecast errors as:

$$d_t = e_{1t} - e_{2t} \tag{3.25}$$

where e_{1t} and e_{2t} are the quadratic loss functions of the two competing models as defined in

(3.24) for $t = T_0 + 1, T_0 + 2, \dots, T$. We test the null hypothesis $H_0: E(d_t) = 0$ against the alternative hypothesis $H_1: E(d_t) \neq 0$.

Assuming covariance stationarity and other regularity conditions on the process d_t, we use the standard result that:

$$\sqrt{(T - T_0)}\left(\bar{d} - \mu\right) \xrightarrow{d} N[0, 2\pi f_d(0)] \tag{3.26}$$

where $f_d(\cdot)$ is the spectral density of d_t and $\bar{d}$ is the sample mean of d_t over time. The Diebold-Mariano (*DM*) test statistic is computed as:

$$DM = \frac{\bar{d}}{\sqrt{2\pi\hat{f}_d(0)/(T - T_0)}} \sim N(0,1) \tag{3.27}$$

where $\hat{f}_d(0)$ is a consistent estimate of $f_d(0)$. This test also corrects for the autocorrelation that multi-period forecast errors usually exhibit.[41] Diebold and Mariano (1995) use a Newey-West type estimator for sample variance of the loss differential to account for this concern.[42]

We apply the Diebold and Mariano (1995) test to forecast errors of two pairs of models and the results are presented in table 3.10. In first step, we compare the forecast accuracy of yields-only and AR(1) model, and subsequently make a comparison of yields-macro model forecast errors with those of yields-only model in the second step for each maturity and each forecast horizon.

The results in the table 3.10 for the first pair point towards the universal significant difference of the RMSE for all the three horizons and all maturities forecasts of the yields-only and AR(1) model, as all *DM-stat* are significantly different from zero. The negative values shows that yields-only model outperforms all the competing forecasts of AR(1) specification of yield (in first pair e_{1t} and e_{2t} are the squared forecast errors functions of yields-only model and AR(1) model respectively).

The *DM-Stat* reported in table 3.10 for the second pair of models (yields-macro and yields-only), indicates a significant difference of the RMSE for one month ahead forecast of yields-macro and yields-only model. The p-value is equal zero for all maturities for $h = 1$. Most notably the negative values indicate superiority of yields-macro model forecasts as we consider e_{1t} and e_{2t} the quadratic loss functions of yields-macro and yields-only models respectively. Comparison of the 6 and 12 months ahead forecasts of both models specify that 11 out of 18 Diebold–Mariano statistics show a statistically significant (at 10% significance level) superiority of yields-macro model over the yields-only model. The results of Diebold and Mariano (1995) test suggest that the resilient predictive power of the yields-macro model at the 1-month-ahead horizon is very attractive for short-term bond trading activities and credit portfolio risk

[41] See, Diebold and Mariano (1995) for detail.
[42] See, Andrews (1991) for detailed econometric applications.

management. Furthermore, it also shows that such extended model (yields-macro model) can form the basis for predicting the stock market performance and state of economy in near future.

Table 3.10: Diebold-Mariano Test-statistic

Maturity	Yields-Only against AR(1) Model			Yields-Macro against Yields-Only Model		
	1 Month Forecast	6 Months Forecast	12 Months Forecast	1 Month Forecast	6 Months Forecast	12 Months Forecast
3	-10.111***	-3.565***	-2.239***	-3.811***	-1.253*	-0.949
6	-9.899***	-3.709***	-2.464***	-4.068***	-1.210*	-0.956
12	-11.510***	-2.356***	-3.544***	-3.994***	-1.299*	-1.398*
24	-12.008***	-2.108***	-1.929**	-4.007***	-1.741**	-2.574***
60	-8.103***	-2.155***	-2.115***	-4.740***	-3.143***	-1.858**
120	-3.881***	-1.396*	-1.822**	-4.117***	-1.193*	-1.515*
180	-2.444***	-2.265***	-1.349*	-3.910***	0.115	-0.934
240	-2.975***	-1.916**	-1.749**	-3.969***	0.827	-0.056
300	-3.387***	-1.268*	-1.257*	-3.985***	-1.531*	0.435

Note: The table presents Diebold–Mariano forecast accuracy comparison test results of the yields-only model against the AR(1) specification of yield and yields-macro model against the yields-only model for 1, 6, and 12 months ahead forecasts. The null hypothesis is that the two forecasts have the same root mean squared error. ***, ** and * show the statistical significance at 1%, 5% and 10% respectively.

Beside the Diebold and Mariano (1995) test to assess the overall quality of the out-of-sample forecasts of the models, we also employ the t-test for the mean equality of absolute forecast errors to evaluate the robustness of our forecast comparison tests results.

3.4.2.3. Mean Equality Test for the Absolute Forecast Errors

The mean equality test for the absolute forecast errors is based on the standard t-test. We employ the t-test for the absolute forecast errors in order to make a direct comparison between the models for each maturity and each forecast horizon. Defining, x_{it} as the absolute forecast errors for model i in period t, where $i = 1,2$ (models) and $t = 1,2,\ldots,T$, we test the null hypothesis $H_0\colon \mu_1 = \mu_2$ against the alternative hypothesis $H_1\colon \mu_1 \neq \mu_2$.

Under the normality assumption and inequality of sample variances for the paired samples, the *t-statistic* for the equality of means is computed as:

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{1}{T}(s_1^2 + s_2^2)}} \sim t_{(2T-2)} \quad (3.28)$$

where $\bar{x}_1$ and $\bar{x}_2$ are the sample means and s_1^2 and s_2^2 are the sample variances for the absolute forecast errors of the two competing models and T is the total number of observations. The *t-statistic* in (3.28) has *t*-distribution with $2(T-1)$ degrees of freedom under the null hypothesis.

Table 3.11 presents the results of testing the equality of absolute mean forecast errors. The absolute mean forecast errors of the yield-only model is universally smaller than that of the

AR(1), as the observed *t-stat* are negative and highly significant for all maturities and forecast horizons, as shown by the left panel of table 3.11. Furthermore, the yields-macro model outperforms the yields-only model for almost all the maturities and forecast horizons, with a few exceptions only for 6 and 12 months ahead forecasts at relatively longer maturities as shown by the right panel. It is worth noting that the t-test in table 3.11 gives almost the same results as those of the Diebold and Mariano test in table 3.10.

Table 3.11: t-test Statistic for Out-of-Sample Forecasts Accuracy Comparisons

Maturity	Yields-Only against AR(1) Model			Yields-Macro against Yields-Only Model		
	1 Month Forecast	6 Months Forecast	12 Months Forecast	1 Month Forecast	6 Months Forecast	12 Months Forecast
3	-9.509***	-8.196***	-3.802***	-6.421***	-2.183**	-2.178**
6	-10.148***	-9.289***	-3.902***	-6.153***	-3.972***	-2.155**
12	-7.254***	-4.467***	-2.222**	-5.538***	-4.215***	-4.662***
24	-8.172***	-5.089***	-2.125**	-5.726***	-2.137**	-3.078**
60	-6.034***	-3.439***	-1.698*	-4.929***	3.837***	-4.095***
120	-6.744***	-4.625***	-3.854***	-6.564***	-2.271**	-0.541
180	-3.601***	-1.662*	-2.114**	-4.082***	1.055	-0.004
240	-4.054***	-2.356**	-2.025**	-3.949***	0.741	-0.265
300	-3.779***	-2.243**	-4.872***	-5.647***	-0.529	-1.697*

Note: The table reports the t-test results of the mean equality of the absolute forecast errors of the yields-only model against AR(1) specification of yield and yields-macro model against the yields-only model for 1, 6 and 12 months ahead forecasts accuracy comparison. The null hypothesis is that the two forecasts errors (absolute values) have the same mean error. The degree of freedom for *t-statistic* is (94), (82) and (70) in one month, 6 months and 12 months ahead forecasts respectively. ***, ** and * show the statistical significance at 1%, 5% and 10% respectively.

The results of all the three tests in this study unanimously suggest that the Nelson-Siegel specifications outperform the competing benchmark forecasts models such as the AR(1) and random walk specifications.[43] Moreover, within the class of Nelson-Siegel models, the yields-macro model has an attractive and greater success than the yields-only model in forecasting the yields for short and medium-term maturities, and the former is still comparable to the latter for a longer horizon forecast.

3.5. Conclusion

The Nelson-Siegel framework of yield curve provides means for an effective time series analysis of yield data. This chapter has examined the role of macroeconomic and stock market variables in the dynamic Nelson-Siegel framework with the purpose of fitting and forecasting the term structure of interest rate on the Japanese government bond market. The yield curve model (proposed extended model) of this study explicitly incorporates both yields factors (level, slope, and curvature) and macroeconomic variables (overall economic activity, exchange rate, stock prices index (TOPIX) and inflation rate).

[43] An alternative benchmark model would be a random walk. We compared the forecast errors of the AR(1) and the random walk models by using the TRMSPE and the DM test (not reported here). The results show that the AR(1) specification is much superior to the random walk for forecasting the yields.

For the in-sample fit, we find that both the yields-macro and yields-only models are capable of distilling the term structure of interest rate quite well, and of describing the evolution and the trends of the Japanese government bonds market. The yields-macro model also leads to a better fit than the yields-only model does in terms of the MAE, RMSE as well as residuals autocorrelation. Regarding the term structure forecasts, the Nelson-Siegel framework of yield curve considerably outperforms the benchmark simple time series models such as an AR(1) and a random walk. For the short horizon forecasts, the out-of-sample predictability of the yields-macro model is superior to the yields-only model for all maturities examined in this study. And for the longer horizons, the former is still compatible to the latter. Moreover, the autocorrelation of forecasts errors of yields-macro model is much smaller and negligible as compared to the yields-only model, and to the persistency of errors in other related studies such as Bliss (1997), de Jong (2000), Diebold and Li (2006) and Ullah (2012).

The overall accuracy in the in-sample fit and the out-of sample forecasts of the proposed yields-macro model ranks it to be a good candidate among the various competing yield curve models in order to forecast the future yield, stock market performance and state of economy. This study also suggests that the correlation problem of residuals across maturities in in-sample fit and persistency of forecast errors could be avoided by incorporating the relevant macroeconomic and equity market factors in the statistical analysis of term structure models.

Chapter 4

Dynamics of the Term Structure of Interest Rates and Monetary Policy

Is Monetary Policy Effective during ZIRP?

4.1. Introduction

A long standing empirical literature has shown that monetary policy is a major factor in the movements of the yield curve. Among other works, Bernanke and Blinder (1992), Estrella and Hardouvelis (1991) and Mishkin (1990) explore the informational content of the spread between long and short-term yields (as an indicator of monetary policy) to forecast the future economic activity and inflation in the US market. They find that the slope of the term structure appears to carry information about future inflation and also provide evidence that an inverted yield curve reflects expectations of a declining rate of real activity. However, monetary transmission is complex and operates through many diverse channels. Other related studies in this regard, include Kozicky and Tinsley (2001), Svensson (2003), and Bernanke *et al.* (2005), have modeled the short-term interest rate as monetary policy instrument into the term structure framework. A common result of this strand is that the relation between the term spread and economic activity may be that the slope of yield curve reflects the stance of monetary policy. If the policymakers raise short-term interest rates, long-term rates are usually not increasing one-to-one with them but slightly less. Hence, the spread tightens and even might become negative. Higher interest rates slow down overall spending and, consequently, stagnates the economic growth. Therefore, a small or negative slope of the yield curve will be an indication for a slower growing economy and a decline in inflation in the future.

These models, however, fail to reflect the stances of monetary policy on real economic activity through the yields spread because of the unusual shape of Japan's yield curve, which is flatter for shorter maturities.[44] Furthermore, the short-term interest rate — the most important and conventional monetary policy instrument — has been bounded at zero in Japan and, hence, is no longer a policy instrument. The alternative may be monetary easing or some other

[44] This reflects the expectation that the Bank of Japan (BOJ) will not raise the target rate for a considerable period.

non-conventional measures.[45] The goal of the central bank is, therefore, to affect the economy across the yield curve, bringing down long-term rates, thereby, boosting the economy. Therefore, it is more appropriate to consider the impact on medium to long-term maturities yields rather than only the term spread, as they are the fundamental conduits for the transmission of monetary policy.[46]

This chapter addresses this issue by formulating a yield curve model that integrates monetary policy as well as real economy factors in the term structure model. The objective is to examine the effectiveness of such policy (monetary easing and non-conventional policy tools) in affecting the yield curve, using the Japanese experience of zero interest rate policy (ZIRP) and quantitative easing monetary policy (QEMP). To be precise, we are interested to figure out the transmission mechanism through which the monetary policy affects the real economy. In evaluating the effectiveness of QEMP during the ZIRP period, we focus on the expectation channel. The effectiveness of such channel depends on the credible commitment of the central bank of maintaining the future policy rate at zero.[47] The desired intermediate effect of the monetary policy will be that the reduction of expected future short-term rates will be transmitted to the long end of the yield curve. The decline in the long-term interest rates will, in turn, lead to increased expectations of inflation and will stimulate activity.

Given that the conventional monetary policy instrument, the short-term interest rates, is not effective in combating the deflationary cycle and debt deflation after the short-term interest rate has reached zero because the policy instrument cannot be lowered further, we aim to understand:

— Should the central bank just watch things deteriorating in the cyclical process and wait for the invisible hand, expecting that things will get better soon (economy will get recovered)?
— Should the central bank use tools that are beyond conventional policy instruments to get the economy out of a deflationary cycle?

To respond to these questions, we have to evaluate the transmission mechanism of policy shocks through the short, medium, and long-term bond yields to the real economy. For this purpose, we use a three-factor term structure model, based on the classic contribution of Nelson and Siegel (1987). We incorporate three macroeconomic variables, i.e., the level of economic activity, money supply and inflation rate in the state-space representation along with stochastic volatility component in the yield curve model.

---

[45] Non-conventional tools include sale/purchases of foreign currency denominated bonds, equities (indexed equity funds listed on the exchange) and real estate (funds), and non-sterilized interventions.

[46] We are aware of the previous studies which show that policy duration has a significant effect on the yield curve (Oda and Ueda, 2007; Okina and Shiratsuka, 2004 and Baba *et al.* 2005), but they do not examine the transmission of this positive effect to the real economy. Furthermore, in Ang and Piazzesi (2003), Hördahl *et al.* (2006), Wu (2002), Evans and Marshall (2007), and Nakajima *et al.* (2010) only few specific yields are taken into consideration and the impact, therefore, does not transmit to the real economy. Moreover, in Bernanke *et al.* (2005), the period of the BOJ's policy is not considered and their empirical results provide only limited insights on monetary policy.

[47] In the commitment regime of monetary policy, the central bank does not optimize the quadratic loss function every period, but, makes the credible commitments about what it will do in the future. By promising to take a certain action in future, the central bank can influence the public's expectation about future policy rates and inflation.

The motivation to add a common stochastic volatility component in the yield curve model is to allow the model to capture latent exogenous shocks that affect the entire yield curve and are not captured by the three factors structure of level, slope and curvature. This extension increases the flexibility of the term structure model to accurately estimate the impact of macroeconomic policies shocks on yield curve after controlling for the influence of latent exogenous shocks.[48] Furthermore, considering the Japanese case is particularly interesting, because its economy has become very unstable and experienced significant institutional and monetary strategy changes during the last decade. In addition, taking into account the entire term structure instead of using only few yields in the observation equation might contribute to the economic interpretation of curvature factor of yield curve.[49]

The study contributes to the existing literature in three ways. The first one is methodological. In calibration the multi-factors Nelson-Siegel model, we include the common stochastic volatility component that follows the EGARCH process, while adopting the state-space approach. The second, instead of using the conventional monetary policy tool, i.e., the short rates, we include the growth rate of money supply as indicator of QEMP to figure out the transmission mechanism of monetary policy to real economy through the yields of entire maturity spectrum. The third, we attribute an economic interpretation to the third unobservable component of the term structure, i.e., the curvature factor.

We proceed as follows. Section 4.2 deals with the term structure model that integrates the yield curve factors with macroeconomic variables and explains the estimation method, while the data structure and estimation results are presented in 4.3. In section 4.4, we relate our framework to the expectation hypothesis. Finally, section 4.5 presents the conclusion of the chapter.

4.2. Term Structure Model and Estimation Method

The macro-finance literature has convincingly advocated the case for the existence of bidirectional link between the term structure and rest of the economy. To this end, we design a dynamic Nelson-Siegel (DNS) yield curve model with macroeconomic variables in the state-space framework that also allows for the time-varying stochastic volatility in yields for various maturities. We use the standard EGARCH specification to describe the volatility process of a common shock in the yields, while adopting the state-space approach. Adding a common stochastic volatility component increases the flexibility of the term structure model and enables it to fit attractively the more complex shapes of the yield curve.[50] In this section, we discuss the

[48] The inclusion of common volatility latent factor also enables the three factors structure of yield curve of Nelson and Siegel (1987) to fit attractively more complex shapes of the yield curve. Moreover, the existence of a common volatility component can be of great importance to interest rate option traders, who manage risk in an entire book of interest rate volatility positions. Knowledge of a common component that determines volatilities in different parts of the yield curve allows traders to mitigate overall risk in the trading book by taking offsetting positions in different yields along the curve.

[49] Given that the yield curve is summarized by three latent factors, i.e., level, slope and curvature, which can be obtained through Kalman filtering.

[50] As Koopman *et al.* (2010) has shown by plotting some fitted curves, they find that allowing for time-varying volatility significantly increases the likelihood value relative to the traditional DNS model.

concept of time-varying factors and volatility in the DNS model. First, in subsection 4.2.1, we describe the model that incorporates macroeconomic variables as well as the common stochastic volatility term in state-space representation. The latent factors model is considered, since it will be a convenient vehicle for introducing the state-space representation. Second, subsection 4.2.2 presents the estimation procedure of the model in the state-space framework using the Kalman filter algorithm.

4.2.1. Yields-Macro Factors Model

An intuitive way to represent our model is to cast the Nelson-Siegel (1987) functional form into state-space framework, which assumes that information about the term structure of interest rates can be summarized by three factors, i.e., the level, slope and curvature of the yield curve, as:

$$R_t(m_i) = \beta_{1t} + \beta_{2t}\left[\frac{1-\exp(-m_i/\tau)}{m_i/\tau}\right] + \beta_{3t}\left[\frac{1-\exp(-m_i/\tau)}{m_i/\tau} - \exp\left(\frac{-m_i}{\tau}\right)\right] + \varepsilon_t \qquad (4.1)$$

where $R_t(m_i)$ is the zero-coupon yield for maturity m at time $t, i = 1,2, \dots, N;\ t = 1,2, \dots, T$, $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the unobservable vector of three latent factors of level, slope and curvature respectively. The constant parameter τ is the decay parameter of the factor loading of the yield curve slope and also determines the optimum point of the curvature factor loading.

