Term Structure Forecasting of Government Bond Yields with Latent and Macroeconomic Factors: Does Macroeconomic Factors Imply Better Out-of-Sample Forecasts?

著者	Ullah Wali, Tsukuda Yoshihiko, Matsuda Yasumasa
journal or publication title	Discussion Papers (Tohoku Economics Research Group)
number	287
year	2012-10-05
URL	http://hdl.handle.net/10097/55764

Discussion Paper Series No. 287

Term Structure Forecasting of Government Bond Yields with Latent and Macroeconomic Factors: Does Macroeconomic Factors Imply Better Out-of-Sample Forecasts?

by

Wali Ullah*, Yoshihiko Tsukuda**, and Yasumasa Matsuda***

October 5, 2012

Keywords: Term structure of interest rates; Macroeconomic fundamentals; Latent factors model; State-space model; Kalman Filter

JEL Codes: G1, E4, C5, C32

---

* Ph.D. student, Graduate School of Economics and Management, Tohoku University, 27-1 Kawauchi, Aoba-ku, Sendai 980-8576, Japan. (E-mail: wali76@yahoo.com)

** Professor and Executive Vice President for Finance and Campus Planning at Tohoku University, Sendai, Japan. (E-mail: tsukuda@econ.tohoku.ac.jp)

*** Professor of Statistics, Graduate School of Economics and Management, Tohoku University, Sendai, Japan (E-mail: matsuda@econ.tohoku.ac.jp)

We would like to thank the anonymous referee and the Departmental Editor Prof. Terence Mills for making useful comments to improve this paper. The research of the second author is partially supported by Grant-in-aid No. 22530205 from Japan Society for the Promotion of Science.

Abstract:

This study 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. We find that incorporating the macroeconomic indicators in yield curve model leads to a better in-sample fit of the term structure. The out-of-sample predictability also improves significantly for all maturities for the short horizon forecasts, however regarding the longer horizons forecasts, the forecast performance of yields-macro and yields-only models is same for maturities beyond 5 years. The one-step state-space estimation approach employed to the yields-macro model produces accurate forecasts and outperforms the results of earlier related studies. Especially, the autocorrelation of the forecasts errors and in-sample residuals persistency across maturities, which is a common phenomenon in the statistical class of term structure models, can be reduced to a greater extent by inclusion of macroeconomic factors in the yield model.

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 a 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 behavior of economic agents, 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 (1986, 1993, 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 (Wali, 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. The stylized facts that form the basis for the statistical class of models of the yields for various maturities are:

- Even though interest rates to a large extent vary randomly, they remain in an interval that is bounded below by 0% and above by around 15%, where the upper bound is less sharp and may be temporarily less volatile. This reflects stabilizing influence by central banks and other economic interactions. Modelers of the dynamics of interest rates believe that this calls for a mean reversion force to be incorporated when modeling the trend in the model.
- Although the interest rate term structure seems to be a high dimensional object, there is strong dependency between the various maturities rates.
- Furthermore, in market we observe three main shapes of term structure curves i.e. upward sloping (so-called normal yield curve), downward sloping (inverse yield curve) or mainly constant (flat yield curve). Although, some argue for curves with multiple humps but they are rarely observed.

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 cross sectional aspects of yield curve, Nelson and Siegel (1987) imposes a parsimonious three-factor structure on the links 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. 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 imposition of absence of arbitrage improves its empirical tractability and predictive performance of the Nelson-Siegel specification. Diebold and Li (2006) show that forecasts obtained from the Nelson-Siegel model outperform competing statistical models, while Diebold *et al.* (2006) argue that the Nelson-Siegel model in state-space form is capable of explaining observed time series with latent factors. Furthermore, Wali (2012) shows that Nelson-Siegel model outpace the competing economical models (affine class of models) in both aspects i.e. in-sample fit as well as forecasts for various horizons.

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 the 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*. (2002), Wu (2002), and Diebold *et al.* (2006), who explicitly incorporate macroeconomic determinants into multi-factors yield curve models. In these studies two assumptions of yield and macroeconomic factors interaction—one-way yields-to-macro or macro-to-yields links—are testable hypotheses, however Diebold *et al.* (2006) have analyzed the bidirectional feedback from the yield curve to the economy and back again. 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.

To assess the role of macroeconomic variables in the yield curve dynamics and forecasting, 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. This model has the substantial flexibility, required to match the changing shape of the yield curve, yet it is parsimonious and easy to estimate. We explicitly incorporate three macroeconomic variables i.e. the level of economic activity, exchange rate and inflation rate and one stock market activity indictor (Stock Market Index) 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. 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. Although, the studies that focus on the forecast performance of statistical class of models come with encouraging results, particularly in term 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 Wali, 2012). In order to overcome such problems in the forecasts errors, we show that inclusion of various yield curve related macroeconomic and stock market variables leads to the reduction/elimination of lags autocorrelation in the forecasts errors in dynamic Nelson-Siegel model.

The remainder of the study 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 the estimation method for both models. We use the Kalman filter method to estimate both the models. This one-step approach improves upon the two-step estimation procedure of Diebold and Li (2006) and provides a unified framework in which to examine the yield curve and the macroeconomy dynamic interaction. The third section deals with the data structure and compares estimation results for the two competing models. In fourth section, we present the out-of-sample forecast performance and the results of various tests to compare the forecast errors evolution over time and maturities of the two models. Finally, the fifth section presents the conclusion of the study.

2. TERM STRUCTURE MODELS AND ESTIMATION METHODS

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 around the world 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 forward rates.

In this section, we incorporate the Nelson-Siegel spot rates model in latent factor framework with and without macroeconomic variables to describe the dynamic evolution of yield curve. The latent factor model is considered as it will be a convenient vehicle for introducing the state-space representation. In the next two sub-sections 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.

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:[1]

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

[1] These types of exponential functions to fit and forecast the observed yield curve become popular as they reconcile the following characteristics:

— Sufficient flexibility to reflect the important and typical patterns of the observed market data.
— Relatively robust against disturbances from individual observations.
— Applicable with only a few observations.
— Results in more stable yield curves.

with the time varying parameter vector $\beta_t = [\beta_{1t}, \beta_{2t}, \beta_{3t}]'$ and time invariant parameter τ. The forward rate representation chosen by Nelson-Siegel belongs to a class of Laguerre functions. These functions are characterized by a polynomial time 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 (1), resulting in:

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

for $i = 1, 2, 3, \ldots, N$ and $t = 1, 2, 3, \ldots, T$.

The Nelson-Siegel specification of yield in (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 (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. Since, the range of shapes the curve can take is dependent on τ, it can be interpreted as the shape parameter. The small values of τ tend to fit low maturities interest rates quite well and larger values of τ 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.

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

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

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

where $t = 1,2, \ldots, T$, and ε_t is $(N\times 1)$ vector of measurement errors, i.e. deviations of the observed yields in period t and 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 factor dynamics. We follow the dynamic frame-work of Diebold and Li (2006) and 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:

$$\begin{bmatrix} \beta_{1t} - \mu_1 \\ \beta_{2t} - \mu_2 \\ \beta_{3t} - \mu_3 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} \beta_{1,t-1} - \mu_1 \\ \beta_{2,t-1} - \mu_2 \\ \beta_{3,t-1} - \mu_3 \end{bmatrix} + \begin{bmatrix} \eta_{1t} \\ \eta_{2t} \\ \eta_{3t} \end{bmatrix} \tag{4}$$

where $t = 1, \ldots ..,T$ is the sample period, 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 may be written as:

$$R_t(m) = \Lambda(\tau)\beta_t + \varepsilon_t \tag{5}$$

$$\xi_t = \mathrm{A}\xi_{t-1} + \eta_t \tag{6}$$

The measurement equation in (5) specify the vector of yields, which contains N different maturities, $R_t(m) = [R_t(m_1) \ldots \ldots R_t(m_N)]'$, 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 $\Lambda(\tau)$ the (N × 3) matrices of factor loadings which are potentially time-varying if the shape parameter τ are 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:

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim WN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma \end{bmatrix}\right) \tag{7}$$

where Σ is (3×3), the co-variance matrix of innovations of the transition system and is assumed to be unrestricted, while the co-variance 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 co-variance matrix of the measurement errors necessary for computational tractability (Diebold *et al.*, 2006). 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.

