3. THE ROLE OF NO-ARBITRAGE RESTRICTION IN TERM STRUCTURE MODEL IN THE CONTEXT OF AN EMERGING MARKET

Wali ULLAH[*]

Abstract

The precise estimation and forecasting of the term structure of interest rate is of vital importance in the context of macroeconomics and finance as the yield curve is considered the fundamental conduit of the monetary policy signal to the real sector. This study examines the extent to which the so called Dynamic Nelson-Siegel model (DNS) and its extended version that impose the no-arbitrage restriction in the standard DNS (AFNS) can fit the term structure of interest rates and forecast its future path in the context of an emerging economy. Both models are illustrated in the state-space framework and empirically compared in terms of in-sample fit and out-of-sample forecast accuracy. For the in-sample fit, both models fit the curve remarkably well even in emerging markets. However, the AFNS model fits the curve slightly better than the DNS model. Regarding the out-of-sample forecasts, the results indicate that the affine based extended model comes with more precise forecasts than the DNS for medium and long term maturities, while the standard DNS outperforms the AFNS at the short end of the yield curve for all three forecast horizons, i.e., 1-, 6- and 12-months. Overall, the results show that there is no single forecast model that dominates its competitors.

Keyword: Yield curve, Forecasting, Emerging markets, No arbitrage, Kalman filter.

JEL Classification: C32, C53, C51, E43, G12, G17

1. Introduction

The yield curve is a two dimensional graphical representation of the yields on zero-coupon bonds and time to maturity. The set of interest rates used to plot the yield curve are generally derived from zero-coupon bonds. However, the non-availability of zero-coupon bonds beyond one-year maturity in the market, makes it necessary to calculate zero rates on coupon-bearing bonds to compute the term structure of interest rates. This plot reflects the

[*] *Department of Economics and Finance, Institute of Business Administration (IBA), Karachi, Pakistan. E-mail: waliullah@iba.edu.pk*

information regarding investors' sentiments and future expectations. Therefore, interest rate forecasts can be used to uncover profitable investment opportunities and serves as a general guide line for economic policy.

Over the last four decades, the yield curve modeling approach has moved from one factor models to multi-factor models with the stochastic volatility component. Since the seminal work of Duffie and Kan (1996), the majority of term structure models are characterized as affine, meaning that zero-coupon bond prices have closed-form solutions with an exponential affine relationship to the state variables. In this regard, the one-factor models include Vasicek (1977), Cox *et al.* (1985) and Hull and White (1990) and for the multi-factor scenario Longstaff and Schwartz (1992), Chen (1996), Chen and Scott (1992), Balduzzi *et al.* (1996) and Dai and Singleton (2000) are used. This class of models is based on the theoretical assumption of the expectation hypothesis that imposes the no arbitrage restriction in the market.

A different type of term structure representation belongs to the statistical class of models, including the Nelson and Siegel (1987) and Svensson (1995) factors models. Nelson and Siegel (1987) formulated the three-factor structure to construct the term structure of interest rates. This model is empirically attractive and does not impose theoretical assumptions. Despite theoretical shortcomings, the Nelson-Siegel (NS) class of models is able to successfully capture yield curve dynamics in all different economic scenarios. Diebold and Li (2006) extended the original Nelson-Siegel model to the dynamic three factors model with factor loadings identical to NS, where the factors can be interpreted as level, slope and curvature respectively, known as the Dynamic Nelson-Siegel (DNS) model. Svensson (1995) increases the flexibility of the original specification by adding a fourth term, leading to the Svensson formulation of the model. All together, the entire NS family of models is often considered superior to affine models for curve fitting (Kim and Orphanides, 2005) and out-of-sample forecasting (Duffee, 2002).

In modeling the yield curve, the assumption of no-arbitrage seems to hold powerful appeal, however, this conflicts with the empirics that the NS specification outperforms the theoretical affine term structure models. Keeping in view this drawback, Christensen *et al.* (2011) derived an affine version of the Nelson-Siegel model based on the theoretical assumption of no-arbitrage restriction. The beauty of their model is that it combines the theory and empirics in a single statistical framework, known as affine arbitrage-free NS term structure models (AFNS). This is done by solving a system of ordinary differential equations (ODEs) derived under the no-arbitrage assumption that has factor loadings identical to the DNS model. Using the U.S. Treasury data, they conclude that this model is able to outperform the DNS model in both in-sample and out-of-sample fit. However, many studies such as Ullah (2016, 2017), Christensen (2012) and Sim and Ohnishi (2015) show that the results derived in Christensen *et al.* (2011) are highly data dependent and the superiority of the AFNS cannot be generalized to all markets.

Although the implementation of yield curve models in the context of emerging economies is recent, there has been little evidence supporting the usefulness of the affine models to forecast yields in emerging markets. This may be due to the lack of good quality and short time span data, which makes it very difficult to reach sound conclusions. In this context, Araujo and Cajueiro (2013) and Caldeira *et al.* (2016) highlight that it is not possible to determine an individual model that consistently produces superior forecasts for all maturities and all forecast horizons. Nevertheless, empirical results suggest that the traditional DNS model has good out-of-sample forecasting performance when compared to the RW, AR(1), and VAR(1), especially when we consider 1- and 3-month ahead horizons. Overall, the results in the context of emerging markets show that there is no single forecast model that dominates all competitors. This is due to the fact that different models outperform the others, depending on time horizon ahead, maturity and forecast period.

This led us to evaluate the performance of Christensen *et al.* (2011) framework in the context of an emerging economy, where the markets are highly regulated and suffer from lack of liquidity. The bond markets in emerging economies are highly segmented and not as efficient as in developed countries. The arbitrage-free version of the DNS, therefore, may not be able to fit and forecast the term structure as accurately as in developed markets. This study applies the Christensen *et al.* (2011) framework to Pakistan's market data to evaluate its performance in comparison to the standard DNS and identify a more appropriate term structure model in the emerging market context. The study may also be helpful to point out the signals of arbitrage opportunity in the market if the standard DNS model comes with more appropriate fitting of the data than the AFNS model.

Against this background, we consider the affine term structure model described in Christensen *et al.* (2011) and compare its performance in terms of in-sample fit and out-of-sample forecasts with the standard dynamic Nelson-Siegel (DNS) model. To address the performance of AFNS model in the environment of the Pakistani bond market, this study seeks answers to the following main questions.

i. How do these two models perform in terms of in-sample fit and out-of-sample forecast in the Pakistani government bonds market?

ii. Is the dominance of the affine Nelson-Siegel model over DNS generalizable in the context of emerging markets?

The remainder of the paper is structured as follows. Section 2 briefly discusses recent trends and reforms in the bond market of Pakistan. Section 3 discusses the two models that are used to estimate and forecast the yield curve in the state space framework and explains the estimation methodology. Section 4 describes the data used in the empirical part of the study and the in-sample fit results. The forecasting performance of the models is evaluated in section 5 and section 6 summarizes and concludes the study.

2. Bond Market in Pakistan

The perpetual cycles of financial crisis during the last three decades have strongly indicated the importance of diversifying the risk profile within the financial system. This can be done by including the fixed income market as an alternate source of funding other than banking and equity sectors. With this emerging perspective, the development of a well-functioning bond market has attained high importance in the new financial market architecture. In Pakistan, the development of the bond market was initiated after the liberalization reforms in the late 1990s, however, Pakistan's bond market has developed at a much slower pace as compared to other countries. Like other emerging markets, most of the debt financing is done through bank borrowings. The outstanding domestic bonds stood at 30% of the GDP, equivalent to PKR 5.8 trillion as of June 2012 (State Bank of Pakistan, 2012). This consists mainly of government bonds, as the corporate market is yet to develop. The Government bond market gained momentum after the introduction of Pakistan Investment Bonds (PIBs) in 2000, which helped to streamline the auction of Government Securities and develop a secondary market for the Government Paper. Outstanding Treasury Bills (TBs) were roughly PKR 2.4 trillion as of June 2012 out of which banks were holding 75% worth of short term paper. Outstanding PIBs amounted to PKR 974 billion, out of which 52% of the holdings were with banks. By June 2017, the short term TBs had piled up to PKR 4.213 trillion, while the total outstanding domestic bonds were PKR 9.5 trillion (State Bank of Pakistan, 2017). The corporate debt market is very shallow and only 88 securities worth PKR 601.23 billion are traded (Pakistan Economic Survey, 2018). Over the last five years the bond market has witnessed 60% growth with about 75% growth in TBs and and 350% in PIBs. The major reason behind these growth trends is that over the last five years, sustained budget deficits have necessitated large issuance of government securities in Pakistan. Taking advantage of this, there is an immense need to develop a deep and liquid bond market because efficient debt markets can help to mitigate the adverse impact of financial crises by providing an alternative source of financing.

