Dynamic Nelson-Siegel Model

Chelsea Anne Go, Brian Godwin Lim

Contents

1 Introduction

Interest rate term structure, or more commonly known as yield curve, refers to the relationship between interest rates or bond yields of similar quality at different maturities. Determining the interest rates, and consequently the yield curve, is important as it plays a significant role in the financial markets. It affects the prices of different financial instruments such as derivatives. It also affects economic decisions such as monetary policies and debt policies. The commonly used benchmark interest rates in the market are derived from the yields of treasury bills and treasury bonds issued by the government. These are generally considered "risk-free" instruments since governments have a low probability of default. Furthermore, the yield curves derived from these instruments are typically upward sloping curves which reflects the premium for greater exposure to market risk for longer maturities. Nevertheless, it is also important to note that the shape of yield curves change throughout time. It moves up and down and changes in shape, which are typically monotonic, humped, or S-shaped.

Furthermore, the yield curve's shape indicates market participants' outlook regarding economic activity. Policy makers and investors are also interested in examining the difference between short-term and long-term interest rates, which is referred to as the term spread. This is often used to forecast economic activity or financial instability. In particular, inversions of the yield curve, indicated by a negative term spread, is considered as an early warning indicator for recessions.

One type of interest rate used in the pricing of different financial instruments is the short rate. Several models in the literature attempt to capture the short rate (stochastic) process. One-factor models are a commonly used models to capture the dynamics of the interest rate process in a parsimonious manner. In these models, the process for r_t involves only one source of uncertainty. The process is also assumed to be stationary. Some common one-factor models are the Vasicek model (1977) and the Cox-Ingersoll-Ross (CIR) model (1985).

In 1987, Nelson & Siegel (1987) introduced another interest rate model – the Nelson-Siegel model (1987). This is a parametrically parsimonious three-factor model for yield curves that can model the shapes generally associated with the yield curve. This model features the level, slope, and curvature parameters (Kumar et al., 2021). The level component conveys market participants' inflation expectations. The slope component represents business cycle circumstances and may thus be modified by the central bank monetary policies. The curvature component covers the short-term rate swings as a result of central bank activities.

As an extension to the original model, the Dynamic Nelson-Siegel model addresses the key practical problem of forecasting the yield curve. To forecast future yield curves, standard time series models are fitted to the level, slope, and curvature parameters. The forecast of these parameters from the models are then used to predict the future yield curve.

2 Conceptual Discussion

2.1 Second Order Ordinary Differential Equation

Consider the second order non-homogeneous ordinary differential equation

$$x'' + a_1 x' + a_2 x = b(t).$$

To solve this ordinary differential equation, consider the homogeneous equation

$$x'' + a_1 x' + a_2 x = 0$$

and construct the characteristic polynomial

$$\lambda^2 + a_1 \lambda + a_2 = 0.$$

Suppose the roots of the characteristic equation are λ_1 and λ_2.

Case 1. $\lambda_1 \neq \lambda_2$. The solution to the homogeneous equation is

$$x = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}.$$

Case 2. $\lambda_1 = \lambda_2 = \lambda_0$. The solution to the homogeneous equation is

$$x = C_1 e^{\lambda_0 t} + C_2 t e^{\lambda_0 t}.$$

Case 3. $\lambda_1 = \alpha + \beta i$, $\lambda_2 = \alpha - \beta i$. The solution to the homogeneous equation is

$$x = C_1 \cos(\beta t)\, e^{\alpha t} + C_2 \sin(\beta t)\, e^{\alpha t}.$$

Suppose further that a particular solution to the second order non-homogeneous ordinary differential equation is $p(t)$. Then, the solution to the non-homogeneous ordinary differential equation is given by

$$x = p(t) + \left[\text{general solution to } x'' + a_1 x' + a_2 x = 0\right].$$

To illustrate, if the homogeneous equation has two equal characteristic roots $\lambda_1 = \lambda_2 = \lambda_0$, then the solution is given by

$$x = p(t) + C_1 e^{\lambda_0 t} + C_2 t e^{\lambda_0 t},$$

where the parameters C_1 and C_2 depend on the initial conditions.

2.2 Forward Rates and Yield Curve

Suppose $P_t(\tau)$ is the price at time t of a zero-coupon bond that pays 1 at time $t+\tau$, *i.e.*, τ is the maturity of the bond. Let $y_t(\tau)$ be the corresponding continuously compounded zero rate. Then, the price of the zero-coupon bond is

$$P_t(\tau) = e^{-y_t(\tau)\tau}.$$

Moreover, the continuously compounded forward rate for the interval $[t+\tau', t+\tau]$ contracted at time t is given by

$$R_t(\tau', \tau) = -\frac{\ln P_t(\tau) - \ln P_t(\tau')}{\tau - \tau'}.$$

Furthermore, if $f_t(\tau)$ is the instantaneous forward rate with maturity $t+\tau$ contracted at time t, then

$$f_t(\tau) = \lim_{\tau' \to \tau} R_t(\tau', \tau) = -\frac{\partial \ln P_t(\tau)}{\partial \tau}.$$

It then follows that

$$y_t(\tau) = \frac{1}{\tau} \int_0^\tau f_t(u) du.$$

The yield curve at time t is the curve that describes the relationship between spot interest rates $y_t(\tau)$ for different maturities τ. These curves are usually monotonic, humped, or S-shaped.

2.3 The Nelson-Siegel Model

The Nelson-Siegel model assumes that the forward rate f_t is a solution to the second order ordinary differential equation

$$x'' + \frac{2}{\lambda_t} x' + \frac{1}{\lambda_t^2} x = \frac{\beta_{0t}}{\lambda_t^2}.$$

According to Nelson & Siegel (1987), experimentation suggests that among the class of solutions, the case of equal roots $\lambda_1 = \lambda_2$ is a more parsimonious model that can still generate the same range of shapes for the yield curve. Hence, the forward rate can be expressed as

$$f_t(\tau) = \underbrace{\beta_{0t}}_{p(t)} + \underbrace{\beta_{1t} e^{-\tau/\lambda_t}}_{C_1 \exp(\lambda_0 t)} + \underbrace{\beta_{2t} \left(\frac{\tau}{\lambda_t}\right) e^{-\tau/\lambda_t}}_{C_2 t \exp(\lambda_0 t)},$$

where β_{0t}, β_{1t}, β_{2t}, and λ_t are constants and $\lambda_t \neq 0$.