Examination of the limits of the Nelson-Siegel model shows, where the interpretations of the factors come from, when time to maturity goes to infinity, we find the infinitely long end of the curve, which is given by $lim_{m\to\infty} R_t(m_i) = \beta_{1t}$. Given the fact that the first factor loading is equal to 1, the β_{1t} gets the interpretation of the level factors. Letting time to maturity goes towards zero, the infinitely short end of the curve can be obtained as, $lim_{m\to 0} R_t(m_i) = \beta_{1t} + \beta_{2t}$, meaning that the short rate is influenced by the first and second factors. Defining the slope of the yield curve as the long minus the short end, it can be seen that it is given by $-\beta_{2t}$. The third factor loading in (4.1) approaches zero in both cases, when time to maturity goes to zero or infinity and is positive for intermediate values of m. Therefore, β_{3t} affects the middle part of the yield curve and, hence, is interpreted as the curvature factor in the DNS model.

Regarding the error term, ε_t, in the Nelson-Siegel model, the earlier studies assume that $\varepsilon_t \sim N(0, \sigma^2 I_N)$. However, the interest rates are the result of trading at financial markets, therefore, the volatility in the series may have changed over time as well. That's why, we assume that:

$$\varepsilon_t = \Gamma_\varepsilon \varepsilon_t^* + \varepsilon_t^+ , \qquad \varepsilon_t^+ \sim N(0, \Omega) \qquad (4.2)$$

where Γ_ε and ε_t^+ are (N×1) vectors of loadings and noise component respectively, and ε_t^* is a scalar representing the common disturbance term. In this model ε_t^* and ε_t^+ are independent. The loading factor, Γ_ε, determines how sensitive the different yields are to the common shock.

The distribution of the common volatility component, ε_t^*, given the information up to time $t-1$ (denoted by ζ_{t-1}) is:

$$\varepsilon_t^* | \zeta_{t-1} \sim N(0, h_t) \tag{4.3}$$

where h_t follows the EGARCH specification, which is given by:[51]

$$\log(h_t) = \gamma_0 + \gamma_1 \frac{\varepsilon_{t-1}^*}{\sqrt{h_{t-1}}} + \gamma_2 \log(h_{t-1}) + \psi \left(\left| \frac{\varepsilon_{t-1}^*}{\sqrt{h_{t-1}}} \right| - \mathbb{E} \left[\left| \frac{\varepsilon_{t-1}^*}{\sqrt{h_{t-1}}} \right| \right] \right) \tag{4.4}$$

where $\mathbb{E}\left(\left|\varepsilon_{t-1}^* / \sqrt{h_{t-1}}\right|\right)$ is the expectation of the absolute value of a standard normally distributed random variable, which is equal to $\sqrt{2/\pi}$. The volatility at $t = 1$ is set equal to the unconditional expectation of the log variance, which is $\mathbb{E}[\log(h_t)] = \gamma_0(1 - \gamma_2)^{-1}$. This specification for variance dynamics enable the common volatility component in the DNS model to account for asymmetric response to positive and negative shocks.

As far as the macro variables are concerned, we include three key variables: the annual growth rate in industrial production (IP_t) and money supply (MS_t) and annual price inflation (INF_t). These variables represent, respectively, the level of real economic activity, monetary policy stances and the inflation rate, which are widely considered to be the minimum set of fundamentals needed to capture basic macroeconomic dynamics.[52]

Diebold *et al.* (2006) and many others find that the time series of estimated factors of Nelson-Siegel model are highly persistent, which implies that these can easily be modeled as AR(1) or VAR(1).[53] Furthermore, it allows to model the macroeconomy related factors in Nelson-Siegel model to assess its dynamic interaction with the yield curve factors. We assume that the yield curve latent factors vector β_t along with the three macroeconomic factors follow a vector autoregressive process of first order, which allows us to formulate the yield curve latent factor model in the state-space form and to use the Kalman filter for obtaining maximum-likelihood estimates of the hyper-parameters and the implied estimate of β_t. Furthermore, the time-varying variance, h_t, depends on the past values of the unobserved common disturbance term, ε_t^*, which, therefore, has to be treated as a latent variable and should be included in the state vector.

In the state-space representation the complete model with observation equation (4.5) and state equation (4.6) can be written as:

$$\begin{bmatrix} R_t(m) \\ Z_t \end{bmatrix} = \begin{bmatrix} \Lambda(\tau) & \Gamma_\varepsilon & 0 \\ 0 & 0 & I_3 \end{bmatrix} \begin{bmatrix} \beta_t \\ \varepsilon_t^* \\ \tilde{Z}_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t^+ \\ 0 \end{bmatrix} \tag{4.5}$$

$$\alpha_{t+1} = \begin{bmatrix} (I_6 - A)\mu \\ 0 \end{bmatrix} + \begin{bmatrix} A & 0 \\ 0 & 0 \end{bmatrix} \alpha_t + \begin{bmatrix} v_{t+1} \\ \varepsilon_{t+1}^* \end{bmatrix} \tag{4.6}$$

[51] Financial markets respond in different ways to positive and negative shocks and it is a common knowledge that volatility tends to increase quickly when negative news reaches to traders and investors, whereas, positive news usually has a much less pronounced effect.

[52] For the expected theoretical relation between macroeconomic and yield curve factors, see Ullah *et al.* (2013).

[53] Using the Japanese market data Ullah *et al.* (2013) and Ullah (2012) find that the three latent factors of yield curve are highly persistent and VAR(1) specification is more appropriate than the AR(1) as well as random walk specifications.

$$\begin{bmatrix} \varepsilon_t^+ \\ v_{t+1} \\ \varepsilon_{t+1}^* \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 & 0 \\ 0 & \Sigma_v & 0 \\ 0 & 0 & h_{t+1} \end{bmatrix}\right) \tag{4.7}$$

where $\alpha_t = \left(\beta_{1t}, \beta_{2t}, \beta_{3t}, \widetilde{IP}_t, \widetilde{MS}_t, \widetilde{INF}_t, \varepsilon_t^*\right)'$ is (7×1) latent vector, $R_t(m)$ is (N×1) vector of zero-coupon yields, $Z_t = (IP_t, MS_t, INF_t)'$ is (3×1) vector of macroeconomic factors, β_t is (3×1) vector of Nelson-Siegel factors, $\Lambda(\tau)$ is (N×3) matrix of factors loadings, A is (6×6) matrix of parameters, μ is (6×1) mean vector of factors, I_6 and I_3 are (6×6) and (3×3) identity matrices respectively and Γ_ε is (N×1) vector. Σ_v is (6×6), the covariance matrix of innovations of the transition system and is assumed to be unrestricted, while the covariance matrix Ω of the innovations to the measurement system of (N×N) dimension is assumed to be diagonal. The latter assumption means that the deviations of the observed yields from those implied by the fitted yield curve are uncorrelated across maturities and time.[54] Furthermore, the variance of ε_{t+1}^* is h_{t+1} and will be modeled as EGARCH process, specified in (4.4).

Moreover, in (4.7), we assume that the innovations, ε_t^+ and v_t, as well as common volatility component, ε_t^*, have Gaussian distribution. The model in equations (4.4 – 4.7) provides a flexible framework for analyzing the interaction between the yield curve and macroeconomy, while simultaneously accounts for the time-varying stochastic volatility in yields for all maturities. In addition, the proposed specification guarantees positive forward rates at all horizons and a discount factor that approaches to zero as maturity increases.

4.2.2. State-space Estimation of the Model

In this subsection, the estimation procedure based on the Kalman filter for the dynamic Nelson-Siegel model with time-varying volatility is explained. For convenience, we introduce some new notations and rewrite the signal and state equations in (4.5) and (4.6) respectively, to obtain the generalized form of DNS model with time-varying volatility in state-space form.

$$y_t = H\alpha_t + w_t \tag{4.8}$$

$$\alpha_{t+1} = C + K\alpha_t + Gu_{t+1} \tag{4.9}$$

$$\begin{bmatrix} w_t \\ u_t|\zeta_{t-1} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} R & 0 \\ 0 & Q_t \end{bmatrix}\right) \tag{4.10}$$

where the expressions of y_t, α_t, H, C, K, G, u_t and w_t are given in appendix E.

The Kalman filter algorithm consists of two steps to find a minimum mean squared error estimate of the latent vector α_t, namely the prediction and the update steps. At a given time t, we form an optimal prediction of y_t based on all information available up to time $t-1$, denoted by $y_{t|t-1}$. This prediction can be made using (4.8) and $\hat{\alpha}_{t|t-1}$, which can be calculated using (4.9) and $y_{t-1|t-1}$. After obtaining the prediction on y_t, the prediction error $\eta_{t|t-1}$ and its covariance

[54] Given the large number of observed yields used, the diagonality assumption of covariance matrix of the measurement errors is necessary for computational tractability.

matrix $F_{t|t-1}$ can be calculated to obtain information on α_t that is not yet contained in $y_{t|t-1}$. In the update step the estimate of α_t at time t using information up to time $t-1$, $\hat{\alpha}_{t|t-1}$ is updated by incorporating the new information from the prediction error to obtain $\hat{\alpha}_{t|t}$. The estimate $\hat{\alpha}_{t|t}$ contains information up to time t. The prediction step is summarised by the following four equations:

$$\hat{\alpha}_{t|t-1} = C + K\hat{\alpha}_{t-1|t-1} \quad (4.11)$$
$$P_{t|t-1} = KP_{t-1|t-1}K' + GQ_tG' \quad (4.12)$$

with

$$\eta_{t|t-1} = y_t - H\hat{\alpha}_{t|t-1} \quad (4.13)$$
$$F_{t|t-1} = HP_{t|t-1}H' + R \quad (4.14)$$

and the update step is described by the two equations given as follows:

$$\hat{\alpha}_{t|t} = \hat{\alpha}_{t|t-1} + P_{t|t-1}H'F_{t|t-1}^{-1}\eta_{t|t-1} \quad (4.15)$$
$$P_{t|t} = P_{t|t-1} - P_{t|t-1}H'F_{t|t-1}^{-1}HP_{t|t-1} \quad (4.16)$$

where P_t is the covariance/MSE matrix of $\hat{\alpha}_t$ in the prediction and update steps. These equations enable the Kalman filter to estimate all latent variables recursively for $t = 1,2,\dots,T$.

Matrix Q_t contains h_{t+1} that is modeled by EGARCH process and relies on latent shocks at time t, which are unobservable. Kim and Nelson (1999) shows that taking conditional expectation of the latent variables in (4.4) gives:

$$\log(h_t) = \gamma_0 + \gamma_1\mathbb{E}\left(\frac{\varepsilon_{t-1}^*|\zeta_{t-1}}{\sqrt{h_{t-1}}}\right) + \gamma_2\log(h_{t-1}) + \psi\mathbb{E}\left(\left|\frac{\varepsilon_{t-1}^*|\zeta_{t-1}}{\sqrt{h_{t-1}}}\right| - \mathbb{E}\left[\left|\frac{\varepsilon_{t-1}^*|\zeta_{t-1}}{\sqrt{h_{t-1}}}\right|\right]\right) \quad (4.17)$$

where the estimate of $\mathbb{E}(\varepsilon_{t-1}^*|\zeta_{t-1})$ is the last element of $\hat{\alpha}_{t-1|t-1}$ from the filtering/update step. In order to start the recursion, the initial value for α_t is set equal to the unconditional mean, $\alpha_{1|0} = \mathbb{E}(\alpha_t) = 0$, and the initial covariance matrix of the state vector, $P_{1|0}$, is:

$$P_{1|0} = \begin{bmatrix} \Upsilon & 0 \\ 0 & h_1 \end{bmatrix} \quad (4.18)$$

where Υ is chosen such that $\Upsilon - W\Upsilon W' = \Sigma_v$ and h_1 is the unconditional expectation of the log variance defined in section 4.2.1.[55] This initiation enables the Kalman filter to provide a minimum mean squared error estimate of α_t at every time $t = 1,2,\dots,T$, given information up to time $t-1$ and given the hyper-parameters.

The Kalman filter provides estimates for the latent variables and the unknown hyper-parameters have to be estimated using maximum likelihood method. Collecting all

[55] We define vector ξ_t consists of the first six elements of α_t vector and model it as:

$$\xi_t = W\xi_{t-1} + v_t, \qquad v_t \sim N(0,\Sigma_v)$$

where $\xi_t = (\beta_{1t}-\mu_1, \beta_{2t}-\mu_2, \beta_{3t}-\mu_3, IP_t-\mu_4, MS_t-\mu_5, INF_t-\mu_6)'$, W is (6×6) and Σ_v is (6×6) covariance matrices of error term v_t. We derive the unconditional mean and covariance of ξ_t, which is summarized as $\xi_t \sim N(0,\Upsilon)$. For detail of initializing the Kalman filter, see Hamilton (1994).

unknown parameters of the measurement and state equations into $\theta = (\tau, \mu, A, \Omega, \Sigma_v, \Gamma_\varepsilon, \gamma_0, \gamma_1, \gamma_2, \psi)$, and assuming that ε_t^+ and v_t are normally distributed, the distribution of y_t conditional on ζ_{t-1} is also Gaussian as $y_t|\zeta_{t-1} \sim N(y_{t|t-1}, F_{t|t-1})$; hence, the Gaussian log likelihood is given by:

$$\log L\ (\theta) = -\frac{NT}{2}\log(2\pi) - \frac{1}{2}\sum_t \log(|F_{t|t-1}(\theta)|) - \frac{1}{2}\sum_t \left(\eta'_{t|t-1}[F_{t|t-1}(\theta)]^{-1}\eta_{t|t-1}\right) \quad (4.19)$$

Numerical optimization of the log likelihood function (4.19) yields maximum likelihood estimates of the hyper-parameters. The process to find the latent factors and consistent estimates of the hyper-parameters is recursive one. The procedure is started by initiating the recursion using certain starting values for the hyper-parameters (θ^0) that enable the Kalman filter to obtain estimates of the latent factors (α_t^0), conditional on the initial choice for the parameters. Subsequently, given (α_t^0), the likelihood function (4.19) is maximized in the optimization step to obtain new estimates of the hyper-parameters, (θ^1), that yield a higher likelihood. These estimates are used in the Kalman filter again to obtain new estimates of latent factors, (α_t^1) and the corresponding likelihood value and so on. These recursive steps in the algorithm continue until the estimates of the hyper-parameters converge and we find the optimum of the likelihood function.[56]

4.3. Empirical Results

We have constructed a monthly time series panel of unsmoothed Fama-Bliss zero-coupon yields for Japanese treasuries of different maturities between 2000 and 2011. We combine this panel with a data set of macroeconomic time series for the same sample period. The details of the data set are provided in section 4.3.1. The estimation results for the joint interaction of macro and yield curve factors along with the EGARCH results are presented in section 4.3.2. Section 4.3.3 presents the results of some formal statistical tests of contemporaneous and lagged interaction between macro and yield curve factors. Finally, in section 4.3.4 and 4.3.5, we discuss the estimation results for macroeconomic and yield curve factors impulse response functions and variance decompositions respectively.

4.3.1. Data Description

We consider the Japanese government bond yields with maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240 and 300 months. The yields are derived from bid/ask average price quotes, from January 2000 through December 2011, using the Fama and Bliss (1987) methodology. Yields extracted from pricing data by Fama-Bliss method exactly (precisely and accurately) price the included bonds and facilitate to avoid having negative spot rates during the lower short-term interest rate period. For the macroeconomic variables, we use data on the following three variables: the annualized growth of industrial production (IP_t), the

[56] We set the convergence criterion of (10^{-6}) for the change in the norm of the parameter vector θ from one iteration to the next.

growth rate of $M2$ money supply (MS_t) as an indicator of monetary policy, and inflation rate (INF_t), measured as annualized monthly changes in the consumer price index.

The data for the price quotes of Japanese government bonds is taken from the Japan Securities Dealers Association (JSDA) bonds files, while for the three macroeconomic variables is obtained from the International Financial Statistics (IFS) published by International Monetary Fund (IMF). Summary statistics of the yields for various maturities along with the three dimensional plot of the data set and the descriptive statistics of the macroeconomic variables is provided in section 1.4.

4.3.2. Estimation Results of the Model

To estimate the dynamic factors model, we use the Kalman filter algorithm suggested in Hamilton (1994). For given values of the system matrices, the Kalman filter is used to evaluate the log likelihood function via the prediction error decomposition. The maximum likelihood estimates of the unknown parameters are obtained via the numerical optimization of the Gaussian log likelihood function by iterating the Marquardt algorithm, using numerical derivatives. The Kalman filter is initialized using the unconditional mean (zero) and unconditional covariance matrix of the state vector, which are derived from the Gaussian distribution for the first 6 components in state vector, given that the innovations of both signal and state equations are normally distributed.

The estimation results of the parameters of state equation are presented in the first panel of table 4.1. High persistency in the yield curve latent factors can be seen from the diagonal elements of the coefficient matrix, all being close to one, however the lagged own dynamics of macroeconomic variables are not as strong as of the yield curve factors.[57] Moreover, the lagged value of the second factor, which proxies for the slope of yield curve, has a significant influence on the level factor along with the statistically significant lagged impact of money supply and inflation rate on industrial production and of MS_{t-1} on inflation rate. This significant relation encourages the use of a VAR model to describe the dynamics of the latent factors in the dynamic Nelson-Siegel model instead of the more parsimonious AR(1) specification. Regarding the yield macro dynamics, industrial production is positively while the growth rate of money supply is negatively related to the overall yield level. Furthermore, the slope factor is also affected by these two variables, i.e., positively by IP_{t-1} and negatively by MS_{t-1}. This suggests that monetary policy shocks account for significant fluctuations in the yield curve shape and policy shocks to short-term interest rates are likely to affect the medium to long-term interest rates. One important channel, through which monetary policy works, is the long end of yield curve, shaping them so that, in turn, they affect the level of economic activity. This relation is consistent with the expectation hypothesis of the yield curve theory. Moreover, the inflation rate has a positive statistically significant impact on the yield curve curvature.

[57] The own-lag coefficient of the second and third factor of yield curve (β_{2t} and β_{3t}) are greater than 0.9, while of β_{1t} and MS_t are greater than 0.8. However, stationarity is assured, because the largest eigenvalue of the matrix A is 0.920. No value lies outside the unit circle.