These statistics indicate that the bond market in Pakistan has expanded rapidly over the last 6 years. It is currently the main source of financing for the government. The central bank is also trying to promote the development of the bond market by generating new products and the establishment of a group of primary dealers responsible for stimulating the market. However, the bond market is marked by a significantly lower volume of trading as it is smaller and illiquid as compared to the bond markets of developed countries. In fact, the sparseness or infrequency of daily Treasury bonds transactions explains the inaccuracy of the interest rates yield curve. The drawback is that there is no specific term structure model of interest rates and market operators devise a proxy of the yield curve based only on the liquid bonds. Thus, the unevenly distributed maturities of different bonds makes the estimation very difficult and the market is less likely to form an entire and smooth yield curve. This study attempts to fill this gap by investigating the performance of the Nelson-Siegel model (with and without no-arbitrage restriction) in generating a smooth yield curve which replicates the stylized facts of the various interest rates in the context of the Pakistan bond market.

3. Term Structure Models

The yield curve presents the relationship between the yields and maturity of zero-coupon bonds at a specific time point t. Because of the limited observed yield points, the yield curve estimation requires the assumption of some model, so that the gaps may be filled in by analogy with the yields seen in the observed maturities. In this section, we briefly explain the Nelson-Siegel spot rates model with and without imposing the arbitrage free restriction. The former is labeled as AFNS, while the latter as DNS model. The next two sub-sections present the dynamic version of the yield curve model (specified as DNS) and the affine version of the dynamic Nelson-Siegel model (AFNS). Lastly, in the third sub-section both specifications are presented in the state space framework along with the estimation procedure.

3.1. The Dynamic Nelson-Siegel Model (DNS)

The term structure model proposed in Nelson and Siegel (1987) is based on the expectation hypothesis that long term spot rates is the weighted sum of the expected future short rates, which are observed in the form of current forward rates in the market at time t. The mathematical function capable of generating the typical forward rates curve shape is related with the solution of the differential equation. However, the specification for the yield curve presented in Nelson and Siegel (1987) is static and can be used to fit the cross-section of yields. By properly re-factorizing the model, Diebold and Li (2006) have reinterpreted the Nelson–Siegel model for the yield curve to model the bond market evolution over time. In the dynamic framework, given the Nelson–Siegel loadings, the term structure of interest rates can be summarized by three factors, i.e., the level, slope and curvature of the yield curve, as:

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

where $R_t(m)$ is the zero-coupon yield for maturity m at time t, $m = 1,2,\dots,N;\ t = 1,2,\dots,T$. $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the unobservable vector of the three latent factors of level, slope and curvature respectively. The constant parameter λ is the decay parameter of the factor loading of the yield curve slope. This representation of the original Nelson-Siegel model interprets the elements in β_t vector as latent factors with different factor loadings and assigns a strong intuitive interpretation to them (Diebold and Li, 2006).

The formulation of the dynamic Nelson-Siegel (DNS) model is parsimonious and easy to estimate. However, for modelling the entire yield curves simultaneously, we need a state-space representation of the model. Since, we assume that the yield curve latent factors vector β_t follow a vector autoregressive process of the first order, which allows us to formulate the yield curve latent factors model in the state-space form, with observation and transition equations (2 and 3 respectively) as:

$$R_t = \Lambda(\lambda)\beta_t + \varepsilon_t \quad (2)$$

$$\beta_{t+1} = (I_3 - F)\mu + F\beta_t + v_{t+1} \quad (3)$$

$$\begin{bmatrix}\varepsilon_t \\ v_{t+1}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\Omega & 0 \\ 0 & \Sigma_v\end{bmatrix}\right) \quad (4)$$

where R_t is (N×1) vector of zero-coupon yields, $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the (3×1) vector of latent factors of the yield curve, $\Lambda(\lambda)$ is (N×3) matrix of loadings, μ is the (3×1) vector of factors mean, and F is (3×3) full-matrix of parameters. The ε_t and v_t are (N×1) and (3×1) innovations vectors of the observation and state equations respectively, Ω is (N×N) covariance matrix of the measurement equation innovations, and Σ_v is the (3×3) covariance matrix of the state innovations.

3.2. The Affine Nelson-Siegel model (AFNS)

While not having a strong theoretical foundation, especially lacking the assumption of no-arbitrage, the term structure representation in DNS is superior in capturing yield curve dynamics. However, to ensure the arbitrage-free restriction explicitly, Christensen *et al.* (2011) have extended the DNS model by adhering to the standard continuous time affine diffusion processes developed in Duffie and Kan (1996). The approach is very similar to other multifactor affine term structure models except for the solutions to the Ricatti ordinary differential equations (ODEs). The solutions of the ODEs are chosen to match the DNS yield function specified in (1).

Defining the state variables $\beta_t \in \mathbb{R}^3$, to be the Markov process, that solve the stochastic differential equation under the $\mathbb{Q}$-measure, which is:

$$d\beta_t = K(t)[\theta(t) - \beta_t]dt + \Sigma(t)\begin{bmatrix}\sqrt{\gamma^1(t) + \delta^1(t)X_t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sqrt{\gamma^3(t) + \delta^3(t)X_t}\end{bmatrix}dW_t \quad (5)$$

where W is a standard Brownian motion in $\mathbb{R}^3$ with independent components, the drifts $\theta \in \mathbb{R}^3$, dynamics $K \in \mathbb{R}^{3\times3}$, the covariance matrix $\Sigma \in \mathbb{R}^{3\times3}$, $\gamma^i \in \mathbb{R}$ and $\delta^i \in \mathbb{R}^{3\times3}$(where δ^i denotes the i^{th}row of the δ matrix).

Moreover, the instantaneous risk-free rate r_t is assumed to be an affine function of the state variables as:

$$r_t = \rho_0(t) + \rho_1(t)'\beta_t \tag{6}$$

where ρ_0 and ρ_1 are bounded continuous functions. Using the Duffie and Kan (1996) formulation that zero-coupon bond prices are exponential affine functions of the state variables, Christensen *et al.* (2011) construct the partial differential equations and then solve the Ricatti ODEs, to find out the zero-coupon yields function, as:

$$R(t,T) = -\frac{1}{T-t}\log P(t,T) = -\frac{\boldsymbol{B}(t,T)'}{T-t}\beta_t - \frac{A(t,T)}{T-t} \tag{7}$$

where $\boldsymbol{B}(t,T)$ and $A(t,T)$ are the solutions to the system of ODEs. However, to incorporate the empirical attractiveness of the DNS model, the solution to such ODEs must be matched to the yield function in the DNS model. To obtain the closest match, Christensen *et al.* (2011) impose the following structure on $K \in \mathbb{R}^{3\times3}$ and $\rho_1 \in \mathbb{R}^{3\times1}$ when solving the ODEs.

$$K = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{bmatrix}, \qquad \rho_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \tag{8}$$

The formulation in (8) ensures that the instantaneous risk free rate is affine in the form $r_t = \beta_1 + \beta_2$. With dynamics from (5), the restrictions in (8) and by limiting the volatility to be constant, i.e., $\Sigma(t) = \Sigma$, the ODEs are:

$$\begin{bmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{bmatrix} = \rho_1 + K'\boldsymbol{B}(t,T) = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{bmatrix} \begin{bmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{bmatrix} \tag{9}$$

$$\frac{dA(t,T)}{dt} = -\boldsymbol{B}(t,T)'K\theta - \frac{1}{2}\sum_{j=1}^{3}[\Sigma'\boldsymbol{B}(t,T)\boldsymbol{B}(t,T)'\Sigma]_{j,j} \tag{10}$$

The solutions to the system of ODEs defined in (9) and (10) are:

$$B^1(t,T) = -(T-t) \tag{11}$$

$$B^2(t,T) = -\frac{1-\exp[-\lambda(T-t)]}{\lambda} \tag{12}$$

$$B^3(t,T) = (T-t)\exp[-\lambda(T-t)] - \frac{1-\exp[-\lambda(T-t)]}{\lambda} \tag{13}$$

$$A(t,T) = (K\theta)_2 \int_t^T B^2(s,T)ds + (K\theta)_3 \int_t^T B^3(s,T)ds + \frac{1}{2}\sum_{j=1}^{3}\int_t^T [\Sigma' \boldsymbol{B}(s,T)\boldsymbol{B}(s,T)'\Sigma]_{j,j} ds \tag{14}$$

Finally, inserting the above solutions (11-14) in the zero-coupon yields equation (7) implies:

$$R(t,T) = \beta_{1t} + \frac{1-\exp[-\lambda(T-t)]}{\lambda(T-t)}\beta_{2t} + \left[\frac{1-\exp[-\lambda(T-t)]}{\lambda(T-t)} - \exp[-\lambda(T-t)]\right]\beta_{3t} - \frac{A(t,T)}{(T-t)} \tag{15}$$

The yield curve representation in (15) is identical to (1) when disregarding the adjustment term $A(t,T)/(T-t)$. This implies that the short rate is a function of the level and slope of the curve, but not curvature. This is in line with other three-factor models (Dai and Singleton, 2000). Moreover, the adjustment term $A(t,T)/(T-t)$ is the main feature that distinguishes the AFNS model from the DNS model, and is therefore of vital importance. Based on this adjustment term Christensen *et al.* (2011) conclude that their AFNS model outperforms the DNS model.

