Modeling and forecasting the term structure of government bond yields

Dedi Rosadi
Department of Mathematics, Gadjah Mada University, Indonesia
Email: dedirosadi@ugm.ac.id

Abdurakhman
Department of Mathematics, Gadjah Mada University, Indonesia
Email: rachmanstat@ugm.ac.id

ABSTRACT. In this paper, we discuss the problem of modeling and forecasting the yield curve of government bond. We first model the yield curve using using Nelson-Siegel (NS) and the Nelson-Siegel-Svenson (NSS or 4 factors) and further estimate the parameters of the model using the hybrid-Genetic Algorithm approach. In Muslim, Rosadi, Gunardi and Abdurakhman (2014) we show this particular estimation method is found to have the best performance for estimating the in-sample empirical yield curve. For comparison purpose, we also estimate the model using full estimation algorithm of nonlinear regression constrained optimization method (see e.g. Rosadi, 2011). Then, we forecast various parameters of all models using Vector Autoregression order 1 (VAR(1)) model (see also alternative approach in Rosadi, Nugraha and Dewi, 2011). The forecasted parameters are then used to calculate the yield curve of the government bonds. The empirical studies are provided using Indonesian Government Bond data, obtained from Indonesia financial market. All of the computations are done using open source software R, where in particular we also use R-GUI package RcmdrPlugin.Econometrics (Rosadi, 2010).

Keywords: Yield Curve, Nelson-Siegel model , Vector Autoregressive, Autoregressive

1. Introduction

The term structure of interest rates is a relation of the yield and the maturity of default free zero-coupon securities and provides a measure of the returns that an investor might expect for different investment periods in a fixed income market. One potential candidate for a standard model of the yield curve is the Nelson and Siegel (1987) (hereafter NS) class of yield curve models. Models of this class are already very popular with practitioners, researchers, and academics in finance and economics (see e.g. Stander, 2005; and Rosadi, Gunardi, Abdurakhman, Utami, and Wulansari (2008), for an overview of various models of yield curve). However, in emphasizing simple relationships by maturity and the fit to observed yield curve data at a given point in time, NS model (and all static yield curve models) overlook the dynamics of interest rates over time. Hence, they are not inter-temporally consistent (i.e interest rates and yield curves at different points in time cannot be mapped to each other via an underlying stochastic time-series process). As the result, they fail to model the dynamic relationship between the parameters of a term-structure model and therefore, the forecasting ability of these models is poor. This has motivated us to apply dynamic model Vector Autoregressive for the purposes of forecasting the yield curve (see also alternative approach in Diebold and Li, 2006; Rosadi, Nugraha and Dewi, 2011).

The rest of this paper can be described as follows. In the next subsection we shortly describe the yield curve and nelson siegel model for the yield curve. In section three and four, we shortly overview the neural networks model and VAR methodologies that will be used for the forecasting. Section five describes the empirical results and last section concludes.
2. Yield Curve and Estimation

Here we introduce the fundamentals of the yield curve. If \( P(t,T) \) denotes the price of zero coupon bond at time \( t \) with time to maturity \( T \), often called as discount function and \( R(t,T) \) denotes its continuously compounded zero-coupon nominal yield to maturity (YTM), also called spot rate, then the discount curve can be obtained from the yield curve as: \( P(t,T) = e^{-(T-t)R(t,T)} \). Therefore, we obtain relation between spot rate \( R(t,T) \) with the price of zero coupon bond \( P(t,T) \) as:

\[
R(t,T) = -\frac{1}{T-t} \ln P(t,T).
\]

Let’s denote \( f(t,T,T+\Delta t) \) be the forward rate with the contract determined at time \( t \) with the settlement date at time \( T \) and maturity at time \( T+\Delta t \). The instantaneous forward rate \( f(t,T,T+\Delta t)_{\text{inst}} \) as defined as forward rate with \( \Delta t \to 0 \), therefore, instantaneous forward rate can be written as \( f(t,T) \). The relation between the price of zero coupon bond with the instantaneous forward rate can be written as:

\[
P(t,T) = \exp\left(-\int_t^T f(t,u)du\right) \quad \text{or} \quad f(t,T) = -\frac{d}{dt} \ln P(t,T)
\]

and therefore the relationship between the yield to maturity and the forward rate is therefore given as:

\[
R(t,T) = \int_t^T f(t,u)du / (T-t)
\]

This means that the zero-coupon yield is an equally weighted mean of the forward rates. It is therefore possible to price any coupon bond as the sum of the present values of future cash flows (coupon payments and the principal payment), if the yield curve or the forward-rate curve is given.

