

## Modelling the term structure of Japanese bond yields with the Nelson-Siegel model

A.K. Tsui<sup>a</sup>, J. X. Wu<sup>a</sup> and Z.Y. Zhang<sup>b</sup>

<sup>a</sup> *Department of Economics, National University of Singapore, Singapore*

<sup>b</sup> *School of Business, Edith Cowan University, Australia*

*Email: [zhaoyong.zhang@ecu.edu.au](mailto:zhaoyong.zhang@ecu.edu.au)*

**Abstract:** The Nelson-Siegel (1987) (NS) model has been credited for its high efficacy in the in-sample fitting and out-of-sample forecasting of the term structures of interest rates. The term structure of interest rates, popularly known as the yield curve, is a static function that relates the time-to-maturity to the yield-to-maturity for a sample of bonds at a given point in time. The conventional way of measuring the term structure is by means of the spot rate curve, or yield curve, on zero-coupon bonds. Yet in reality, the entire term structure is not directly observable, which gives rise to the need to estimate it using several approximation techniques. Over the last three decades, various methods to estimate term structures from bond prices have been proposed. In recent years most of the existing studies (as well as major central banks around the globe) have been employing the class of NS models to estimate and construct zero-coupon yield curves.

This paper aims to study the term structure of the Japanese bond yields by employing the NS model vs other non-NS models using five different sets of zero-coupon bond yield rates data obtained from the Bank of Japan covering the period spanning from January 2000 to November 2007. This period has been chosen because it clearly exhibits the liquidity trap problem, which forces all bond yields to remain close to zero for an extended period. We propose 18 different NS models, each with different decay components and time series appendages, against 14 other non-NS models ranging from the simple random-walk model to complicated specifications like the VAR and VECM models. A h-period(s)-ahead out-of-sample expanding window forecast is conducted for each of these 32 different models, using daily, weekly and monthly bond yields of 15 different maturities.

This study has demonstrated that due to the presence of liquidity trap in Japan, out-of-sample expanding window forecasts in general perform inferiorly vis-à-vis other non-NS models, and this is coupled with the other problem of obtaining negative yield forecasts for bonds with shorter maturities. Moreover, the results show that the NS class of models can be useful in forecasting shorter horizons like weeks and days, works better with a decay rate other than the conventional way of treating it as the value that maximizes the loading on the medium-term factor at exactly 30 months, and can work well with time series models such as GARCH and EGARCH in terms of volatility forecasting. It is also found that, when the NS models are used for yield forecasts, the NS-VAR model should be considered since it is up to par against the competitor models, even with liquidity trap at work.

**Keywords:** *In-sample fitting and out-of-sample forecasting; Japanese bond yields; the Nelson-Siegel (NS) model*

## 1. INTRODUCTION

The term structure of interest rates, popularly known as the yield curve, is a static function that relates the time-to-maturity to the yield-to-maturity for a sample of bonds at a given point in time. The conventional way of measuring the term structure is by means of the spot rate curve, or yield curve, on zero-coupon bonds. Yet in reality, the entire term structure is not directly observable, which gives rise to the need to estimate it using several approximation techniques. There are a wide variety of diverse yield models in the literature. The major goal of these models is to describe the future yield curve structure as best as possible. However, the modelling of a yield curve is more complicated than the modelling of a stock derivative, and the models for the yield curve structure alone are more complex as well. In recent years, the Nelson-Siegel (1987) (NS) model and its extended models have been credited for its high efficacy in the in-sample fitting and out-of-sample forecasting of the term structures of interest rates. Many existing studies (as well as some major central banks around the globe) have been employing the class of NS models including the Svensson-Soderlind model to estimate and construct zero-coupon yield curves.

The primary goal of this paper is to investigate the Nelson-Siegel and the extended NS models on their forecasting performances, and to examine the term structure of the Japanese bond yields by employing the NS model vs other non-NS models using five different sets of zero-coupon bond yield rates data obtained from the Bank of Japan covering the period spanning from January 2000 to November 2007. In particular, we propose 18 different NS models, each with different decay components and time series appendages, against 14 other non-NS models ranging from the simple random-walk model to complicated specifications like the VAR and VECM models. A  $h$ -period(s)-ahead out-of-sample expanding window forecast is conducted for each of these 32 different models, using daily, weekly and monthly bond yields of 15 different maturities. The sample period was chosen because it clearly exhibits the liquidity trap problem, which forces all bond yields to remain close to zero for an extended period.