Christensen *et al.* (2011) show that (14) has an analytical solution, when $\theta = 0$. This involves solving six integrals, demonstrating that the $A(t,T)$ term is a complex function of the volatility matrix Σ, decay parameter λ and maturity $m = T - t$. Here we simply apply their closed-form solution of (14) for the correlated AFNS.

The AFNS is a continuous time model, the time dimension should be modeled in terms of dynamics instead of a time-series model such as an AR or random walk. However, it does not mean that similar models cannot be estimated to ensure comparison. For the AFNS, the dynamics of the state variables can be modeled as:

$$d\beta_t = K(\theta - \beta_t)dt + \Sigma dW_t \tag{16}$$

with $\beta_t = [\beta_{1t}, \beta_{2t}, \beta_{3t}]'$, $\theta = [\theta_1, \theta_2, \theta_3]'$, $dW_t = [dW_{1t}, dW_{2t}, dW_{3t}]'$, and dynamics $K \in \mathbb{R}^{3\times3}$,and the covariance matrix $\Sigma \in \mathbb{R}^{3\times3}$.

The specification for yield curve in (15) and state dynamics in (16) for the AFNS resemble to the DNS measurement equation (2) and state equation (3) respectively. In order to simplify the mathematical notations, the state-space form of the AFNS model under $\mathbb{P}$-measure can be written as:

$$R_t = \Lambda(\lambda)\beta_t - A + \varepsilon_t \tag{17}$$

$$\beta_{t+1} = (I_3 - e^{-K\Delta t})\theta + e^{-K\Delta t}\beta_t + v_{t+1} \tag{18}$$

$$\begin{bmatrix}\varepsilon_t \\ v_{t+1}\end{bmatrix} \sim N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix}\Omega & 0\\ 0 & \Sigma_v\end{bmatrix}\right) \tag{19}$$

where β_t is (3×1) vector of latent factors, K is the (3×3) full-matrix of parameters, θ is (3×1) vectors of factors mean like μ in DNS and A is the (N×1) vectors of yield-adjustment term (which is a function of maturity). The definitions and dimensions of all remaining matrices and vectors are same as discussed in the DNS model specification.

3.3. Statistical formulation of the models and estimation method

The estimation procedure is based on the Kalman filter. For convenience, we introduce some new notations and rewrite the signal and state equations to obtain the generalized form for the models in the state-space framework. For both models, the specification is given as:

$$R_t = BX_t + \tilde{A} + w_t \ , \quad \forall\, t = 1,2,\dots,T \tag{20}$$

$$X_t = C + HX_{t-1} + u_t \tag{21}$$

$$\begin{bmatrix} w_t \\ u_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega & 0 \\ 0 & Q_t \end{bmatrix}\right) \tag{22}$$

where $\tilde{A} = 0$ for the DNS, the expressions of $B, \tilde{A}, X_t, C, H, \Omega, Q_t, w_t$ and u_t in case of both DNS and AFNS are given in Appendix-I. In both specifications, the matrix Ω is assumed to be diagonal for computational traceability, while the covariance matrix Q_t is non-diagonal. Moreover, the transition and the measurement errors are assumed orthogonal to the initial state. The state-space form of the affine based model under $\mathbb{P}$-measure is similar to its counterpart DNS model with the exception of yield-adjustment term $\tilde{A}$ in the affine model.

The Kalman filter algorithm is implemented along the lines of Harvey (1989) and Welch and Bishop (2006) to evaluate the Gaussian likelihood function to obtain the latent factors and estimates of the hyper-parameters in both frameworks. The filter is initialized at the unconditional mean and variance of the state variables under the $\mathbb{P}$-measure. The optimal estimate in Kalman filter is the conditional mean of X_t dependent on information set, denoted as ζ_t. Using the transition equation, the recursive prediction step can be calculated as:

$$X_{t|t-1} = \mathbb{E}(X_t|\zeta_{t-1}) = C + HX_{t-1} \tag{23}$$

$$P_{t|t-1} = \mathbb{E}\left[\left(X_t - X_{t|t-1}\right)\left(X_t - X_{t|t-1}\right)'|\zeta_{t-1}\right] = HP_{t-1}H' + Q_t \tag{24}$$

where $P_{t|t-1}$ is the mean square error (MSE) matrix at the prediction step. In the prediction step for the AFNS model,$C = [I_3 - \exp(-K\Delta t)]\theta$ and $H = \exp(-K\Delta t)$. Furthermore, due to the continuous nature of the affine model, the covariance matrix Q_t of transition equation for the AFNS is computed as:

$$Q_t = \int_0^{\Delta t} \exp(-Ks)\,\Sigma\Sigma' \exp(-K's)ds \tag{25}$$

Using the measurement equation the estimates in (23) and (24) are updated by observing R_t , thus in the update step:

$$X_{t|t} = \mathbb{E}(X_t|\zeta_t) = X_{t|t-1} + P_{t|t-1}B'S_t^{-1}\eta_t \tag{26}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}B'S_t^{-1}BP_{t|t-1} \tag{27}$$

where η_t is the forecast error vector computed as: $\eta_t = R_t - BX_{t|t-1} - \tilde{A}$ (where $\tilde{A} = 0$ for the DNS specification) and S_t is the MSE matrix of η_t computed as: $S_t = BP_{t|t-1}B' + \Omega$. The difference in statistical model building becomes quite clear when comparing the prediction step for the DNS and affine model in the Kalman filter. However, because of the model freedom under the $\mathbb{P}$-measure the two setups can be made as identical as possible by specifying similar models. The update step is identical in both setups with the exception of $\tilde{A}$ (which is the difference between the two measurement equations). Therefore, only the forecasting errors are different.

The Kalman filter iterative process is started with X_0 and P_0 being set at the unconditional mean and covariance as discussed in Hamilton (1994). The beginning of the Kalman filter iteration also depends on parameters vector ψ. Denoting $\psi = (\lambda, \mathrm{B}, C, B, \Omega, \mathrm{Q}_t)$ as the vector of unknown parameters, and assuming that the forecasting errors η_t are Gaussian, the Gaussian log likelihood is computed as:

$$\log L\ (\psi) = \sum_{t=1}^{T} \left(-\frac{N}{2}\log(2\pi) - \frac{1}{2}\log|S_t| - \frac{1}{2}\eta_t' S_t^{-1} \eta_t \right) \tag{28}$$

The matlab built-in function `fminsearch`, which is based on Nelder and Mead numerical optimization routine is used to maximize the log likelihood function (28) and obtain the estimates of the parameters. The var-covariance matrix of the estimates is computed using the outer product of the numerically computed gradient vector.

4. Empirical Results

Considering Pakistan's government bond yields of 15 different maturities between 2002 and 2016, we estimate the DNS and AFNS using the Kalman filter algorithm to obtain the estimates of the latent factors and the maximum likelihood method to estimate the parameters. The details of the data-set are provided in section 4.1. The estimation results regarding the in-sample fitting of both models are presented in section 4.2.

4.1. Data description

We use the Pakistan yield data published by the Mutual Fund Association of Pakistan (MUFAP) and Pak Brunei Investment Company. We collect monthly observations for the period August 2002 until December 2016 on yields for 15 maturities of 3, 6, 9, 12, 18, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months.

The descriptive statistics for the yields are presented in table 1.