There are various models of yield curve available in literature, see e.g. Stander (2005) and Rosadi, et.al (2008) for overview. NS models the instantaneous forward-rate curve \( f(t,m) \) is given as (Nelson and Siegel, 1987)

\[
f(t,m) = \beta_0 + \beta_1 \exp\left(-\frac{m}{\tau_1}\right) + \beta_2 \frac{m}{\tau_1} \exp\left(-\frac{m}{\tau_1}\right)
\]

This implies that Nelson-Siegel yield curve is given as:

\[
R(t,m) = \frac{1}{m} \int_0^m f(x)dx = \beta_0 + \beta_1 \left[1 - \exp\left(-\frac{m}{\tau_1}\right)\right] \frac{m}{\tau_1} + \beta_2 \left[1 - \exp\left(-\frac{m}{\tau_1}\right)\right] \frac{m}{\tau_1} - \exp\left(-\frac{m}{\tau_1}\right)
\]

The rate of exponential decay is governed by the parameter \( \tau \) and \( \beta_0, \beta_1, \beta_2 \) are interpreted as the short-term component, long-term component, and the medium-term component, respectively, and \( m \) denotes the maturity time. The parameters \( \beta_0 \) and \( \tau \) need to be positive and \( \beta_0 + \beta_1 > 0 \).

Svensson (1994) has expanded the original 3-factors model by adding the second hump such that it becomes 4-factors or Nelson-Siegel-Svensson/NSS model. The NSS model has the following form

\[
f_{t}(\lambda; \hat{\beta}, \tau) = \beta_0 + \beta_1 \exp\left(-\frac{\lambda}{\tau_1}\right) + \beta_2 \left[1 - \exp\left(-\frac{\lambda}{\tau_1}\right)\right] \frac{\lambda}{\tau_1} - \exp\left(-\frac{\lambda}{\tau_1}\right) + \beta_3 \left[1 - \exp\left(-\frac{\lambda}{\tau_2}\right)\right] \frac{\lambda}{\tau_2} - \exp\left(-\frac{\lambda}{\tau_2}\right)
\]

where \( \hat{\beta} \) is linear parameter vector that \( \hat{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3)' \) and \( \tau = (\tau_1, \tau_2)' \) nonlinear parameter that determines the position of the first and second hump and the other parameters are the same as NS model. \( \beta_0 \) must be positive, it is the asymptotic value of \( f_t(\lambda; \hat{\beta}, \tau) \). The curve will tend towards the asymptote as the \( \lambda \) approaches infinity. \( \beta_1 \) determines the starting (short-term) value of the curve in terms of deviation from the asymptote. It also defines the basic speed with which the curve tends towards its long-term trend. The curve will have a negative slope if this parameter is positive and vice versa. Note that the
sum of $\beta_0$ and $\beta_1$ is the vertical intercept. $\tau_1$ must be positive, specifies the position of the first hump or U-shape on the curve. $\beta_2$ determines the magnitude and direction of the hump. If $\beta_2$ is positive, a hump will occur at $\tau_1$ whereas, if $\beta_2$ negative, a U-shaped value will occur at $\tau_1$. $\tau_2$ must also be positive, specifies the position of the second hump or U-shape on the curve. And $\beta_3$ analogous to $\beta_2$, determines the magnitude and direction of the second hump (Bolder and Streliski, 1999).