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 (Wali, 2012), it is of immense interest to relate the estimated 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*. (2002) 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 rates will be interpreted as a rise in inflation expectation 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 conditions in 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 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 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 effects 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 rates 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 too should 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.

Since, the term structure includes significant amount of information about the market's expectation 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 these studies of 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 a flattening of the yield curve associated with changes in the curvature factor 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 in 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 for the stock market is concerned, the annual growth rate of stock market aggregate index (SI_t) is considered in the model as indicator of the capital market performance. Though, the stock market aggregate index (SI_t) is equity market indicator, in this study we call all the four variables (IP_t, EX_t, INF_t and 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 following system of equations.

$$\begin{bmatrix} R_t(m) \\ Z_t \end{bmatrix} = \begin{bmatrix} \Lambda(\tau) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \beta_t \\ \tilde{Z}_t \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \end{bmatrix} \tag{8}$$

$$\xi_t = \mathrm{A}\xi_{t-1} + \eta_t \tag{9}$$

$$\begin{bmatrix} \varepsilon_t \\ \eta_t \end{bmatrix} \sim WN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 \\ 0 & \Sigma \end{bmatrix} \right) \tag{10}$$

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 and μ (7×1) vectors of estimated mean of factors. 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 (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 guarantee 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 yields at various maturities.

Furthermore in (7) and in (10), we assume that the innovations of both, the measurement 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 a central limit theorem.

2.3. Estimation Method

There are several approaches to estimating 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 month 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 (5-6) and (8-9) as linear Gaussian state-space models. The state vector of unobserved factors β_t can be estimated conditional on the past and current observations $R_1, R_2, \ldots, R_t$ via the Kalman filter. Defining $\hat{\xi}_{t|s}$ as the minimum mean square linear estimator (MMSLE) of β_t given $R_1, R_2, \ldots, R_s$ 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:

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + W_{t|t-1}\Lambda(\tau)' F_t^{-1} v_t \tag{11}$$

$$W_{t|t} = W_{t|t-1} - W_{t|t-1}\Lambda(\tau)'F_t^{-1}\Lambda(\tau)W_{t|t-1} \tag{12}$$

with

$$v_t = R_t - \Lambda(\tau)\hat{\xi}_{t|t-1} \tag{13}$$

$$F_t = \Lambda(\tau)W_{t|\mathrm{t-1}}\Lambda(\tau)' + \Omega \tag{14}$$

where v_t is the prediction error vector and F_t is the prediction error co-variance matrix. The MMSLE of the state vector for the next period $t + 1$, conditional on $R_1, R_2, \ldots, R_t$, is given by the prediction step as:

$$\hat{\xi}_{t+1|t} = A\hat{\xi}_{t|t} \tag{15}$$

$$W_{t+1|t} = AW_{t|t}A' + \Sigma \tag{16}$$

For a given time series of $R_1, R_2, \ldots, R_t$, the Kalman filter computations are carried out recursively for $t = 1, \ldots, T$ with initializations $\hat{\xi}_{1|0} = \mu$ (the unconditional mean) and $W_{1|0} = \Pi$, where Π is the co-variance 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 factors 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 (12) and (16) and storing $\hat{\xi}_{t|t}$ and $\hat{\xi}_{t+1|t}$ as obtained in (11) and (15) respectively for t=1,2,…,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 co-variance matrix.

The sequence of smoothed estimates $\hat{\beta}_{t|T}$ is then calculated in reversed order by iterating on:

$$\hat{\xi}_{t|T} = \hat{\xi}_{t|t} + W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\left(\hat{\xi}_{t+1|T} - \hat{\xi}_{t+1|t}\right) \qquad (17$$

for t=T-1,T-2,…,1. The corresponding co-variance matrix is similarly found by iterating on:

$$W_{t|T} = W_{t|t} - \left(W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\right)\left(W_{t+1|T} - W_{t+1|t}\right)\left(W_{t|t}\Lambda(\tau)'W_{t+1|t}^{-1}\right)' \tag{18}$$

in reverse order for t=T-1,T-2,…,1.

The parameters in the VAR(1), the constants vector μ, coefficients matrix A and both the co-variance matrices (Ω and Σ) along with 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:

$$ln\, L\,(\theta) = -\frac{NT}{2}ln(2\pi) - \frac{1}{2}\sum_{t} \ln\left[\left|F_{t|t-1}(\theta)\right|\right] - \frac{1}{2}\sum_{t} v_{t|t-1}'\left[F_{t|t-1}(\theta)\right]^{-1} v_{t|t-1} \tag{19}$$

The specification in (19) is a function of the parameter set $\theta = (\tau, A, \Omega, \Sigma)$. 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 $F_{t|t-1}$.[2] The shape parameter τ is assumed to be constant over time.

As a result, $ln\, L\,(\theta)$ in (19) can be evaluated by the Kalman filter for a given value of θ. Marquardt non-linear optimization algorithm is employed for the purpose of maximization based on the numerical evaluation. The state-space framework allows that the co-variance matrices Ω and Σ can be full or diagonal. Commonly, Ω is assumed to be diagonal (that for given β_t, the equations for the different yield maturities are uncorrelated) to reduce the number of coefficients and to obtain computational tractability.

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 representations provide a powerful framework for analysis

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

and estimation of dynamic models. The recognition that the Nelson–Siegel function is easily put in state-space form is particularly useful because application of the Kalman filter then delivers maximum-likelihood estimates and optimal filtered and smoothed estimates of the underlying factors. 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.

3.1. Data

The data we use are monthly spot rates for zero-coupon and coupon bearing bonds, generated using pricing data of Japanese bonds and treasury bills. The standard way of measuring the term structure of interest rates is by means of the spot rates on zero-coupon bonds. However, due to the limited maturities spectrum and lack of market liquidity of treasury bills, it is inevitable to derive the longer maturity zero-coupon rates from coupon-bearing treasury notes and bonds. 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[3], in the first stage, each month we calculate one day continuously compounded forward rates for the available maturities from the price data, and in second stage, we sum the daily forward rates to generate end of month term structure of yield 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 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).

<<Table 1>>

[3] Fama and Bliss (1987) provide a description of the methodology to derive the zero-coupon yield from the observed bond prices. They estimate forward rates at the observed maturities from the bond pricing data. Their method sequentially constructs the forward rates necessary to price successively longer-maturity bonds, called unsmoothed Fama-Bliss forward rates, and then constructs unsmoothed Fama-Bliss yields by averaging the appropriate unsmoothed Fama-Bliss forward rates.

Table I provides summary statistic for the dataset. For each maturity, we report mean, standard deviation, minimum, maximum, skewness, kurtosis and autocorrelation coefficients at various displacements. The summary statistic reveals that the average yield curve is upward sloping. Unconditional volatility decreases by maturity and yields for all maturities are highly persistent. 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.

In addition to the findings in table I, we see a few interesting characteristics in figure 1, that plots cross-section of yields over time. 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. Furthermore, in our data set on average we observe the upward sloping yield curves.

<<Figure 1>>

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). The data for former three 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 four 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 were following *I(1)* process. Secondly, for the consistency purpose with the interest rate data, as our yields for all maturities is measured in annual percent format. 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. The descriptive statistics of the macroeconomic variables and capital market indicator are depicted in table II.

<<Table 1I>>

3.2. Estimation of the Models

We apply the Kalman filter to the state-space representation for yields-only model (5-7) and yields-macro model (8-10) to compute optimal yields predictions and the corresponding 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 co-variance 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.[4] The log-likelihood function as specified in (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 co-variance 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 square method to estimate the measurement equations 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 3.72 (the median value) given in Wali (2012).