Table 4.1: Latent Factors VAR(1) and EGARCH Models Parameters Estimates

Panel 1: Latent Factors VAR(1) Model Parameters Estimates

	μ	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	IP_{t-1}	MS_{t-1}	INF_{t-1}
β_{1t}	**3.011**	**0.871**	**0.021**	0.018	**0.019**	**-0.070**	-0.019
	(0.222)	(0.041)	(0.005)	(0.214)	(0.001)	(0.008)	(0.017)
β_{2t}	**-2.794**	-0.227	**0.920**	-0.002	**0.031**	**-0.038**	-0.017
	(0.210)	(0.163)	(0.286)	(0.703)	(0.001)	(0.004)	(0.093)
β_{3t}	**-2.127**	-0.285	-0.189	**0.908**	-0.161	0.014	**0.126**
	(0.825)	(0.314)	(0.459)	(0.056)	(0.217)	(0.016)	(0.050)
IP_t	**2.500**	-0.314	**-0.838**	-0.147	**0.573**	**0.457**	**0.333**
	(0.267)	(0.239)	(0.217)	(0.104)	(0.080)	(0.193)	(0.217)
MS_t	**4.573**	0.187	**1.182**	**-0.645**	-1.076	**0.889**	0.305
	(0.305)	(0.135)	(0.627)	(0.214)	(0.917)	(0.185)	(0.877)
INF_t	**-0.012**	**-0.640**	0.172	0.232	-0.599	**0.258**	**0.689**
	(0.002)	(0.073)	(0.872)	(0.976)	(0.931)	(0.037)	(0.039)

Panel 2:Test for the Joint-Significance of Individually Insignificant Coefficients

Test Statistic	Value	df	P-value
Wald Statistic	41.557	17	0.000
LR Statistic	39.262	17	0.000

Panel 3: EGARCH Model Parameters Estimates

γ_0	γ_1	γ_2	ψ
-0.217	**0. 861**	**0. 229**	**0.973**
(0.018)	(0.128)	(0.018)	(0.407)

Note: The table reports the estimates for the parameters of the transition equation of yields-macro factors dynamics and of EGARCH specification in the dynamic Nelson-Siegel model. Panel 1 presents the estimates for the vector μ and matrix A, while panel 2 shows the results of the Wald-test and likelihood ratio (LR) test for the joint significance of individually insignificant coefficients in matrix A. The null hypothesis is that insignificant coefficients are simultaneously equal to zero. Both the test statistics are Chi-square with their respective degrees of freedom (df). P-value is the probability value of the test statistic. Panel 3 shows the parameters' estimates of the volatility processes (EGARCH) of the common component in the yield curve model. The standard errors are in parenthesis. Bold entries denote parameter estimates significant at the 5 percent level.

It is interesting to observe that the impact of yield curve factors on macroeconomy is much stronger than of macro on yield curve factors. The spread factor, often used as a predictor of economic recessions, has a negative significant effect on the level of economic activity and positive effect on money supply, suggesting that a decrease in the slope of yield curve (becoming flat or negatively sloped) can be considered as a signal of economic slowdown. It also confirms that the BOJ responds with more expansionary monetary policy in subsequent period, when the yield curve becomes flat and vice versa. Furthermore, the curvature factor is negatively related to money supply, implying that the curvature factor contains information about the expected stance of monetary policy and could also be informative about the evolution of the economy (Monch, 2006).[58] The results also confer that the level of economic activity and money supply do not responds to the level factor rapidly. They may take more time to respond. Moreover, the negative lagged impact of level factor on inflation rate suggests that as the long end of yield curve goes

[58] It shows that the BOJ responds with more expansionary policy as the yield curve becomes flat (curvature falls).

down, inflationary expectations become stronger as a consequence of rise in aggregate demand. It indicates that the long end of yield curve contains important information about the future inflation. Finally, the parameter τ is estimated at 59.795 with a standard error of 0.959, indicating that the estimate is highly significant. It implies that the loading on the curvature factor is maximized at a maturity of about 5 years.

As many of the coefficients in matrix A are statistically insignificant, Wald-test and LR (likelihood ratio) test for their joint significance are employed and the results are presented in the second panel of table 4.1. Both the test statistics reject the null-hypothesis of joint insignificance of the 17 individually insignificant coefficients in the state equation. This suggests that inclusion of macroeconomic factors in the Nelson-Siegel specification of yield curve improves the model's overall fit and prediction power (Ullah *et al.* 2013).

Financial market volatility in many prior studies is characterized by asymmetric volatility rather than symmetric. For example, stock market volatility tends to surge when indices are falling and revert back to normal levels only gradually when prices increase. This phenomenon is also present in interest rate markets, as studied by Dungey *et al.* (2009), who find US treasuries to increase (and, hence, the yields fall) in volatile times. Yet, the standard GARCH model is not able to allow for different responses of volatility to negative and positive shocks and, therefore, implies a perfectly symmetric structure on the volatility process. In order to allow for asymmetric dynamics, we estimate the EGARCH specification of the volatility process for the common component in the Nelson-Siegel model. The third panel of table 4.1 presents the estimates of the parameters for the EGARCH specification given by equation (4.4). The results support the hypothesis of asymmetric volatility dynamics in the common shock component as all parameters, including ψ, are statically significant, supporting the finding of Dungey *et al.* (2009).

The high and significant estimate of the γ_1 indicates that much weight is put on recent shocks. The lag volatility coefficient γ_2 in the EGARCH equation is low but statistically different from zero. Therefore, the volatility of the common component is highly sensitive to the latest innovations; it increases quickly with large shocks and reverts back soon thereafter.

In order to obtain a better insight, in panel (a) of figure 4.1, common volatility (h_t) is plotted over time along with the estimate of Γ_ε vector (indicating the sensitivity of various maturities yields to volatility process of the common component) in panel (b). Some historical events are clearly illustrated in the graph. Firstly, the spike in May 2002 coincides exactly with the sharp fall in the slope of yield curve during this period (see figure 4.2). This is the turning point of monetary policy regime in the Japanese economy as discussed in Ito and Mishkin (2004), Ito (2005) and Ullah *et al.* (2013). During this period, the Bank of Japan (BOJ) launched quantitative easing (QEMP) to affect long-term interest rates in order to stimulate the economy.

The second hike corresponds to the jump in forward rates. Forward rates jumped up in mid-2004, because of higher expectation of an exit from deflation in the near future. However, it fell in the latter half of 2004 because of the BOJ released "Outlook for Economic Activity and Prices" in October 2004, in which slight positive inflation rate in terms of change in CPI was expected by policy board members. During this period, the prices of JGBs (Japanese Government Bonds) declined rapidly and the long end of yield curve rotated upward, because of (i) economic recovery as the growth rate increased close to 6% (quarter to quarter rate, annualized) in 2003:IV and first 2 quarters of 2004[59] and (ii) the BOJ adopted some of the non-conventional measures of monetary policy including purchases of long bonds to boost the economy. So, during the period of an increase in forward rates and rapid economic recovery, the interest rate on long-term bond rises and, hence, the volatility in bond market as well.[60]

During the global financial crisis of 2008, the time series shows an increase in volatility too.[61] Yet, in the two small spikes, the latest corresponds to the start of the problems concerning the Eurozone in early 2010. Furthermore, the response to global financial crisis of 2008 is a bit stronger than to recent Eurozone sovereign debt crisis.

Overall, the estimated stochastic volatility pattern over time shows that volatility is high during the quantitative easing monetary policy regime while drops sharply in the post QEMP periods. Apparently, the panic following the two former domestic policy shocks caused extreme surges in volatility of the common component in the dynamic Nelson-Siegel EGARCH model as compared to the last two global events. Yet, this should not be surprising as these events actually concerned sovereign debt of the global financial markets and, therefore, cannot be directly translated into the Japanese government treasury yields.

Panel (b) of figure 4.1 plots the loadings in the vector Γ_{ε} against maturity. For the 3 months maturity the loading is fixed at the value of 1 in order to overcome identification issues. The overall pattern of loadings across maturities is roughly similar to that of Koopman *et al.* (2010), who find a remarkably lower sensitivity of the 12 months and 9 years maturities. The 9 years is the before last maturity in their sample and we, therefore, compare it to the 20 years yield, for which the estimated loading is nearly zero. Hence, it seems that regardless of the choice of maturities included in the sample, the dynamic Nelson-Siegel EGARCH model fits the loadings to the common component in such a manner that the shortest and longest maturities are more sensitive to common shocks and the maturities in between much less.[62] Overall, the plot shows that the short rates are more sensitive to shocks than the long rates. The estimated sensitivity of

---

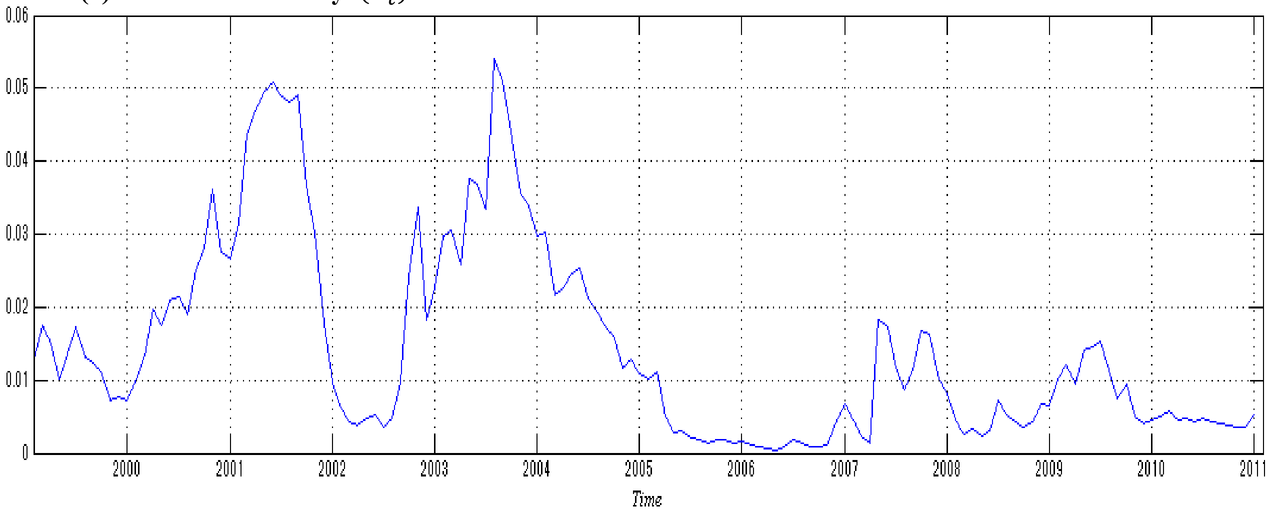
[59] For details, see the report "Japanese Economy 2004 Prospect for Continuous Recovery and Risks" (Summary) Directorate General for Economic Research, Cabinet office, December 2004.

[60] After a significant fall in stock prices in 2002 and again in the spring of 2003, the stock market regained confidence for the rest of the year, as did the economy as a whole. The Nikkei 225 index recorded a low of 7,600 in April 2003, less than one fifth of the peak at the end of 1989.The growth rate increased close to 6% (quarter to quarter rate, annualized) in 2003:IV and 2004:I. Optimism spread to the economy. The size of deflation shrank from about 1% to near zero by the end of 2004. Furthermore, in mid-2004, the BOJ adopted some of the non-conventional measures proposed by critics, including purchases of long bonds (Ito and Mishkin, 2004 and Ito, 2005).

[61] A financial crisis that arose in the US mortgage market in late-2007 after a sharp increase in mortgage foreclosures, mainly subprime, collapsed numerous mortgage lenders and hedge funds. The meltdown spilled over into the global credit market as risk premiums increased rapidly and capital liquidity was reduced. The sharp increase in foreclosures and the problems in the subprime mortgage market were largely blamed on loose lending practices, low interest rates, a housing bubble, and excessive risk taking by lenders and investors.

[62] Here the shortest rate means yield of maturity of one year, whereas, the longest rates corresponds to yields of maturities of 25 years and beyond.

Panel (a): Common Volatility (h_t)



Panel (b): Common Volatility Component Loadings (Γ_ε)

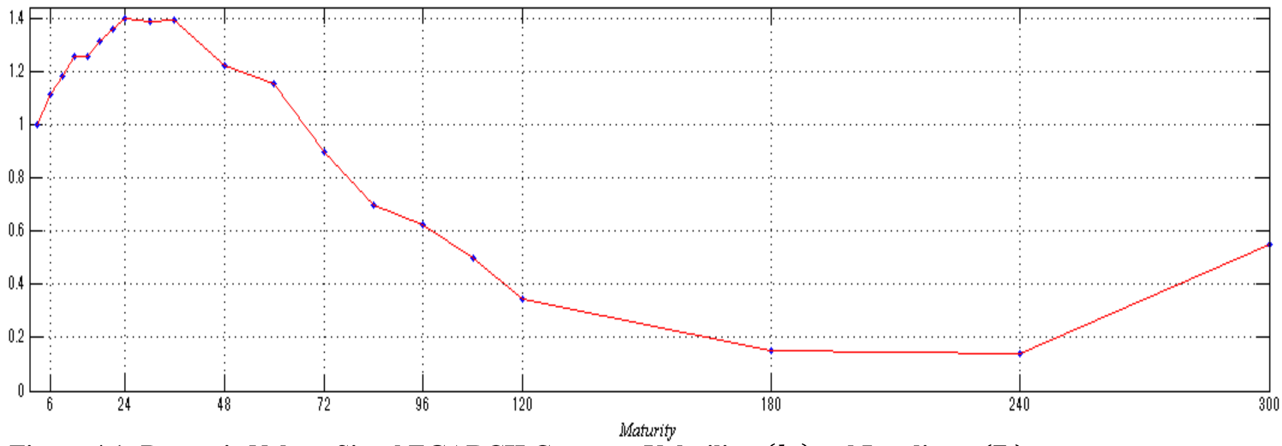


Figure 4.1: Dynamic Nelson-Siegel EGARCH Common Volatility $(h_t)$and Loadings (Γ_ε)
The figure shows a plot of the volatility (h_t) of the common shock component (ε_t^*) over time for the dynamic Nelson-Siegel EGARCH model in panel (a). Panel (b) plots the loadings for the different yields against maturity (in months). The loadings are the elements of the vector (Γ_ε) and are defined as the sensitivities of the various yields to the common shock.

short rates such as 3, 6, and 9 months maturities yields is smaller than of the results of earlier empirical studies, particularly for the US market and, therefore, the shape of the curve is a bit different as compared to the one in Koopman *et al.* (2010). It may be due the sticking of the very short rates to zero in the Japanese market during much of the sampled period. Therefore, we compare the sensitivity of one year maturity rate to the initiating point of curve in Koopman *et al.* (2010).

Furthermore, the time series of the yield curve factors' estimates with potentially related macroeconomic variables are plotted in figure 4.2.[63] The time series path of level factor along with the slope factor is closely related to the annual growth of money supply as depicted in the left panel of figure 4.2. It confers that shocks to monetary policy are important sources of variation in the long end of the yield curve and pricing the long-term maturity bonds. The variation in the slope of yield curve, because of the shift in the long end, carries information

[63] One should be aware that rather than smoothed estimates, we plot the update step (filtered) estimates of the yield curve factors.

about the state of economy. Moreover, the variation in inflation is closely explained by the curvature factor of the yield curve. The CPI based inflation rate closely follows the pattern of the curvature factor of the yield curve as depicted in the right panel of figure 4.2. The correlation between inflation rate and curvature factor is $\rho(\hat{\beta}_{3t}, INF_{t-1}) = 0.182$, $\rho(\hat{\beta}_{3t-1}, INF_t) = 0.364$ and $\rho(\hat{\beta}_{3t}, INF_t) = 0.316$. In addition, the curvature factor is also closely related to the growth rate in money supply.

The figure shows that the three latent factors of yield curve, i.e., level, slope and curvature, are closely related to the monetary policy regimes of the Japanese economy during the last 11 years. After a prolonged stagnation and financial instability because of 1990s assets bubble burst and bad debt crisis of 1997, the Japanese economy was heading towards another severe recession in early 1998. The average growth rate had been extremely low, at around 1% since 1992, and the financial institutions had become very weak. Many financial institutions failed in November 1997 and the psychology of the financial market turned extremely negative in the spring of 1998. The growth rate turned negative in the first quarter of 1998, the fragile financial institutions were downgraded by credit rating agencies and obliged to pay a higher interest rate in the interbank market (known as Japan premium), and prices started to decline.[64] In order to avoid the severe recession, the so-called zero interest rate policy (ZIRP) was introduced.[65] The economy did not respond quickly, however, it started to show some signs of recovery in the spring of 2000, as a consequence, the ZIRP was lifted in August 2000. Almost as soon as the interest rate was raised, the Japanese economy entered into another recession and many changes are urged in monetary policy, including the return to ZIRP.[66] In February 2001, the Bank of Japan (BOJ) introduced the Lombard lending facility as well as cut the official discount rate from 0.5% to 0.35%.[67] However, these measures did not show any significant impact and further steps to easing in monetary policy were taken. The target inter-bank rate was lowered immediately to 0.15%, and would go down to zero, as conditions warranted. The official discount rate was sharply cut to 0.1%. Furthermore, the non-conventional quantitative easing was also decided to be used as a tool for an easy monetary policy. During this regime, we observe that the long rates as well as the slope of yield curve have a downward trend. The curvature factor and inflation rate also fall during this period, but, a closer look at the right of figure 4.2 reveals that the inflation rate follows the path of curvature with one period lag. The drop in curvature factor confers that the fall in long rates was larger than the decline in short rates. Hence, the slope was falling along with alteration in the shape of yield curve (becoming flat).

During the last quarter of 2002 the regime switched as in September 2002, the BOJ started

[64] Three events contributed to the weakening of the Japanese economy in 1997-98: (i) the consumption tax (VAT) rate was increased in April 1997, (ii) Japan's banking crisis erupted in November 1997 and continued to the spring of 1999, and (iii) the Asian currency crisis started in July 1997 and continued to the spring of 1998. The Japanese financial institutions had to pay the Japan premium, when they borrowed dollars in the London offshore markets (Ito and Harada, 2005). The growth rates were quickly going down and so was the inflation rate.

[65] The overnight call rate was radically reduced to 0.25% in September 1998 and to 0.15% in early 1999 from 0.5%.

[66] See Ullah *et al.* (2013) for the detailed description.

[67] Detailed description of the Lombard lending facility is given in Ullah *et al.* (2013).

to purchase equities that the commercial banks held. The action was intended as monetary policy tool (Ito and Mishkin, 2004). Furthermore, commitment to ZIRP was declared until the deflationary concerns subside along with an additional policy measure of quantitative easing (QE) in March 2001. The new policy was a big improvement over the last regime and economy showed some signs of recovery in the latter half of 2003 and early 2004. Despite the good performance in the GDP growth rate in 2003:IV, the financial and capital market participants were expecting that the ZIRP will continue for a long time. During this period (known as recovery regime), the long end is gradually rotating and, hence, the slope is on increasing trend. The curvature of yield curve increases only until the end of 2003 along with a minor increase in the inflation rate. The process completes around late 2004.