The Nelson Siegel model class can be estimated using constrained nonlinear regression (NLR) optimization procedure and using genetic algorithm (GA), based on YTM function (eq. (1) and (2)) above (see Muslim et al., 2014). For NLR optimization method, there are two alternative estimation procedures available in literature, namely Full estimation algorithm and Partial-estimation algorithm (e.g., Bolder and Streliski, 1999). In Partial-estimation algorithm, we are first fixing the $\tau$s and $\beta$s are estimated, or the other ways, fixing the $\beta$s and $\tau$s are estimated. For the full estimation algorithm, we may apply, for example, the sequential Quadratic Programming (SQP), Nelder-Mead Simplex method, or other constrained optimization methods (see e.g., Lange, 2010; Lagarias, Reedsz, Wrightx and Wright, 1998). In this paper, we use full estimation algorithm as implemented in function `nlminb` in R (implemented PORT optimisation method, see http://netlib.bell-labs.com/netlib/port/). Alternatively we may use function `constrOptim` (implemented Nelder-Mead and several other optimization procedures). See e.g. Rosadi (2011) for further detail. In Nugraha and Rosadi (2011), we apply the OLS method for the starting values of the parameters, and further use these starting values for Nelder Mead estimation of the parameters using R-CLI. The R-GUI implementation of the NS model is also available in package RcmdrPlugin.Econometrics (Rosadi, 2010). The alternative GA method is implemented in R package GA, and shown in Muslim et al. (2014) that this estimation method is found to have the better performance than NLR method for estimating the in-sample empirical yield curve. Unlike NLR method, the latter also shown will not suffer from the problem of sensitivity to the starting values in the optimization procedure.

3. Forecasting Yield Curves

A good approximation to yield-curve dynamics should not only fit well in-sample, but also forecast well out-of-sample. Because the NS-class yield curve models depends on their parameters $\beta$s and $\tau$s, forecasting the yield curve is equivalent to forecasting $\beta$s and $\tau$s. In this section we outline the method of forecasting using Vector autoregression (see also Rosadi, Nugraha and Dewi, 2011 for alternative forecasting method based on neural network approach).

The VAR($p$) model with k endogeneous variables $y_t = (y_{1t}, \ldots, y_{kt})$ can be written as

$$ y_t = A_1 y_{t-1} + \ldots + A_p y_{t-p} + CD_t + u_t $$

where $A_i, i = 1, \ldots, p$ is the coefficient matrix with dimension $(k \times k)$, $u_t$ is white noise process with dimension $k$ and the time invariant and positive definite matriks $E u_t u_t = \Sigma_u$. Matrix $C$ is the coefficient matrix of $m$-independent random variables $D_t$ with dimension $k \times m$, where the matrix $D_t$ contains all possible random variables, such as constanta, trend components, dummy variables and/or seasonal dummy variables.

The parameters of the VAR model can be estimated using ordinary least square method, where the optimal order can be found using information criteria, such as AIC (Akaike Information Criteria), HQC (Hannan Quinn Criteria), SBC (Schwarz Bayesian Criteria).

In this paper, for forecasting the NS and NSS yield curve, we apply the VAR model for the endogenous variables $\beta$s and $\tau$s of the NS model. Although the optimal order of the model probably larger that 1, for parsimonious approach we only consider VAR(1) method. Various econometrics models as the alternatives model for forecasting yield curve are discussed in e.g. Diebold and Li (2006) including random walk model, forward curve regression, autoregressive model, error correction model, etc. See also Rosadi, Nugraha and Dewi (2011) for alternative forecasting method based on neural network approach.
4. Empirical results

4.1. Data

For practical application, we use the daily transaction data of price zero coupon bond of Indonesia Government Bond from May 2, 2008 until July 31, 2008, which is the data during the global financial crisis in 2008. The BI-rate and 3-month SBI are used as the data with the smallest maturity time where we use the most recent observation available in each month. The other observations are the transaction price of bond (clean price) on the respective dates. Several series of bonds are considered as the important series and will be used as the anchor for curve estimation. If no transaction series are available for the “anchor” data on a particular day, we will use the quote price of HIMDASUN (where the historical data are available from Bloomberg).

Table 1. Important bond for yield curve estimation

<table>
<thead>
<tr>
<th>TTM</th>
<th>Series</th>
<th>1-month</th>
<th>3-months</th>
<th>1-Y</th>
<th>3-Y</th>
<th>5-Y</th>
<th>7-Y</th>
<th>10-Y</th>
<th>15-Y</th>
<th>20-Y</th>
<th>30-Y</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BI rate</td>
<td>SBI-3</td>
<td>FR0002</td>
<td>FR0016</td>
<td>FR0049</td>
<td>FR0027</td>
<td>FR0048</td>
<td>FR0046</td>
<td>FR0047</td>
<td>FR0050</td>
<td></td>
</tr>
</tbody>
</table>

From the raw data, we calculate the time to maturity (TTM) and yield to maturity (YTM) for each observation.