This study has demonstrated that due to the presence of liquidity trap in Japan, out-of-sample expanding window forecasts in general perform inferiorly vis-à-vis other non-NS models, and this is coupled with the other problem of obtaining negative yield forecasts for bonds with shorter maturities. Moreover, the results show that the NS class of models can be useful in forecasting shorter horizons like weeks and days, works better with a decay rate other than the conventional way of treating it as the value that maximizes the loading on the medium-term factor at exactly 30 months, and can work well with time series models such as GARCH and EGARCH in terms of volatility forecasting. It is also found that, when the NS models are used for yield forecasts, the NS-VAR model should be considered since it is up to par against the competitor models, even with liquidity trap at work.

The rest of this study is organized as follows. Section 2 discusses the NS model and the extended NS models as well as the methodology used in this study. Section 3 analyzes the data sets, namely, the daily, weekly and monthly spot rates to be used in this study, and evaluates the various forecast results. The last section concludes with implication drawn from our findings.

## 2. METHODOLOGY AND THE MODEL

We discuss the Nelson and Siegel model in this section. Assume that spot rates are obtained from a second-order differential equation with real and unequal roots. The instantaneous forward rate function  $f_t(\tau)$  at maturity  $\tau$  is defined as:

$$f_t(\tau) = b_{1,t} + b_{2,t}e^{-\tau/\lambda_{1,t}} + b_{3,t}e^{-\tau/\lambda_{2,t}} \quad (1)$$

Where  $\lambda_{1,t}$  and  $\lambda_{2,t}$  are time constants associated with the equation, and the parameters  $\beta_{1,t}$ ,  $\beta_{2,t}$  and  $\beta_{3,t}$  are determined by the initial conditions. Yet, this model was deemed to be over-parameterized by Nelson and Siegel (1987), and having too many parameters makes it difficult for any standard nonlinear estimation software to converge. For this reason, equation (1) was further streamlined by setting  $\lambda_t = \lambda_{1,t} = \lambda_{2,t}$  and adding  $\frac{\tau}{\lambda_t}$  to the factor loading on  $b_{3,t}$ , thereby transforming the forward rate curve into the following form:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t}e^{-\tau/\lambda_t} + \beta_{3,t}\left(\frac{\tau}{\lambda_t}\right)e^{-\tau/\lambda_t} \quad (2)$$

With this transformation, the forward rate curve can now be seen as a constant plus a Laguerre function, which is basically a product of a polynomial and an exponential decay term. Intuitively, this particular forward rate curve can be assumed to be the solution to a second-order differential equation with equal roots for spot rates. By taking the average of forward rates over different maturities, the corresponding spot rate or yield curve can then be obtained:

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1 - e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right) + \beta_{3,t}\left(\frac{1 - e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right) \quad (3)$$

As the NS yield curve corresponds to a discount curve that begins at one at zero maturity and approaches zero at infinite maturity, following Fabozzi *et al.* (2005), Diebold and Li (2006) and Diebold, Rudebusch, and Aruoba (2006b), we treat  $\lambda_t$  as a multiplier and hence transform equations (2) and (3) into:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t}e^{-\lambda_t\tau} + \beta_{3,t}\lambda_t\tau e^{-\lambda_t\tau} \quad (4)$$

$$y_t(\tau) = \beta_{1,t} + \beta_{2,t}\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3,t}\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) \quad (5)$$

which are used for all in-sample fitting and out-of-sample forecasting exercises in the study. In the NS model, the exponential decay rate ( $\lambda_t$ ) controls the speed of decay for the NS yield function  $y_t(\tau)$ . Theoretically speaking, a smaller value of  $\lambda_t$  is supposed to produce slow decay and can thus better fit the yield curve at long maturities; whereas a greater  $\lambda_t$  exhibits the direct opposite of producing an accelerated decay which results in the better fitting of the curve at short maturities. However, a challenge faced is how to choose a suitable decay rate  $\lambda_t^*$  for each single point in time. As regards the three coefficients  $\beta_{1,t}$ ,  $\beta_{2,t}$  and  $\beta_{3,t}$ , which are called the “latent level, slope and curvature factors” in Diebold and Li (2006), they each have their own idiosyncratic traits. The long-term factor,  $\beta_{1,t}$ , governs the yield curve level since an increase in this coefficient raises all short- and long-term yields equally, thereby changing the level of the yield curve.  $\beta_{2,t}$  relates to the short-term factor and is closely associated with the yield curve slope.  $\beta_{3,t}$  is closely linked to the yield curve curvature and it may be viewed as a medium-term factor, akin to its loading  $\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}$ .