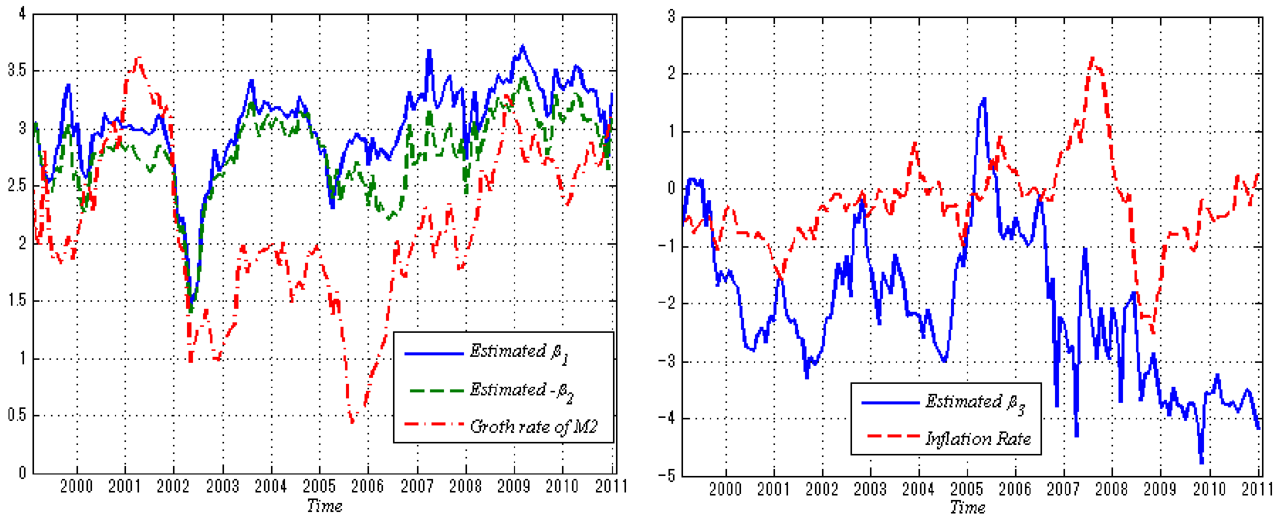


Figure 4.2: Time Series Plot of Nelson-Siegel Estimated Factors with Macroeconomic Variables
Model-based level and slope (estimated β_{1t} and $-\beta_{2t}$) are plotted with annual growth of the $M2$ (Money Supply) and the curvature factor of yield curve (β_{3t}) with the Inflation rate. Inflation rate is the 12-month percent change in the consumer price index.

As the signs of economic recovery became clearer, an exit from QEMP became a popular topic in 2005 in the BOJ meetings. The market started to expect that the Bank would make a move to exit from QEMP and raise the interest rate. The BOJ finally terminated QEMP in March 2006, citing that the QEMP conditions had been satisfied and decided to lower the excess reserves gradually. Thus, in early 2006, there is a rapid decrease in the growth rate of money supply. This phenomenon is reflected with an increase in the level factor and the slope of the yield curve along with a sharp decline in the curvature factor. The fall in curvature factor is accompanied with an increase in inflation rate. The fall in curvature was a signal of entering into another recession as the yield curve was becoming flat once again. The behavior of curvature factor is consistent with the signal of the slope factor in early 2006. It suggests that the economy was responding to an exit from the QEMP and the yield curve was signaling out another recession; however, the ZIRP was still effective.

Furthermore, during the global financial crisis of 2008 and Eurozone debt crisis of 2010, the long as well as the short end go down, but the response of short end is not as strong as of long rates. Thus, the slope falls and the curvature factor moves to right (apparent in the right panel of figure 4.2), means that the yield curve was becoming flat. During the financial crisis of 2008, the Japanese exports fall rapidly and the inflation rate goes to negative once again (it is like a supply side shock for the domestic economy). The fall in curvature and, hence, the slope of yields curve was signaling another future recession, and in response the inflation rate was falling.

In summary, during the initial period of adopting the ZIRP, the QEMP and the world financial crisis, we observe a decline in the yields of long-term bonds and the slope of the yield curve. While in the period of recovery, the yield curve long end as well as the slope are on the increasing trend. In particular, the curvature reflects the cyclical fluctuations of the economy too. Like the yield curve spread, a decrease in curvature is signaling towards economic slowdown and vice versa. It is worth noting that the fall in curvature appears to complement the transition from an upward sloping yield curve to a flat one. Furthermore, the curvature factor seems either to anticipate the future inflation or complemented by the inflation rate, suggesting that the curvature factor is the main driving force of the inflation rate, and transmits the stance of monetary policy in the yield curve shape and, hence, the economy.

The estimate of covariance matrix of the state innovations, as depicted by Σ_v in (4.7), along with the results of Wald and LR tests for its diagonality are shown in table 4.2. There are only 8 out of 15 individually significant covariance terms (whereas 7 are insignificant) at the 5% level of

Table 4.2: Estimate of Covariance Matrix $\boldsymbol{\Sigma_v}$ and its Diagonality Test

Estimates of Covariance Matrix $\boldsymbol{\Sigma_v}$

	$\hat{\eta}_{1t}$	$\hat{\eta}_{2t}$	$\hat{\eta}_{3t}$	$\hat{\eta}_{4t}$	$\hat{\eta}_{5t}$	$\hat{\eta}_{6t}$
$\hat{\eta}_{1t}$	**1.461** (0.094)					
$\hat{\eta}_{2t}$	**-0.029** (0.003)	**1.232** (0.705)				
$\hat{\eta}_{3t}$	-0.047 (0.050)	**0.041** (0.009)	**0.347** (0.031)			
$\hat{\eta}_{4t}$	**0.159** (0.015)	0.069 (0.008)	0.329 (0. 740)	**1.978** (0.003)		
$\hat{\eta}_{5t}$	-0.217 (0.194)	**-0.184** (0.041)	0.034 (0.036)	-0.049 (0.040)	**0.822** (0.027)	
$\hat{\eta}_{6t}$	**0.038** (0.007)	**0.027** (0.015)	**-0.050** (0.004)	0.096 (0.107)	**0.115** (0.044)	**1.441** (0.281)

Test for Diagonality of Covariance Matrix $\boldsymbol{\Sigma_v}$

Test Statistic	Value	df	P-value
Wald Statistic	31.409	15	0.000
LR Statistic	36.874	15	0.000

Note: The upper panel of table reports the estimate of the covariance matrix of innovations of the transition equation. The standard errors are in parenthesis. The lower panel presents the results of the Wald-test and LR-test for the null hypothesis that the covariance matrix $\boldsymbol{\Sigma_v}$ is diagonal. Both the test statistics are Chi-square with their respective degrees of freedom (df). P-value is the probability value of the test statistic. Bold entries denote parameters estimates significant at the 5 percent level.

significance. We perform the Wald and LR tests for the joint significance of the off-diagonal elements of the matrix and both the test statistics reject the null-hypothesis of the diagonality of the Σ_v matrix with very high probability. The result is consistent with our prior expectation that the innovations of transition system are cross correlated.

4.3.3. Formal Tests for Macro and Yield Curve Factors Interactions

The coefficient matrix A and the covariance matrix Σ_v shown in table 4.1 and 4.2 respectively are crucial for assessing the interactions between the yield curve factors and the macroeconomic variables. The (6×6) matrices A and Σ_v are partitioned into four (3×3) blocks as:

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} \qquad \Sigma_v = \begin{bmatrix} \Sigma_1 & \Sigma_2' \\ \Sigma_2 & \Sigma_4 \end{bmatrix} \tag{4.20}$$

where A_1 and A_4 show the yield curve factors and macroeconomic variables dynamics with its own lags respectively. Furthermore, A_2 and A_3 show the extent of lagged linkage from macro-to-yields and yields-to-macro factors respectively. Moreover, we attribute all the covariance terms given by the block Σ_2 to the contemporaneous effect of yield curve factors on the macro variables in accordance to the order of yield and macro factors, employed in the state equation (4.6). As such, there are two links from yields to the macroeconomy in our setup: the contemporaneous link given by Σ_2 and the effects of lagged yields on the macroeconomy are embodied in A_3. Conversely, links from the macroeconomy to yields are symbolized in A_2.

Table 4.3: Tests for Yields-Macro Factors Interactions

Null Hypothesis	Number of Restrictions	Wald Test Statistic		LR Test Statistic	
		Test Statistic	P-value	Test Statistic	P-value
$A_2 = 0$	9	41.126	0.000	31.071	0.000
$A_3 = 0$	9	32.665	0.000	26.369	0.000
$\Sigma_2 = 0$	9	21.207	0.000	34.471	0.000
$A_2 = A_3 = 0$	18	54.956	0.000	61.178	0.000
$A_2 = \Sigma_2 = 0$	18	49.189	0.000	42.502	0.000
$A_3 = \Sigma_2 = 0$	18	47.279	0.000	53.612	0.000
$A_2 = A_3 = \Sigma_2 = 0$	27	96.647	0.000	117.019	0.000

Note: The table presents the results of the Wald-test and LR-test for the no lagged and/or contemporaneous yields-macro factors interaction. A_2, A_3 and Σ_2 refers to the relevant blocks of A and Σ_v matrices. A_2 and A_3 show the extent of lagged linkage from macro-to-yields and yields-to-macro factors respectively, and Σ_2 refers to the contemporaneous effect of yield curve factors on the macro variables. Both the test statistics are Chi-square with the degrees of freedom equal to the number of restrictions. P-value is the probability value of the test statistic.

We employ the likelihood ratio (LR) and Wald tests for the various restrictions of yield and macro dynamics (on the matrix A and Σ_v) and the results of both tests are reported in table 4.3. Both the tests reject the no individual contemporaneous as well as the lagged interaction hypothesis (as the null hypothesis of (i) $A_2 = 0$, (ii) $A_3 = 0$, and (iii) $\Sigma_2 = 0$ are rejected).

Furthermore, the null hypothesis of no interaction of two joint restrictions and three joint restrictions are also rejected with very high probability (as the null hypothesis of (i) $A_2 = A_3 = 0$, (ii) $A_2 = \Sigma_2 = 0$, (iii) $A_3 = \Sigma_2 = 0$, and (iv) $A_2 = A_3 = \Sigma_2 = 0$ are rejected). The results suggest that both hypotheses, of "no macro to yields" depicted by A_2 and "no yields to macro for contemporaneous as well as lagged impact" depicted by A_3 and Σ_2 respectively, should be rejected at a very high level of significance. It confers that there is clear statistical evidence in favor of a bi-directional link between the macroeconomy and the yield curve factors.

4.3.4. Macroeconomic and Yield Curve Impulse Response Functions

Following Diebold *et al.* (2006), we consider the dynamic relationships between the macro and the yield curve factors through impulse response analysis. From an estimated VAR, we compute variance decomposition (VDCs) and impulse response functions (IRFs) which serve as tools for evaluating the dynamic interactions and strength of causal relations among variables in the system.[68] In simulating IRFs and VDCs, it should be noted that VAR innovations may be contemporaneously correlated. This means that a shock in one variable may work through the contemporaneous correlation with innovations in other variables. Therefore, the responses of a variable to shocks in another variable of interest cannot be adequately represented and isolated shocks to individual variables cannot be identified (Lutkepohl, 1991). Therefore, we use Cholesky factorization which orthogonalizes the innovations as suggested in Sims (1980) to solve this identification problem. This strategy requires a pre-specified causal ordering of the variables, because the results from VDCs and IRFs may be sensitive to the variables' ordering. The ordering of variables suggested in Sims (1980) starts with the most exogenous variable in the system and ends with the most endogenous variable.[69]

To see whether the ordering could be a problem, the contemporaneous correlations of VAR error terms are checked (results are not reported to conserve space). The results show that there are high correlations among the three yield curve factors $(\beta_{1t}, \beta_{2t}, \beta_{3t})$ and between the yield curve factors and growth rate of money supply (MS_t). Other correlations are mostly less than 0.25. Based on the strength of correlation, we arrange the variables according to the following order: $(IP_t, INF_t, MS_t, \beta_{3t}, \beta_{2t}, \beta_{1t})$.[70]

[68] In estimating VAR model, one should be aware of the stationarity consideration of the variable in the system, otherwise the results may be suffered from spurious relationship. A regression involving the levels of *I(1)* series will produce misleading results (Phillips, 1986). We use all the six factors in deviation form in order to ensure stationarity. Furthermore, all the roots of transitional matrix lie inside the unit circle. See appendix F for the derivation of VAR model to simulate the IRFs and VDCs.

[69] To avoid the subjective criteria of pre-specified ordering of variables, we also computed the generalized impulses (GIRF) as described in Pesaran and Shin (1998). The resulting responses (not reported here to save space) are almost similar to the one obtained from Cholesky factorization.

[70] In Diebold *et al.* (2006), the order is reverse as they put yield curve factors before macroeconomic variables. The intuition behind their ordering is that the yield curve observations are dated at the beginning of the month, whereas for the macroeconomic variables, the end of month data is used. Under this identification scheme, yield factors are assumed to be contemporaneously unaffected by the macro factors. But in our case, we do not assume that macroeconomic factors do not contemporaneously cause variation in the yield factors as we use the end of month's price quotes to calculate the zero-coupon yields and macroeconomic data is also collected at the end of each month.

There are four blocks of impulse responses, i.e., yield curve factors responses to macro shocks, macro variables responses to yield factors shocks, yield to yield factors shocks, and macro to macro variables shocks, but given the focus of this chapter, here we consider only the former two blocks. Particularly, we focus on the yield curve factors response to monetary policy shocks and back to real activity and inflation rate. The results of impulse response functions of the two blocks along with 90% confidence band are presented in figure 4.3.[71] Overall the results convey an interesting message that the response of macro variables to yield factors is much stronger than the response of latter to the former variables shocks.

Considering the responses of the yield curve to the macro shocks, the slope and level factors show very little response than the curvature factor to the shocks in all the three macroeconomic variables. It attributes to the prominent role and economic interpretation of the curvature factor of term structure. The results show that a stochastic positive shock in the industrial production immediately push down the long end with an increase in the slope factor (the yield curve become less positively sloped or more negatively sloped), suggesting that the yield curve becomes flatter in response to the supply side shocks. However, the curvature factor moves to left with a 4 to 5 months delay, indicating that inflation expectation rises as a result of expansionary monetary policy in subsequent periods during recession. After 15 months, the long end goes up and the yield curve becomes steeper (as β_{2t} falls and β_{3t} reaches to its maximum). This behavior of long rates is consistent with the inflationary expectation hypothesis of Fisher (1896). Furthermore, the behavior of all three yield curve factors in subsequent periods is consistent with the idea that during recessions, premia on long-term bonds tend to be high and yields on short bonds tend to be low. Hence, during recessions, upward sloping yield curve not only indicates bad times today, but also better times tomorrow. Shocks in inflation rate immediately push up β_{2t} (decreases the slope) and down the level factor, however the curvature factor immediately moves to left. The reaction of slope as well as level factor is consistent with the behavior of Japanese economy during the first decade of 21st century. The inflation rate is almost zero and the surprise to actual inflation cannot give a boost to long rates but it falls in accordance to the long-run expectations. However, in response to a change in the level factor, the slope as well as curvature factors react.

Positive shocks to money supply induce the long rates to rise and, hence, the slope increases (meaning β_{2t} falls), however, the curvature factor reacts much stronger than the former two. The fall in curvature factor is associated with a rise in inflationary expectation, consistent with the expected positive impact of money supply on inflation rate. Recalling that the ultimate objective of the Japanese monetary policy during the decade is to affect the yield curve level in order to stimulate the economy, the success of monetary policy could be defined as a decrease in the long end of the yield curve either via expected short-term rates (policy-duration effect), term premium (portfolio-rebalancing channel) or both of them. However, this represents only an intermediate target in an attempt to generate economic recovery and to stop deflation. The final

---

[71] VAR satisfies the stability condition as all roots lie inside the unit circle.

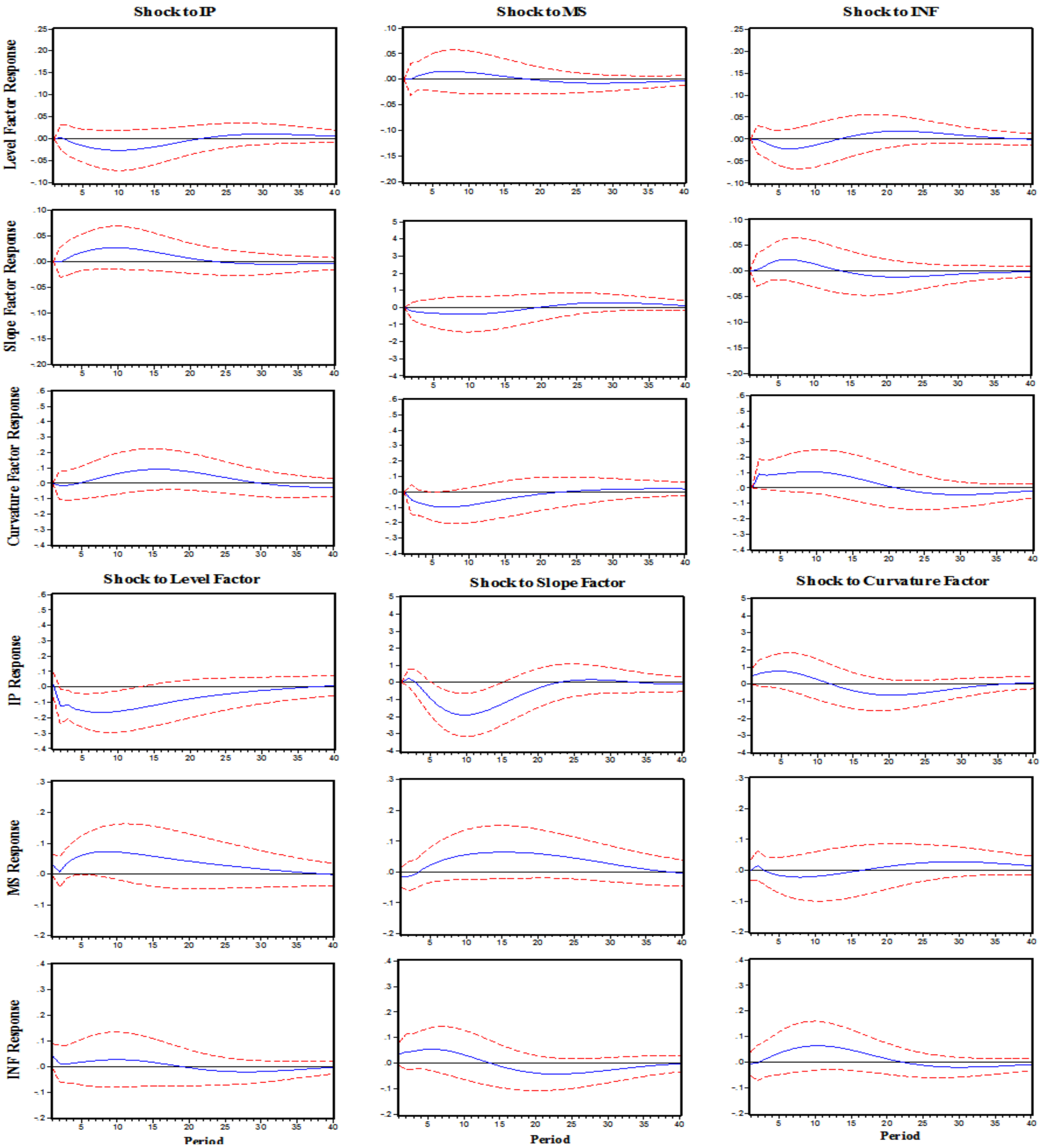


Figure 4.3: Impulse Responses of Yields-Macro Factors to Cholesky one S.D. Innovation

The figure shows the reactions of three yield curve factors (i.e., $\beta_{1t}, \beta_{2t}, \beta_{3t}$ denoted by level, slope and curvature factors respectively) and three macroeconomic variables (i.e., IP_t, MS_t, INF_t) to a shock in each exogenous variables in the VAR(1) model over 40 months. We simulate the VAR(1) model of yield and macro factors and compute response of each factor to Cholesky one standard deviation innovation. The solid blue line denotes the estimated response, while the red dashed line shows $\pm 2(SE)$ (plus-minus two standard error) confidence band.

goal of the BOJ is expected to be, to increase inflation expectations and thus future short-term interest rates, which, in turn, will raise the long-term interest rates. As argued by Nagayasu (2004), monetary policy mechanisms take one to two years to achieve their full effects. It seems appropriate to expect that the effectiveness of QEMP, if any, would result in an increase in the level factor. Therefore, during the ZIRP and QEMP, the long end immediately jumps up in response to a shock in monetary policy. The rise in the level factor reflects the strengthening credibility of the BOJ and, thus, the effectiveness of its policy. Indeed, as argued in Diebold *et al.* (2006) and Bianchi *et al.* (2009), if monetary policy is credible, the level factor, other things being equal, should fall after a positive shock to call rate, because the expectation of future inflation declines. Since, the BOJ commits itself during the QEMP period to maintain the short-term rates to a zero level, the decline in the level factor after an increase in the call rate is by analogy equivalent to a rise in this factor to a monetary policy expansion. This can be due to an expectation of the economic recovery and rise in inflation, indicating a monetary policy success.