Table 1

Descriptive statistics of yields data across maturities

Maturity	Mean	SD	Max	Min	SK	Kurtosis	$\hat{\rho}$ (1)	$\hat{\rho}$ (6)	$\hat{\rho}$ (12)
3	8.552	3.426	13.449	1.118	-0.561	2.427	0.990	0.890	0.703
6	8.660	3.436	13.751	1.131	-0.533	2.439	0.990	0.887	0.694
9	8.744	3.411	13.914	1.233	-0.522	2.445	0.990	0.883	0.686
12	8.828	3.388	14.109	1.335	-0.511	2.449	0.990	0.880	0.678
18	9.054	3.309	14.570	1.726	-0.489	2.381	0.989	0.867	0.663
24	9.279	3.245	15.063	2.117	-0.456	2.313	0.987	0.848	0.642
30	9.424	3.205	15.271	2.316	-0.441	2.272	0.987	0.844	0.635
36	9.569	3.171	15.478	2.515	-0.422	2.232	0.986	0.838	0.625
48	9.831	3.036	15.841	3.096	-0.397	2.276	0.985	0.825	0.618
60	10.010	2.956	15.866	3.474	-0.381	2.275	0.984	0.816	0.605

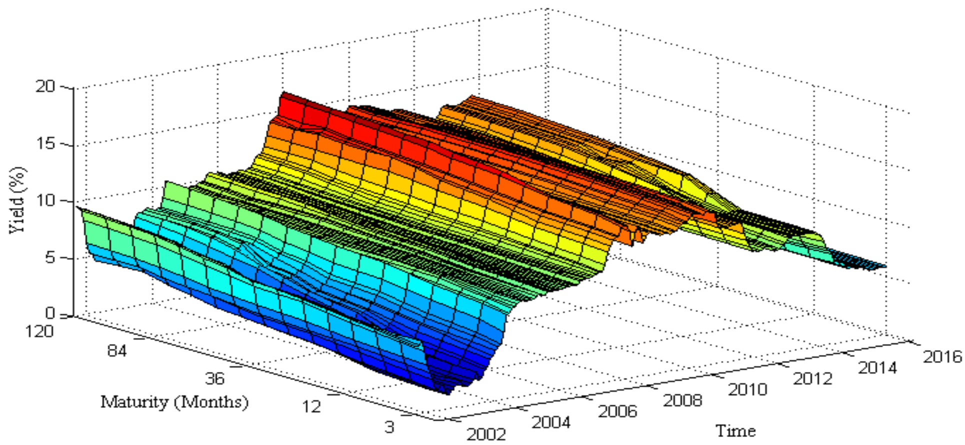
Maturity	Mean	SD	Max	Min	SK	Kurtosis	$\hat{\rho}(1)$	$\hat{\rho}(6)$	$\hat{\rho}(12)$
72	10.236	2.803	16.087	3.845	-0.375	2.391	0.983	0.798	0.583
84	10.389	2.711	16.183	4.192	-0.356	2.409	0.982	0.792	0.579
96	10.512	2.621	16.266	4.584	-0.309	2.350	0.981	0.790	0.580
108	10.572	2.598	16.389	4.683	-0.326	2.392	0.981	0.790	0.585
120	10.628	2.604	16.531	4.531	-0.345	2.528	0.980	0.782	0.570

Note: The table shows descriptive statistics for monthly yields at different maturities. The last three columns contain sample autocorrelations at displacements of 1, 6 and 12 months. The sample period is 2002:08–2016:12. The number of observations is 173.

The results show the average yield curve to be upward sloping as the mean yield increases with maturity. Moreover, the short rates are found to be more volatile and persistent than long rates. Skewness exhibits an upward trend with maturity. Kurtoses of the short rates are lower than those of the long rates. The yields for all maturities are also highly persistent. Figure 1 presents a three-dimensional plot of the yield curve data. The visual inspection indicates that the yield curves have an upward slope at all points in time considered in this study. Moreover, the shapes are almost stable except early 2006 and 2010. The figure shows that the yield curves have shifted down in the current episode of monetary policy ranging from early 2015 till date.

Figure 1

Yield Curves- data plot



Note. The figure shows the yield curves, 2002:08–2016:12. The sample consists of monthly yield data from August 2002 to December 2016 (173 months) for maturities of 3, 6, 9, 12, 18, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months (15 maturities).

4.2. Estimation results of the models

The two versions of the Nelson-Siegel model, i.e., AFNS and DNS are estimated using the Kalman filter algorithm. For given values of the system matrices, the Kalman filter is used to evaluate the log likelihood function via the prediction error decomposition. The maximum-likelihood estimates of the unknown parameters are obtained by optimizing the Gaussian log likelihood function. The filtering process is initialized using the unconditional mean (equal to μ in the DNS and θ in the AFNS framework) and unconditional covariance matrix of the state vector. Moreover, we assume that both innovations (signal and state equations) are independent to the initial state vector, i.e., $E(w_t X_0) = 0$ and $E(u_t X_0) = 0$. The Kalman filter algorithm is also sensitive to the initializing values of parameters, we use the estimates of the parameters reported in Ullah *et al.* (2015) as the initial guess for the hyper parameters.

The estimation results for the vector autoregressive (VAR) of the latent factors in the standard DNS and AFNS models are presented in table 2. The mean vector of the three factors in both setups are statistically significant and almost similar in terms of estimated coefficients. The transition matrix (denoted by F in DNS and K in AFNS setup) results show that own lag dynamics play a significant role in projecting the latent factors in both frameworks as all estimated coefficients are highly significant and explain a significant portion of variation in current value of the estimated factors. Regarding, the cross factors dynamics, the lagged level factor has statistically significant impact on the slope and curvature factors in both setups. Moreover, we observe that the level factor is affected significantly by the lagged slope and curvature factors in the AFNS and DNS framework respectively.

Concerning the comparison between the two models, the mean vectors and λ are directly comparable, while the matrix K should be transformed using $exp(-K\Delta t)$ to be compared with the transition matrix F in the DNS. The mean vector and decay parameter in both models are almost similar (with very minor difference in terms of magnitudes).

The estimate of decay parameter λ in both setups is also similar, being 0.0376 with standard error of 0.0017 in the DNS model, and 0. 0344 with standard error of 0.0021 in the AFNS model, indicating that both estimates are highly significant. This shows that the curvature factor loading hits its maximum point at the maturity of about 48-month in the DNS model, while at 52-month maturity in the case of AFNS model. The difference may be due to the yield adjustment term in the AFNS framework, which plays a significant role at the long end of the curve as shown in figure 2.

The state transition matrix K and covariance matrix Σ in AFNS are modeled continuously and need to be converted to be comparable. The continuous state transition matrix K and volatility matrix Σ are converted into one-month conditional matrices, for the mean-reversion matrix K that is done by computing $\exp(-K\Delta t)$ and for the volatility matrix Σ, equation (25) is used with $\Delta t = 1/12$. Results for the conditional one month K (denoted as $\widetilde{K}$) are shown in (29)

$$\widetilde{K} = \begin{pmatrix} 1.0019 & -0.0031 & -0.0216 \\ -0.0114 & 0.9925 & -0.0259 \\ 0.0075 & -0.0033 & 0.9637 \end{pmatrix} \quad (29)$$

Comparing the matrix $\widetilde{K}$ in (29) and F in table 2, the results indicate that all three factors are more persistent in the AFNS than the DNS. Cross-factors dynamics seems unimportant in the AFNS model, while reasonably stronger in DNS framework. This suggests that AR specification will be more appropriate to model the latent factor in the transition equation as compared to VAR specification in the affine framework. Since, it will be of immense interest to evaluate the likelihood for the AR specification of the state vector and compare with the dependent factors DNS. Based on the log-likelihood value for the in-sample fit, the AFNS seems to be superior to the DNS model.

Lastly, we apply the Wald test to evaluate the joint significance of individually insignificant coefficients in both models. The results are presented in panel 2 of table 2. The results indicate that the individually insignificant coefficients are jointly significant as the p-values are extremely small in both setups.

Table 2

Latent factors parameters estimate of the AFNS and DNS models

Model	**AFNS**				**DNS**			
Panel 1: Estimates of transition matrix and mean vector								
	θ	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$	μ	$\beta_{1,t-1}$	$\beta_{2,t-1}$	$\beta_{3,t-1}$
$\beta_{1,t}$	**9.657** (0.007)	**-0.023** (0.004)	**0.037** (0.001)	0.265 0.1901	**9.672** (0.090)	**0.624** (0.005)	0.034 (0.021)	**0.179** (0.013)
$\beta_{2,t}$	**-7.646** (0.008)	**0.135** (0.018)	**0.092** (0.005)	0.3190 0.2401	**-7.877** (0.362)	**0.082** (0.028)	**0.967** (0.015)	0.271 (0.032)
$\beta_{3,t}$	**-2.408** (0.001)	**-0.091** (0.001)	0.039 (0.068)	**0.443** 0.001	**-1.866** (0.526)	**-0.145** (0.079)	0.031 (0.021)	**0.729** (0.036)
λ	0.034	(0.002)			**0.038**	(0.002)		
Log L	1439				1372			
Panel 2:Test for Joint-Significance of the Insignificant Coefficients (Wald Test)								
Wald	Value	df	P-Value		Value	df	P-Value	
Test Statistic	54.517	3	0.000		39.157	3	0.000	

Note: The table reports the estimates for the parameters of the transition equation of the AFNS and DNS models. Panel 1 presents the estimates for the mean vector and transition matrix along with the decay parameter λ estimate. The standard errors are in parenthesis. Bold entries denote that parameter estimates are significant at the 5% level. The second panel presents the results of the Wald-test for the joint significance of individually insignificant coefficients in transition matrix. The null hypothesis is that insignificant coefficients are simultaneously equal to zero. The test statistic is Chi-square with their respective degrees of freedom (df). P-Value is the probability value of the test statistic.