The NS model was later extended by researchers including Litterman and Scheinkman (1991), Björk and Christensen (1999), Bliss (1997) and Svensson (1994) to explore more flexible NS specifications, either through the use of additional factors, further decay parameters, or by a composite of both. Some of these extensions have been tested in other papers, though the results were mixed. It is generally found that extensions like Litterman-Scheinkman’s (1991) two-factor model are overly simplistic and yield inaccurate results, whereas models such as the Svensson extension are currently welcomed by major banks around the globe.

We now turn to the discussion of general specification of the models. Let  $Y_t$  be a  $N \times 1$  vector of observed spot rates or yields which is dependent on  $N$  different maturities, such that  $Y_t = [y_t(\tau_1) \dots y_t(\tau_N)]'$ , at any one point in time. In order to facilitate term structure forecasting, we incorporate dynamics to create time-series models for the collected  $\beta_t$  factor estimates, and then assess the models performance. Models selected for the factor forecast include RW, AR( $p$ ), VAR( $p$ ), GARCH and EGARCH specifications. Given the

voluminous literature on ARCH, models built around this concept have become an indispensable tool for financial analysts, bankers and fund managers throughout the world. McAleer (2005) gives a recent exposition of a wide range of univariate and multivariate, conditional and stochastic models of volatility, and other recent surveys have examined multivariate extensions of the original GARCH framework (Bauwens *et al.*, 2006; Maasoumi and McAleer, 2006; and Caporin and McAleer, 2012). Unlike Diebold and Li (2006) and de Pooter (2007) who conveniently assumed first orders for all their time series models, in this study we adopt a more cautious approach by first finding the respective optimum orders with the help of various information criteria. Thus, we define the general specification of the NS model as follows:

$$Y_t = X_t \beta_t + \varepsilon_t \quad (6)$$

$$\beta_t = \mu + \Phi \beta_{t-1} + \nu_t \quad (7)$$

where  $X_t$  is a  $(N \times K)$  matrix of factor loadings which are potentially time-varying if the decay parameter(s) are estimated together with the factors,  $\beta_t$  is a  $(K \times 1)$  vector of factors and its order of integration is assumed to be  $I(0)$  at this point in time, and  $\varepsilon_t$  is a  $N \times 1$  vector of estimation errors and each component within it is assumed to be independent across maturities and have different variance terms  $\sigma^2(\tau)$ .  $\mu$  is a  $(K \times 1)$  vector, while  $\Phi$  is a  $(K \times K)$  square matrix, and  $\nu_t$  is also a  $(K \times 1)$  vector of residuals for the time series model. It is also assumed that the measurement equation and state equation error vectors are both orthogonal and normally distributed, such that:

$$\begin{bmatrix} \varepsilon_t \\ \nu_t \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} 0_{N \times 1} \\ 0_{K \times 1} \end{bmatrix}, \begin{bmatrix} H & 0 \\ 0 & Q \end{bmatrix} \right) \quad (8)$$

where  $H$  represents a diagonal  $(N \times N)$  measurement equation covariance matrix, and  $Q$  denotes a state equation covariance matrix, which may be assumed to be either a diagonal  $(K \times K)$  matrix or a full matrix, depending on the estimation procedure. For instance,  $Q$  (and also  $\Phi$ ) are diagonal if we consider separate AR( $p$ ) models for each factor; on the contrary,  $Q$  (and  $\Phi$ ) are full matrices when a joint VAR( $p$ ) estimation is carried out instead. The NS GARCH (1,1) model used for forecast is specified as follows:

$$\Delta \hat{y}_{t+h} = \Delta \hat{\beta}_{1,t+h} + \Delta \hat{\beta}_{2,t+h} \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \Delta \hat{\beta}_{3,t+h} \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \quad (9)$$

$$\Delta \beta_{i,t+h} = \mu_i + \varepsilon_{t+h}, \quad \varepsilon_{t+h} \sim \mathcal{N}(0, \sigma_{\tau,t+h}^2), \quad i = 1, 2, 3 \quad (10)$$

$$\sigma_{\tau,t+h}^2 = \omega + \alpha \varepsilon_{t+h-1}^2 + \beta \sigma_{\tau,t+h-1}^2 \quad (11)$$