The lower block of figure 4.3 summarizes impulse response functions of macroeconomic variables to unexpected increase in the yield curve factors. The level shock has a negative effect on industrial production, although its impact seems small but is statistically significant. It reinforces the idea that the contribution of macroeconomic variables to level factor variation, if any, comes from the level of economic activity. Furthermore, a positive surprise change in the level factor indicates an increase of inflation and monetary expansion. However, the inflation increases a little and reverts to zero immediately, but the response of money growth is more prolonged. It suggests that the BOJ adopts an expansionary monetary policy in response to a decline in aggregate spending, results from a sudden increase in the long-term interest rate.[72]

The responses to an unexpected positive change in the slope factor are consistent with the monetary policy stances in the Japanese economy. An increase in the slope factor means a reduced spread between long-term and short-term bonds, which indicates a monetary policy tightening and, thus, economic activity declines within the upcoming 3 to 4 months.[73] The direction of reaction of the IP_t contributes to the view that the yield curve slope acts as an indicator of the future state of economy. The reaction of inflation looks qualitatively similar to the response to the level shock. An unexpected increase of the slope factor is followed by an initial increase in inflation rate, but it is short lived and very small. The money growth rate falls in response to the slope shock but reverts to zero immediately and then increases. It confers that the BOJ implements the expansionary monetary policy, as the spread between the long and short end tightens, to avoid the upcoming recession.

Unlike Diebold *et al.* (2006), the macroeconomic variables have significant reactions to the positive change in curvature factor. The increase in curvature means transition from flat yield curve to a steeper one. The economic activity expands along with an increase in inflation rate in response to an unanticipated positive shock in the curvature of yield curve. It suggests that the curvature is a leading indicator/main driving force of future inflation and also reflects the cyclical fluctuations of the economy. It advocates that the curvature factor also presents the stances of monetary policy and can predict the future path of economy and inflationary expectations. The

---

72 The increase in the long term interest rate (as a result of an increase in the short rates, according to the expectation hypothesis) causes a decline in aggregate spending.

73 Normally a decrease in yield curve slope announces an economic slowdown. But the loading of the slope factor in our model decreases with maturity and corresponds to the difference between short and long-term yields, therefore an increase in this factor corresponds to a decrease in the term spread.

reaction of monetary policy is virtually zero in response to a change in curvature factor, consistent with the prevailing economic situation during the decade in the Japanese economy.

Summarizing, it turns out that the contribution of macroeconomic variables, though small in magnitude but does not quickly shift to low levels, suggest a significant role of the macroeconomic variables in influencing the yield curve during the long-lasting economic stagnation in the Japanese economy. The lower block that sums up the reaction of the macroeconomic variables in response to the shocks in the yield curve factors suggests that, after the short-term interest rate has reached zero, the monetary policy signals can be transmitted significantly and with higher probability (as all the responses are statistically significant) to the real sector through the yield curve three factors.

4.3.5. Macroeconomic and Yield Curve Variance Decompositions

Variance decompositions (VDCs) is an alternative method to IRFs for examining the effects of exogenous shocks on dependent variables. It shows how much of the forecast error variance for any variable in a system is explained by innovations to each explanatory variable over a series of time horizons. Usually, own series shocks explain most of the error variance, although the shock will also affect other variables in the system. From table 4.4, the VDC substantiates that the spread factor of yield curve plays an influential role in the variation of yield level. Furthermore, a significant role is also played by the entire three macroeconomic variables in fluctuating the yield level factor; however, the relative role of industrial growth is more prominent than the other two. It confers that, rather than monetary policy, the supply side shocks also contribute significantly to the variation of long-term interest rate. This indicates that news about the future evolution of output might be more important for the dynamics of the yield curve than inflationary concerns for that period.

The variation in slope factor mainly comes from the level factor and monetary policy. The impact of monetary policy is consistent with the behavior of spread and monetary growth in figure 4.2. Furthermore, the changes in the curvature factor are attributed to the shift of long end of yield curve and variation in inflation rate. However, at the longer horizon forecasts, the slope and the other two macroeconomic variables (i.e., money supply and industrial growth) play a significant role as well. But the slope factor contributes a bit more than the rest of two (IP_t and MS_t) to explain the variation in β_{3t}.

Regarding the variance decomposition of the extant of economic activity (represented by the growth rate of industrial production), it is apparent that the slope factor plays a crucial role at all horizons of forecasts, followed by the curvature factor in contributing the variation in IP_t. The variance appears to be explained about 4 to 6 percent by each of the level factor and inflation innovations. It highlights the idea that the slope of yield curve and its curvature factor signals out the state of economy in near future. This indicates that the information about the slope and the curvature of the yield curve might be an important signal about the future evolution of output than the long rates and inflationary concerns for that period.

Table 4.4: Variance Decompositions of Yield Curve Factors and Macroeconomic Variables

Period	β_{1t}	β_{2t}	β_{3t}	IP_t	MS_t	INF_t
Variance Decomposition of $\boldsymbol{\beta_{1t}}$						
1	75.864	11.977	0.572	5.193	4.128	2.187
12	71.876	13.818	0.818	5.985	4.296	3.207
24	65.614	18.725	1.942	6.297	4.524	3.798
40	63.696	15.819	1.134	9.754	5.749	3.856
Variance Decomposition of $\boldsymbol{\beta_{2t}}$						
1	3.493	88.214	0.757	1.676	5.453	0.406
12	5.986	85.078	1.347	2.238	5.006	0.345
24	5.236	83.94	1.469	2.759	6.265	0.326
40	5.356	82.762	1.879	3.068	6.426	0.508
Variance Decomposition of $\boldsymbol{\beta_{3t}}$						
1	25.723	4.426	62.844	0.061	1.220	5.726
12	27.111	10.676	37.831	2.634	5.987	8.695
24	27.824	12.859	32.197	5.772	7.664	13.683
40	26.105	10.801	31.888	6.002	8.751	16.451
Variance Decomposition of $\boldsymbol{IP_t}$						
1	4.473	4.465	0.428	86.755	0.728	3.150
12	4.056	18.531	7.996	63.676	1.482	4.259
24	6.137	23.592	9.124	54.649	1.568	4.930
40	6.731	25.185	10.482	50.365	2.239	4.935
Variance Decomposition of $\boldsymbol{MS_t}$						
1	3.224	9.514	6.844	0.236	79.764	0.418
12	10.729	23.725	12.811	4.85	43.118	4.767
24	12.899	27.357	10.360	4.977	34.119	10.288
40	12.815	26.125	11.003	6.455	32.323	11.279
Variance Decomposition of $\boldsymbol{INF_t}$						
1	1.034	3.963	12.001	15.197	0.603	67.202
12	1.239	3.732	13.910	18.018	1.460	61.640
24	1.306	5.215	14.007	17.246	2.627	59.599
40	1.873	6.385	14.409	16.570	2.771	57.992

Note: The table reports the results of variance decompositions of all the six variables in the system. We simulate the VAR(1) model of yield and macro factors and compute the contribution of innovations of each explanatory variable over a series of time horizons. Each entry is the proportion of the forecast variance (at the specified forecast horizon) for a 1, 12, 24 and 40 months' time horizons that is explained by the particular factor.

Looking at the variance decomposition of money supply, it shows that the slope factor is the dominant factor, followed by the level and curvature factors. Productivity shocks also contribute after one year but less than the inflation rate in explaining the variance of monetary growth. The result is consistent with the idea that the shape and particularly the curvature of yield curve represent the stances of monetary policy to affect the level of economic activity and inflation rate in the economy.

Finally, the variation in inflation is explained by industrial production to a greater extent. It suggests that supply side shocks are more influential in determining the path of inflation rather than the demand side during the first decade of 21st century in the Japanese economy; because during the QEMP, the contribution of monetary policy shock is negligible. This is consistent with the ineffectiveness of the QEMP in affecting the expectation about future inflation as well as the long end of yield curve, because the inflation response veers to zero and becomes insignificant more rapidly.[74] Regarding the contribution of yield curve factors, the variation in the inflation is largely due to the curvature factor of yield curve. The remaining factors such as β_{1t}, β_{2t} and MS_t contribute marginally in the variance decomposition of inflation rate.

4.4. Evidence on the Expectation Hypothesis and Time-varying Term Premium

The crucial link between the central bank's instrument and long-term interest rates is the expectation hypothesis (EH) of the yield curve theory.[75] It provides an indication about how anticipation of future monetary policy decisions affects the economy. This issue is especially important for the Japanese economy because the principal channel suggested by either ZIRP or QEMP is the expectation channel. One important channel through which monetary policy works is the long-term interest rates, shaping them so that in turn they affect the level of economic activity. The expectation that the policy of low short-term interest rates may be maintained for a substantial period of time will likely lower medium to long-term interest rates, which in turn will rise inflationary expectations and boost economic activity. However, the empirical support for the EH and the effectiveness of the policy commitment is rather mixed. Thornton (2004) applies a bivariate VAR for long-term and short-term interest rates for the period from March 1981 to January 2003. He shows that the EH does not hold for the Japanese case. One possible explanation for the empirical failure of the EH is the presence of a time-varying term premium (liquidity and risk premia). Time-variation in term premia might arise because of changes in preferences of market participants toward risk. In addition, the term premium could also vary with the business cycle, as investors might be more risk-averse in recessions than in booms.

In this study, we use a time-varying risk premium latent factor model to review the validity of traditional EH and the effects of the BOJ's expectations management on the JGBs yield curve. The implicit assumption behind our time-varying latent factors model is that the agents review their expectations about uncertainty regarding inflation, real activity and monetary policy in each period and, hence, the term premia is time-varying. This assumption allows us to perform accurate predictions and makes the model more flexible and realistic.

In terms of our notations in section 1.2 and 4.2, which pertain to the pure discount bond yields in our data set, the theoretical bond yield is denoted as $R_t(m)^{EH}$. The model is:

---

[74] It suggests that rather than the demand side, the supply shocks (oil price, fall in exports due to East Asian crisis, US crisis and Eurozone crisis) should also be considered possible reasons for the prolonged deflationary period in Japan.

[75] The traditional expectation hypothesis of the term structure states that movements in long rates are due to movements in expected future short rates. Any term or risk premia are assumed to be constant through time.

$$R_t(m) = R_t(m)^{EH} + \phi_t(m) + \varepsilon_t \tag{4.21}$$

where $\phi_t(m)$ is a time-varying term premium that may also vary with the maturity and $R_t(m)^{EH}$ is the theoretical bond yields which is defined as:

$$R_t(m)^{EH} = \left(\frac{1}{m}\right) \sum_{i=0}^{m-1} E_t R_{t+i}(1) \tag{4.22}$$

$$\phi_t(m) = \phi_{t-1}(m) + v_t \tag{4.23}$$

$$\begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma \end{bmatrix}\right) \tag{4.24}$$

Specifically, we compare the theoretical bond yields $R_t(m)^{EH}$, constructed via (4.22), plus $\phi_t(m)$ under the assumption that the expectation hypothesis does not hold, with the actual bond yields $R_t(m)$. We construct the expected future 1-month maturity yields by iterating forward the estimated yield model using the measurement equation (4.5) and the state equation (4.6) for $E_t R_{t+i}(1)$; and then compute the theoretical bond yields at each point in time using (4.22). Furthermore, the term premium is computed by using the state-space model with the signal equation specified in (4.21) and state equation in (4.23), assuming that the latent factor $\phi_t(m)$ follows the random walk process. We use the Kalman filter algorithm as discussed in section 4.2.2 to estimate the time-varying term premia.

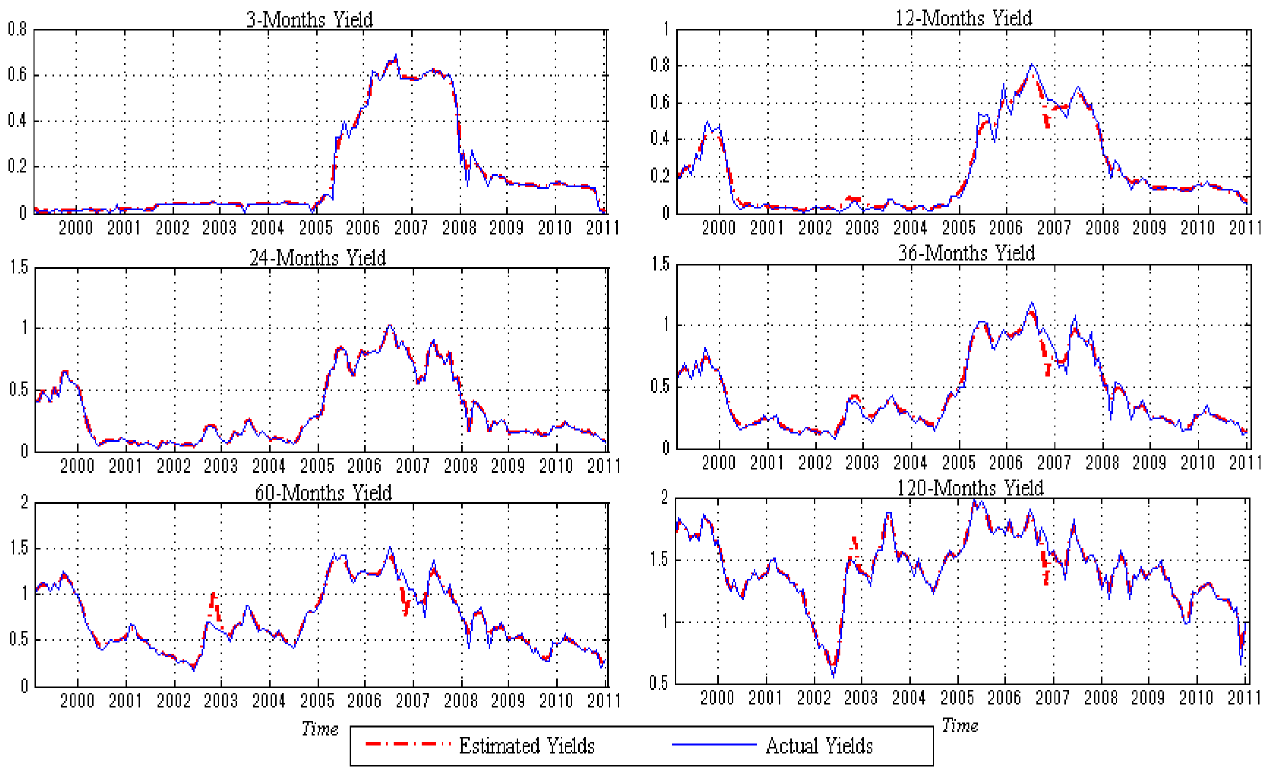


Figure 4.4: Estimated Yields $\left[R_t(m)^{EH} + \hat{\phi}_t(m)\right]$

The figure presents time series plot of the estimated yields $\left[R_t(m)^{EH} + \hat{\phi}_t(m)\right]$ along with the actual bond yields $R_t(m)$. The estimated yields are depicted by dashed red line, whereas the actual yields by solid blue line. The number of observations is 144.

Figure 4.4 provides some selected maturities estimated yields together with their actual counterpart. The estimated yields $[R_t(m)^{EH} + \hat{\phi}_t(m)]$ are tracking actual yields very well, despite limited deviation that occurs during the last 3 months of 2003 and late 2007. However, overall the results indicate that the expectation hypothesis of the term structure of interest rates does not hold during the ZIRP/QEMP period, because the estimated term premia vary considerably during the sampled period.[76]

Furthermore, the estimated term premia of some selected maturities is shown in figure 4.5. There is substantial variation over time in the behavior of the term premium for all the maturities. In particular, it is interesting to note that during the ZIRP and the QEMP period (2001-2006), the term premia decline to a lower level. This could be due to the heavy demand for JGBs from the BOJ during that period.

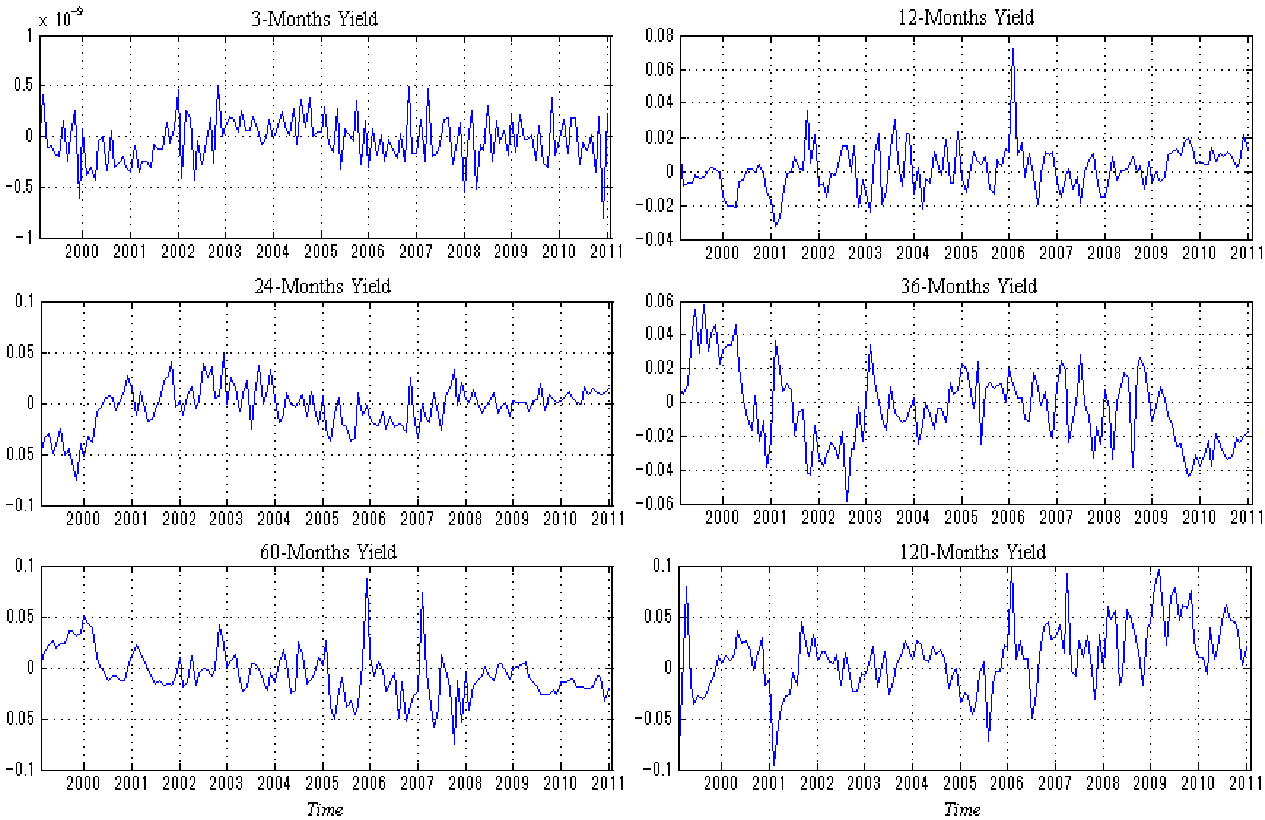


Figure 4.5: Estimated Term Premium
The figure presents time series plot of the estimated term premium $\phi_t(m)$ under the assumption that the expectation hypothesis does not hold. The number of observations is 144.