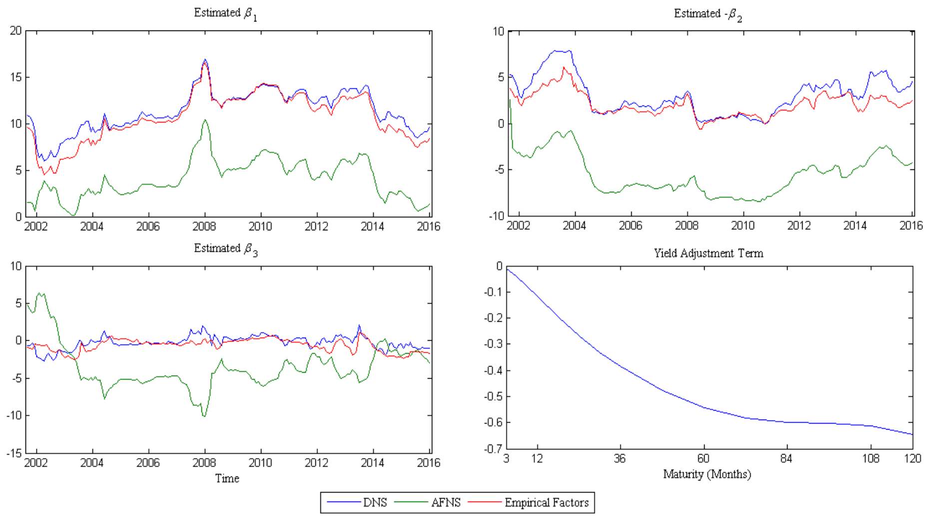
To compare the transition errors, the estimates of covariance matrix Σ of both models are reported in table 3. The covariance matrix Σ of the AFNS is transformed into one-month conditional volatility matrix Q computed by approximating the integral in (25) numerically and the computed matrix is presented in (30). The variance as well as most of the covariance terms in volatility matrix of the AFNS seem considerably smaller than the DNS case. This may be due to the yield-adjustment term in the affine model, because the yield adjustment term is the key difference in the update step between the affine and non-affine classes. While the statistical framework in both classes are very different, the yield-adjustment term is the main economical difference.

$$Q = 10^{-5} \times \begin{pmatrix} 2.1219 & 1.2961 & -2.7538 \\ 1.2961 & 1.0003 & -1.7397 \\ -2.7538 & -1.7397 & 3.5913 \end{pmatrix} \qquad (30)$$

The corresponding estimate of yield-adjustment term A, which differentiates the affine class from the standard Nelson-Siegel family of models, is plotted in the bottom-right panel of figure 2. The yield-adjustment term increases the flexibility of the curve fitting, especially at longer maturities. The yield-adjustment term is monotonically decreasing and plays a significant role at the long end of the curve. The shape is almost similar to the one reported in Christensen *et al.* (2011) for the correlated AFNS model.

Figure 2

Time series plot of estimated factors and their empirical proxies



Note: *Model-based level, slope and curvature (i.e., estimated factors) are plotted against the data based level, slope and curvature (i.e., empirical proxies), where level is defined as the 10-year yield, slope as the difference between the 10-year and 3-month yields and curvature as two times the 2-year yield minus the sum of the 10-years and 3- month zero-coupon yields. Rescaling of estimated factors is based on Diebold and Li (2006). The bottom right pane shown the estimate of the yield-adjustment term in AFNS model.*

Table 3

Estimates of covariance matrix Σ

	$\Sigma(.,1)$	$\Sigma(.,2)$	$\Sigma(.,3)$	$\Sigma(.,1)$	$\Sigma(.,2)$	$\Sigma(.,3)$
		AFNS			**DNS**	
$\Sigma(1,.)$	**0.184** (0.005)			**1.292** (0.093)		
$\Sigma(2,.)$	**0.003** (0.001)	**0.691** (0.0161)		0.002 (0.014)	**0.549** (0.108)	
$\Sigma(3,.)$	0.251 (0.180)	**-0.147** (0.002)	**0.022** (0.000)	0.735 (0.622)	**0.341** (0.125)	**0.971** (0.063)

Note: The table shows the estimates of the covariance matrices of the innovations in the state equations for both models, i.e., AFNS and DNS. The standard errors are in parenthesis. Bold entries denote that parameter estimates are significant at the 5% level.

The three latent factors β_{1t}, β_{2t} and β_{3t} categorized as level, slope and curvature factors respectively are plotted in figure 2 for both AFNS and DNS models along with their empirical proxies, i.e., level (*L*), slope (*S*) and curvature (*C*), which are computed from the observed zero-coupon yield data as: (i) the level factor is defined as the 10-year yield (ii) the slope is the difference between the 10-year and 3-month zero rates, and (iii) the curvature is two times the two-year yield minus the sum of the 10-year and 3-month zero coupon yields. The pairwise correlation of empirically defined level factor and $\hat{\beta}_{1t}$ (model based) is $\hat{\rho}\left(L_t, \hat{\beta}_{1t}^{AFNS}\right) =$

$0.8812,$ and $\hat{\rho}(L_t, \hat{\beta}_{1t}^{DNS}) = 0.9876$. The estimated pairwise correlation between the slope and $\hat{\beta}_{2t}$ is $\hat{\rho}(S_t, \hat{\beta}_{2t}^{AFNS}) = -0.8569$ and $\hat{\rho}(S_t, \hat{\beta}_{2t}^{DNS}) = -0.9541$, while for the curvature (C) and $\hat{\beta}_{3t}$ is $\hat{\rho}(C_t, \hat{\beta}_{3t}^{AFNS}) = 0.4846$ and $\hat{\rho}(C_t, \hat{\beta}_{2t}^{DNS}) = 0.4727$. The pairwise correlations for the DNS for all three factors is higher than that of the AFNS, which can be attributed to the yield adjustment term. There is very small difference in estimated correlation for the $\hat{\beta}_{1t}$ and $\hat{\beta}_{2t}$, but somewhat larger difference for the $\hat{\beta}_{3t}$ with their empirical proxies. Overall, the analysis suggests that the estimated factors and the empirically defined factors follow the same pattern across time, and therefore, $\hat{\beta}_{1t}$, $\hat{\beta}_{2t}$ and $\hat{\beta}_{3t}$ can be called level, slope and curvature factors, respectively.

To further investigate the in-sample performance of the models, the descriptive statistics of the residuals, including root mean square errors (RMSE), mean absolute errors (MAE) and autocorrelation at various displacements are reported in table 4 for some specific maturities.

The results show that the residuals autocorrelations across time for all maturities are considerably small of the DNS model as compared to AFNS model. Based on table 4 alone it is very difficult to rank the models, but it does underscore the evidence that the more sophisticated AFNS model outperforms the DNS in terms of MAE as well as RMSE except the 3-month maturity, however, the improvement is very small for short maturities. The AFNS has lowest RMSEs for 13 out of 15 maturities. Overall, the results in table 4 suggest that a more flexible model is required to fit the yield curve accurately. Besides the residuals analysis, in figure 3 we plot the average observed yield curve and the fitted curves by the two competing models.

Table 4

Descriptive statistics of the yield curve residuals

Maturity	MAE	RMSE	$\hat{\rho}\ (1)$	$\hat{\rho}\ (6)$	MAE	RMSE	$\hat{\rho}\ (1)$	$\hat{\rho}\ (6)$
	AFNS Model				**DNS- Model**			
3	0.614	0.789	0.843	0.435	0.592	0.757	0.605	0.219
6	0.570	0.739	0.793	0.358	0.577	0.739	0.605	0.243
9	0.527	0.708	0.733	0.275	0.552	0.716	0.597	0.272
12	0.525	0.694	0.661	0.192	0.540	0.711	0.602	0.314
18	0.360	0.503	0.423	0.048	0.550	0.723	0.641	0.380
24	0.298	0.452	0.268	-0.093	0.601	0.785	0.693	0.380
30	0.317	0.489	0.350	-0.060	0.644	0.833	0.749	0.446
36	0.390	0.576	0.480	0.029	0.699	0.895	0.788	0.481
48	0.469	0.680	0.516	0.036	0.716	0.915	0.801	0.464
60	0.561	0.789	0.541	0.029	0.744	0.936	0.812	0.479
72	0.508	0.754	0.526	0.013	0.730	0.924	0.812	0.428
84	0.451	0.695	0.484	-0.020	0.718	0.910	0.806	0.432
96	0.458	0.690	0.458	-0.078	0.699	0.891	0.795	0.436
108	0.491	0.708	0.446	-0.042	0.715	0.893	0.803	0.479
120	0.436	0.690	0.401	-0.068	0.708	0.900	0.810	0.486

Note: The table presents summary statistic of the residuals for different maturity times of the measurement equation of AFNS and DNS models, using monthly data 2002:08–2016:12. RMSE and MAE are the root mean squared errors and mean absolute error respectively. $\hat{\rho}\ (i)$ denotes the sample autocorrelations at displacements of 1 and 6 months. The number of observations is 173.