And the EGARCH(1,1,1) with AR(1) in the mean equation can be expressed as follows when combined with the NS model:

$$\Delta \hat{y}_{t+h} = \Delta \hat{\beta}_{1,t+h} + \Delta \hat{\beta}_{2,t+h} \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} \right) + \Delta \hat{\beta}_{3,t+h} \left( \frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \quad (12)$$

$$\Delta \beta_{i,t+h} = \mu_i + \phi_i \Delta \beta_{i,t+h-1} + \varepsilon_{t+h}, \quad \varepsilon_{t+h} \sim \mathcal{N}(0, \sigma_{\tau,t+h}^2), \quad i = 1, 2, 3 \quad (13)$$

$$\ln(\sigma_{\tau,t+h}^2) = \omega + \alpha \left| \frac{\varepsilon_{t+h-1}}{\sigma_{\tau,t+h-1}} \right| + \beta \ln(\sigma_{\tau,t+h-1}^2) + \gamma \left( \frac{\varepsilon_{t+h-1}}{\sigma_{\tau,t+h-1}} \right) \quad (14)$$

### 3. EMPIRICAL ANALYSIS

#### 3.1 Data Description

For the purpose of in-sample fitting and out-of-sample forecasting, we collected five different sets of zero-coupon bond yield rates data from the Bank of Japan archives, including daily yield (1888 observations), weekly yield (408), and monthly yield (94), spanning almost eight years from January 2000 to November 2007. This period was chosen because it clearly exhibits the liquidity trap problem, which forces all bond yields to remain close to zero for an extended period. Similar to the United States, Japan applies a particular variant of the “smoothing splines” method in the estimation of zero-coupon yield rates. It means that the instantaneous forward rate curves, expressed as a linear combination of cubic B-splines, are constructed from price quotes on selected risk-free fixed income assets: 3-, 6-, 120- and 240-month bonds. Following that, the forward rate curves are interpolated by using smoothing splines, after which the spot rates can then be computed by taking the average over the forward rates. The purpose for us to use the daily and weekly yield rates is to test the robustness of the NS model at the daily and weekly levels.

Also, utilizing yield rates on daily and weekly bases allow for a larger sample, which in turn makes the estimates more stationary. We also pool all the data sets into fifteen fixed maturities, i.e., 3, 6, 12, 24, 36, 48, 60, 72, 84, 96, 108, 120, 180, 240, and 360 months, which will greatly simplifies our estimation and forecasting sequences.

To save space, we show in Figure 1 the time-series of zero-coupon yield curves for 3-, 12-, 60-, 120- and 360-month bonds, based on monthly raw data from January 2000 to November 2007. Two observations can be noticed: there is a low-yield period between 24 to 36 months for all types of bonds from January 2000; and bonds with maturities of 12 months or less have yields close to zero from 12 months onwards until 72 months. The former characteristics may suggest a possible structural break, corresponding to the privatization of the Postal Savings System as well as a complete overhaul of the existing financial structure in Japan. The latter is a classic example of “liquidity trap”. Table 1 reports the summary of the descriptive statistics for monthly yields from January 2000 to November 2007. As it can be seen in Table 1, the mean values confirm that the average yield curve is characterized as upward sloping and concave, and the standard deviation reveals that it is most volatile for medium and long-term maturities, but more stable for the shorter-term maturities. One possible explanation is that the presence of liquidity trap forces short-term yields to converge towards zero, thereby “stabilizing” it. It is also noted that sample serial autocorrelations at a displacement of 1 month for all maturities, as well as pairwise correlations between yields that have close maturities are extremely high. In addition, the descriptive statistics on daily and weekly data lead to more or less the same conclusions, except for the fact that serial autocorrelations remained very high even up to a displacement of 24 months for all maturities, and the Jarque-Bera probability converges to zero as the number of observations gets larger.