4.5. Conclusion

This chapter explores the evolution between the yield curve and the Japanese economy with a special focus on examining the effects of the quantitative easing strategy in Japan on the yield curve and the possible feed-back effect on the real sector by applying a macro-finance model.

76 We tried the model with fixed term premia, however the theoretical yield deviate too much from the actual yield. It suggests the supportive evidence in regard of time-varying term premia. Results are not reported to conserve space.

The yield curve model of this study explicitly incorporates both yields factors (level, slope, and curvature) and macroeconomic variables (overall economic activity, money supply and inflation rate). We also extend the model in Diebold *et al.* (2006) to include time-varying stochastic volatility in the yield model (observation equation) in state-space framework.

Empirical results from the yields-macro factors model show that there is statistically significant bidirectional causality between the macroeconomic and yield curve factors; however by contrast with conventional wisdom, macro variables play a less prominent role in explaining the yield factors as compared to the strength of effect from latter to the former. Furthermore, volatility in bond markets is found to be asymmetrically affected by positive and negative shocks and more sensitive to recent innovations rather than the lag volatility. Short maturities tend to be more sensitive to the common shocks in interest rate market than the long maturities.

The structural decomposition indicates that it is the entire term structure of interest rate that transmits the policy shocks to the real economy rather than the only yields spread (as considered in the previous studies regarding the Japanese economy). The monetary policy signals passes through the yield curve level and slope factors to stimulate the economic activity. The curvature factor, besides reflecting the cyclical fluctuations of the economy, acts as the leading indicator for future inflation. The curvature factor seems either to anticipate or is accompanied by inflation rate. One can infer from the overall results that the slope and curvature factors (in our framework) serve as leading counter-cyclical and pro-cyclical indicators respectively.

In addition, the proposed yields-macro factors model does not empirically substantiate the traditional expectation hypothesis during the ZIRP and QEMP regimes.

Chapter 5

Conclusion and Policy Implications

Pricing assets in the fixed-income market is an important field in financial econometrics. The most basic asset in the fixed-income market is the zero-coupon bond. The complete set of zero-coupon bonds of all maturities results in the term structure of interest rates that forms the basis of the fixed-income market. The term structure of interest rates is also an important element in macroeconomics and finance. For macroeconomics, in a monetary policy context, forward rates are potentially useful as indicators of market expectations of future interest rates, inflation rate and exchange rate as discussed by Svensson (1994) and Sodelind and Svensson (1997) and the yield curve carries information about future GDP growth as shown by Estrella and Mishkin (1996, 1998). For finance, fixed income portfolio managers and risk managers make use of the yield curve for pricing derivatives (e.g., interest rate futures and options) and performing hedging operations. However, the market does not provide us securities at all the desired maturities and what we observe is only an incomplete set of yields across the maturity spectrum. This way, to overcome this problem, it is necessary to use some model to observe the yields for the entire maturity spectrum.

5.1. Summary

This thesis deals with the analysis of term structure theory and its application, with a particular focus on the modeling and forecasting aspects of the yield curve. Furthermore, the proposed models are related to the state of economy and monetary policy issues. Advanced time series models and sophisticated mathematical tools are applied to analyze the time-varying relationship between yields and time to maturity. Four major research objectives are addressed: (i) to compare the different modeling and estimation techniques' ability to characterize and predict the time-varying nature of the yield curve, (ii) to evaluate the out-of-sample forecast performance of the statistical class of models with and without macroeconomic variables in the state-space representation, (iii) to examines the effect of monetary policy stances on term structure and the possible feed-back effect on the real sector during zero interest rate policy (ZIRP) period, and (iv) do the state-space representation and estimation through Kalman filter (one-step procedure) of the model lead to better in-sample fit and more accurate forecasts than the two-step estimation approach?

Regarding the first research objective, in chapter 2, the short-term CIR model is calibrated using the maximum likelihood estimation method and the dynamic Nelson and Siegel (1987) model using non-linear least squares procedure. Furthermore, the linearized version of Nelson-Siegel model is also estimated for comparison purposes. For the in-sample fit, there is a significant lack of information on the short rates CIR model. Contrary to CIR model, the Nelson-Siegel model is capable to distill the term structure of interest rate quite well and describe the evolution and the trends of the market for time-varying τ_t estimation process. Fixing the shape parameter τ_t to the median value leads to fit the yield curve better than the CIR model but not than the non-linear estimation process of the Nelson-Siegel model. Regarding the term structure forecast, the CIR model should be considered too poor to describe the term structure in the simulation based context. It generates a downward slope average yield curve. On the other hand, the Nelson-Siegel model is not only compatible to fit attractively the yield curve but also accurately forecast the future yield for various maturities. The Nelson-Siegel model is capable to replicate most of the stylized facts of the Japanese market yield curve.

Targeting the second question, in chapter 3, the dynamic Nelson-Siegel model with and without macroeconomic variables in the state-space representation is evaluated through Kalman filter. The proposed extended model explicitly incorporates both yields factors (level, slope, and curvature) and macroeconomic variables (overall economic activity, exchange rate, stock prices index and inflation rate). The Nelson-Siegel type models in the state-space framework considerably outperform the benchmark simple time series forecast models such as AR(1) and random walk. The yields-macro model incorporating macroeconomic factors leads to a better in-sample fit of the term structure than the yields-only model. The out-of-sample predictability of the former for the short horizon forecasts is superior to the latter for all maturities examined, and for the longer horizons the former is still compatible to the latter. Inclusion of macroeconomic factors can dramatically reduce the autocorrelation of the forecasts errors, which has been a common phenomenon of statistical analysis in the previous term structure models.

Regarding the monetary policy and the practical significance of the model, in chapter 4, a joint macro-finance latent factors model is formulated that integrates monetary policy as well as real activity factors in the term structure model, to examine the effect of the monetary policy stances on term structure and the possible feed-back effect on the real sector. The analysis indicates that it is the entire term structure that transmits the policy shocks to the real economy rather than the yields spread only. The monetary policy signals passes through the yield curve level and slope factors to stimulate the economic activity. The curvature factor, besides reflecting the cyclical fluctuations of the economy, acts as a leading indicator for future inflation. In addition, policy influence tends to be low as the short end becomes segmented toward medium/long-term of the yield curve. Furthermore, the volatility in the bond markets is found to be asymmetrically affected by the positive and negative shocks and short end tends to be more sensitive to stochastic shocks than the long maturities. We also related our yield curve modeling approach to the traditional macroeconomic approach based on the expectations hypothesis. The analysis shows that the expectation hypothesis does not hold during the ZIRP and QEMP regimes.

While answering the question, which of the selected estimation approaches is more suitable to predict the path of time-varying term structure of interest rate for the sample under consideration, one can compare the results of in-sample fit and out-of-sample forecast accuracy of Nelson-Siegel model in chapter 2 and 3. The results show that one-step estimation approach leads not only to better in-sample fit but also produces out-of-sample forecasts with greater accuracy. More importantly, the residuals of in-sample fit and the out-of-sample forecast errors based on the Kalman filter are less persistent than those of the two-step approach. Therefore, simultaneous estimation through Kalman filter based on the conditional expectation and conditional variance can be considered superior to the latter in terms of in-sample fitting as well as out-of-sample predictability.

5.2. Policy Implications

The results in this dissertation have several implications for policy. First, the analysis of the dynamic Nelson and Siegel (1987) yield curve model is relevant for how central banks and financial institutions analyze the term structure. If financial institutions and central banks are looking for a model to study the evolution of the yield curve in Japanese market, the Nelson-Siegel family of models could be a good candidate.

Second, the inclusion of macroeconomic and equity market related variables in the standard yield curve model improves the fit of the model and comes up with more accurate forecasts and less persistent forecast errors.

Third, the dynamic extension of standard Nelson-Siegel model (inclusion of macroeconomic factors) provides a framework to predict the future path of expected inflation, stock market and other macroeconomic variables that are of a particular relevance and considered as leading indicators of the future state of economy.

Fourth, the modeling approach in chapter 4 provides a more flexible framework to analyze the effectiveness of monetary policy in ZIRP regime and figure out the transmission mechanism of monetary policy. The study highlights the importance of the yield curve factors for policy analysis which can serve as leading pro-cyclical or counter-cyclical indicators.

5.3. Directions for Further Research

The study offers many directions and opens many interesting challenges for future research. Though, the statistical class of models comes up with encouraging results in in-sample fit as well as forecasts, it is necessary to introduce extensions in the standard Nelson-Siegel model that incorporate additional flexibility in the yield curve model in order to fit accurately the curves with multiple maxima and/or minima.[77] We are aware of the popular extension by Svensson (1994), but it leads to high degree of correlation between the loadings (multicolinearity) and makes it difficult to estimate the parameters precisely. These correlations are affected by the shape parameter τ_t value as illustrated in Annaert *et al.* (2000).

[77] De Pooter (2007) also notes that the shape parameter τ_t cannot handle the complete set of shapes the curve takes.

Another key aspect of the term structure is time-varying stochastic volatility. The interest rate volatility for various maturities can be modeled in many different ways. The alternative specification (to EGARCH model used in chapter 4) for the asymmetric effect of positive and negative shocks can be TARCH and TGARCH. Moreover, the volatility can also be linked with macroeconomic events as one can proceed with GARCHX specification. For this reason, this study suggests further extensions in state-space framework of yield curve model in such a way to introduce TARCH, TGARCH or GARCHX effects, and compare the in-sample fit as well as future yields predictability performance among the various extensions.

Furthermore, the lag-lead (causality) analysis between the yield factors and the stock market will be of a higher significance for the efficiency analysis of both markets, the bond and stock.

References

Almeida C, Gomes R, Leite A, Vicente J. 2007. Movements of the term structure and criteria of minimization of prediction errors in an exponential parametric model. EPGE Working Paper, Getulio Vargas Foundation.

Andrews DWK. 1991. Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59:817–858.

Ang A, Piazzesi M, Wei M. 2006. What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131:359–403.

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

Annaert J, Claes AGP, De Ceuster MJK, Zhang H. 2000. Estimating Nelson-Siegel - A Ridge regression approach, University of Antwerp, Belgium. [online available at <http://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2010-Aarhus/Smoothed_Bootstrap_-_Nelson-Siegel_Revisited_June_2010.pdf>; accessed Sept. 5, 2011].

Baba N, Nishioka S, Oda N, Shirakawa M, Ugai H. 2005. Japan's deflation, problems in the financial system, and monetary policy. *Monetary and Economic Studies* 23(1):47–111.

Bernanke B, Blinder A. 1992. The Federal Funds Rate and the channels of monetary transmission. *American Economic Review* 82:901–921.

Bernanke B, Reinhart V, Sack B. 2005. Monetary policy alternative at the zero bound: an empirical assessment. *Brookings Papers on Economic Activity* 2:1–78.

Bianchi F, Mumtaz H, Surico P. 2009. The great moderation of the term structure of UK interest rates. *Journal of Monetary Economics* 56(6):856–871.

Bikbov R, Chernov M. 2010. No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics* 159 (1):166–182.

Bjork T, Christensen BJ. 1999. Interest rate dynamics and consistent forward rate curves. *Mathematical Finance* 9:323–348.

Black F, Derman E, Toy E. 1990. A one-factor model of interest rates and its application to Treasury bond options. *Financial Analyst Journal* 46:33–39.

Bliss R. 1997. Testing term structure estimation methods. *Advances in Futures and Options Research* 9:97–231.

Byeon YH, Ogaki M. 1999. An empirical investigation of exchange rates and the term structure of interest rates. Working Paper 99-20, Department of Economics, Ohio State University.

Campbell JY, Shiller RJ. 1991. Yield spreads and interest rate movements: a bird's eye view.

Review of Economic Studies 58:495–514.

Campbell JY. 1986. Bond and stock returns in a simple exchange model. *Quarterly Journal of Economics* 101:785–804.

Campbell JY. 1996. Understanding risk and return. *Journal of Political Economy* 104:298–345.

Campbell JY. 1999. Asset prices, consumption and the business cycle. In *The Handbook of Macroeconomics*. Taylor JB, Woodford M (eds.). Elsevier: North-Holland, Amsterdam;1231–1303.

Campbell, JY. 1993. Intertemporal asset pricing without consumption data. *American Economic Review* 83:487–512.

Christensen JHE, Diebold FX, Rudebusch GD. 2011. The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics* 164:4–20.

Clostermann J, Schnatz B. 2000. The determinants of the Euro-Dollar exchange rate synthetic fundamentals and a non-existing currency. *Applied Economics Quarterly* 46(3):274–302.

Cochrane JH, Piazzesi Monika. 2005. Bond risk premia. *American Economic Review* 95:138–160.

Cox JC, Ingersoll JE, Ross SA. 1985. A theory of the term structure of interest rates. *Econometrica* 3:385–407.

Dai Q, Singleton KJ. 2002. Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics* 63:415–441.

de Jong F. 2000. Time series and cross section information in affine term structure models. *Journal of Business and Economic Statistics* 18:300–314.

De Pooter M. 2007. Examining the Nelson-Siegel class of term structure models. Tinbergen Institute Discussion paper, No. 2007-043/4.

Dewachter H, Lyrio M. 2002. Macroeconomic factors in the term structure of interest rates. Manuscript, Catholic University Leuven, Erasmus University, Rotterdam.

Dickey DA, Fuller WA. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74:427–431.

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

Diebold FX, Mariano RS. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13:253–263.

Diebold FX, Rudebusch GD, Aruoba B. 2006. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics* 131:309–338.

Diebold FX, Rudebusch GD. 2012. *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton University Press.

Duffee G. 2002. Term premia and interest rate forecasts in affine models. *Journal of Finance* 57: 405–443.

Duffee GR. 2011. Information in (and not in) the term structure. *Review of Financial Studies* 24: 2895–2934.

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

Dungey M, McKenzie M, Tambakis D. 2009. Flight-to-quality and asymmetric volatility responses in US Treasuries. *Global Finance Journal* 19:252–267.

Dunn K, Singleton KJ. 1986. Modeling the term structure of interest rates under non-separable utility and durability of goods. *Journal of Financial Economics* 17:27–55.

Elen van EALJ. 2010. Term structure forecasting: does a good fit imply reasonable simulation results? A thesis submitted to Tilburg School of Economics and Management, Tilburg University. [online available at < http://arno.uvt.nl/show.cgi?fid=112924> ; accessed August 11, 2011].

Estrella A, Hardouvelis G. 1991. The term structure as a predictor of real economic activity. *Journal of Finance* 46:555–576.

Estrella A, Mishkin F. 1998. U.S. recessions: financial variables as leading indicators. *The Review of Economics and Statistics* 80(1):45–61.

Estrella A, Mishkin FS. 1996. The yield curve as a predictor of U.S. recessions. Federal Reserve Bank of New York, Current Issues in Economics and Finance.

Evans CL, Marshall DA. 2007. Economic determinants of the nominal treasury yield curve. *Journal of Monetary Economics* 54(7):1986–2003.

Fabozzi FJ, Martellini L, Priaulet P. 2005. Predictability in the shape of the term structure of interest rates. *Journal of Fixed Income* 15(1):40–53.

Fama E, Bliss R. 1987. The Information in long-maturity forward rates. *American Economic Review* 77:680–692.

Fama EF, French KR. 1992. The cross-section of expected stock returns. *Journal of Finance* 47 (2):427–465.

Fama EF, French KR. 1993. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33 (1):3–56.

Fama EF. 1990. Term structure forecasts of interest rates, inflation, and real returns. *Journal of Monetary Economics* 25:59–76.

Fisher M, Nychka D, Zervos D. 1995. Fitting the term structure of interest rates with smoothing splines. Board of Governors of the Federal Reserve System, Federal Reserve Board Working Paper 95-1.

Fisher, I. 1896. *Appreciation and Interest*, New York: Macmillan.

Gürkaynak RS, Sack B, Jonathan HW. 2007. The U.S. treasury yield curve: 1961 to the present. *Journal of Monetary Economics* 54(8):2291–2304.

Hamilton JD, Kim D. 2002. A re-examination of the predictability of the yield spread for real economic activity. *Journal of Money,Credit,and Banking* 34:340–360.

Hamilton JD. 1994. State-space models. In *Handbook of Econometrics*. Engle RF, McFadden DL. (eds.), Elsevier: Amsterdam; 3041–3080.

Harvey CR. 1988. The real term structure and consumption growth. *Journal of Financial Economics* 22:305–333.

Heath D, Jarrow R, Morton A. 1990. Bond pricing and the term structure of interest rates: a discrete time approximation. *Journal of Financial and Quantitative Analysis* 25:419–440.

Hördahl P, Tristani O, Vestin D. 2006. A joint econometric model of macroeconomic and term structure dynamics. *Journal of Econometrics* 131:405–444.

Hull J, White A. 1990. Pricing interest rate derivative securities. *Review of Financial Studies* 3: 573–592.

Inci AC, Lu B. 2004. Exchange rates and interest rates: can term structure models explain currency movements? *Journal of Economic Dynamics & Control* 28:1595–1624.

Ito T, Harada K. 2005. Japan premium and stock prices: two mirrors of Japanese banking crises. *International Journal of Finance & Economics* 10:195–211.

Ito T, Mishkin FS. 2004. Two decades of Japanese monetary policy and the deflation problem. NBER Working Paper 10878, National Bureau of Economic, Inc.

Ito T. 2005. Japanese monetary policy: 1998-2005 and beyond. BIS Papers No 31, Bank for International Settlements.