Both models are almost coinciding and capture yield dynamics well, however, the DNS model suffers consistently from underestimating the average yield curve almost for all maturities. The AFNS seems to be more attractive to fit the curve as it coincides at some points with the empirical observed average yields.

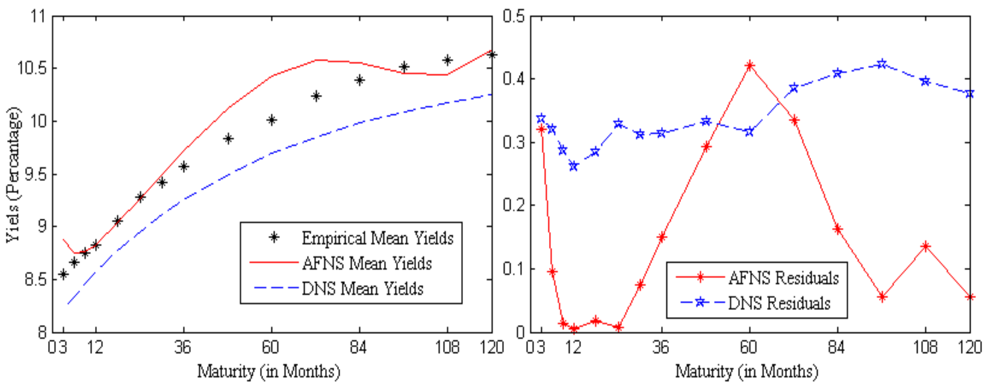
The overall conclusion from in-sample fit is consistent with the findings in Christensen *et al.* (2011) and Ullah (2017) as the affine model outperforms the counterpart non-affine model. The results show that the AFNS is able to fit the yield curve more accurately than the DNS. The great success of the affine model may be due to the yield-adjustment term in the observation equation of the yield curve and higher persistency of factors in the state equation.

5. Out-of-sample Forecasting

The attractive performance in terms of in-sample fitting of the yield curve of the affine model based on the Nelson-Siegel framework as compared to the standard DNS provides motivation to evaluate the future predictive accuracy of the two models in comparison to some benchmark forecast model. As a benchmark we consider AR model of yields to compute the h-period ahead forecasts.

Figure 3

Mean yield curves and absolute residuals



Note: The figure shows the empirical yield curve and the mean estimated yield curves and averaged absolute residuals of the AFNS and DNS yield curve models. The average fitted curves are computed by substituting the smoothed estimates of the yield curve factors and the estimated λ in the corresponding signal equations. The left pane shows the mean yield curves, while averaged absolute residuals are shown in right pane. The averaged absolute residual are computed as the mean of absolute residuals across time for 15 distinct maturities.

For computing the forecasts with the standard DNS and AFNS, the estimation period runs from August 2000 until December 2011, leaving an out-of-sample evaluation period from January 2012 until December 2016. Three forecast horizons, $h = 1{,}6$ and 12 months ahead, are considered.

As the yield curve in the Nelson-Siegel frameworks is mainly governed by the three latent factors, therefore, we compute the h-step ahead forecasts of the yield curve factors and subsequently insert the forecasted factors in the yield curve function to predict the yields for the desired horizon. In the first stage, we estimate each model over a subsample using the state space specification presented in (20-22) and in next stage, predict the h-period ahead latent factors at each point of time in the out-of-sample forecast period by iterating forward the transition equation h-period ahead using the filtered state factors obtained in the previous stage. The h-period ahead predicted state vector is computed as:

$$\mathbb{E}(X_{t+h}|\zeta_t) = \left(\sum_{i=0}^{h-1} \widehat{H}^i\right)\hat{C} + \widehat{H}^h \hat{X}_{t|t} \quad (31)$$

where $\hat{C}$ and $\widehat{H}$ are the parameters estimates of the state and $\hat{X}_{t|t}$ is the most recent available estimated factors vector in the update step. For the AFNS model $\left(\sum_{i=0}^{h-1} \widehat{H}^i\right)\hat{C} = \left[I - \exp\left(-\widehat{K}h\right)\right]\hat{\theta}$ being (3×1) vector and $\widehat{H}^h = \exp\left(-\widehat{K}h\right)$ is (3×3) matrix. Subsequently, the result from (31) is inserted in (20) to compute the h-month ahead forecasted yield dented as $\hat{R}_{t,t+h}$. The forecast errors for the h-step ahead forecast are calculated as: $e_{t,t+h} = R_{t+h} - \hat{R}_{t,t+h}$, where R_{t+h} is the actual observed yield vector at $t + h$ and the $\hat{R}_{t,t+h}$ is the (N×1) vector of the h-month ahead forecasted yields in period t.

5.1. Term structure forecast results

The results of forecast accuracy of the three competing models for the forecast horizons $h = 1, 6$ and 12 months are presented in table 5. We examine a number of descriptive statistics for the forecast errors, including MAE, RMSE and autocorrelation.

The results of one month ahead forecast show that both Nelson-Siegel specifications outperform the AR(1) yield forecasts in terms of all descriptive features of the forecasts errors, whereas, the standard DNS performs slightly better than the affine based extended model until 24-month maturities. The MAE and RMSE for most of the maturities of the AFNS are a bit smaller than the DNS beyond 24-month maturities, however, in terms of the errors autocorrelation the DNS outperforms the affine setup.

The results for the 6 months and one year ahead forecasts show that the forecast errors become larger as we lengthen the forecast horizon. Similar to the one-month ahead forecast, the Nelson-Siegel type models outpace the benchmark AR(1) forecasts in all three descriptive features for $h = 6,$ and 12. Between the two Nelson-Siegel specifications, the order of superiority runs from AFNS to DNS for the long and medium term maturities, whereas DNS outpaces the AFNS until 6-month and 24-month maturities for $h = 6$ and 12 months ahead forecast horizons respectively. For $h = 6$ months, the MAE and RMSE for the AFNS are reasonably smaller than DNS except for first two maturities, while beyond 24-month, for $h = 12$. However, the errors persistency is similar in both setups.

In summary, the out-of-sample forecast results of the two Nelson-Siegel specifications seem reasonably accurate in terms of lower forecast errors than the benchmark AR(1) model of yield. Moreover, for all three forecast horizons the simple DNS model outperforms the AFNS model at the short end of the curve, whereas the AFNS dominates the former for medium and long term maturities. The poor performance of the affine framework in comparison to DNS at the short end of the curve may be attributed to the yield adjustment term.

5.2. Out-of-sample forecast accuracy comparisons

To evaluate the overall quality of the out-of-sample forecasts of the models, we use the trace root mean squared prediction error (TRMSPE) criterion, which summarizes the forecasting performance of each model. TRMSPE is not a formal statistical test but rather is a standard criterion, which is widely used to evaluate forecast accuracy. To assess the forecast accuracy in terms of standard statistical test of all three models, we employ the Diebold and Mariano (1995) test for loss differential quadratic errors.

Table 5

Out-of-sample forecasting results

	AFNS model			DNS model			AR(1) model		
Maturity	MAE	RMSE	$\hat{\rho}\,(1)$	MAE	RMSE	$\hat{\rho}\,(1)$	MAE	RMSE	$\hat{\rho}\,(1)$
1 month ahead forecasting									
3	1.197	1.416	0.885	0.419	0.456	0.850	1.072	1.274	0.937
6	1.079	1.271	0.872	0.470	0.516	0.764	1.092	1.315	0.943
12	0.918	1.094	0.856	0.422	0.487	0.945	1.117	1.342	0.944
24	0.573	0.670	0.658	0.611	0.717	0.988	1.140	1.405	0.949
36	0.475	0.597	0.583	0.802	0.928	0.809	1.160	1.394	0.950
60	0.515	0.650	0.619	0.855	0.967	0.806	1.169	1.385	0.954
96	0.594	0.724	0.702	0.816	0.935	0.799	1.191	1.419	0.951
120	0.561	0.689	0.691	0.766	0.895	0.792	1.011	1.253	0.950
6 months ahead forecasting									
3	2.001	2.274	0.847	1.754	2.061	0.780	2.111	2.300	0.937
6	1.754	2.069	0.853	1.703	2.013	0.794	3.166	2.357	0.947
12	1.603	1.948	0.859	1.612	1.923	0.816	2.175	2.373	0.948
24	1.417	1.832	0.866	1.847	2.280	0.880	2.217	2.422	0.942
36	1.438	1.860	0.879	1.995	2.447	0.906	2.242	2.432	0.944
60	1.462	1.907	0.879	2.054	2.465	0.917	2.309	2.500	0.959
96	1.390	1.722	0.875	2.035	2.408	0.926	2.290	2.505	0.953
120	1.379	1.738	0.876	1.974	2.358	0.925	3.127	2.355	0.954
12 months ahead forecasting									
3	4.079	4.436	0.954	1.927	2.506	0.881	3.128	3.292	0.937
6	3.731	4.095	0.956	1.957	2.417	0.892	3.175	3.349	0.941
12	3.388	3.749	0.959	1.942	2.268	0.906	3.218	3.394	0.944
24	2.742	3.269	0.956	1.852	2.304	0.935	3.225	3.416	0.931
36	1.489	2.032	0.960	1.977	2.463	0.952	3.227	3.411	0.936
60	1.278	2.021	0.952	1.974	2.447	0.956	3.266	3.456	0.946
96	1.670	2.033	0.941	1.955	2.400	0.958	3.246	3.457	0.936
120	1.630	2.037	0.939	1.887	2.334	0.961	3.156	3.380	0.948