We present estimation results of vector μ and matrix A for the yields-macro model in the first panel of table III, 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 errors of factors are sufficiently small, compared to the estimated coefficients.[5] 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 factor 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)

---

4 For detail of initializing the Kalman filter see Hamilton (1994).

5 The p-value for all intercept terms are small than 0.020.

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. Since, 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.

<<Table II1>>

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 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 positively related to level of economic activity while the impact on stock market performance is statistically insignificant. Since, 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 inflation rate. The inflation rate is negatively related to the curvature factor.[6]

<<Table 1V>>

---

[6] We also considered the state space model, considering the macroeconomic variable as exogenous in the transitional equation (9) that can be expressed as:

$$\xi_t = \mathrm{A}\xi_{t-1} + \Gamma Z_{t-1} + \eta_t$$

$$\eta_t \sim WN(0, \Sigma)$$

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) co-variance matrix of the error term (η_t) of state equation. We estimate and forecast with observation equation (8) and above mentioned state equation and the results of estimated state vector β_{t} and its forecast values are almost similar to our earlier representation of yields-macro model.

The estimates of co-variance matrix of the state innovations as depicted by Σ in (7 and 10) and Wald-test of its diagonality in both models is shown in table IV. 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 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 reject 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.

<<Table V>>

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 co-variance matrix using (17) and (18) respectively. Table V shows the descriptive statistics of the three time varying Kalman filter smooth estimates factors along with averaged smoothed residuals for both the models i.e. the yields-macro and yields-only models.

The estimated vector of parameters $\hat{\beta}_t$ is highly statistically significant for both the models.[7] Comparing the mean, standard deviation and other descriptive features of the estimated factors across models shows that both the models give rather similar estimates for the level, slope and curvature factors in magnitude. From the autocorrelations in the table 6 of the estimated factors, we can see that the β_{1t} is the 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 average residuals 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.

Furthermore, the time-series of the factors smoothed estimates with various empirical proxies and potentially related macroeconomic variables are plotted in figure 2 and 3. Figure 2 depicts the estimated factors of both the models with their empirical proxies. The level of

[7] The p-value of individual t-stat of all the estimated factors is less than 0.01% in both the competing models.

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 factor estimates for both the models give rather close similar estimates for the level, slope and curvature factors. The pairwise correlation of empirically defined factors and estimated factors of the yields-macro model are $\rho(L_t, \hat{\beta}_{1t}) = 0.768$, $\rho(S_t, \hat{\beta}_{2t}) = -0.902$ 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 the defined factors seem to follow the same pattern and hence, may truly be called level, slope and curvature factors, respectively.

<<Figure 2>>

The level factor is closely related to annual growth of money supply, namely MS_t= 100 ×[(M_t − M_{t-12})/ M_{t-12}] as depicted in figure 3.[8] The correlation between $\hat{\beta}_{1t}$ and annual growth of money supply is -0.352, consistent with inflationary expectations as suggested by the Fisher equation. It suggests the effectiveness of monetary policy in affecting future expectations about the long end of yield curve. Shocks to monetary policy are important sources of variation in bonds of long term maturity maturities. Monetary policy surprises that act to drive up short rates will alter expectations about future interest rates by shifting the level of the yield curve up in a persistent way and thereby stimulate the Japanese economy. It confer that the shift of long end and hence the shape of yield curve has an important information of the state of economy. The figure show that monetary policy shocks account for a substantial fraction of the variance in the long end of yield curve consistent with the costly price adjustment hypothesis of monetary policy. 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 right panel of figure 3. It suggests that monetary policy is an important source of variation of the shape of the yield curve and hence the macro-economy, even in the zero interest rate policy regime.

One can observe that $\hat{\beta}_{1t}$ and $-\hat{\beta}_{2t}$ follow almost the same pattern in figure 2.

[8] where M is level of seasonally adjusted money supply M2 and its data has been retrieved from International Financial Statistics, IMF.

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 regime during the decade.

In early 1998 in Japanese economy the demand was falling and the economy was heading into a recession and financial instability.[9] In order to avoid the severe recession, the so-called zero interest rate policy (ZIRP) was introduced and an easy monetary policy was adopted.[10] The economy did not respond quickly, however, it started to show some sign 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.[11]

<<Figure 3>>

In February 2001, the Bank introduced the Lombard lending facility as well as cutting the official discount rate from 0.5% to 0.35%.[12] 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.

---

[9] 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 In November 1997 (Ito and Mishkin, 2004).

[10] The overnight call rate was radically reduced to 0.25% in September, 1998 and to 0.15% beginning of 1999 from 0.5%.

[11] First, the ICT bubble ended and stock prices in the Japan 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.

[12] 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.

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 Bank on the ground that it would reduce the risk of commercial banks, and it was made clear that it was not intended as monetary policy, 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 expect that ZIRP will continue for a long time. Since, during the recovery regime the long end is gradually rotating and hence the slope is on increasing trend. The process completes around late 2003.

Since, during the initial period of ZIRP and severe recession, we observe a sharp decline in the yields of long term bonds and shape of yield curve and during the period of recovery the yield curve long end as well as 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 is an important leading indicator of the business condition and state of economy.

Furthermore, table VI and figure 4 present the descriptive statistics and the three dimensional plot of the smoothed residuals for the all the maturities. Both the models fit the yield curve remarkably well. Table VI 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 across maturities of yields-macro model is lower than of yields-only model almost for all maturities.

<<**Table 1V**>>

<<**Figure 4**>>

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 VII presents four different criterions to compare the in-sample fit of the yield curve. Table VII 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 the inclusion of macroeconomic factors leads to the estimation of 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 VI1>>

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 across maturities of the signal system innovations in the yields-macro model is lower than of the yields-only models and leads to 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 spreads 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.

4. OUT-OF-SAMPLE FORECASTING

A good approximation to yield curve dynamics should not only fit well in-sample, but also produces satisfactorily 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 two models i.e. the yields-macro and yields-only 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. both 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 month as well as 6 and 12 months ahead are considered. The h-month ahead factors forecasts, β_{t+h}, are iterated forecasts which follow from forward iteration of the state equations in (6 for yields-only model and 9 for yields-macro model) as:

$$\hat{\xi}_{t+h|t} = \hat{A}^h \hat{\xi}_{t|t} \tag{20}$$

where $\hat{A}^h$ denotes the matrix $\hat{A}$ multiplied by itself h times. The first three elements of $\hat{\xi}_{t+h|t}$ in (20) is subsequently substituted in the observation equations (that are $\hat{\beta}_{t+h|t}$) , results in:

$$\hat{R}_{t+h|t}(m_i) = \Lambda(\hat{\tau})\big(\hat{\beta}_{t+h|t}\big) \tag{21}$$

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

4.1. Term Structure Forecast Evaluation

In tables VIII, IX and X, we compute the descriptive statistics of *h-month-ahead* out-of-sample forecasting results of yields-macro and yields-only models, 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 time t for $t+h$ as $[R_{t+h}(m_i) - \hat{R}_{t,t+h}(m_i)]$, where $\hat{R}_{t,t+h}(m_i)$ is the forecasted yield in period t for $t+h$ period (for i^{th} maturity) and is not the Nelson–Siegel fitted yield. $R_{t+h}(m_i)$ 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 forecasts of yields-macro and yields-only models are reported in table VIII. The one month ahead forecasting results for the yields-only model appear suboptimal as the forecasts errors appear 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 model. The mean, MAE and RMSE of forecast errors of yields-macro model are slightly smaller than that of yields-only model for all maturities and much smaller than of the related work such as Bliss (1997), de Jong (2000) and Diebold and Li (2006). In relative terms, the results indicate that yields-macro model outperform the yields-only model for the one month ahead forecast horizon.

<<Table VII1>>

The results of 6 months and one year ahead forecast in table IX and X respectively reveal that matters worsen radically with longer horizon forecasts. For 6 months ahead forecast the yields-macro model outperform the yields-only model in term of mean forecast errors, MAE, RMSE and lag autocorrelation for all maturities. The 6 month ahead forecasts results seem not good as the one month ahead forecasts in term of lag autocorrelation. However, the forecast errors in terms of MAE, RMSE and lag autocorrelation are much better than the related studies for the yields-macro model.