Therefore, Neslon-Siegel model for yield $y_t(\tau)$ can be expressed as

$$\begin{aligned} y_t(\tau) &= \frac{1}{\tau} \int_0^\tau \left[\beta_{0t} + \beta_{1t} e^{-u/\lambda_t} + \beta_{2t} \left(\frac{u}{\lambda_t}\right) e^{-u/\lambda_t}\right] du \\ &= \beta_{0t} + \beta_{1t} \left(\frac{1 - e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right) + \beta_{2t} \left(\frac{1 - e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right), \end{aligned}$$

where β_{0t}, β_{1t}, and β_{2t} are the **factors** and their coefficients are the **factor loadings**.

The factor loading of β_{0t} is 1, a constant that does not decay to zero even when $\tau \to \infty$. Hence, it has significant contribution to the yields for all maturities. Thus, it is referred to as the long-term factor.

The factor loading of β_{1t} is $\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right)$, which decreases to 0 when $\tau \to \infty$. However, it has significant contribution to the yields at shorter maturities (smaller values of τ). Thus, it is referred to as the short-term factor.

Lastly, the factor loading of β_{2t} is $\left(\frac{1-e^{\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right)$, which starts at 0, then increases, and finally decreases to 0. More precisely, it approaches 0 as $\tau \to 0^+$ and as $\tau \to \infty$. Hence, it has significant contribution to the yields at medium-term maturities. Thus, it is referred to as the medium-term factor.

The factors β_{0t}, β_{1t}, and β_{2t} can also be interpreted in terms of level, slope, and curvature of the yield curve, respectively (Diebold & Li, 2006). The factor β_{0t} is related to the yield curve level since $\lim_{\tau\to\infty} y_t(\tau) = \beta_{0t}$. Thus, this factor is responsible for parallel shifts in the yield curve. Moreover, the factor β_{1t} is related to the yield curve slope, defined as $\lim_{\tau\to\infty} y_t(\tau) - \lim_{\tau\to 0^+} y_t(\tau) = -\beta_{1t}$. Lastly, the factor β_{2t} is related to the yield curve curvature, defined as $2y_t(2) - y_t(0.25) - y_t(10) = 0.00053\beta_{1t} + 0.37\beta_{2t}$.

2.4 Parameter Estimation

The Nelson-Siegel model for yield is

$$y_t(\tau) = \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right) + \beta_{2t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right),$$

where λ_t is the shape parameter. For each t, when λ_t is specified, the model becomes linear in the parameters β_{0t}, β_{1t}, and β_{2t}. Mathematically,

$$\begin{aligned} y_t(\tau) &= \beta_{0t} + \beta_{1t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t}\right) + \beta_{2t}\left(\frac{1-e^{-\tau/\lambda_t}}{\tau/\lambda_t} - e^{-\tau/\lambda_t}\right) \\ Y_\tau^t &= \beta_{0t} + \beta_{1t}X_{1\tau}^t + \beta_{2t}X_{2\tau}^t, \end{aligned}$$

where $Y_\tau^t = y_t(\tau)$, $X_{1\tau}^t = \dfrac{1-e^\tau/\lambda_t}{\tau/\lambda_t}$ and $X_{2\tau}^t = \dfrac{1-e^\tau/\lambda_t}{\tau/\lambda_t} - e^{-\tau/\lambda_t}$, with specific values of τ giving specific values of Y_τ^t, $X_{1\tau}^t$, and $X_{2\tau}^t$.

Suppose at time t, the interest rates $y_t(\tau)$ for different maturities τ is given. Then, for each t, the parameters β_{0t}, β_{1t}, and β_{2t} in the equation above are estimated using the ordinary least squares (OLS) method for multiple linear regression.

Several methods can be used to obtain the value(s) of λ_t. One method is the fixed lambda method which considers the same value of λ_t for all t (de Lara-Tuprio et al., 2017). Thus, a series of estimates $\{\beta_{0t}\}$, $\{\beta_{1t}\}$, $\{\beta_{2t}\}$ is obtained based on a single value of λ_t. In this method, a grid search is performed for λ_t values ranging from 2 to 40 in small increments. For each value of λ_t, the average R^2 across all t is calculated. Finally, the value of λ_t that produced the highest R^2 is chosen.

2.5 Forecasting Yield Curve

In forecasting the yield curve under the Dynamic Nelson-Siegel model, consider the following. Suppose a series of estimates $\{\beta_{0t}\}$, $\{\beta_{1t}\}$, $\{\beta_{2t}\}$ is obtained using a single value of λ_t (fixed lambda method). For each β_{it}, $i = 0, 1, 2$, several time series models are fitted. The ARIMA models are considered for the mean equation while Exponential GARCH models are considered for the variance equation. Mathematically, an ARIMA(p, d, q)-EGARCH(m, s) model is given by

$$(1-B)^d X_t = \phi_0 + \sum_{i=1}^{p} \phi_i (1-B)^d X_{t-i} + a_t - \sum_{i=1}^{q} \theta_i a_{t-i},$$

$$a_t = \sigma_t \epsilon_t,$$

$$\ln(\sigma_t^2) = \alpha_0 + \sum_{i=1}^{m} \alpha_i \frac{|a_{t-i}| + \gamma_i a_{t-i}}{\sigma_{t-i}} + \sum_{i=1}^{s} \beta_i \ln(\sigma_{t-i}^2),$$

where B is the back-shift operator, ϕ_i is the coefficient of the autoregressive (AR) component, θ_i is the coefficient of the moving average (MA) component, a_t is the error term at time t, σ_t is the standard deviation at time t, $\epsilon_t \sim (0, 1)$ is a white noise series, α_i is the coefficient of the moving average (MA) component of the residual of the log variance, γ_i is the leverage effect of a_t, i periods ago, and β_i is the coefficient of the autoregressive (AR) component of the residual of the log variance.

The ARIMA-EGARCH model that has the lowest information criterion is chosen. Afterwards, the β_{0t}, β_{1t}, β_{2t} forecasts are obtained using these models. Finally, using the ℓ-step ahead forecasts $\beta_{0(t+\ell)}$, $\beta_{1(t+\ell)}$, and $\beta_{2(t+\ell)}$, the yield curve in ℓ periods is calculated as

$$y_{(t+\ell)}(\tau) = \beta_{0(t+\ell)} + \beta_{1(t+\ell)} \left(\frac{1 - e^{-\tau/\lambda}}{\tau/\lambda} \right) + \beta_{2(t+\ell)} \left(\frac{1 - e^{-\tau/\lambda}}{\tau/\lambda} - e^{-\tau/\lambda} \right),$$

where λ is the value of λ_t obtained using the fixed lambda method.

3 Assumptions

The following assumptions are made in forecasting the yield curve using the Dynamic Nelson-Siegel model.