Note: The table reports the results of out-of-sample forecasting using state-space specification for the AFNS and DNS models along with the AR(1) forecasts of yields for various maturities. We estimate the models recursively from 2002:08 to the time that the forecast is made, beginning in 2012:01 and extending through 2016:12. We define forecast errors at $t+i$ as $R_{t+i}(m) - \hat{R}_{t,t+i}(m)$, where $\hat{R}_{t,t+i}(m)$ is the $t+i$ month ahead forecasted yield at period t, and we report the mean absolute errors (MAE) and root mean squared errors (RMSE) of the forecast errors, as well as autocorrelation coefficient at first displacement.

5.2.1. Trace root mean squared prediction error

For evaluating the forecast performance, the full sample of forecast errors, i.e., for all 15 maturities and forecast periods, is considered to compute the trace root mean squared prediction error (TRMSPE). TRMSPE combines the forecast errors of all maturities and summarizes the performance of each model, thereby allowing for a direct comparison between the models. Detailed description of TRMSPE is given Ullah *et al.* (2013). The results of TRMSPE for all three models are given in table 6. At first sight the results shows that the forecasts becomes worse as the forecast horizon becomes longer, and the Nelson-Siegel based yield curve specifications outperform the benchmark AR(1) yield forecasts for all forecast horizons. Moreover, the performance of AFNS based specification is better than the DNS at 1- and 6-month ahead forecast horizons, while the latter outpaces the former at $h = 12$.

Table 6

TRMSPE results for out-of-sample forecasts accuracy comparisons

Models	1-Month Forecasts	6-Month Forecasts	12-Months Forecasts
AFNS	0.8200	1.5015	1.7829
DNS	0.6248	1.5706	1.6963
AR(1)	1.3922	2.4372	3.4251

Note: The table reports the Trace Root Mean Squared Prediction Error (TRMSPE) results of out-of-sample forecasts accuracy comparison for horizons of one, 6 and 12 months for the AFNS, DNS and AR(1) yield curve models. In computing the TRMSPE, the full sample of forecast errors, i.e., all 15 maturities are considered.

5.2.2. Diebold–Mariano test

The drawback of using TRMSPE statistics is that these are single statistics summarizing individual forecasting errors over an entire sample. Although frequently used, unfortunately they do not give any insight as to where in the sample models make their largest and smallest forecast errors. We, therefore, employ the Diebold and Mariano (1995) test to assess the forecast performance for each maturity for the different pairs of models. Based on the TRMSPE criterion, we make three pairs of models to make the comparison, i.e., (i) the DNS against the AR(1) model, (ii) AFNS against the AR(1) model, and (iii) AFNS against the DNS model.

To test for statistically significant differences in forecasting accuracy between competing models we apply the Diebold and Mariano (DM) test. Given a pair of two competing forecasting models, i.e., 1 and 2, the difference between the two quadratic loss functions is computed as $d_t = e_{1t}^2 - e_{2t}^2$, where e_{1t}^2 and e_{2t}^2 are the quadratic loss functions of the two competing models, the DM test statistic is computed as:

$$DM = \frac{\bar{d}}{\sqrt{2\pi \hat{f}_d(0)/T}} \sim N(0,1) \qquad (32)$$

where $\hat{f}_d(.)$ is the consistent estimate of the spectral density of d_t and $\bar{d}$ is the sample mean of d_t for $t = 1,2,\dots,T$. We apply the Diebold and Mariano (1995) test to forecast errors of three pairs of models and the results are presented in table 7 for all three forecast horizons.

Table 7

Diebold-Mariano test-statistic

Maturity	DNS against the AR(1)			AFNS against the AR(1)			AFNS against the DNS		
	$h = 1$	$h = 6$	$h = 12$	$h = 1$	$h = 6$	$h = 12$	$h = 1$	$h = 6$	$h = 12$
3	-2.931	-3.588	-2.220	-1.942	1.449	0.092	5.265	0.437	1.973
6	-3.283	-4.323	-3.079	-0.367	2.781	1.092	5.321	0.111	1.444
9	-2.427	-2.170	-5.990	-0.318	-0.563	-1.088	5.216	-0.025	1.388

Maturity	DNS against the AR(1)			AFNS against the AR(1)			AFNS against the DNS		
12	-3.349	-2.945	-2.874	-2.348	-0.453	-1.110	4.815	-0.049	1.372
18	-2.084	-4.329	-2.097	-1.176	-0.206	-1.993	2.828	-0.414	1.018
24	-1.988	-5.669	-3.314	-2.746	-0.261	-2.381	-1.327	-0.711	0.726
30	-6.250	-3.306	-2.438	-6.865	-3.672	-2.205	-3.298	-0.784	-2.571
36	-2.198	-2.768	-2.373	-4.747	-4.359	-2.296	-4.354	-0.846	-0.429
48	-2.402	-2.919	-3.484	-5.192	-3.336	-2.221	-4.378	-0.874	-0.329
60	-2.188	-1.997	-3.482	-5.805	-5.390	-2.245	-4.025	-0.822	-0.239
72	-2.159	-2.850	-4.478	-6.364	-2.347	-2.182	-4.034	-0.916	-3.057
84	-4.589	-2.005	-3.492	-7.462	-2.267	-3.000	-4.001	-1.982	3.184
96	-3.952	-2.832	-3.492	-5.753	-3.130	-2.807	-2.715	-2.206	3.424
108	-2.725	-3.951	-3.506	-2.484	-2.324	-2.748	-1.791	-2.246	-0.541
120	-3.427	-3.891	-3.441	-4.107	-2.380	-2.819	-2.530	-3.103	-0.355

Note: The table presents Diebold–Mariano forecast accuracy comparison test results for the three different pairs of models, i.e., the AFNS, DNS models and AR(1) forecasts for 1, 6 and 12 months ahead forecasts. The null hypothesis is that the two forecasts have the same root mean squared error. Bold entries denotes that the test-statistic is statistically insignificant at 10% significance level.

Besides the statistical significance, the sign of the test statistic in each column conveys important information about the performance and priority of one model over the other. In each pair of models e_{1t} is the vector of forecast errors of the model mentioned first in each column (in top row) of table 7, while e_{2t} corresponds to the errors of the model mentioned later, i.e., in the first column e_{1t} denotes the forecast errors vector of DNS model and e_{2t} corresponds to the forecast errors of AR(1) model. Therefore, the negative sign in the first column shows the priority of DNS over AR(1) and positive sign indicates the superiority of AR(1) over DNS model. All other columns should be read in the same way.

The results of the first and second pair point towards a significant difference in squared errors of the Nelson-Siegel based models, i.e., DNS and AFNS, than the AR(1) model for all three forecast horizons except for few maturities in the case of AFNS model at the short end of the curve. The DNS unanimously outperforms the AR(1) as all test-statics are negative and statistically significant. In case of the AFNS model, some of the test-statistics at the short maturities are insignificant but still negative and for medium and long maturities, the AFNS outpaces the AR(1) model. The Nelson-Siegel based yield curve specifications therefore, come with more accurate forecasts than the AR model of yields.

Between the two Nelson-Siegel frameworks, the DNS perform better than the AFNS until 18-month maturity, while the latter outperforms the former beyond 24-month maturities for the one-month ahead forecasts. For $h = 6$ and 12 both are equally well for the short and medium term maturities, while AFNS has significantly lower forecast errors than the DNS at the long end of the yield curve.

Overall the results of the preceding two criterions indicate that: (i) Nelson-Siegel class either with or without arbitrage restriction outperforms the naïve time series forecasts, such as AR and random walk, (ii) affine based model outperforms their counterpart DNS model at all forecast horizons for medium and long term maturities, whereas DNS has slightly lower RMSEs at the short end of the yield curve. This should not really come as a surprise when remembering the conclusion in the analysis of in-sample fit that AFNS has some problems in fitting the curve for some short maturities. Results are in line with Christensen *et al.* (2011), who conclude that affine term structure models can have very good performance for out-of-sample.