#### 3.2 Empirical Results

We conducted in-sample fitting exercise by estimating equation (5) using different values for  $\lambda_i$ . In particular, following Diebold and Li (2006) we performed OLS estimation of equation (5) by assigning  $\lambda_i$

Figure 1: Japanese Zero-Coupon Yields for Various Maturities for Monthly Data

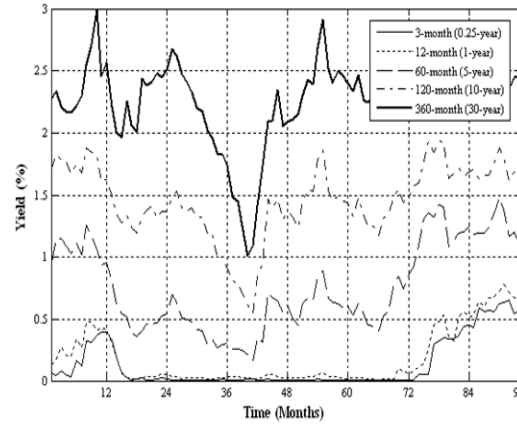


Table 1: Descriptive Statistics for Monthly Data

Maturity (Months)	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis	Jarque-Bera Probability	$\rho_1$	$\rho_{12}$	$\rho_{24}$
3	0.129	0.010	0.650	0.001	0.201	1.358	3.334	0.000	0.892	0.108	0.090
6	0.141	0.011	0.700	0.001	0.214	1.327	3.295	0.000	0.905	0.109	0.095
12	0.181	0.044	0.785	0.008	0.230	1.186	2.955	0.000	0.899	0.118	0.094
24	0.319	0.153	1.025	0.030	0.296	0.951	2.445	0.000	0.927	0.123	0.116
36	0.444	0.302	1.124	0.077	0.307	0.685	2.002	0.004	0.934	0.100	0.093
48	0.597	0.482	1.294	0.128	0.328	0.505	1.869	0.011	0.935	0.087	0.087
60	0.760	0.652	1.478	0.164	0.355	0.314	1.834	0.032	0.935	0.073	0.083
72	0.911	0.859	1.595	0.224	0.353	0.145	1.951	0.098	0.930	0.031	0.084
84	1.081	1.062	1.752	0.268	0.355	-0.068	2.181	0.259	0.922	-0.032	0.094
96	1.238	1.235	1.855	0.379	0.341	-0.277	2.514	0.345	0.914	-0.086	0.065
108	1.360	1.376	1.924	0.464	0.324	-0.521	3.011	0.120	0.908	-0.148	0.115
120	1.456	1.472	1.930	0.540	0.300	-0.832	3.683	0.002	0.906	-0.080	0.082
180	1.668	1.671	2.219	0.654	0.314	-1.006	4.388	0.000	0.901	-0.072	0.110
240	1.987	2.040	2.500	0.904	0.292	-1.666	6.595	0.000	0.896	-0.130	0.035
360	2.261	2.342	2.995	1.014	0.337	-1.553	6.353	0.000	0.876	-0.154	-0.077

equal to 0.0598 and 0.0299, and also NLS estimation assuming  $\lambda_t$  is time varying with daily, weekly and monthly yield data from January 2000 to November 2007. The results (not reported due to space limitation but available upon request) show that in general the long-term factor ( $\hat{\beta}_{1,t}^{OLS}$ ) and the medium-term factor ( $\hat{\beta}_{3,t}^{OLS}$ ) are found more persistent than the short-term factor ( $\hat{\beta}_{2,t}^{OLS}$ ), which is consistent with the stylized facts that yield dynamics are persistent and longer rates are more persistent than the short rates. It is also found that the pairwise correlation between  $\hat{\beta}_{1,t}^{OLS}$  and  $\hat{\beta}_{2,t}^{OLS}$  is the largest in all the three cases and also significantly negative.

We have conducted the  $h$ -period(s)-ahead expanding window forecasting for both the NS and non-NS models, but report in Table 2 the results for one-week ahead expanding window only due to space limitation. The rest of the results are available upon request.

Table 2: Out-of-Sample 1-Month-Ahead and 3-Months-Ahead Expanding Window Forecasting Results)