1. An Actual/360 day count convention is used in all calculations.

2. The PDST-R2 rates are considered in this paper. The data consists of the 1-month, 3-month, 6-month, 1-year, 2-year, 3-year, 4-year, 5-year, 7-year, 10-year, and 20-year interest rates. Interest rates with tenors less than or equal to 1 year are assumed to be zero rates. However, the interest rates with tenors greater than 1 year are not zero rates. The zero rates are obtained using the bootstrapping method.

3. The latest one year data (June 29, 2015 to June 30, 2016) is considered the out-of-sample dataset. Meanwhile, the dataset from January 14, 2009 to June 28, 2015 is considered the in-sample dataset. For simplicity, data prior to January 14, 2009 are dropped from the analysis since they produced negative discount factors.

4. The zero rates every Wednesday (or the next available date) is selected to represent the zero rate each week.
5. The grid search fixed lambda method is used to obtain the optimal value of the shape parameter λ_t. This is performed using increments of 0.001.
6. In modeling the beta parameters β_{0t}, β_{1t}, and β_{2t}, the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test is performed to check whether the series is stationary while the ARCH LM test is then performed to test for ARCH effect.
7. The Akaike information criterion (AIC) is used to determine the best time series model for each beta.

4 Python Implementation

The implementation of the Dynamic Nelson-Siegel model using Python is as follows.

4.1 Bootstrapping for Zero-Rates

The data consists of the 1-month, 3-month, 6-month, 1-year, 2-year, 3-year, 4-year, 5-year, 7-year, 10-year, and 20-year interest rates. The interest rates for tenors less than or equal to 1 year are considered zero rates. For interest rates with tenors greater than 1 year, the bootstrapping method is performed to obtain the corresponding zero rates.

Let n_j be the number of years corresponding to tenor j. First, interest rates for increments of 0.5 years are obtained using linear interpolation. Afterwards, the discount factors DF_{n_j} are obtained. For $j \leq 1$,

$$\text{DF}_{n_j} = \frac{1}{1 + r_j n_j},$$

since the 6-month and 1-year PDST-R2 rates are already zero rates. For $j > 1$,

$$\text{DF}_{n_j} = \frac{1 - r_j \displaystyle\sum_{i=1}^{2j-1} \left(n_{0.5i} - n_{0.5(i-1)}\right) \times \text{DF}_{n_{0.5i}}}{1 + \left(n_j - n_{j-0.5}\right) \times r_j}.$$

Finally, the continuously compounded zero rates are obtained as

$$z_{n_j} = -\frac{\ln\left(\text{DF}_{n_j}\right)}{n_j}.$$

Additionally, the interest rates with tenors of 1-month and 3-month are also converted to continuously compounded rates using the formula

$$z_{n_j} = \frac{\ln(1 + r_j n_j)}{n_j},$$

where r_j is the simple interest rate. An Actual/360 day count convention is used in all calculations. The Python codes are found in Listings 1 and 2 in the Appendix.

The data is also filtered to remove invalid entries from bootstrapping. Hence, the final dataset consists of 1819 dates from January 13, 2009 to June 30, 2016.

Date	1M	3M	6M	1Y	2Y	3Y	4Y	5Y	7Y	10Y	20Y
12/7/2007	4.1426	4.1284	4.7912	5.4464	5.8302	5.8963	5.9334	6.0878	6.1322	6.6328	9.9247
12/10/2007	4.0928	4.0789	4.6691	5.4464	5.8060	5.9300	5.9318	6.0877	6.1305	6.6975	9.8777
12/11/2007	4.0928	4.0789	4.7912	5.4464	5.7295	5.8287	5.9041	5.9850	6.1143	6.6474	9.9414
12/12/2007	4.0928	3.9799	4.7180	5.4464	5.7301	5.7783	5.8628	5.9087	6.0992	6.5163	9.9264
12/13/2007	4.0928	4.0789	4.6936	5.3991	5.6560	5.6574	5.7381	5.7574	6.0611	6.4902	9.9949
12/14/2007	4.0928	4.0789	4.6447	5.3991	5.6312	5.6546	5.6819	5.7201	5.9731	6.5071	10.0153
12/17/2007	4.1924	4.1779	4.6936	5.4464	5.6145	5.6629	5.6738	5.8399	6.0623	6.4871	9.9243
12/18/2007	4.1924	4.2768	4.8156	5.3991	5.6299	5.6574	5.6956	5.8940	6.1153	6.5055	9.8232
12/19/2007	4.2422	4.2273	4.8156	5.3991	5.6299	5.7091	5.6599	5.8937	6.1068	6.5407	10.0645

Figure 1: Daily Zero Rates

Lastly, the weekly data is extracted by taking the zero rates every Wednesday (or the next available date) of each week. Thus, the final dataset consists of 390 dates from January 14, 2009 to June 29, 2016.

Date	1M	3M	6M	1Y	2Y	3Y	4Y	5Y	7Y	10Y	20Y
1/14/2009	5.0318	5.0924	4.9883	5.2219	5.5443	5.6586	5.8204	6.2255	6.8826	7.3928	21.2755
1/21/2009	4.5269	4.5363	4.8203	4.9487	5.2471	5.6535	5.9096	6.0345	7.1198	7.5835	27.4660
1/28/2009	4.5909	4.2128	4.5062	4.7107	5.4484	5.7433	6.0115	6.2750	6.9693	7.5779	27.2703
2/4/2009	4.4921	4.1860	4.4498	4.5971	5.4264	5.8070	6.1732	6.3260	7.1398	7.8840	18.0043
2/11/2009	4.4921	4.4752	4.8162	4.8536	5.3670	5.8546	6.1671	6.3547	7.2346	8.0717	17.1947
2/18/2009	4.4921	4.4752	4.2938	4.4912	5.4049	5.5713	6.1878	6.3205	7.2732	8.1816	16.9522
2/25/2009	4.6914	4.7223	4.8162	4.7662	5.3189	5.7136	6.1770	6.3443	7.2465	8.4731	16.1511
3/4/2009	4.4664	4.4496	4.3305	4.5570	5.0960	5.5678	5.8718	6.2721	7.2738	8.2916	16.8247
3/11/2009	4.4913	4.1628	4.4840	4.6006	5.1502	5.4433	5.9835	6.2685	7.3529	8.2748	12.6380

Figure 2: Weekly Zero Rates

4.2 Determining the Shape Parameter λ_t

Using the fixed lambda method (de Lara-Tuprio et al., 2017), the optimal shape parameter λ_t is determined. A grid search is performed for values of λ_t from 2 to 40 with increments of 0.001. For each value of λ_t, multiple linear regressions are performed, one for each date. Afterwards, the average R^2 across all dates is calculated. Table 1 below shows the summary of the grid search. The results show that the optimal shape parameter is $\lambda_t = 5.434$ with an average R^2 of 0.964466. It is also worth noting that the average R^2 for λ_t near this value are almost equal.