Japanese economy 2004 prospect for continuous recovery and risks (summary). Directorate General for Economic Research, Cabinet office, December 2004. [Available online at <http://www5.cao.go.jp/keizai3/2004/1219nk/je2004_summary.pdf>, accessed Sept. 5, 2012]

Kim CJ, Nelson CR. 1999. State-space Models with Regime Switching. MIT Press, Cambridge, Massachusetts.

Koopman S, Mallee M, van der Wel M. 2010. Analyzing the term structure of interest rates using dynamic Nelson-Siegel model with time-varying parameters. *Journal of Business and Economic Statistics* 28:329–343.

Kozicki S, Tinsley PA. 2001. Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics* 47:613–652.

Linton O, Mammen E, Nielsen J, Tanggard C. 2001. Estimating yield curves by kernel smoothing methods. *Journal of Econometrics* 105(1):185–223.

Litterman R, Scheinkman J. 1991. Common factors affecting bond returns. *Journal of Fixed Income* 1:54–61.

Lutkepohl H. 1991. *Introduction to Multiple Time Series Analysis*. Springer-Varlag, Berlin.

McCulloch JH. 1971. Measuring the term structure of interest rates. *The Journal of Business* 44: 19–31.

McCulloch JH. 1975. The tax-adjusted yield curve. *The Journal of Finance* 30:811–830.

Mishkin F. 1990. The information in the longer maturity term structure about future inflation. *Quarterly Journal of Economics* 105:815–828.

Monch E. 2006. Forecasting the yield curve in a data-rich environment: a no-arbitrage factor-augmented VAR approach. Working paper, Humboldt University Berlin.

Nagayasu J. 2004. The term structure of interest rates and monetary policy during a zero interest rate period. *Monetary and Economic Studies* 22(2):19–43.

Nakajima J, Shiratsuka S, Teranishi Y. 2010. The effects of monetary policy commitment: evidence from time-varying parameter VAR analysis. IMES Discussion Paper Series. Institute for Monetary and Economic Studies, Bank of Japan.

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

Oda N, Ueda K. 2007. The effects of the Bank of Japan's zero interest rate commitment and quantitative monetary easing on the yield curve: a macro-finance approach. *The Japanese Economic Review* 58(3):303–328.

Okina K, Shiratsuka S. 2004. Policy commitment and expectation formation Japan's experience under zero interest rates. *The North American Journal of Economics and Finance* 15(1):75–100.

Pesaran MH, Shin Y. 1998. Impulse response analysis in linear multivariate models. *Economics Letters* 58:17–29.

Phillips P. 1986. Understanding spurious regressions in econometrics. *Journal of Econometrics* 33:311–340.

Piazzesi M, Schneider M. 2006. Equilibrium yield curves. In *NBER Macroeconomics Annual*. Woodford M, Acemoglu D, Rogoff K. (eds.). MIT Press, Cambridge; 389–442.

Piazzesi M. 2005. Bond yields and the Federal Reserve. *Journal of Political Economy* 113:311–344.

Rudebusch GD, Wu T. 2003. A macro-finance model of the term structure, monetary policy and the economy. Manuscript, Federal Reserve Bank of San Francisco.

Schich S. 1999. What the yield curves say about inflation: does it change over time? Working Paper 227, OECD Economics Department.

Schumacher JM. 2009. Syllabus financial models, Tilburg University, Tilburg, the Netherlands.

Shea G. 1984. Pitfalls in smoothing interest rate term structure data: equilibrium models and spline approximations. *Journal of Financial and Quantitative Studies* 19:253–269.

Sims CA. 1980. Macroeconomics and reality. *Econometrica* 48:1–48.

Soderlind P, Svensson L. 1997. New techniques to extract market expectations from financial instruments. *Journal of Monetary Economics* 40 (2):383–429.

Stock J, Watson M. 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 16:788–829.

Svensson LEO. 1994. Estimating and interpreting forward interest rates: Sweden 1992-1994. NBER Working Paper Series. No. 4871.

Svensson LEO. 2003. Escaping from a liquidity trap and deflation: the foolproof way and others. NBER Working Paper 10195.

Thornton DL. 2004. Testing the expectations hypothesis: some new evidence for Japan. *Monetary and Economic Studies* 22(2):45–69.

Ullah W, Tsukuda Y, Matsuda Y. 2013. Term structure forecasting of government bond yields with latent and macroeconomic factors: do macroeconomic factors imply better out-of-sample forecasts? *Forthcoming in Journal of Forecasting*.

Ullah W. 2012. Term structure modeling and forecasting of government bond yields. Conference proceedings for 2012 meeting of World Finance & Banking Symposium, Shanghai, China, December 17–18, 2012.

Vasicek OA, Fong HG. 1982. Term structure modeling using exponential splines. *Journal of Finance* 37:339–348.

Vasicek OA. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5:177–188.

Wu T. 2002. Monetary policy and the slope factors in empirical term structure estimations. Working Paper 02-07, Federal Reserve Bank of San Francisco.

APPENDICES

Appendix A

Implied Forward and Spot Rates

This appendix deals with the derivation of implied forward rates and spot rates functions defined in (1.3) and (1.4) respectively.

1. Implied Forward Rates

Suppose that $R_t(1)$ and $R_t(2)$ are spot rates with $P_t(1)$ and $P_t(2)$ prices for one and two years maturities discount bonds respectively in period t. Assuming a unit face value of both bonds implies that $P_t(1) > P_t(2)$. One can consider the following two alternatives:

1. Investing the amount $P_t(2)$ in the two years maturity bond and receiving $\exp[2R_t(2)]$ annually;
2. Investing $P_t(1)$ in one year maturity bond today and receiving $\exp[R_t(1)]$ at the end of period t and reinvesting this amount in the one year bond with price $P_{t+1}(1)$ having yield $R_{t+1}(1)$, which is the period $(t+1)$ spot rate for the bond with maturity of one year.

The no-arbitrage condition implies that $P_{t+1}(1)$ and $R_{t+1}(1)$ must satisfy the following two conditions, respectively.

$$P_t(2) = P_t(1).P_{t+1}(1)$$
$$\exp[-2.R_t(2)] = \exp[-R_t(1)].\exp[-R_{t+1}(1)]$$

However, in spot market at t both $R_{t+1}(1)$ and $P_{t+1}(1)$ are not observable and can, therefore, be referred to as implied forward rate $f_t(1)$ and implied forward price respectively. The yield satisfying these conditions is also referred to as the implied forward rate of a one year bond belonging to the period $t+1$.

$$P_t(2) = P_t(1).\exp[-R_{t+1}(1)]$$
$$P_t(2) = P_t(1).\exp\left(-f_t(1)\right)$$
$$f_t(1) = -\log[P_t(2)/P_t(1)]$$
$$f_t(1) = -\log\left[1 + \frac{P_t^{'}(1)}{P_t(1)}\right]$$
$$f_t(1) = -\frac{P_t^{'}(1)}{P_t(1)}$$

Extending the argument to n-period case, the no-arbitrage condition implies that:

$$P_t(n) = P_t(1).P_{t+1}(1).P_{t+2}(1) \dots P_{t+n}(1)$$
$$\exp[-nR_t(n)] = \exp[-R_t(1)].\exp[-R_{t+1}(1)].\exp[-R_{t+2}(1)] \dots \exp[-R_{t+n}(1)]$$

The solution yields to:

$$f_t(m) = -\frac{P_t^{'}(m)}{P_t(m)}$$

2. Spot Rates

The discount function in (1.1) implies that:

$$\log[P_t(m)] = -mR_t(m)$$

Differentiating with respect to m yields:

$$\frac{d}{dm}\log[P_t(m)] = -\left[R_t(m) + m\frac{dR_t(m)}{dm}\right]$$

From (1.3)

$$-f_t(m) = -\left[R_t(m) + m\frac{dR_t(m)}{dm}\right]$$
$$\int_0^m f_t(u)\,du = \int_0^m \left[R_t(u) + u\frac{dR_t(u)}{du}\right] du$$
$$\int_0^m f_t(u)\,du = [u.R_t(u)]_0^m$$
$$R_t(m) = \frac{1}{m}\int_0^m f_t(u)\,du$$

Appendix B

CIR Model Results for 3, 6, 12 and 18 Months Maturity Data

The results of initial estimates of OLS along with the MLE optimal estimates using the dataset for 3 months, 6 months, 12 months, and 18 months maturity periods for the entire sample (2000:01–2011:12) are depicted in table A-1. The results of MLE show that the average fitted yield curve is upward sloping. Figure A-1 plots the average observed yields and the estimated yield curves for all the four maturities data. It shows that the CIR model plots an upward sloping yield curve like the observed positively sloped average yield curve. However, the discrepancy between the estimated curves for all the four data sets and average observed yield curve is very high.

Table A-1: Results of the MLE Estimation of the CIR Model

Maturity		$\hat{\kappa}$	$\hat{\mu}$	$\hat{\sigma}$	log L
3 Months	Initial (OLS)	0.8729	0.0017	0.0983	
	MLE	1.4762	0.0017	0.0743	5969.100
6 Months	Initial (OLS)	1.1527	0.0017	0.1350	
	MLE	1.9030	0.0017	0.0821	5794.000
12 Months	Initial (OLS)	0.7615	0.0021	0.0982	
	MLE	1.5163	0.0022	0.0788	6190.400
18 Months	Initial (OLS)	0.8642	0.0026	0.0876	
	MLE	1.6859	0.0027	0.0807	6542.700

Note: The table presents the initial OLS and MLE estimated results of $\hat{\xi}$ vector using the time series data of 3 months, 6 months, 12 months, and 18 months maturities from 2000:01–2011:12. log L denotes the log likelihood value of the MLE estimation. The number of observations is 144.

Furthermore, we estimate the CIR model for the two sub-periods, sub-period 1(2000:01–2006:12) and sub-period 2 (2007:01–2011:12) to observe the yield curve behavior during the prolonged period of zero policy rates. In table A-2, we provide the initial estimates and MLE estimated parameters for the two subsets of data, i.e., the zero interest rate period (2000– 2006) and the non-zero interest rate period (2007–2011). Furthermore, the estimated yield curves for both the sub-periods are depicted in figure A-2.

The maximum likelihood estimates for the first sub-period shows that the fitted yield curve is negatively sloped, however for the second sub-period the estimated yield curve has an upward slope for all the four maturities data sets.

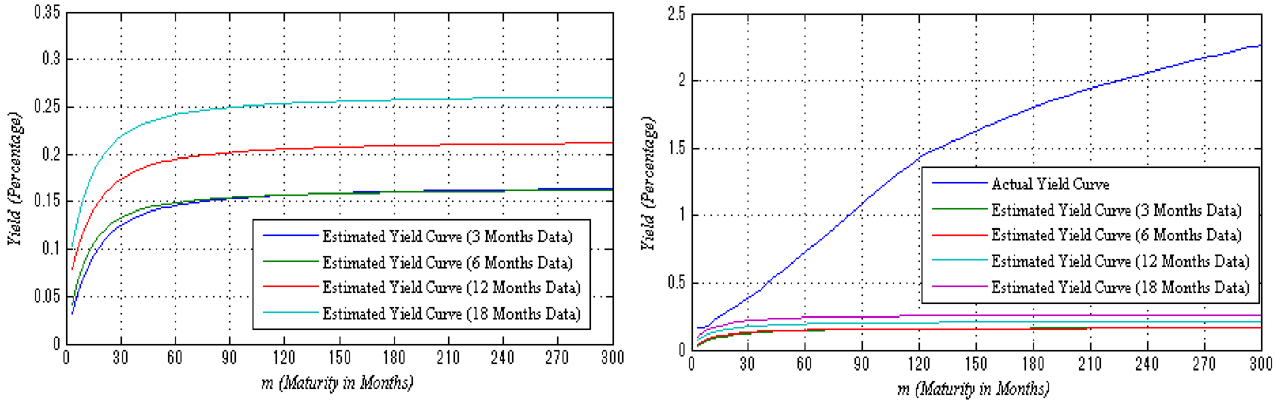


Figure A-1: Fitted Yield Curves with the CIR Model
Actual average (data-based) and fitted (model-based) yield curves for various maturities are plotted. The fitted yield curves are obtained by evaluating the CIR function at the MLE estimated $\hat{\kappa}, \hat{\mu}$ and $\hat{\sigma}$ from table A-1.

Overall the results of 3 months, 6 months, 12 months, and 18 months maturities data sets generate the same yield curve as we have estimated using the two years maturity data for the overall sample as well as for the two sub-periods, however, the 24 months yield data fits the estimated yield curve slightly better than the 3 months, 6 months, 12 months, and 18 months at short maturities.

Table A-2: Results of the MLE Estimation of the CIR Model for Sub-Periods

Maturity		$\hat{\kappa}$	$\hat{\mu}$	$\hat{\sigma}$	log L
		Sub-Period I (2000:01– 2006:12)			
3 Months	Initial (OLS)	1.3230	0.0011	0.1119	
	MLE	3.7744	0.0008	0.0825	3344.000
6 Months	Initial (OLS)	1.3230	0.0011	0.1119	
	MLE	3.7744	0.0008	0.0825	4121.000
12 Months	Initial (OLS)	0.6122	0.0024	0.1116	
	MLE	2.2455	0.0017	0.0885	3371.000
18 Months	Initial (OLS)	0.6577	0.0028	0.0899	
	MLE	1.9660	0.0023	0.0829	3708.600
		Sub-Period II (2007:01– 2011:12)			
3 Months	Initial (OLS)	0.8952	0.0020	0.0771	
	MLE	1.0691	0.0022	0.0696	2676.200
6 Months	Initial (OLS)	1.1414	0.0023	0.0779	
	MLE	1.3698	0.0024	0.0700	2709.500
12 Months	Initial (OLS)	1.4333	0.0027	0.0770	
	MLE	1.8056	0.0028	0.0701	2856.600
18 Months	Initial (OLS)	1.5149	0.0031	0.0852	
	MLE	2.0789	0.0032	0.0807	2910.600

Note: The table presents the initial OLS and MLE estimated results of $\hat{\xi}$ vector using the time series data of 3 months, 6 months, 12 months, and 18 months maturities for two sub-periods, i.e., sub-period 1 (2000:01 – 2006:12) and sub-period 2 (2007:01 – 2011:12). log L denotes the log likelihood value of the MLE estimation. The number of observations for the first sub–period and second sub–period is 84 and 60 respectively.

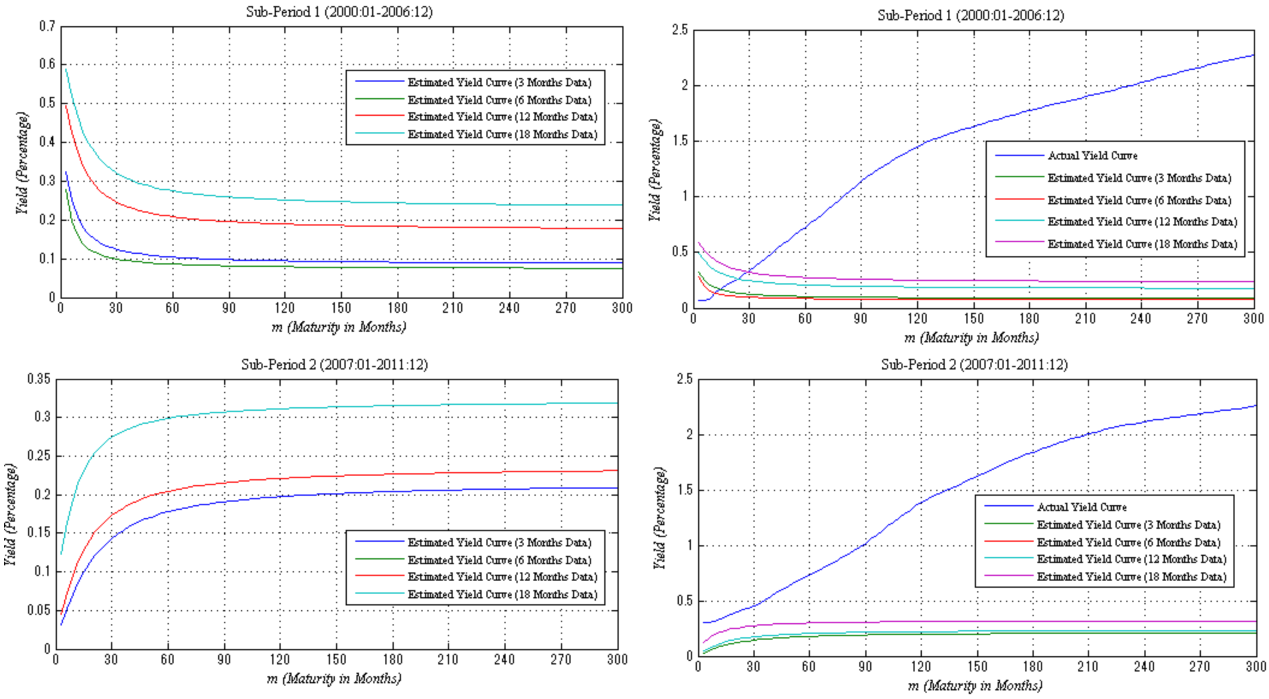


Figure A-2: Fitted Yield Curve with the CIR Model for Two Sub-Periods

Actual average (data-based) and fitted (model-based) yield curves for two sub-periods, i.e., sub-period 1 (2000:01 – 2006:12) and sub-period 2 (2007:01 – 2011:12) using the time series data of 3 months, 6 months, 12 months, and 18 months maturities are plotted. The fitted yield curves are obtained by evaluating the CIR function at the MLE estimated $\hat{\kappa}, \hat{\mu}$ and $\hat{\sigma}$ from table A-2.

Appendix C

Derivation of Analytical Gradient $\nabla F(\boldsymbol{\beta}, \boldsymbol{\tau})$ for the Non-Linear Ordinary Least Square of the Nelson-Siegel Model

To minimize the sum of squared zero-coupon yield errors, the objective function $F(\beta, \tau)$ is as given in (2.16):

$$F(\beta, \tau) = [R(m) - \Lambda(\tau)\beta]^2 \tag{A-1}$$

Differentiate the objective function in (A-1) w.r.t β and τ,

$$\frac{\partial F}{\partial \beta_1} = [-2(R(m) - b)] = 0 \tag{A-2}$$

$$\frac{\partial F}{\partial \beta_2} = [-2\tau a.(R(m) - b)] = 0 \tag{A-3}$$

$$\frac{\partial F}{\partial \beta_3} = [-2[\tau a - \exp(-m/\tau)].(R(m) - b)] = 0 \tag{A-4}$$

$$\begin{aligned}\frac{\partial F}{\partial \tau} = \Bigg[2\Bigg\{&-\beta_2\left(a + \frac{\exp(-m/\tau)}{\tau}\right) \\ &- \beta_3\left(a - \frac{\exp(-m/\tau)}{\tau} - \frac{\exp(-m/\tau).m}{\tau^2}\right)\Bigg\}.(R(m) - b)\Bigg] = 0\end{aligned} \tag{A-5}$$

where

$$a = \frac{[1 - \exp(-m/\tau)]}{m}$$

$$b = -\beta_1 - \beta_2 a\tau - \beta_3[a\tau - \exp(-m/\tau)]$$

The system of equations derived analytically in (A-2), (A-3), (A-4) and (A-5) is non-linear and can be solved numerically. The numerical solution of the system implies to the Nelson-Siegel estimated factors vector $\hat{\psi}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\tau}_t)'$.