Yield Forecasting																
MODEL	1-Month-Ahead Forecast							3-Months-Ahead Forecast								
	ARMSPE	TRMSPE	RMSPE				ARMSPE	TRMSPE	RMSPE							
			3m	6m	12m	60m			120m	360m	3m	6m	12m	60m	120m	360m
NS-OLS( $\lambda=0.0299$ )-RW-L	0.14086	0.14924	0.11727	0.04839	0.12531	0.12692	0.12793	0.26263	0.19736	0.20345	0.19161	0.08950	0.11620	0.20258	0.18805	0.16160
NS-OLS( $\lambda=0.0598$ )-AR-C	0.14222	0.15093	0.11507	0.04677	0.12687	0.12853	0.12572	0.26906	0.20283	0.20943	0.13510	0.08834	0.16791	0.20397	0.19342	0.32856
NS-OLS( $\lambda=0.0598$ )-VAR-C	0.10228	0.10674	0.04858	0.04880	0.05742	0.11523	0.12452	0.12346	0.17821	0.18453	0.08869	0.08996	0.11298	0.21172	0.20045	0.17892
NS-OLS( $\lambda=0.0299$ )-RW-L	0.10758	0.11245	0.06501	0.05178	0.04951	0.11654	0.13058	0.13449	0.17418	0.18000	0.10768	0.09302	<b>0.09772</b>	0.20617	0.19945	0.17793
NS-OLS( $\lambda=0.0299$ )-AR-C	0.11042	0.11704	0.06262	0.05507	<b>0.04722</b>	0.12310	0.13880	0.14031	0.16256	0.16916	0.10550	0.09311	0.09154	0.21652	0.21379	0.18687
NS-OLS( $\lambda=0.0299$ )-VAR-C	0.10282	0.10740	<b>0.04821</b>	0.04856	0.05737	0.11563	0.12552	0.13001	0.17893	0.18543	0.08774	0.08922	0.11290	0.21251	0.20188	0.17938
NS-NLS( $\lambda=0.0330$ )-RW-L	0.22836	0.24245	0.05696	0.06443	0.10891	0.28108	0.28874	0.19302	0.25456	0.26799	0.09593	0.10071	0.13217	0.31625	0.30670	0.20255
NS-NLS( $\lambda=0.0330$ )-AR-C	0.22959	0.24245	0.05762	0.06692	0.10931	0.28904	0.28822	0.19312	0.25934	0.27194	0.09562	0.10202	0.13072	0.32015	0.31103	0.22368
NS-NLS( $\lambda=0.0330$ )-VAR-C	0.18919	0.19350	0.05718	0.04897	0.05171	0.11771	0.12711	0.12711	0.15255	0.16175	0.08756	0.08756	0.11169	0.22525	0.21298	0.15298
RW-L	<b>0.09812</b>	<b>0.10196</b>	0.04982	0.04847	0.05663	<b>0.10002</b>	0.11679	0.12643	<b>0.16925</b>	<b>0.17470</b>	0.08834	0.09054	0.11127	<b>0.19917</b>	0.18684	0.17102
SR	0.09929	0.10317	0.04982	0.04771	0.06114	0.12327	0.11785	0.12575	0.18340	0.18720	0.08834	0.10558	0.12628	0.21000	0.19319	0.17744
AR-L	0.09943	0.10302	0.05156	0.04978	0.06194	0.11273	<b>0.11672</b>	<b>0.12418</b>	0.17420	0.17933	0.09406	0.09419	0.11778	0.20995	<b>0.19678</b>	<b>0.16711</b>
AR-C	0.10042	0.10436	0.04860	0.05080	0.05993	0.11229	0.11895	0.12596	0.17558	0.18180	<b>0.08691</b>	<b>0.08726</b>	0.11354	0.20837	0.19447	0.17806
VAR-L	0.11164	0.11651	0.05162	0.05300	0.06162	0.13838	0.13058	0.14286	0.25237	0.26550	0.10470	0.11247	0.12033	0.31291	0.29217	0.30624
VAR-C	0.10071	0.10364	0.05746	0.05502	0.07101	0.12293	0.11838	0.13192	0.16273	0.16791	0.10267	0.10250	0.12334	0.21159	0.20225	0.19205
VECM: 1 Common Trend	0.91934	1.13405	0.05481	<b>0.04994</b>	0.08236	0.79262	1.40151	2.20065	6.15185	7.56024	0.17677	0.12925	0.48929	5.18180	9.32125	14.40202
VECM: 2 Common Trends	0.91937	1.13424	0.05530	0.05540	0.08229	0.79102	1.40187	2.20284	6.10571	7.47118	0.29485	0.30329	0.49242	5.13801	9.18780	14.24487