λ_t	Average R^2
5.434	0.964466
5.435	0.964466
5.433	0.964466
5.436	0.964466
5.432	0.964466
5.431	0.964466

Table 1: Fixed Lambda Method Results

Lastly, the time series of β_{0t}, β_{1t}, and β_{2t} corresponding to $\lambda_t = 5.434$ are obtained for each date. Table 2 below shows the summary of the β_{0t}, β_{1t}, β_{2t}, and R^2 obtained.

	Mean	Min	Max	Standard Deviation
β_{0t}	9.788597	1.390245	54.749518	6.205514
β_{1t}	-7.683124	-48.813521	-1.176916	5.211242
β_{2t}	-1.154726	-69.032021	11.154890	8.405596
R^2	0.964466	0.783727	0.999479	0.032677

Table 2: Summary of β_{0t}, β_{1t}, β_{2t}, and R^2

4.3 Time Series Models for Beta Parameters

After obtaining the time series $\{\beta_{0t}\}$, $\{\beta_{1t}\}$, $\{\beta_{2t}\}$, several ARIMA-EGARCH models are fitted to each series. The model with the lowest AIC is then selected. The details of each time series model are described below.

4.3.1 β_{0t} Factor

First, the optimal ARIMA-EGARCH model for β_{0t} is determined. A KPSS test is first performed on $\{\beta_{0t}\}$ with $\alpha = 0.05$. The results show a critical value of 1.550948 with a p-value of 0.00. Hence, the null hypothesis is rejected and the series is non-stationary. Thus, the differenced series $\{d\beta_{0t}\}$ is considered. Figure 3 shows the time series plot of $\{d\beta_{0t}\}$, its autocorrelation function (ACF), and its partial autocorrelation function (PACF). Both the ACF and PACF plots show a significant autocorrelation for lag 2.

Figure 3: $d\beta_{0t}$ Plot, ACF, and PACF

Afterwards, the possible ARIMA models are explored and summarized in Table 3. Using AIC as the criteria, the results show that the best mean model is ARIMA(2,1,2) with an AIC of 1324.16. The BIC and HQIC also supports this conclusion.

Model	AIC	BIC	HQIC
ARIMA(2, 1, 2)	1324.16	1343.25	1331.77
ARIMA(2, 2, 2)	1339.96	1359.04	1347.57
ARIMA(2, 1, 1)	1341.69	1356.96	1347.78
ARIMA(1, 1, 2)	1343.57	1358.84	1349.65
ARIMA(2, 2, 1)	1343.63	1358.89	1349.71

Table 3: ARIMA Models for β_{0t}

After fitting the mean model, the residuals then are studied. An ARCH LM test is performed on the model residuals. The results show a critical value of 116.985703 with a p-value of 0.00. Hence, this indicates the presence of an ARCH effect. The residuals are then fitted into several EGARCH models. The ϵ_t distribution considered includes normal, Student's t, and the generalized error distribution (GED). Table 4 shows the possible models and their summary. Using AIC as the criteria, the results show that the best model is ARIMA(2,1,2)-EGARCH(1,1) with an AIC of 873.39. The BIC also supports this conclusion.

Model	ϵ_t distribution	AIC	BIC
ARIMA(2, 1, 2) - EGARCH(1, 1)	Student-t	873.39	892.49
ARIMA(2, 1, 2) - EGARCH(2, 1)	Student-t	875.17	898.09
ARIMA(2, 1, 2) - EGARCH(1, 2)	Student-t	875.30	898.22
ARIMA(2, 1, 2) - EGARCH(2, 2)	Student-t	877.17	903.91
ARIMA(2, 1, 2) - EGARCH(1, 1)	GED	888.88	907.98

Table 4: ARIMA-EGARCH Model for β_{0t}

Hence, the time series model for β_{0t} is ARIMA(2,1,2)-EGARCH(1,1).

4.3.2 β_{1t} Factor

Afterwards, the optimal ARIMA-EGARCH model for β_{1t} is determined. A KPSS test is first performed on $\{\beta_{1t}\}$ with $\alpha = 0.05$. The results show a critical value of 1.445343 with a p-value of 0.00. Hence, the null hypothesis is rejected and the series is non-stationary. Thus, the differenced series $\{d\beta_{1t}\}$ is considered. Figure 4 shows the time series plot of $\{d\beta_{1t}\}$, its autocorrelation function (ACF), and its partial autocorrelation function (PACF). Both the ACF and PACF plots show a significant autocorrelation for lag 2.

Figure 4: $d\beta_{1t}$ Plot, ACF, and PACF

Afterwards, the possible ARIMA models are explored and summarized in Table 5. Using AIC as the criteria, the results show that the best mean model is ARIMA(2,1,2) with an AIC of 1281.64. The BIC and HQIC also supports this conclusion.

Model	AIC	BIC	HQIC
ARIMA(2, 1, 2)	1281.64	1300.72	1289.25
ARIMA(2, 2, 2)	1298.99	1318.06	1306.59
ARIMA(2, 1, 1)	1301.25	1316.52	1307.34
ARIMA(1, 1, 2)	1305.80	1321.07	1311.89
ARIMA(2, 2, 1)	1307.97	1323.23	1314.05

Table 5: ARIMA Models for β_{1t}

After fitting the mean model, the residuals then are studied. An ARCH LM test is performed on the model residuals. The results show a critical value of 128.906283 with a p-value of 0.00. Hence, this indicates the presence of an ARCH effect. The residuals are then fitted into several EGARCH models. The ϵ_t distribution considered includes normal, Student's t, and the generalized error distribution (GED). Table 6 shows the possible models and their summary. Using AIC as the criteria, the results show that the best model is ARIMA(2,1,2)-EGARCH(1,1) with an AIC of 820.18. The BIC also supports this conclusion.

Model	ϵ_t distribution	AIC	BIC
ARIMA(2, 1, 2) - EGARCH(1, 1)	Student-t	820.18	839.28
ARIMA(2, 1, 2) - EGARCH(2, 1)	Student-t	821.97	844.89
ARIMA(2, 1, 2) - EGARCH(1, 2)	Student-t	822.02	844.94
ARIMA(2, 1, 2) - EGARCH(2, 2)	Student-t	823.62	850.36
ARIMA(2, 1, 2) - EGARCH(1, 1)	GED	841.31	860.41

Table 6: ARIMA-EGARCH Model for β_{1t}

Hence, the time series model for β_{1t} is ARIMA(2,1,2)-EGARCH(1,1).