<<Table 1X>>

For 12 months ahead the yields-macro model performs well than the yields-only model in terms of lower RMSE. However, the autocorrelation of the forecasts errors for both models is almost same for all the maturities. It is worth noting, moreover, that related papers such as Bliss (1997) and de Jong (2000) also find serially correlated forecast errors, often with persistence much stronger than ours.

<<Table X>>

In summary, the out-of-sample forecasts results of the yields-macro model seem reasonably well in term of lower forecasts errors and lags autocorrelation. These results are slightly different from Dieobld and Li (2006) for the yields-only model. In term of lower RMSE, our results for all the three horizons forecast are preferred than that of related studies 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.

4.2. Out-of-Sample Forecast Accuracy Comparisons

To assess the overall quality of the out-of-sample forecasts of the two competing models, we use a number of standard forecasts errors evaluation criteria. In particular, we report the Trace Root Mean Squared Prediction Error (TRMSPE), t-test for mean equality of squared forecast errors along with ANOVA F-test. Furthermore, we also employ Diebold and Mariano (1995) forecast accuracy comparison test to evaluate the models' overall predictive accuracy.

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 out-of-sample forecasts with h−months ahead forecast horizon, we compute the RMSE for a m_i maturity yield, with m = 1, 2,. . . ,N, as follows:

$$RMSE(m_i) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left[R_{t+h}(m_i) - \hat{R}_{t,t+h}(m_i)\right]^2} \tag{22}$$

where $[R_{t+h}(m_i) - \hat{R}_{t,t+h}(m_i)]$ is the forecast errors at $t+h$ for yield of maturity i and $\hat{R}_{t,t+h}(m_i)$ is the forecasted yield in period t for $t+h$ period.

The TRMSPE is an aggregate over all yield maturities for m = 1, 2,. . . ,N, as follows:

$$TRMSE = \sqrt{\frac{1}{NT}\sum_{m=1}^{N}\sum_{t=1}^{T}\left[R_{t+h}(m_i) - \hat{R}_{t,t+h}(m_i)\right]^2} \tag{23}$$

In table XI, we report the Trace Root Mean Squared Prediction Error (TRMSE) for both the models, i.e. yields-macro and yields-only model for all the three forecasts horizons.

<<Table X1>>

The results of TRMSE in table XI show that the forecasts worsen (the forecasts errors are getting larger with lengthening the forecast horizon), as we lengthen the forecast horizons for both the models. The forecasts of yields-macro model are a bit better than of yields-only model for horizon of one and 6 months, whereas for 12 month ahead forecast horizon the forecast errors of yields-macro model are much smaller than that of yields-only model.

4.2.2. Diebold-Mariano Test

In order to get a more deep insight, we employ the Diebold and Mariano (1995) test for the squared forecast errors in order to make a direct comparison between the two models for each maturity and each forecast horizon.

The main feature of Diebold and Mariano (DM, 1995) model-free test of forecast accuracy is that it is directly applicable to quadratic loss functions, multi-period forecasts, and forecast errors that are non-Gaussian with non-zero-mean and serially and contemporaneously correlated (correlated across maturities as well as over time). Assuming the forecast errors as:

$$e_t = \left[R_{t+h}(m) - \hat{R}_{t,t+h}(m)\right]^2 \tag{24}$$

where $[R_{t+h}(m) - \hat{R}_{t,t+h}(m)]$ is the forecast errors at t for $t+h$ yield and $\hat{R}_{t,t+h}(m_i)$ is the forecasted yield in period t for $t+h$ period. The basis of the test is the sample mean of the observed differential of quadratic loss series as:

$$d_t = e_{1t} - e_{2t} \tag{25}$$

where e_{1t} and e_{2t} are the quadratic loss functions of the bench mark model and competing model respectively for t=1,2,…,T. Assuming covariance stationarity and other regularity conditions on the process d_t, we use the standard result that:

$$\sqrt{T}\left(\bar{d} - \mu\right) \xrightarrow{d} N[0, 2\pi f_d(0)] \tag{26}$$

where $f_d(\cdot)$ is the spectral density of d_t and $\bar{d}$ is the sample mean of differential of quadratic loss function.

$$f_d(\lambda) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_d(k)\, exp[-ik\lambda] \qquad \text{for} - \pi \leq \lambda \geq \pi \tag{27}$$

$$\bar{d} = \frac{1}{T}\sum_{t=1}^{T}[e_{1t} - e_{2t}] \tag{28}$$

where $\gamma_d(k)$ in (27) is the auto-covariance of d_t sequence at displacement *k*:

$$\gamma_d(k) = E[(d_t - \mu)(d_{t-k} - \mu)] \tag{29}$$

The Diebold-Mariano test statistic is:

$$DM = \frac{\bar{d}}{\sqrt{2\pi\hat{f}_d(0)/T}} \tag{30}$$

$$DM \sim N(0,1) \tag{31}$$

where $\hat{f}_d(0)$ is a consistent estimate of $f_d(0)$. The null hypothesis as $H_0: E(d_t) = 0$ is rejected in favour of the two sided alternative hypothesis that $H_1: E(d_t) \neq 0$, when DM, in absolute value, exceeds the critical value of a standard unit Gaussian distribution. This function also corrects for the autocorrelation that multi-period forecast errors usually exhibit. Note that an efficient *h*-period ahead forecast will have forecast errors following *MA*(*h-1*) processes. Diebold and Mariano (1995) use a Newey-West type estimator for sample variance of the loss differential to account for this concern.

Furthermore, consistent estimators of $f_d(0)$ can be of the form:

$$\hat{f}_d(0) = \frac{1}{2\pi}\sum_{k=-m(T)}^{m(T)} w\left[\frac{k}{m(T)}\right]\hat{\gamma}_d(k) \tag{32}$$

where

$$\hat{\gamma}_d(k) = \frac{1}{T}\sum_{t=k+1}^{T}\left(d_t - \bar{d}\right)\left(d_{t-k} - \bar{d}\right) \tag{33}$$

and *m(T)* in (32) is the bandwidth or lag truncation that increases with *T* but at a slower rate, and $w(\cdot)$ is the weighting scheme or kernel.[13] One weighting scheme, called the truncated rectangular kernel and is used in Diebold and Mariano (1995), is the indicator function that takes the value of unity when the argument has an absolute value less than one.

$$\mathrm{w}(x) = I(|\mathrm{x}| < 1) \tag{34}$$

We apply the Diebold and Mariano (1995) test to forecast errors of both models for each maturity and each forecast horizon. The Diebold and Mariano (1995) statistic (*DM-Stat*) reported in table XII, indicates universal significance of the RMSE differences 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-only model and yields-macro model respectively. Comparison of the 6 and 12 months ahead forecasts of both models specify that five out of 11 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 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 XI1>>

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

4.2.3. Mean Equality Test for the Squared Forecast Errors

The mean equality test for the squared forecast errors is based on analysis of variance (ANOVA). The basic idea is that if the two models have the same mean for forecast errors, then the variability between the sample means of forecast errors (between models) should be the same as the variability of forecast errors within any model. Denote the x_{it} as the forecast

---

[13] See, Andrews (1991) for detailed econometric applications.

errors for model i in period t, where i=1,2,…,G (models) and t=1,2,…,T. The between and within group sums of squares of forecast errors are defined as:

$$SS_B = \sum_{i=1}^{G} (\bar{x}_i - \bar{x})^2 \tag{35}$$

$$SS_W = \sum_{i=1}^{G} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)^2 \tag{36}$$

where $\bar{x}_i$ is the sample mean within group, $\bar{x}$ is the overall sample mean, SS_B is the between groups sum of squares and SS_W is within the group sum of squares. The *F-statistic* for the equality of means is computed as:

$$F = \frac{SS_B/(G-1)}{SS_W/(GT-G)} \tag{37}$$

where *GT* is the total number of observations. The *F-statistic* in (37) has F-distribution with *G-1* numerator degrees of freedom and *GT-G* denominator degrees of freedom under the null hypothesis of independent and identical normal distribution, with equal means and variances in each model.