Appendix D

Out-of-Sample Forecast Performance of the Nelson-Siegel Model

We follow the Diebold and Li (2006) method and model the estimated four time-varying factors of Nelson-Siegel model as first order auto-regressive and vector auto-regressive and make out of sample forecast for one month, 6 months and 1 year horizons.[78] The yield forecasts based on underlying univariate AR(1) factor specifications are:

$$\hat{R}_{t+h}(m) = \Lambda(\hat{\tau}_{t+h})\hat{\beta}_{t+h} \quad \text{(A-6)}$$

$$\hat{\psi}_{t+h} = A_0 + A_1\hat{\psi}_t + \varepsilon_{t+h} \quad \text{(A-7)}$$

where A_0 is (4×1) vector of constants, A_1 is (4×4) diagonal matrix, $\hat{\psi}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\tau}_t)'$ and $\varepsilon_{t+h} \backsim N(0, \Sigma)$ is (4×1) error vector. A_0 and A_1 are obtained by regressing $\hat{\psi}_t$ on $\hat{\psi}_{t-h}$. The multivariate VAR(1) model specification is same as in (A-7) but we modify A_1 to be (4×4) full matrix rather than a diagonal matrix.

We estimate and forecast recursively, using data from January 2000 to the time that the forecast is made, beginning in January 2008 and extending through December 2011. Subsequently, we substitute the forecasted factors $\hat{\psi}_{t+h}$ at time t in (A-6) to get the forecasted yield denoted as $\hat{R}_{t,t+h}(m)$.

In tables A-3, A-4 and A-5, we compute the descriptive statistics of *h-month-ahead* out-of sample forecasting results of the dynamic Nelson–Siegel models of AR(1) and VAR(1) representation of $\hat{\psi}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\tau}_t)'$, for maturities of 3, 6, 12, 18, 24, 36, 60, 120, 180, 240 and 300 months for the forecast horizons of $h = 1, 6$ and 12 months.

We define forecast errors at $t + h$ as $[R_{t+h}(m) - \hat{R}_{t,t+h}(m)]$, where $\hat{R}_{t,t+h}(m)$ is the forecasted yield in period t for $t + h$ period and is not the Nelson–Siegel fitted yield. $R_{t+h}(m)$ is the actual yield in period $t + h$. We examine a number of descriptive statistics for the forecast errors, including mean, standard deviation, mean absolute error (MAE), root mean squared error (RMSE) and autocorrelations at various displacements.

The results of one month ahead forecast of AR(1) and VAR(1) representation are reported in table A-3. The one month ahead forecasting results appear suboptimal as the forecast errors appear serially correlated. The average forecast errors and RMSE are much smaller than that of

[78] Diebold and Li (2006) model the three estimated factors of Nelson-Siegel model $\hat{\beta}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t})'$ as they assume the shape parameter τ_t is constant. Contrarily, we model the four estimated factors of Nelson-Siegel model $\hat{\psi}_t = (\hat{\beta}_{1t}, \hat{\beta}_{2t}, \hat{\beta}_{3t}, \hat{\tau}_t)'$, assuming a time-varying τ_t.

the related work such as Bliss (1997), de Jong (2000) and Diebold and Li (2006). In relative terms, RMSE comparison at various maturities reveals that AR(1) forecasts are slightly better than the VAR(1), however in term of serial correlation of errors the VAR(1) outperform the AR(1) specification.

Table A-3: Out-of-Sample 1 Month Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
Forecast Summary for AR(1) Specification							
3	0.049	0.152	0.102	0.047	0.865	-0.067	0.000
6	0.022	0.143	0.102	0.039	0.850	-0.076	0.000
12	-0.022	0.148	0.126	0.028	0.892	-0.073	0.000
18	-0.046	0.181	0.163	0.037	0.870	-0.059	0.000
24	-0.079	0.197	0.187	0.039	0.856	-0.029	0.000
36	-0.120	0.227	0.227	0.050	0.821	0.000	0.000
60	-0.152	0.251	0.255	0.077	0.768	0.059	0.000
120	-0.003	0.188	0.145	0.065	0.547	0.093	0.000
180	0.098	0.169	0.157	0.052	0.495	0.045	0.000
240	0.128	0.162	0.172	0.045	0.550	-0.018	0.000
300	0.087	0.145	0.139	0.028	0.643	-0.067	0.000
Forecast Summary for VAR(1) Specification							
3	0.048	0.208	0.140	0.079	0.829	-0.048	0.000
6	-0.053	0.321	0.245	0.143	0.825	0.063	0.000
12	-0.222	0.552	0.147	0.515	0.841	0.112	0.000
18	-0.348	0.764	0.235	0.426	0.846	0.119	0.000
24	-0.460	0.928	0.383	0.073	0.847	0.123	0.000
36	-0.612	1.166	0.396	0.527	0.848	0.124	0.000
60	-0.735	1.375	1.080	1.012	0.847	0.126	0.000
120	-0.515	1.213	0.596	0.557	0.83	0.128	0.000
180	-0.277	0.944	0.761	0.376	0.826	0.120	0.000
240	-0.141	0.744	0.593	0.806	0.816	0.117	0.000
300	-0.107	0.598	0.477	0.478	0.826	0.111	0.000

The table presents the results of out-of-sample 1-month-ahead forecasting using AR (1) and VAR (1) specification of the estimated factors. We estimate all models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+1$ as $R_{t+1}(m) - \hat{R}_{t,t+1}(m)$, where $\hat{R}_{t,t+1}(m)$ is the $t+1$ month ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors and root mean squared errors of the forecast errors, as well as their first, 12th and 24th order sample autocorrelation coefficients.

The results in table A-4 and A-5 of 6 months and one year ahead forecast respectively, reveal that matters worsen radically with longer horizons forecast. For 6 months ahead forecast, the AR(1) forecasts are slightly better than the VAR(1), while for the 12 months ahead, the VAR(1) performs better than the AR(1) in terms of lower RMSE. However, in regard of auto-correlation of the forecast errors, VAR(1) outperforms AR(1) for all maturities in both 6 and 12 months ahead forecasts.

Table A-4: Out-of-Sample 6 Months Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
Forecast Summary for AR(1) Specification							
3	0.096	0.184	0.122	0.077	0.883	-0.067	0.000
6	0.078	0.172	0.116	0.066	0.867	-0.071	0.000
12	0.050	0.177	0.126	0.061	0.889	-0.065	0.000
18	0.037	0.208	0.154	0.082	0.872	-0.046	0.000
24	0.013	0.224	0.173	0.088	0.852	-0.017	0.000
36	-0.018	0.254	0.205	0.102	0.809	0.012	0.000
60	-0.049	0.278	0.234	0.107	0.755	0.070	0.000
120	0.070	0.210	0.169	0.080	0.571	0.104	0.000
180	0.145	0.185	0.191	0.070	0.527	0.060	0.000
240	0.159	0.176	0.200	0.059	0.587	-0.026	0.000
300	0.110	0.166	0.163	0.044	0.689	-0.081	0.000
Forecast Summary for VAR(1) Specification							
3	-0.660	0.451	0.698	0.095	0.820	0.040	0.000
6	-0.647	0.435	0.680	0.063	0.808	0.065	0.000
12	-0.622	0.434	0.661	0.057	0.813	0.075	0.000
18	-0.589	0.454	0.647	0.144	0.828	0.083	0.000
24	-0.573	0.463	0.642	0.132	0.830	0.090	0.000
36	-0.539	0.484	0.633	0.142	0.828	0.092	0.000
60	-0.474	0.502	0.597	0.140	0.820	0.101	0.000
120	-0.194	0.413	0.366	0.247	0.724	0.127	0.000
180	-0.008	0.363	0.265	0.271	0.666	0.116	0.000
240	0.083	0.312	0.228	0.239	0.636	0.100	0.000
300	0.086	0.270	0.201	0.183	0.643	0.070	0.000

The table presents the results of out-of-sample 6-month-ahead forecasting using AR (1) and VAR (1) specification of the estimated factors. We estimate all models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+6$ as $R_{t+6}(m) - \hat{R}_{t,t+6}(m)$, where $\hat{R}_{t,t+6}(m)$ is the $t+6$ months ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors and root mean squared errors of the forecast errors, as well as their first, 12th and 24th order sample autocorrelation coefficients.

Furthermore, we also compute the Trace Root Mean Squared Prediction Error (TRMSPE) which combines the forecast errors of all maturities and summarizes the performance of each model, thereby allowing for a direct comparison between the models.[79] In table A-6, we report the TRMSPE for both the specifications of yield curve factors, i.e., AR(1) and VAR(1) for all the three forecasts horizons.

[79] Given a sample of T out-of-sample forecasts of N distinct maturities with $h-$months ahead forecast horizon, we compute the TRMSPE as follows:

$$TRMSPE = \sqrt{\frac{1}{NT}\sum_{m=1}^{N}\sum_{t=1}^{T}\left[R_{t+h}(m) - \hat{R}_{t,t+h}(m)\right]^2}$$

where $\hat{R}_{t,t+h}(m)$ is the forecasted yield in period t for $t+h$ period.

Table A-5: Out-of-Sample 12 Months Ahead Forecasting Results

Maturity	Mean	Std. Dev.	MAE	RMSE	$\hat{\rho}$ (1)	$\hat{\rho}$ (12)	$\hat{\rho}$ (24)
Forecast Summary for AR(1) Specification							
3	0.093	0.197	0.130	0.083	0.848	-0.003	0.000
6	0.075	0.183	0.125	0.070	0.874	-0.074	0.000
12	0.046	0.188	0.138	0.061	0.897	-0.070	0.000
18	0.032	0.216	0.167	0.075	0.896	-0.054	0.000
24	0.006	0.230	0.187	0.077	0.882	-0.029	0.000
36	-0.028	0.258	0.221	0.084	0.881	-0.068	0.000
60	-0.065	0.281	0.246	0.091	0.798	0.048	0.000
120	0.050	0.211	0.169	0.074	0.612	0.078	0.000
180	0.129	0.186	0.183	0.066	0.557	0.034	0.000
240	0.150	0.179	0.196	0.057	0.612	-0.049	0.000
300	0.108	0.170	0.163	0.045	0.716	-0.098	0.000
Forecast Summary for VAR(1) Specification							
3	-0.081	0.150	0.135	0.041	0.590	-0.070	0.000
6	-0.091	0.135	0.129	0.034	0.510	-0.080	0.000
12	-0.106	0.135	0.144	0.031	0.599	-0.059	0.000
18	-0.109	0.162	0.171	0.033	0.699	-0.024	0.000
24	-0.127	0.185	0.198	0.038	0.722	0.026	0.000
36	-0.149	0.236	0.251	0.059	0.746	0.066	0.000
60	-0.172	0.309	0.305	0.111	0.758	0.087	0.000
120	-0.035	0.299	0.229	0.175	0.676	0.015	0.000
180	0.064	0.271	0.199	0.177	0.639	-0.049	0.000
240	0.100	0.236	0.191	0.139	0.609	-0.073	0.000
300	0.067	0.208	0.164	0.099	0.621	-0.073	0.000

The table presents the results of out-of-sample 12-month-ahead forecasting using AR (1) and VAR (1) specification of the estimated factors. We estimate all models recursively from 2000:1 to the time that the forecast is made, beginning in 2008:1 and extending through 2011:12. We define forecast errors at $t+12$ as $R_{t+12}(m) - \hat{R}_{t,t+12}(m)$, where $\hat{R}_{t,t+12}(m)$ is the $t+12$ months ahead forecasted yield at period t, and we report the mean, standard deviation, mean absolute errors and root mean squared errors of the forecast errors, as well as their first, 12th and 24th order sample autocorrelation coefficients.

The performances of AR(1) is to some extent superior to that of the VAR(1) model of factors in terms of TRMSPE for the one month and six months ahead forecasts horizons, while the VAR(1) outperform the AR(1) for twelve months ahead forecasts. It suggests that for longer horizons forecasts the multivariate VAR(1) specification of factors can forecast the future yields with greater accuracy than the univariate AR(1) model of factors.

Table A-6: TRMSPE Results for Out-of-Sample Forecasts Accuracy Comparisons

TRMSPE	1 Month Forecasts	6 Months Forecasts	12 Months Forecasts
AR(1) Model of Factors	0.046	0.076	0.079
VAR(1) Model of Factors	0.054	0.085	0.055

Note: The table reports the Trace Root Mean Squared Prediction Error (TRMSPE) results of out-of-sample forecasts accuracy comparison for horizons of one, 6, and 12 months for both AR(1) and VAR(1) specification of factors.

In summary, the out-of-sample forecast results of the Nelson-Siegel seem reasonably well in terms of lower forecast errors, however the errors are serially correlated. These results are slightly different from Dieobld and Li (2006). In term of lower RMSE, our results for all the three horizons forecast are preferred than that of related studies. Diebold and Li (2006) have a great success in forecasts, particularly in terms of the errors persistency, using a different dataset with maturities up to 10-year, whereas we have maturities up to 25-year. The original Nelson-Siegel framework might forecast the long maturities sub-optimally. The serial correlation of forecast errors may likely come from a variety of sources, some of which could be eliminated, such as, pricing errors due to illiquidity may be highly persistent and could be reduced by including variables that may explain mispricing as suggested by Dieobld and Li (2006).

Appendix E

Coefficients in the General State-space Form

$$y_t = \begin{bmatrix} R_t(m) \\ Z_t \end{bmatrix} \qquad \alpha_t = \left[\beta_t', \widetilde{IP}_t, \widetilde{MS}_t, \widetilde{INF}_t, \varepsilon_t^*\right]'$$

$$\mathrm{H} = \begin{bmatrix} \Lambda(\tau) & \Gamma_\varepsilon & 0 \\ 0 & 0 & I_3 \end{bmatrix} \qquad K = \begin{bmatrix} A & 0_6 \\ 0_6' & 0 \end{bmatrix}$$

$$C = \begin{bmatrix} (I_6 - A)\mu \\ 0 \end{bmatrix} \qquad G = I_7$$

$$w_t = \begin{bmatrix} \varepsilon_t^+ \\ 0 \end{bmatrix} \qquad u_t = \begin{bmatrix} v_{t+1} \\ \varepsilon_{t+1}^* \end{bmatrix}$$

$$R = \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} \qquad Q_t = \begin{bmatrix} \Sigma_v & 0 \\ 0 & h_{t+1} \end{bmatrix}$$

where $\alpha_t = \left(\beta_t', \widetilde{IP}_t, \widetilde{MS}_t, \widetilde{INF}_t, \varepsilon_t^*\right)' = \left(\beta_{1t}, \beta_{2t}, \beta_{3t}, \widetilde{IP}_t, \widetilde{MS}_t, \widetilde{INF}_t, \varepsilon_t^*\right)'$ is the (7×1) vector of yield curve and macroeconomic factors, $R_t(m)$is (N×1) vector of zero-coupon yield, $Z_t = (IP_t, MS_t, INF_t)'$ is the (3×1) vector of macroeconomic factors, β_t is (3×1) vector of Nelson-Siegel factors, $\Lambda(\tau)$ is (N×3) matrix of factors loadings, A is (6×6) matrix of parameters, μ is (6×1) mean vector of factors, I_6 and I_3 are (6×6) and (3×3) identity matrices respectively and Γ_ε is (N×1) vector. Σ_v is (6×6), the covariance matrix of innovations of the transition system and Ω is the (N×N) dimension covariance matrix of the innovations to the measurement system. Furthermore, ε_t^+ is the (N×1) error vector of measurement equation and v_{t+1} is (6×1) innovation vector of first six state equations.

Appendix F

The VAR Model and Calculation of Impulse Response Functions and Variance Decompositions

The VAR (1) models considered in this study, in general notations can be written as:

$$y_t = \mu + Ay_{t-1} + \varepsilon_t \tag{A-8}$$

where y_t is $(m \times 1)$ vector of endogenous variables, μ is the constant vector and A is the $(m \times m)$ transition matrix. The errors ε_t follow:

$$\varepsilon_t \sim N(0, \Sigma) \tag{A-9}$$

where Σ is $(m \times m)$ covariance matrix and is assumed to be non-diagonal. Letting that the unconditional mean of y_t is $c = (I_m - A)^{-1}\mu$, where I_m is $(m \times m)$ identity matrix. We can write (A-8) as:

$$y_t - c = \mu + A(y_{t-1} - c) + \varepsilon_t \tag{A-10}$$

Assuming that all roots of A lie inside the unit circle, we can write the moving average (MA) representation of the VAR in (A-10) as:

$$y_t - c = \sum_{i=0}^{\infty} A^i \varepsilon_{t-i} = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i} \tag{A-11}$$

For $i = 1,2,....$ and $\Psi_i = A^i$, where A^i denotes the matrix A multiplied by itself i times, but $\Psi_0 = I_m$.

As shown in section 4.3.2 that the covariance matrix of VAR is non-diagonal, therefore we cannot compute the IRFs using the original residuals ε_t and the interpretation of elements in A in (A-8) is not straightforward. We use a Cholesky decomposition to transform the innovations so that the resulting components are uncorrelated. Specifically, we derive a lower-triangular matrix L such that $\Sigma = LDL'$, where D is a diagonal matrix and the diagonal elements of L are unity. The transformed residuals take the form as $v_t = L^{-1}\varepsilon_t$, then $\mathbb{E}(v_t v_t') = D$. But D is diagonal, verifying that the elements v_t are contemporaneously uncorrelated. We compute the response of y_t to a one standard deviation shock in ε_t as:

$$\frac{\partial y_{t+s}}{\partial v_{jt}} = \Psi_s l_j \qquad \text{(A-12)}$$

The variance decompositions (VDCs) provide information about the relative importance of each random innovation in the system in affecting the variables in the VAR. The contribution of the j^{th} variable to the mean squared error (MSE) of the s period ahead forecast, under the assumption of non-diagonal Σ is:

$$MSE^j(s) = Var(v_{jt})\left[l_j l_j' + \Psi_1 l_j l_j' \Psi_1' + \Psi_2 l_j l_j' \Psi_2' + \cdots + \Psi_{s-1} l_j l_j' \Psi_{s-1}'\right] \qquad \text{(A-13)}$$

The MSE is given by:

$$MSE(s) = \sum_{j=1}^{k} MSE^j(s) \qquad \text{(A-14)}$$

where both MSE and MSE^j are $(m \times m)$ matrices. We compute the VDC at forecast horizon s as the fraction of MSE of the i^{th} variable due to shocks to the j^{th} variable:

$$VDC_i^j(s) = \frac{MSE_{ii}^j(s)}{MSE_{ii}(s)} \qquad \text{(A-15)}$$

where MSE_{ii} and $MSE_{ii}^j(s)$ are the (i, i) elements of the MSE and MSE^j matrices respectively.