4.3.3 β_{2t} Factor

Finally, the optimal ARIMA-EGARCH model for β_{2t} is determined. A KPSS test is first performed on $\{\beta_{2t}\}$ with $\alpha = 0.05$. The results show a critical value of 1.024366 with a p-value of 0.00. Hence, the null hypothesis is rejected and the series is non-stationary. Thus, the differenced series $\{d\beta_{2t}\}$ is considered. Figure 5 shows the time series plot of $\{d\beta_{2t}\}$, its autocorrelation function (ACF), and its partial autocorrelation function (PACF). Both the ACF and PACF plots show a significant autocorrelation for lag 2.

Figure 5: $d\beta_{2t}$ Plot, ACF, and PACF

Afterwards, the possible ARIMA models are explored and summarized in Table 7. Using AIC as the criteria, the results show that the best mean model is ARIMA(0,1,2) with an AIC of 1717.23. The BIC and HQIC also supports this conclusion.

Model	AIC	BIC	HQIC
ARIMA(0, 1, 2)	1717.23	1728.68	1721.79
ARIMA(2, 1, 0)	1717.70	1729.15	1722.26
ARIMA(1, 1, 2)	1717.74	1733.00	1723.82
ARIMA(2, 2, 1)	1718.14	1733.39	1724.22
ARIMA(2, 1, 1)	1719.43	1734.70	1725.52

Table 7: ARIMA Models for β_{2t}

After fitting the mean model, the residuals then are studied. An ARCH LM test is performed on the model residuals. The results show a critical value of 52.109060 with a p-value of 0.00. Hence, this indicates the presence of an ARCH effect. The residuals are then fitted into several EGARCH models. The ϵ_t distribution considered includes normal, Student's t, and the generalized error distribution (GED). Table 8 shows the possible models and their summary. Using AIC as the criteria, the results show that the best model is ARIMA(0,1,2)-EGARCH(1,1) with an AIC of 1330.96. The BIC also supports this conclusion.

Model	ϵ_t distribution	AIC	BIC
ARIMA(0, 1, 2) - EGARCH(1, 1)	Student-t	1330.96	1350.06
ARIMA(0, 1, 2) - EGARCH(1, 1)	GED	1331.02	1350.12
ARIMA(0, 1, 2) - EGARCH(2, 1)	Student-t	1332.55	1355.47
ARIMA(0, 1, 2) - EGARCH(2, 1)	GED	1332.78	1355.71
ARIMA(0, 1, 2) - EGARCH(1, 2)	Student-t	1332.96	1355.88

Table 8: ARIMA-EGARCH Model for β_{2t}

Hence, the time series model for β_{2t} is ARIMA(0,1,2)-EGARCH(1,1).

4.4 Forecasting the Yield Curve

After obtaining the time series models for β_{0t}, β_{1t}, and β_{2t} in the previous subsection, the ℓ-step ahead forecasts are determined, $\ell = 1, 2, \ldots, 53$. These are obtained using the built-in `forecast()` function of the ARIMA and EGARCH models. Using the ℓ-step ahead forecasts $\beta_{0(t+\ell)}$, $\beta_{1(t+\ell)}$, and $\beta_{2(t+\ell)}$, the yield curve in ℓ periods is calculated as

$$y_{(t+\ell)}(\tau) = \beta_{0(t+\ell)} + \beta_{1(t+\ell)}\left(\frac{1-e^{-\tau/\lambda}}{\tau/\lambda}\right) + \beta_{2(t+\ell)}\left(\frac{1-e^{-\tau/\lambda}}{\tau/\lambda} - e^{-\tau/\lambda}\right),$$

where $\lambda = 5.434$. Figure 6 shows the forecasted weekly yield curve for one year while Figure 7 shows the actual weekly yield curve for the one year period from June 29, 2016 to June 30, 2016. Figures 15 and 16 in the Appendix shows the corresponding values. An initial look shows that the model captures the general upward trend of the actual yield curve. However, the forecasted yield curve is "smooth" while the actual yield curve is more irregularly shaped. This is expected since the Dynamic Nelson-Siegel model can only model "smooth" yield curve shapes.

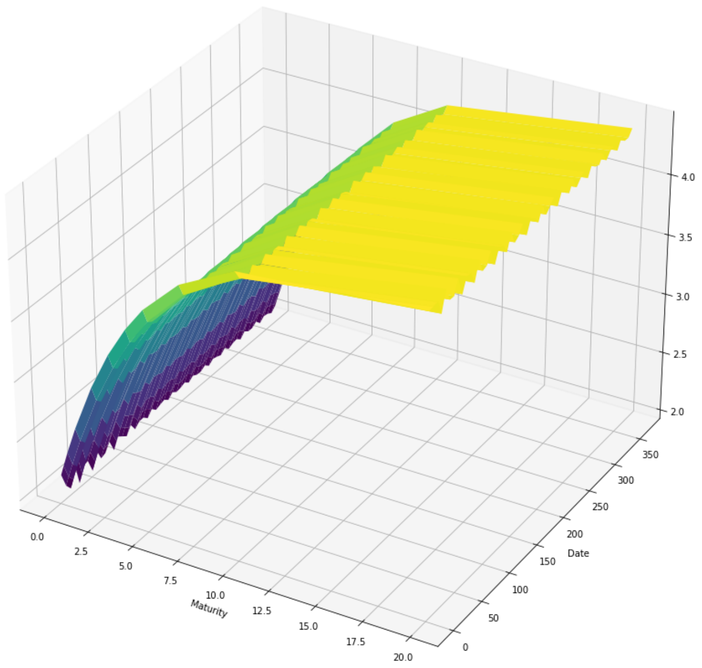

Figure 6: Forecasted Yield Curve

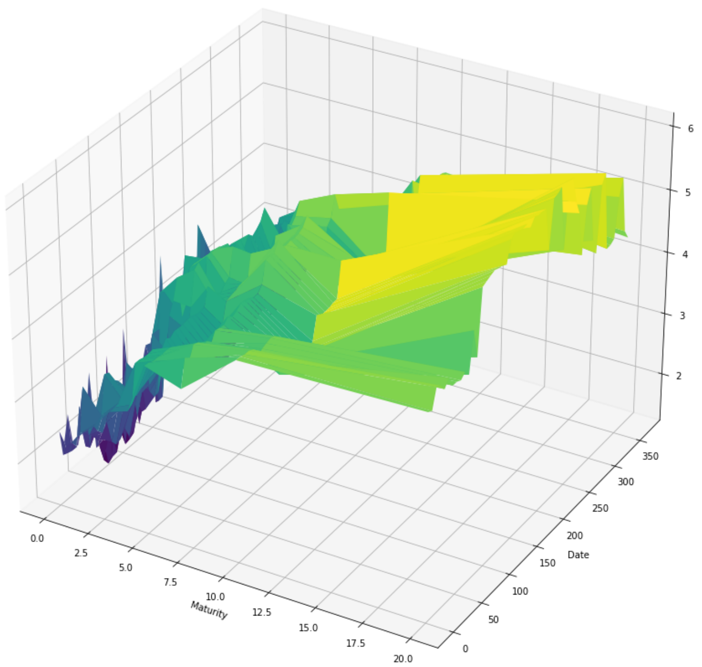

Figure 7: Actual Yield Curve

A closer look at the forecasted yield curve and actual yield curve shows that the model performs well in forecasting near-term yield curves. Figures 8 and 9 show the 1-step and 2-step ahead forecast of the yield curve. Visually, these figures show a good fit to the actual yield curve. However, Figures 10 and 11 show that the model performs poorly in the 52-step and 53-step ahead forecast of the yield curve.