As in our case *G=2*, we also compute the *t-statistic*, which is simply the square root of the *F-statistic* with denominator degree of freedom.

<<Table XII1>>

The results of the mean equality test of the squared forecast errors of both models i.e. yields-macro model and yields-only model for various maturities are presented in table XIII. Both the tests *t-test* and *F-test* results show that there is statistically significance difference in mean of squared forecasts errors of both the models for one month ahead forecast for all maturities. For 6 months ahead forecast the mean of squared forecast errors are not same until 5 years maturities and beyond 5 years maturities the forecast errors are same for the two competing models. Similarly, for 12 months ahead forecast both the models produces same forecast errors beyond 18 months maturities and below 18 months the forecast errors of yields-macro model statistically different (lower) than yields-only model. Overall the results of the mean equality test of the squared forecast errors show that the Diebold and Mariano (1995) test results are consistent and robust.

In sum, the results of the three aforementioned tests suggest that the yields-macro model has an attractive and greater success in forecasting the yields for short and medium term maturities than the yields-only model for a longer horizon forecast. As far as the short horizon forecasts are concerned, the yields-macro model performs very well and outperform the yields-only model for all maturities. Since, the yields-macro model can serve as a benchmark model and can forecasts the future yields with greater accuracy among the various competing yield curve models.

5. Conclusion

The Nelson-Siegel framework of yield curve provides means for an effective time series analysis of yield data. In this paper, we propose to incorporate macroeconomic as well as stock market factors in the state-space representation of the dynamic Nelson-Siegel model to analyze its crucial role in the in-sample fit and out-of-sample forecasts of the term structure of interest rate. We have specified and estimated a yield curve model that incorporates both yields factors (level, slope, and curvature) and macroeconomic variables [overall economic activity, exchange rate, stock prices index (TOPIX) and inflation]. The state-space representation of the models facilitates estimation of the time varying latent factors of yield curve and its interaction with the macroeconomic factors. It leads to an efficient estimation as well as enhancement of the forecast power of the model.

For the in-sample fit, the results show that the both the models i.e. yields-macro and yields-only models based on the Nelson-Siegel framework 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, our yields-macro model leads to slightly better fit than the yields-only model as the residuals of the former are smaller, having lower RMSE and lower residuals correlation across maturities than the later. Furthermore, we find statistically significant evidence of macroeconomic effects on the latent factors of yield curve and back again. It suggests that the market contain important predictive information about the yield curve.

Regarding the term structure forecasts, Nelson-Siegel framework of yield curve seem reasonably well in terms of low forecast errors. In terms of lower RMSE, our results for all the three horizons forecasts are preferred than that of related studies. Among the two competing models forecasts results, the three employed tests [i.e. TRMSE, Diebol–Mariano (1995) test and mean equality of squared forecast errors] suggest that the yields-macro model is capable to produce more accurate forecasts than the yields-only model. Particularly, the correlation of forecasts errors of yields-macro model is much smaller and negligible as compared to the yields-only model and the persistency of errors in other related studies.

Summarizing, it turns out that the yields-macro representation of Nelson and Siegel (1987) model is compatible to fit attractively the yield curve (in sample fit) and to accurately forecast the future yield for various maturities. The overall accuracy of the proposed extended model has been observed in in-sample fit and out-of sample forecasts over various horizons of forecasts and maturities. Furthermore, the study suggests that the correlation problem of residuals across maturities in in-sample fit and persistency (strongly dependent) of forecasts errors can be avoided by incorporating the relevant macroeconomic and equity market factors in the standard yield curve model. Since, the yields-macro model can serve as a benchmark model and can be good candidate among the various competing yield curve models to forecast the future yield, stock market performance and state of economy.

References

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

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.

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

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

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.

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.

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

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

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

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

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.

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.

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

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

Hördahl P, Tristani O, Vestin D. 2002. A joint econometric model of macroeconomic and term structure dynamics. Manuscript, European Central Bank.

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

Kim CJ, Nelson CR. 1999. State Space Models with Regime Switching. MIT Press, Cambridge.

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

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

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

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

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.

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

Wali U. 2012. Term structure modeling and forecasting of government bond yields. Accepted for 2012 meeting of World Finance & Banking Symposium, Shanghai, China, December 17-18.

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

Table I. Descriptive Statistics of Yields Data across Maturities

Maturity	Mean	S. Deviation.	Max	Min	Skewness	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
3	0.167	0.348	0.692	0.002	1.346	3.259	0.892	0.753	0.077
6	0.164	0.345	0.733	0.004	1.367	3.469	0.877	0.744	0.081
9	0.176	0.339	0.77	0.003	1.348	3.412	0.874	0.727	0.092
12	0.224	0.327	0.812	0.004	1.003	2.6	0.878	0.673	-0.001
15	0.25	0.327	0.855	0.003	0.956	2.487	0.87	0.665	0.021
18	0.276	0.304	0.99	0.013	0.974	2.589	0.873	0.657	0.018
21	0.303	0.303	0.99	0.027	0.932	2.475	0.877	0.651	0.022
24	0.327	0.292	1.027	0.019	0.896	2.382	0.875	0.648	0.025
30	0.387	0.284	1.117	0.027	0.871	2.368	0.865	0.624	0.026
36	0.446	0.281	1.186	0.078	0.815	2.315	0.862	0.596	0.035
48	0.594	0.28	1.368	0.121	0.653	2.133	0.855	0.557	0.027
60	0.73	0.273	1.517	0.161	0.509	2.079	0.856	0.531	0.027
72	0.864	0.265	1.627	0.216	0.365	2.137	0.849	0.485	0.025
84	1.011	0.262	1.759	0.285	0.214	2.234	0.842	0.421	0.035
96	1.165	0.26	1.878	0.382	-0.009	2.418	0.83	0.37	0.051
108	1.302	0.246	1.951	0.474	-0.224	2.784	0.832	0.358	0.091
120	1.424	0.231	1.998	0.549	-0.535	3.457	0.83	0.347	0.102
180	1.801	0.217	2.24	0.758	-1.388	6.203	0.841	0.299	0.183
240	2.061	0.209	2.525	0.934	-1.934	8.291	0.85	0.28	0.152
300	2.267	0.207	2.86	1.07	-1.774	7.983	0.874	0.279	-0.045

Note: The table shows descriptive statistics for monthly yields at different maturities. The last four columns contain sample autocorrelations at displacements of 1, 6 and 12 months. The sample period is 2000:01–2011:12. The number of observations is 144.

Table II. Descriptive Statistics of Macroeconomic and Stock Market Variables Data

	IP_t	EX_t	INF_t	SI_t
Mean	0.705	-2.316	-0.225	-0.451
Std. Dev.	7.359	9.503	0.801	4.964
Maximum	16.506	21.233	2.098	12.011
Minimum	-18.476	-21.189	-2.532	-20.258
Skewness	-0.706	0.406	0.246	-0.375
Kurtosis	3.583	3.077	3.979	4.065
$\hat{\rho}$ (1)	-0.276	-0.022	0.113	0.258
$\hat{\rho}$ (6)	0.243	-0.102	-0.229	-0.131
$\hat{\rho}$ (12)	0.795	-0.101	0.448	0.063
ADF-Stat (Intercept)	-2.647	-12.045	-10.558	-9.084
P-Value (ADF-Stat)	0.086	0.000	0.000	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 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 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 two rows contain augmented Dickey–Fuller (ADF) unit root test-statistics and its p-value.

Table III. Latent Factors VAR(1) Model Parameter 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	-0.01
	(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
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**	-0.56	-0.371	**0.866**				
	(0.538)	(0.502)	(0.312)	(0.085)				

Note: The table reports the estimates for the parameters of the transition equation for both, yields-macro and yields-only, models. The upper panel presents estimates for the yields-macro model of vector A_0 and matrix A, while the lower panel for the yields-only model of vector A_0 and matrix A_1. IP_t is annual growth rate in industrial production, EX_t is the (Yen/$) annual growth of the real exchange rate, INF_t is the 12-month percent change in the consumer price index and and (SI_t) is 12 months growth rate of Tokyo Stock Exchange Index (TOPIX). The standard errors are in parenthesis. Bold entries denote parameter estimates significant at the 5 percent level.