Yield Volatility Forecasting																
MODEL	1-Month-Ahead Forecast							3-Months-Ahead Forecast								
	ARMSPE	TRMSPE	RMSPE				ARMSPE	TRMSPE	RMSPE							
			3m	6m	12m	60m			120m	360m	3m	6m	12m	60m	120m	360m
NS-OLS( $\lambda=0.0598$ )-GARCH-C	0.03901	0.04009	0.05642	0.05407	0.05119	0.03664	0.03098	0.03312	0.03905	<b>0.04037</b>	0.05713	0.05514	0.05289	0.03689	0.03005	0.03092
NS-OLS( $\lambda=0.0598$ )-EGARCH-C	0.08717	0.08827	0.07988	0.06491	0.05419	0.09433	0.09403	0.10442	0.08264	0.08371	0.08022	0.06527	0.05153	0.08603	0.08994	0.10610
NS-OLS( $\lambda=0.0299$ )-EGARCH-C	0.05063	0.05110	0.04507	0.04225	0.03903	0.05311	0.05550	0.05741	0.06062	0.06141	0.05368	0.05276	0.04542	0.06009	0.06943	0.07412
NS-OLS( $\lambda=0.0299$ )-GARCH-C	0.06331	0.06474	0.07025	0.07291	0.06776	0.05602	0.04962	0.03840	0.05692	0.06739	0.07869	0.07607	0.07136	0.05732	0.04667	0.03937
NS-OLS( $\lambda=0.0299$ )-EGARCH-C	0.06189	0.06328	0.04647	0.04565	0.04811	0.05884	0.06874	0.09495	0.07315	0.07474	<b>0.05953</b>	0.04968	0.05075	0.07500	0.08520	0.09644
NS-NLS( $\lambda=0.0330$ )-GARCH-C	0.07372	0.07459	0.06240	0.05961	0.05844	0.08187	0.08206	0.07295	0.07542	0.07665	0.06146	0.05882	0.05697	0.08333	0.08326	0.06954
NS-NLS( $\lambda=0.0330$ )-EGARCH-C	0.04262	0.04365	0.06031	0.06008	0.05394	0.04283	0.03503	0.03147	0.04354	0.04484	0.06418	0.06175	0.05703	0.04308	0.03513	0.03167
NS-NLS( $\lambda=0.0330$ )-GARCH-C	0.07927	0.07924	0.04806	0.04604	0.04332	0.06914	0.06956	0.11507	0.09065	0.08958	0.05150	0.04829	0.04997	0.08888	0.12285	0.13711
NS-NLS( $\lambda=0.0330$ )-EGARCH-C	0.31430	0.31759	0.30144	0.29759	0.30747	0.35717	0.31783	0.21626	0.30304	0.30558	0.31440	0.30670	0.33640	0.30307	0.31234	0.21234
GARCH-L	0.19076	0.24619	0.68775	0.42454	0.07963	0.16939	0.12874	0.10620	0.78637	1.92708	7.16089	1.94978	0.20939	0.24232	0.14630	0.17474
GARCH-C	0.03373	0.06147	0.22468	0.02152	0.00849	0.02099	0.02255	0.03522	0.39573	1.42550	5.51825	0.15436	0.00964	0.02122	0.02251	0.03487
GARCH-AR-L	0.06306	0.17781	0.88140	0.34709	0.05166	0.02146	0.02230	<b>0.02887</b>	0.81987	2.96426	11.47280	0.36489	0.21743	0.02219	0.02662	0.03059
GARCH-AR-C	<b>0.02147</b>	<b>0.02320</b>	0.04214	<b>0.01915</b>	<b>0.00793</b>	<b>0.01971</b>	0.02264	0.03692	<b>0.03119</b>	0.04728	0.15964	<b>0.04287</b>	0.00910	0.02012	0.02311	0.04151
EGARCH-L	0.02282	0.02459	<b>0.03707</b>	0.04457	0.01234	0.02116	<b>0.01978</b>	0.03566	0.10541	0.30035	1.14235	0.20968	0.01352	<b>0.01979</b>	<b>0.01778</b>	<b>0.02253</b>
EGARCH-C	0.34191	1.21580	4.70578	0.14558	0.00900	0.02181	0.02253	0.04403	0.68158	2.43918	8.43302	0.50470	<b>0.00886</b>	0.02087	0.02274	0.05320

As regards the yield forecasting, the RW model seems work exceptionally well in most cases, and both the AR and VAR models produced almost similar and relatively fine forecasts. There is also a tendency for these two models to converge towards the RW model, which means that, when the selection criteria of significant coefficients becomes more stringent, many of those coefficients on the lagged regressors as well as the constant term were found to be insignificant and the model ultimately becomes a RW. In contrast, for the weekly and daily forecasts, the NS models (except for NS-VAR) generally performed worse than the non-NS models (except VECM). For monthly forecasts, the NS models appear to function well when it comes to longer horizons, particularly for 12-months-ahead forecasts. On top of that, the NS-VAR model performed roughly the same for all three variations of decay rates. It is also interesting to note that the NS-RW and NS-AR specifications when  $\lambda = 0.0299$  yields the lowest TRMSPE among the three decay rates, which indicates that Diebold's choice of  $\lambda$  is not necessarily the best one available.