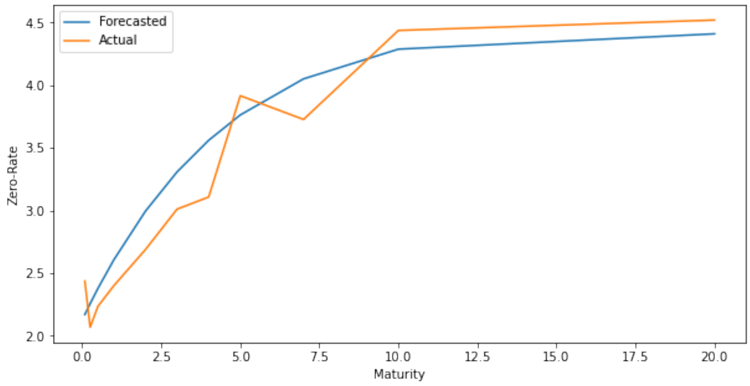

Figure 8: 1-Step Ahead Forecast

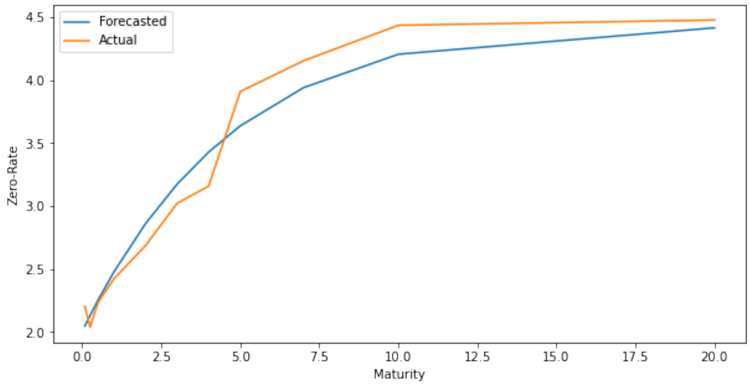

Figure 9: 2-Step Ahead Forecast

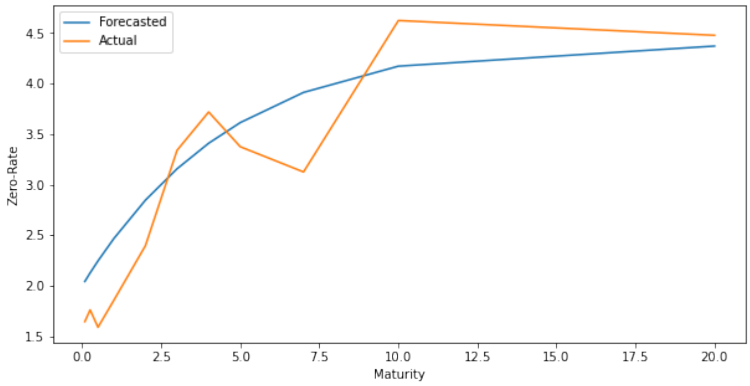

Figure 10: 52-Step Ahead Forecast

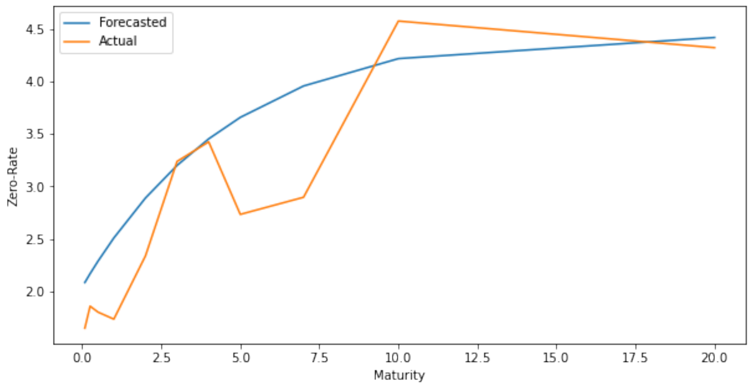

Figure 11: 53-Step Ahead Forecast

Finally, the performance of the model is evaluated using the root mean squared error (RMSE). The results show that the overall RMSE is 0.593478%, suggesting only a moderate overall performance. When broken down per tenor, Table 9 also supports this claim. This indicates that across all weeks, the RMSE is at most 45% of the (1M) interest rate which is quite large.

Tenor	RMSE
1M	0.625646%
3M	0.509131%
6M	0.495121%
1Y	0.602862%
2Y	0.549034%
3Y	0.582999%
4Y	0.336898%
5Y	0.391189%
7Y	0.539529%
10Y	0.401889%
20Y	1.120126%

Table 9: RMSE per Tenor

However, when broken down per week, Table 10 shows that near-term forecasts are generally better with a lower RMSE compared to the longer-term forecasts. The RMSE for the first 10 weeks after the in-sample period (until 2015-09-02) are all below 0.4000%. However, the RMSE after the 10th week (after 2015-09-02) are almost all above 0.4000%. This translates to a near-term RMSE with a maximum of just 25% of the actual (1M) interest rate which is relatively good. Overall, the results indicate that the Dynamic Nelson-Siegel model can reliably forecast the yield curves for the near future.