Table IV. Estimates of Co-variance Matrix $\boldsymbol{\Sigma}$

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

Tests for diagonality of Co-variance Matrix $\boldsymbol{\Sigma}$

	Yields-Macro Model			**Yields-Only Model**		
Wald Test Statistic	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 co-variance 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 co-variance matrix $\boldsymbol{\Sigma}$ 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.

Table V. Descriptive Statistic of the Nelson-Siegel Factors Estimates

Models	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. Deviation.	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	-0.67
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)	0.318	0.454	0.531	0.301	0.300	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 maturity) 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.

Table VI. Descriptive Statistic of the Yield Curve Residuals

	Yield-Macro Model						Yields-Only Model					
Maturity	Mean	S. Dev.	RMSE	MAE	$\hat{\rho}\,(1)$	$\hat{\rho}\,(12)$	Mean	S. Dev.	RMSE	MAE	$\hat{\rho}\,(1)$	$\hat{\rho}\,(12)$
3	-0.014	0.094	0.094	0.068	0.489	0.157	-0.019	0.093	0.095	0.069	0.591	0.141
6	-0.028	0.083	0.087	0.057	0.436	0.056	-0.032	0.083	0.089	0.058	0.594	0.047
9	-0.032	0.081	0.087	0.05	0.545	0.014	-0.035	0.081	0.088	0.052	0.684	0.007
12	-0.003	0.029	0.029	0.022	0.368	0.186	-0.005	0.029	0.03	0.023	0.344	0.203
15	0.001	0.015	0.015	0.012	0.375	0.262	0.001	0.015	0.015	0.012	0.423	0.282
18	0.004	0.013	0.011	0.008	0.156	0.125	0.003	0.012	0.012	0.008	-0.043	0.082
21	0.004	0.012	0.013	0.009	0.037	0.087	0.004	0.013	0.013	0.009	-0.057	0.135
24	-0.001	0.015	0.015	0.011	0.139	0.118	0.001	0.015	0.015	0.011	0.161	0.132
30	-0.002	0.018	0.018	0.014	0.260	0.057	-0.001	0.019	0.019	0.014	0.333	0.069
36	-0.008	0.022	0.022	0.017	0.371	0.033	-0.007	0.022	0.023	0.017	0.519	0.045
48	0.001	0.014	0.014	0.011	0.212	-0.013	0.001	0.014	0.014	0.011	0.333	-0.043
60	-0.009	0.023	0.024	0.02	0.447	-0.017	-0.008	0.023	0.024	0.020	0.475	-0.023
72	-0.016	0.04	0.043	0.034	0.491	0.094	-0.017	0.042	0.045	0.035	0.567	0.105
84	-0.006	0.06	0.06	0.047	0.621	0.227	-0.008	0.063	0.064	0.05	0.682	0.242
96	0.02	0.067	0.07	0.054	0.632	0.293	0.016	0.072	0.074	0.057	0.678	0.308
108	0.038	0.051	0.063	0.051	0.632	0.338	0.033	0.057	0.065	0.051	0.684	0.362
120	0.05	0.051	0.07	0.056	0.455	0.214	0.043	0.056	0.070	0.057	0.558	0.262
180	0.004	0.074	0.063	0.055	0.689	0.411	-0.012	0.066	0.066	0.056	0.742	0.407
240	-0.003	0.097	0.077	0.063	0.688	0.364	-0.024	0.079	0.082	0.066	0.735	0.333
300	0.03	0.081	0.064	0.063	0.484	0.171	0.005	0.083	0.083	0.068	0.516	0.117

Note: The table presents summary statistics 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 is the root mean squared errors and mean absolute error respectively. $\hat{\rho}\,(i)$ denotes the sample autocorrelations at displacements of 1 and 12 months. The number of observations is 144.

Table VII. 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.

Table VIII. Out-of-Sample 1 Month Ahead Forecasting Results

Maturity	Mean	Std. Deviation.	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	0.12	0.058	0.022	0.419	-0.07	-0.008
12	-0.001	0.119	0.023	0.001	0.319	-0.111	0.077
18	0.007	0.135	0.008	0.001	0.354	-0.077	0.019
24	0.004	0.149	0.011	0.001	0.401	0.117	0.185
36	-0.003	0.165	0.017	0.001	0.4	-0.082	0.064
60	-0.006	0.192	0.02	0.001	0.447	-0.124	-0.01
120	0.044	0.176	0.057	0.005	0.476	-0.163	-0.099
180	-0.012	0.18	0.055	0.005	0.537	-0.028	-0.085
240	-0.024	0.184	0.066	0.009	0.599	0.042	-0.03
300	0.005	0.194	0.068	0.01	0.622	-0.03	-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
18	-0.012	0.066	0.083	0.054	0.237	-0.096	0.086
24	0.005	0.015	0.096	0.062	0.443	-0.045	-0.008
36	-0.024	0.079	0.11	0.071	0.586	0.041	-0.085
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	0.04
180	-0.006	0.022	0.136	0.052	0.78	-0.068	-0.019
240	-0.032	0.083	0.146	0.053	0.8	-0.111	-0.042
300	-0.008	0.023	0.15	0.06	0.661	-0.015	0.058

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. We estimate both 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 $R_{t+1}(m_i) - \hat{R}_{t,t+1}(m_i)$, where $\hat{R}_{t,t+1}(m_i)$ 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, 6th and 12th sample autocorrelation coefficients.

Table IX. Out-of-Sample 6 Months Ahead Forecasting Results

Maturity	Mean	Std. Deviation.	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	0.7	-0.237	-0.115
6	-0.018	0.148	0.098	0.037	0.66	-0.254	-0.097
12	0.004	0.169	0.112	0.045	0.671	-0.231	-0.022
18	0.010	0.199	0.132	0.059	0.7	-0.224	0.002
24	0.005	0.216	0.147	0.063	0.72	-0.237	0.016
36	-0.002	0.236	0.163	0.076	0.714	-0.212	0.044
60	-0.008	0.268	0.195	0.087	0.725	-0.188	-0.007
120	0.039	0.23	0.171	0.079	0.726	-0.115	-0.089
180	-0.012	0.24	0.167	0.104	0.766	-0.023	-0.11
240	-0.021	0.251	0.17	0.127	0.79	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	0.11	0.039	0.745	0.262	0.127
6	-0.019	0.14	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
18	0.011	0.185	0.146	0.066	0.733	0.241	0.165
24	0.006	0.201	0.161	0.072	0.742	0.245	0.174
36	-0.002	0.219	0.177	0.086	0.726	0.232	0.193
60	-0.009	0.25	0.21	0.093	0.732	0.178	0.059
120	0.041	0.222	0.176	0.081	0.743	0.19	-0.12
180	-0.014	0.234	0.167	0.118	0.793	0.258	-0.15
240	-0.023	0.245	0.167	0.145	0.822	0.296	-0.098
300	0.012	0.26	0.187	0.16	0.841	0.292	-0.106

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. We estimate both 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 $R_{t+6}(m_i) - \hat{R}_{t,t+6}(m_i)$, where $\hat{R}_{t,t+6}(m_i)$ 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, 6th and 12th sample autocorrelation coefficients.