6. Conclusion

The term structure of interest rates plays a key role in financial markets. The current long end of the yield curve holds useful information for forecasting future short yields and economy activity (Piazzesi, 2010). Market participants use these forecasts for pricing financial assets, taking investment decisions, and managing financial risks. Central banks use them to make monetary policy decisions. Accurate modeling and precise forecasting of the yield curve serves policymakers in evaluating past, current, and future economic conditions and helps market participants in taking better financial decisions.

This study considers the affine based dynamic Nelson-Siegel model of the term structure of interest rate and compares it empirically with the standard dynamic Nelson-Siegel in terms of in-sample fit as well as out-of-sample forecast accuracy in the Pakistani bond market environment. The models are illustrated in the state-space framework and the Kalman filter algorithm is employed to estimate and make out-of-sample forecasts, using the monthly time series of yields between 08:2002 and 12:2016.

The results of the in-sample fit show that introducing the no-arbitrage restriction in the model improves the in-sample fit of the Nelson-Siegel model in emerging markets as it does in the developed markets. The success of the affine based models may be due to the yield-adjustment factor in the observation equation and higher persistency of factors in the state equation. Overall the results suggest that the affine based model is able to fit the yield curve more accurately than the DNS for most of the maturities. Largely, the findings from the in-sample-fit are consistent with the findings in Christensen *et al.* (2011).

Regarding term structure forecasting, we conclude that both specifications of the yield curve based on the Nelson-Siegel functional form can replicate the interest rates' general trends in emerging economies. The out-of-sample forecast results of the Nelson-Siegel specifications seem reasonably accurate in terms of low forecast errors and outperform the benchmark naive time series forecast models of yields, such as AR(1) and random walk. The forecast errors of the AFNS model are reasonably smaller than the standard DNS model for medium and long maturities at all three forecast horizons, whereas the DNS wins the horse race at the short end of the curve. Moreover, we encounter larger forecast errors in the short end of the curve for the both DNS as well AFNS model, however, the DNS forecasts are slightly better than AFNS.

The conclusion that comes to mind from the out-of-sample superior predictive power of the AFNS as compared to the DNS, is the role of yield adjustment term and the richer specification in the AFNS framework. The analysis indicates that the small values of the adjustment term is the contributing factor to the poor forecast performance of the AFNS for the short maturities. Furthermore, the off-diagonal elements in the transition matrix, K, and volatility matrix does not seem as important as they are in the DNS framework. The AFNS under alternative specification of the transition and volatility matrix may imply better out-of-sample forecasts for the short maturities.

Overall, the results show that there is no single forecast model that dominates its competitors. This is due to the fact that different models outperform others, depending on time horizon ahead, maturity and forecast period.

Furthermover, short rates are highly influenced by macroeconomic factors and it seems that AFNS as well as DNS out-of-sample performance may be further boosted by including relevant macroeconomic factors in the model. However, the results of this study can serve as a benchmark for future research in which observable macroeconomic information could be incorporated in the model. We think that both these issues are natural next steps in our future research.

Acknowledgement

We would like to thank the anonymous referees for their useful comments to improve this paper.

References

Araujo, R.C. and Cajueiro, D.O., 2013. Combining term structure of interest rate forecasts: The Brazilian case. *Economia*, 14(2), pp. 102-121.

Balduzzi, P., Das, S. R., Foresi, S. and Sundaram, R. K., 1996. A simple approach to three factor affine term structure models. *Journal of Fixed Income*, 6, pp.43-53.

Caldeira, J.F., Moura, G.V. and Santos, A.P., 2016. Predicting the yield curve using forecast combinations. *Computational Statistics & Data Analysis*, 100(C), pp. 79-98.

Chen L., 1996. Stochastic Mean and Stochastic Volatility: A Three-Factor Model of the Term Structure of Interest Rates and its Applications to the Pricing of Interest Rate Derivatives. Blackwell Publishers, Oxford, U.K.

Chen R. and Scott, L., 1992. Pricing interest rate options in a two factor Cox-Ingersoll-Ross model of the term structure. *Review of Financial Studies,* 5, pp.613-636.

Christensen J. M., 2012. Nelson Siegel and affine processes and their use in term structure modeling. MSc. Finance Thesis, Department of Business Studies, Aarhus School of Business, Aarhus University, Denmark.

Christensen, J. H. E., Diebold, F. X. and Rudebusch, G. D., 2011. The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics*, 164, pp.4-20.

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

Dai, Q. and Singleton, K., 2000. Specification analysis of affine term structure models. *Journal of Finance*, 55, pp.1943-1978.

Diebold, F.X. and Li, C., 2006. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130, pp.337–364.

Diebold, F.X. and Mariano, R. S., 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13, pp.253–263.

Duffee, G.R., 2002, Term premia and interest rate forecasts in affine models. *Journal of Finance*, 53, pp.527-552.

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

Economic Survey of Pakistan, 2017-16, Finance Division, Government of Pakistan.

Hamilton, J.D., 1994. State Space Models. In Handbook of Econometrics, Engle, R.F. and McFadden, D. L., (eds). Elsevier: Amsterdam; pp.3041–3080.

Harvey, A.C., 1989. Forecasting Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Hull, J. and White, A., 1990. Pricing interest-rate derivative securities. *Review of Financial Studies*, 3(4), pp.573-592.

Kim, D.H. and Orphanides, A., 2005. Term structure estimation with survey data on interest rate forecasts. Finance and Economics Discussion Board No. 48, Board of Governors of the Federal Reserve System.

Longstaff, F.A. and Schwartz, E.S., 1992. Interest rate volatility and the term structure: a two factor general equilibrium model. *Journal of Finance*, 47, pp.1259-1282.

Nelson, C.R. and Siegel, A.F., 1987. Parsimonious modeling of yield curves. *Journal of Business*, 60, pp.473-489.

Sim, D. and Ohnishi, M., 2015. A modified arbitrage-free Nelson--Siegel model: an alternative affine term structure model of interest rates. *Asia-Pacific Financial Markets*, 22(1), pp.53-74.

State Bank of Pakistan, Annual Reports 2012-11 and 2017-16.

Svensson, L.E. O., 1995. Estimating forward interest rates with the extended Nelson-Siegel method. *Sveriges Riksbank, Quarterly Review*, 3, pp.13-26.

Ullah, W., 2015. The term structure of government bond yields in an emerging market: empirical evidence from Pakistan bond market. Proceedings of Conference on Structural Reforms for Inclusive Development, Applied Economics Research Centre, March 17-18, 2015, Karachi, Pakistan.

Ullah, W., 2017. Term structure forecasting in affine framework with time-varying volatility: do no-arbitrage restriction and stochastic volatility factor imply better out-of-sample forecasts? *Statistical Methods and Applications, doi: 10.1007/s10260-017-0378-y.*

Ullah, W., 2016. Affine term structure model with macroeconomic factors: do no-arbitrage restriction and macroeconomic factors imply better out-of-sample forecasts? *Journal of Forecasting*, 35(4), pp.329–346.

Ullah, W., Tsukuda, Y. and Matsuda, Y., 2013. Term structure forecasting of government bond yields with latent and macroeconomic factors: do macroeconomic factors imply better out-of-sample forecasts? *Journal of Forecasting*, 32, pp.702–723.

Vasicek, O., 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5, pp.177-188.

Appendix-I: Coefficients and latent variable in the general state-space form

In the statistical formulation of the models in section 3.3, the matrices and vectors for the state and observations equations should be considered as follows. The matrices and vectors in state-space system in (20-22) for the simple DNS model should be defined as:

$B = \Lambda(\lambda)$: (N×3)	$X_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$: (3×1)	$w_t = \varepsilon_t$: (N×1)
$C = (I_3 - F)\mu$: (3×1)	$H = F$: (3×3)	$u_t = v_t$: (3×1)
$\Omega = \Omega$: (N×N)	$Q_t = \Sigma_v$: (3×3)	$\tilde{A} = 0$

Furthermore, for the AFNS model the system in (20-22) should be defined as:

$B = \Lambda(\lambda)$: (N×3)	$X_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$: (3×1)	$w_t = \varepsilon_t$: (N×1)
$C = (I_3 - e^{-K\Delta t})\theta$: (3×1)	$H = e^{-K\Delta t}$: (3×3)	$u_t = v_t$: (3×1)
$\Omega = \Omega$: (N×N)	$Q_t = \Sigma_v$: (3×3)	$\tilde{A} = -A$: (N×1)

In both specifications $\Lambda(\lambda)$ is (N×3) matrix of loadings, $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the (3×1) vector of latent factors of the yield curve, μ and θ are (3×1) vectors of factors mean, and F and K are (3×3) full-matrices of parameters. The ε_t and v_t are (N×1) and (3×1) innovations vectors of the observation and state equations respectively, Ω is (N×N) covariance matrix of the measurement equation innovations, and Σ_v is (3×3) lower triangular covariance matrix of the state innovations.