Week	RMSE
2015-07-01	0.255678%
2015-07-08	0.175686%
2015-07-15	0.340064%
2015-07-22	0.304211%
2015-07-29	0.245809%
2015-08-05	0.238645%
2015-08-12	0.260857%
2015-08-19	0.354889%
2015-08-26	0.368384%
2015-09-02	0.384304%
2015-09-09	0.447147%
2015-09-16	0.554409%
2015-09-23	0.405689%
2015-09-30	0.510510%
2015-10-07	0.613098%

Table 10: RMSE per Week

5 Conclusion

This paper forecasted the term structure of the PDST-R2 rates using the Dynamic Nelson-Siegel model. The zero rates are first obtained using the bootstrapping method which produced 390 weeks of data from January 14, 2009 to June 29, 2016. Afterwards, using the fixed lambda method (de Lara-Tuprio et al., 2017), the optimal shape parameter is estimated to be $\lambda_t = 5.434$ with an average R^2 of 0.964466. Next, the time series $\{\beta_{0t}\}$, $\{\beta_{1t}\}$, $\{\beta_{2t}\}$ corresponding to $\lambda_t = 5.434$ are calculated using multiple linear regressions. For each beta factor, the optimal ARIMA-EGARCH model is determined using AIC as the criteria. The optimal model for β_{0t} and β_{1t} is ARIMA(2,1,2)-EGARCH(1,1) while the optimal model for β_{2t} is ARIMA(0,1,2)-EGARCH(1,1). Finally, the ℓ-step ahead forecasts of the yield curve are determined following the corresponding time series models for the betas and the Nelson-Siegel model.

The model is observed to perform only moderately well with an overall RMSE of 0.5935%. The results show that the model is able to capture the general upward trend of the actual yield curve. Furthermore, the model also performs well in forecasting near-term yield curves, particularly in the 1-step and 2-step ahead forecasts with an RMSE only 10% of the actual (1M) interest rate. Overall, the Dynamic Nelson-Siegel model is accurate and reliable in predicting the yield curve for the near future. This is sufficient for an active market since the model can be regularly updated as new information arrive.

Overall, the three-factor Nelson-Siegel model is a widely used model since it can accurately and reliably forecast the yield curve of interest rates. While the results suggest a good predictive performance for the near future, the model performance can still be improved by finding the joint distribution of the beta residuals via the appropriate copula.

References

de Lara-Tuprio, E. P., Bataller, R. T., Torres, A. D. D., Cabral, E. A., & Fernandez Jr, P. L. (2017). Forecasting the term structure of philippine interest rates using the dynamic nelson-siegel model. *DLSU Business & Economics Review*, *27*(1), 1–1. Retrieved 2022-05-24, from https://ejournals.ph/article.php?id=11592

Diebold, F. X., & Li, C. (2006, February). Forecasting the term structure of government bond yields. *Journal of Econometrics*, *130*(2), 337–364. Retrieved 2022-05-24, from https://linkinghub.elsevier.com/retrieve/pii/S0304407605000795 doi: 10.1016/j.jeconom.2005.03.005

Kumar, R. R., Stauvermann, P. J., & Vu, H. T. T. (2021, February). The relationship between yield curve and economic activity: an analysis of g7 countries. *Journal of Risk and Financial Management*, *14*(2), 62. Retrieved 2022-05-24, from https://www.mdpi.com/1911-8074/14/2/62 doi: 10.3390/jrfm14020062

Nelson, C. R., & Siegel, A. F. (1987, January). Parsimonious modeling of yield curves. *The Journal of Business*, *60*(4), 473. Retrieved 2022-05-24, from https://www.jstor.org/stable/2352957 doi: 10.1086/296409