4.2. Estimating Nelson Siegel Parameter

In NS model, there are four parameters that will be estimated \((\tau, \beta_0, \beta_1, \beta_2)\) for each daily transaction during May 2 to July 31 2008. On the other hand, for estimating NSS model, we need to estimate six parameters, namely \((\tau_1, \tau_2, \beta_0, \beta_1, \beta_2, \beta_3)\).

The estimation is implemented in software R.2.15.2. For estimating the parameter, we apply full estimation algorithm, implemented in function \texttt{nlminb} in R. For Genetic algorithm, we use function \texttt{ga} in the package \texttt{GA}. The GA method is implemented with replication 30 times and we use the average of results as the estimator of the parameters.

It has been known that the NS model is sensitive to the method of estimation and particularly to the starting values used for the parameters. From our studies in Rosadi et.al (2008), for our data, we find that the optimal choice of starting values (at least locally optimal) of \(\beta_0\) is the yields of BI rate and for \(\beta_1, \beta_2\) and \(\tau\) are equal to 1, and for the constrains, we consider \(\beta_0\) is greater or equal than the yields of BI rate and with \(\tau \geq 1\).

4.4. Prediction using VAR

Using the data set of \((\tau, \beta_0, \beta_1, \beta_2)\) and \((\tau_1, \tau_2, \beta_0, \beta_1, \beta_2, \beta_3)\) from GA and NLR approach from May 2 until June 30, 2008 (41 data), we try to fit VAR(1) model (although using AIC, we found that the best model for the data is VAR(2) model). Using the result of VAR(1) model, we predict the future 22 days values of the parameters \((\tau, \beta_0, \beta_1, \beta_2)\) and \((\tau_1, \tau_2, \beta_0, \beta_1, \beta_2, \beta_3)\) (i.e. can be considered as the data testing). Based on these predicted values of parameters, we calculate the predicted yields to maturity, and calculate the mean square error (MSE) of the prediction. We found that the forecasting based on VAR(1) -NLR based method is very volatile. We also further found that the VAR(1) -GA methods will give smaller MSE than VAR(1) -NLR based method, however detail is omitted. For illustrative purpose, we show the following two figures, which compare the prediction of Yield Curve using VAR(1)-GA based method on NS and NSS model and the real data for two trading days, i.e. on July 7, 2008 and July 29, 2008, which is the result of 5 days (a week) and 21 days (5 weeks) forecasting. Based on these pictures.
(and other similar behavior for the other days), we found that the accuracy of prediction based on NSS-VAR(1)-GA will perform better for short term forecasting of the yield curve, where for the long term forecasting, there is no difference between using NS and NSS method.

5. Conclusion

The above results indicate that VAR method can be used to predict the behavior of yield curve, and we found that the accuracy of prediction based on NSS-VAR(1)-GA will perform better for short term forecasting of the yield curve, where for the long term forecasting, there is no difference between using NS and NSS method. However, it seems that the NS model in some cases cannot model the behavior of yield data accurately, and therefore resulting an inaccurate prediction of future behavior of the yield curve. For the future research, we suggest considering more advanced model for yield curve models, such as the extended method by incorporating the financial and macroeconomic variables into NS model (see e.g. the application of this method for corporate bond modeling in Rosadi, Qoyyimi and Wulansari, 2010) which can be used for modeling yield curve of either investment or non-investment grade bonds, or in zero coupon bond modeling in Ludvigson and Ng (2009).

![Yield Curve of Nelson-Siegel Model](image)

Figure 1. Predicted Yield Curve for July 7, 2008

REFERENCES


Figure 2. Predicted Yield Curve for July 29, 2008


Muslim, Rosadi, Gunardi, and Abdurakhman, 2014, Calibrating The Nelson-Siegel Model Classes And Their Estimation Using Hybrid-Genetic Algorithm Approach: Case Study Of Indonesian Government Bonds, submitted


Rosadi, Nugraha and Dewi, 2011, Forecasting the Indonesian Government Securities Yield Curve using Neural Networks and Vector Autoregressive Model, Proceeding ISI 58 th Dublin, Ireland
Rosadi, D., Qoyyimi, D. T. and Wulansari, Y., 2010, Modeling Yield Spread for Indonesian Corporate Bond, Internal Research Report, Department of Statistics - Gadjah Mada University and Indonesian Bond Pricing Agency