Table X. Out-of-Sample 12 Months Ahead Forecasting Results

Maturity	Mean	Std. Deviation.	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	-0.04	-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	-0.044
18	0.008	0.246	0.18	0.077	0.839	-0.008	-0.041
24	0.003	0.266	0.196	0.084	0.847	-0.018	-0.03
36	-0.005	0.288	0.216	0.099	0.841	-0.018	-0.009
60	-0.019	0.318	0.24	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	0.26	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	0.18	0.133	0.047	0.842	0.01	0.174
6	-0.006	0.178	0.135	0.045	0.821	-0.018	0.167
12	0.003	0.205	0.16	0.055	0.837	0.017	0.182
18	0.009	0.24	0.186	0.077	0.845	-0.006	0.155
24	0.003	0.26	0.204	0.084	0.852	-0.024	0.154
36	-0.006	0.281	0.225	0.099	0.843	-0.032	0.163
60	-0.02	0.309	0.25	0.119	0.829	-0.071	0.065
120	0.034	0.252	0.18	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

Note: The table reports the results of out-of-sample 12-month-ahead forecasting using using state-space specification for the yields-macro and yields-only models. We estimate both 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 $R_{t+12}(m_i) - \hat{R}_{t,t+12}(m_i)$, where $\hat{R}_{t,t+12}(m_i)$ 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, 6th and 12th sample autocorrelation coefficients.

Table XI. TRMSE 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
Yield-Only Model	0.003	0.044	0.062

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

Table XII. Diebold-Mariano Test-statistic

Maturity	1 Month Forecast		6 Months Forecast		12 Months Forecast	
	Statistic	P-Value	Statistic	P-Value	Statistic	P-Value
3	-3.811	0.000	-1.151	0.125	-0.949	0.174
6	-4.068	0.000	-1.21	0.113	-0.956	0.171
12	-3.994	0.000	-1.299	0.099	-1.098	0.138
18	-3.926	0.000	-1.469	0.072	-1.627	0.053
24	-4.007	0.000	-1.741	0.041	-2.574	0.005
36	-4.308	0.000	-2.683	0.004	-2.17	0.015
60	-4.74	0.000	-3.143	0.001	-1.858	0.032
120	-4.117	0.000	-1.193	0.117	-1.515	0.064
180	-3.91	0.000	0.115	0.544	-0.934	0.176
240	-3.969	0.000	0.827	0.794	-0.056	0.48
300	-3.985	0.000	-1.531	0.937	0.435	0.666

Note: The table presents Diebold–Mariano forecast accuracy comparison test results of the yields-macro model against the yields-only model for 1, 6 and 12 months ahead forecasts. The bench mark model is the yields-macro model. The null hypothesis is that the two forecasts have the same root mean squared error. Negative values indicate superiority of the yields-macro model forecasts and the p-values are is the probability of asymptotic standard unit Gaussian distribution under the null hypothesis.

Table XIII. Out-of-Sample Forecast Accuracy Comparisons

Maturity	1 Month Forecast		6 Months Forecast		12 Months Forecast	
	t-statistic	F-statistic	t-statistic	F-statistic	t-statistic	F-statistic
3	3.88***	15.055***	2.26**	5.107**	2.219**	4.048**
6	4.069***	16.558***	2.368**	5.608**	2.195**	4.038**
12	3.973***	15.781***	2.159**	4.662**	2.165**	4.027**
18	3.926***	15.413***	1.868*	3.488*	1.712*	2.813*
24	4.007***	16.06***	1.754*	3.078*	0.105	0.011
36	4.307***	18.553***	1.768*	3.125*	0.13	0.017
60	4.742***	22.487***	2.024**	4.095**	0.236	0.056
120	4.125***	17.016***	0.735	0.541	0.145	0.021
180	3.918***	15.352***	0.061	0.004	0.083	0.007
240	3.963***	15.703***	0.515	0.265	0.006	0.005
300	3.979***	15.831***	0.835	0.697	0.064	0.004

Note: The table reports the t-test and F-test results of the mean equality of the squared forecast errors for forecast accuracy comparison of the yields-macro model against the yields-only model for 1, 6 and months ahead forecasts. The null hypothesis is that the two forecasts have the same mean squared error. The degree of freedom for *t-statistic* is 94 while for ANOVA *F-statistic* is (1,94). ***, ** and * shows the statistical significance at 1%, 5% and 10% respectively.

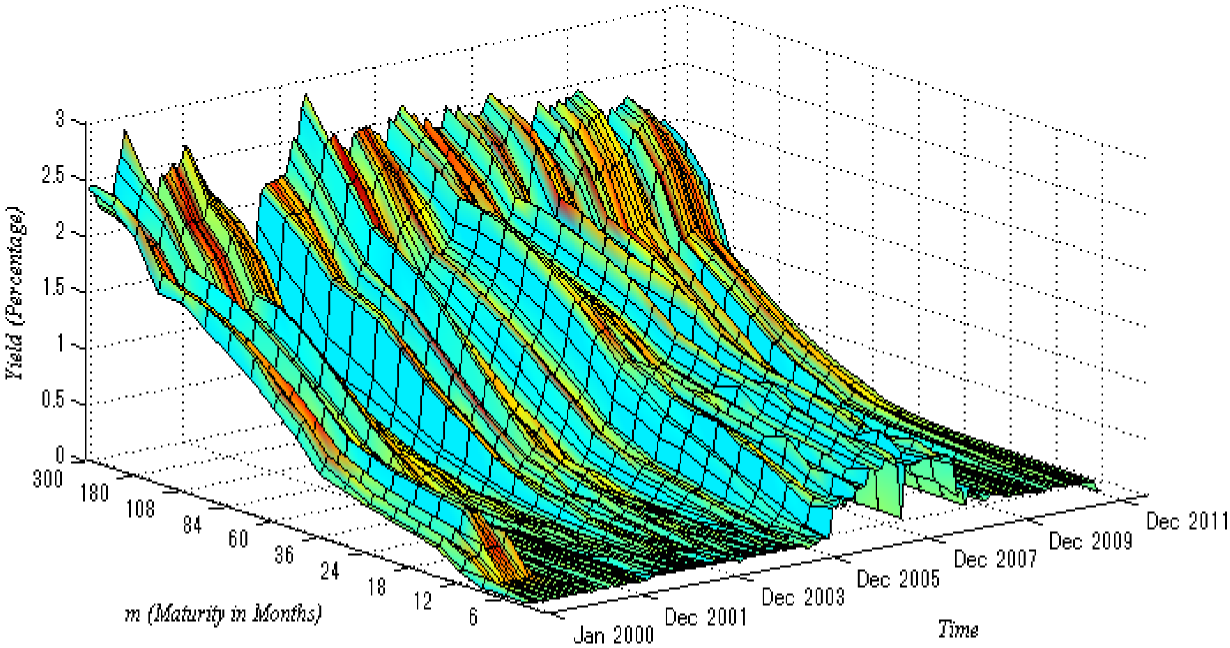

Figure 1. Yield Curves, 2000:01 to 2011:12

The sample consists of monthly yield data 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).