On the other hand, the results indicate that the non-NS models outperform the NS models in volatility forecasts with the presence of liquidity trap. This finding actually is not a surprise as liquidity trap has impelled bond yields towards zero and less volatile. As regards the various decay rates, it appears that the lowest ARMSPEs and TRMSPEs are attained when  $\lambda = 0.0598$  in most cases, while the NLS-estimated model with  $\lambda$  set at 0.033 performed a lot worse. This result confirms that a faster decay rate allows forecasted conditional volatility to converge quickly to zero. As liquidity trap has induced actual volatility close to zero, the NS model with the largest  $\lambda$  is likely to stand out from the rest.

#### 4. CONCLUDING REMARKS

We have demonstrated that due to the presence of liquidity trap in Japan, out-of-sample expanding window forecasts for the NS models in general perform inferiorly vis-à-vis the non-NS models, and this is coupled with the other problem of obtaining negative yield forecasts for bonds with shorter maturities. Moreover, the results show that the NS class of models can be useful in forecasting shorter horizons like weeks and days, works better with a decay rate other than the conventional way of treating it as the value that maximizes the loading on the medium-term factor at exactly 30 months, and can work well with time series models such as GARCH and EGARCH in terms of volatility forecasting. It is also found that, when the NS models are used for yield forecasts, the NS-VAR model should be considered since it is up to par against the competitor models, even with liquidity trap at work

#### ACKNOWLEDGEMENT

The third author wishes to acknowledge the financial support from the Sumitomo Foundation.

#### REFERENCES

- Bauwens, L., Laurent, S., & Rombouts, J. (2006). Multivariate GARCH models: a survey, *Journal of Applied Econometrics*, 21, 79-109.
- Bliss, R. (1997), Testing Term Structure Estimation Methods," *Advances in Futures and Options Research*, 9, 197-231.
- Björk, T. & Christensen, B.J. (1999). "Interest Rate Dynamics and Consistent Forward Rate Curves," *Mathematical Finance*, Wiley Blackwell, vol. 9(4), pages 323-348.
- Caporin, M., & McAleer, M. (2012). Do we really need both BEKK and DCC? A tale of two multivariate GARCH models, *Journal of Economic Surveys*, 26, 736-751.
- de Pooter, Michiel (2007), "Examining the Nelson-Siegel Class of Term Structure Models: In-Sample Fit versus Out-of-Sample Forecasting Performance," Institute Discussion paper, No. 2007-043/4.
- Diebold, F. X., Piazzesi, M. and Rudebusch, G.D. (2005). Modeling bond yields in finance and macroeconomics. *American Economic Review* 95, 415-420.
- Diebold, F. X. and Li, C. (2006), "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, Vol. 130, 2006, pp. 337-364.
- Diebold, F. X., Li, C. and Yue, V. Z. (2008), Global yield curve dynamics and interactions: A dynamic Nelson-Siegel approach. *Journal of Econometrics* 146, 351-363.
- Diebold, F.X., Rudebusch, G.D. and Aruoba, S.B. (2006). The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics* 131, 309-338.
- Fabozzi FJ, Martellini L, Priaulet P. (2005). Predictability in the shape of the term structure of interest rates. *Journal of Fixed Income* 15-1: 40-53.
- Litterman, R., and J. Scheinkman, 1991, Common Factors Affecting Bond Returns," *Journal of Fixed Income*, June, 54-61.
- Maasoumi, E., & McAleer, M. (2006). Multivariate stochastic volatility: An overview, *Econometric Reviews*, 25, 139-144.
- McAleer, M. (2005). Automated inference and learning in modeling financial volatility, *Econometric Theory*, 21, 232-261.
- Nelson, Charles R. and Siegel, Andrew F. (1987), "Parsimonious Modeling of Yield Curves," *Journal of Business*, Vol. 60, No. 4, pp. 473-489.
- Nelson, Daniel B. (1991), "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, Vol. 59, No. 2, pp. 347-370.
- Svensson, L.E.O. (1994). Estimating and interpreting forward interest rates. NBER Working Paper Series,