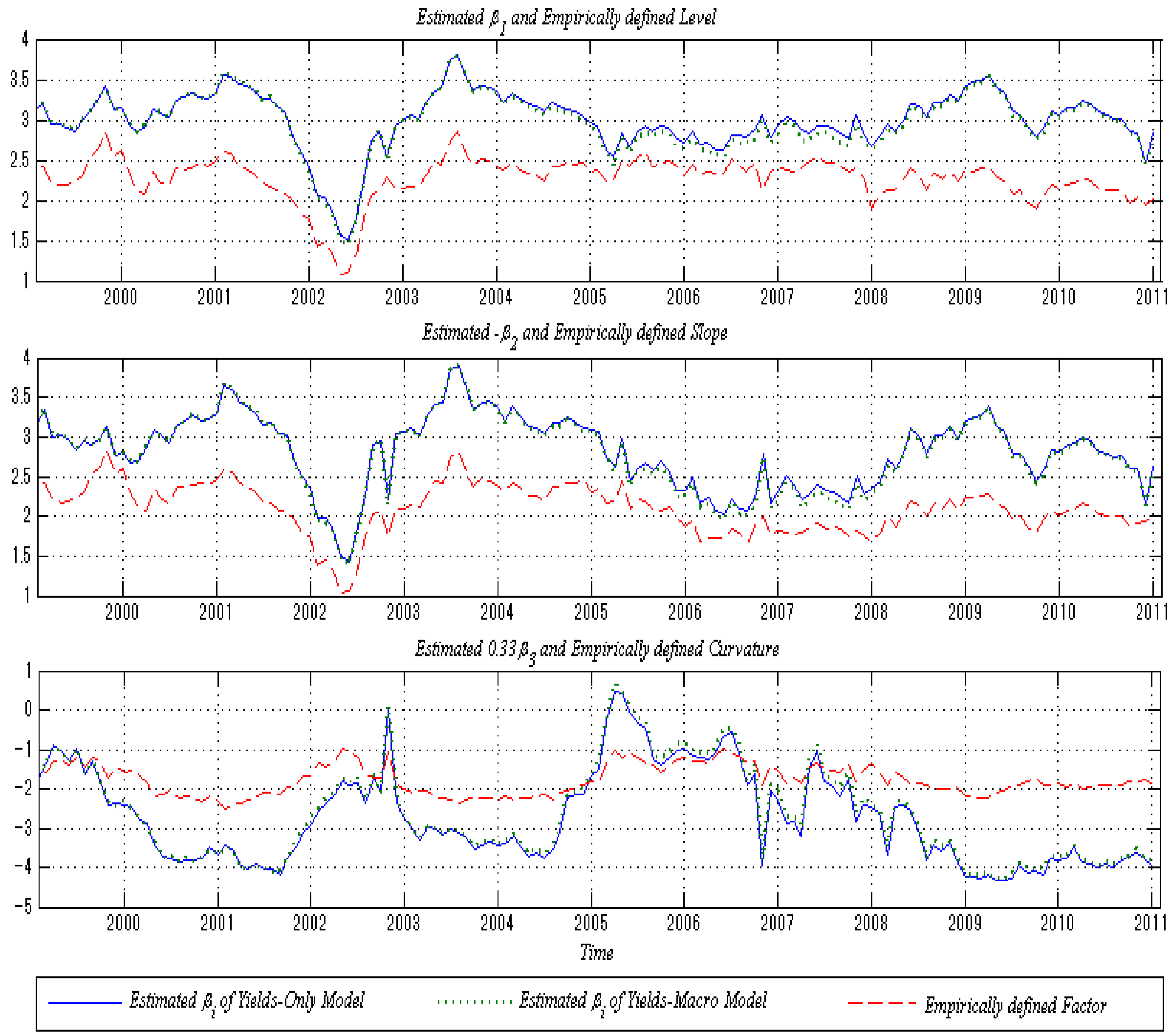

Figure 2. Time Series Plot of Nelson-Siegel Estimated Factors and Empirical Level, Slope and Curvature

Model-based level, slope and curvature (i.e., estimated factors) vs. data-based level, slope and curvature, 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) .

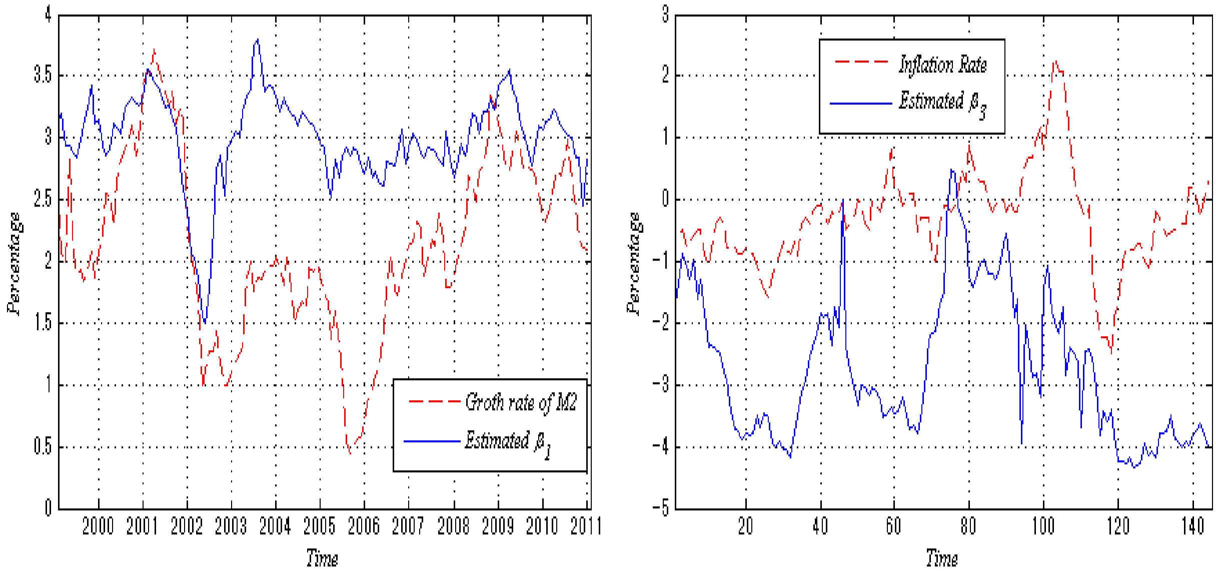

Figure 3. Time Series Plot of Nelson-Siegel Estimated Factors with Macroeconomic Factors

Model-based level and curvature (i.e., estimated factors) are plotted with annual growth of the M2 (Money Supply) and Inflation rate. Inflation rate is the 12-month percent change in the consumer price index.

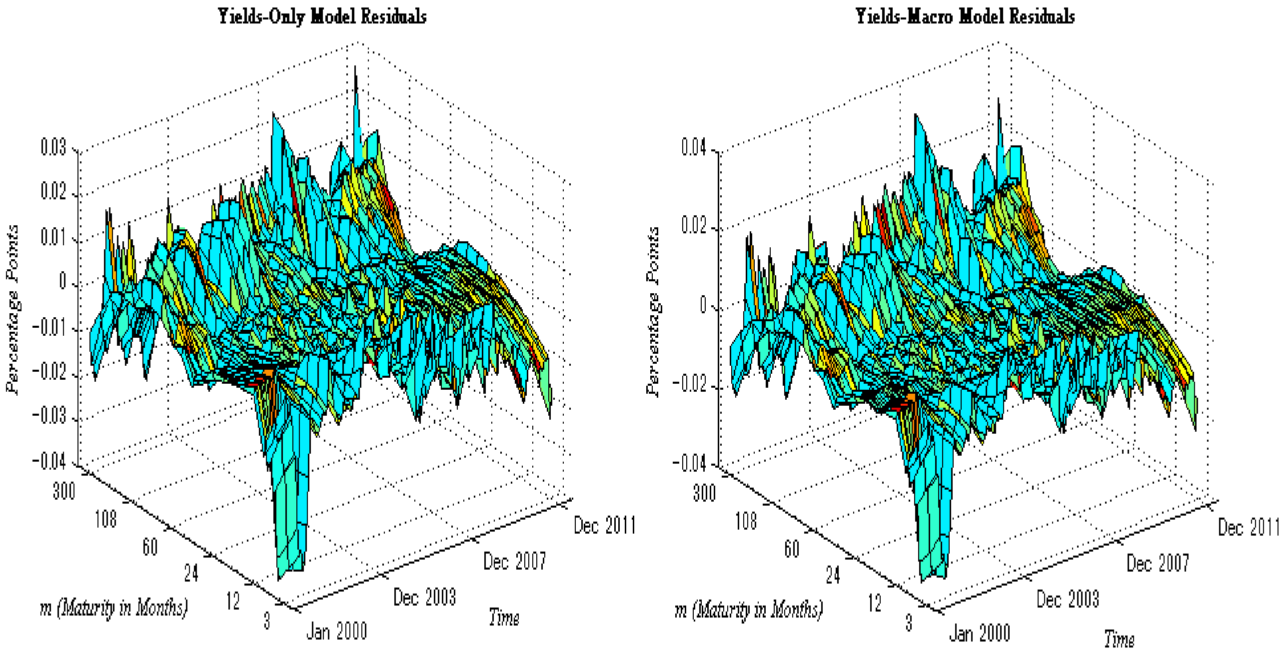

Figure 4. Nelson-Siegel Model based Yield Curves Residuals, 2000:01-2011:12

The sample consists of monthly yield data 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).