Appendix

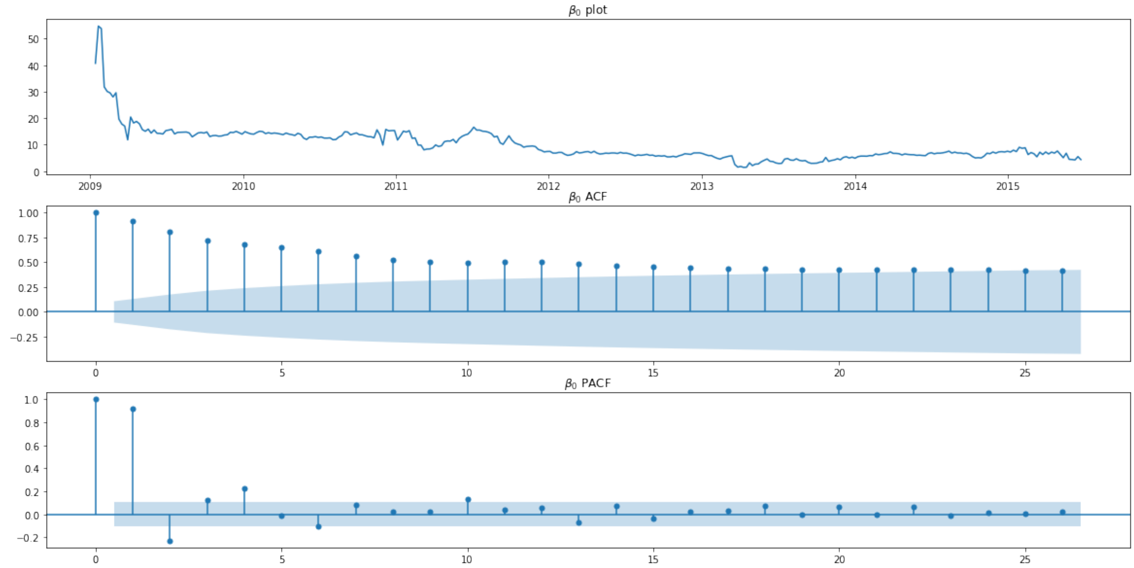

Figure 12: β_{0t} Plot, ACF, and PACF

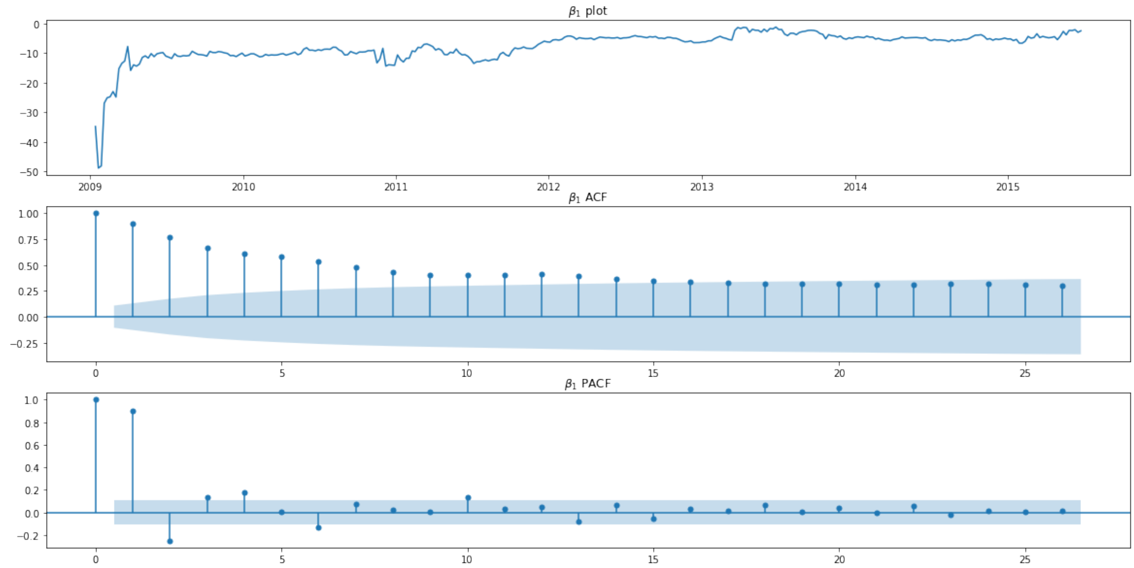

Figure 13: β_{1t} Plot, ACF, and PACF

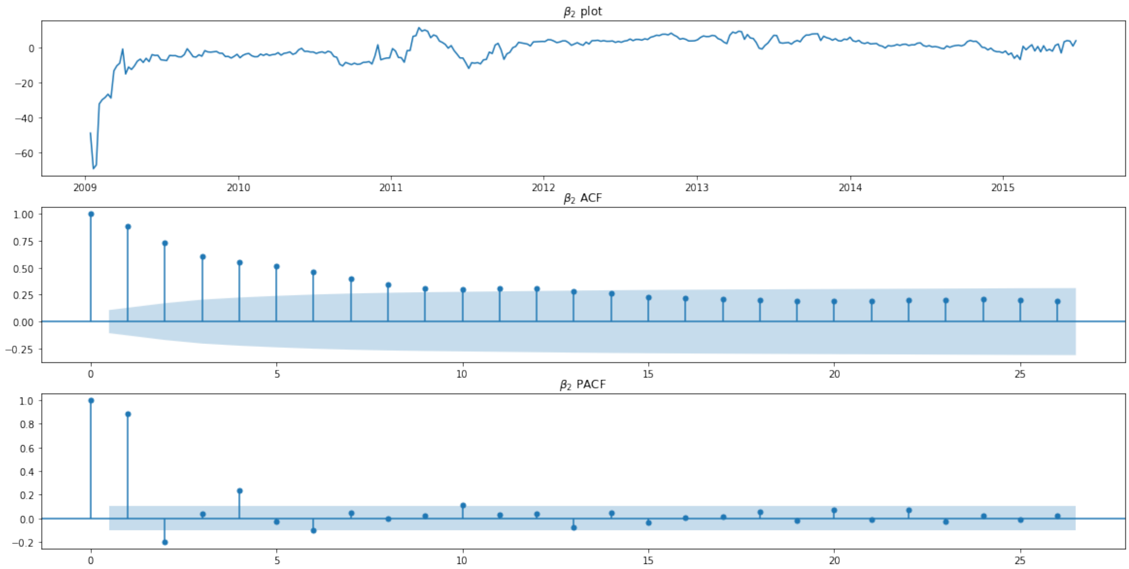

Figure 14: β_{2t} Plot, ACF, and PACF

Date	1M	3M	6M	1Y	2Y	3Y	4Y	5Y	7Y	10Y	20Y
7/1/2015	2.1706	2.2564	2.3803	2.6069	2.9952	3.3088	3.5609	3.7634	4.0513	4.2884	4.4108
7/8/2015	2.0475	2.1314	2.2528	2.4756	2.8601	3.1737	3.4290	3.6369	3.9401	4.2055	4.4150
7/15/2015	1.9915	2.0768	2.2002	2.4267	2.8180	3.1377	3.3985	3.6111	3.9224	4.1971	4.4235
7/22/2015	2.1714	2.2505	2.3647	2.5738	2.9333	3.2249	3.4605	3.6509	3.9246	4.1561	4.3041
7/29/2015	1.9741	2.0606	2.1857	2.4156	2.8131	3.1384	3.4039	3.6210	3.9397	4.2227	4.4644
8/5/2015	2.0886	2.1705	2.2888	2.5056	2.8793	3.1834	3.4303	3.6305	3.9210	4.1718	4.3547
8/12/2015	2.1018	2.1833	2.3011	2.5171	2.8893	3.1920	3.4376	3.6367	3.9253	4.1740	4.3530
8/19/2015	1.9857	2.0719	2.1967	2.4258	2.8219	3.1458	3.4103	3.6263	3.9433	4.2244	4.4625
8/26/2015	2.1330	2.2129	2.3285	2.5401	2.9043	3.2000	3.4393	3.6329	3.9121	4.1500	4.3094
9/2/2015	2.0293	2.1137	2.2346	2.4611	2.8478	3.1634	3.4205	3.6299	3.9357	4.2041	4.4191

Figure 15: Forecasted Yield Curve

Date	1M	3M	6M	1Y	2Y	3Y	4Y	5Y	7Y	10Y	20Y
7/1/2015	2.4374	2.0695	2.2368	2.4006	2.6889	3.0115	3.1083	3.9165	3.7275	4.4378	4.5207
7/8/2015	2.2029	2.0385	2.2338	2.4217	2.6850	3.0210	3.1586	3.9091	4.1537	4.4353	4.4768
7/15/2015	2.5962	2.0167	2.1135	3.1232	2.7432	3.0326	3.7567	3.8760	4.0680	4.4122	4.7809
7/22/2015	1.9988	2.0148	2.7045	2.2735	3.3506	3.0139	3.7565	3.8803	4.2498	4.4128	4.7473
7/29/2015	1.9724	2.3711	2.2124	2.3618	3.3175	3.0355	3.7324	3.8617	4.2256	4.3731	4.6360
8/5/2015	2.2445	1.9095	2.1191	2.2259	3.1371	3.0249	3.7344	3.8754	4.0912	4.1889	4.7468
8/12/2015	2.3088	1.8262	2.0257	2.1358	3.0696	2.9148	3.7453	3.8655	4.1307	4.2712	4.5742
8/19/2015	2.9811	2.2800	1.9003	2.2523	2.5923	2.9673	3.6328	3.3586	4.0425	4.2999	4.5724
8/26/2015	1.6871	1.7577	1.8611	1.8698	3.2490	2.9899	3.6176	3.3781	4.1895	4.3247	4.5370
9/2/2015	1.6805	1.4397	1.7803	1.7638	3.1088	3.0281	3.7469	3.3779	4.1717	4.2994	4.6051

Figure 16: Actual Yield Curve

```python
def maturity(date, years, years_prev=0):
    return ( (date + relativedelta(months=int(12*years)))
    - (date + relativedelta(months=int(12*years_prev))) ).days / 360
```

Listing 1: Maturity Python Code

```python
def bootstrapping(df, dt=0.5):
    interp = {}; discount = {}; zero_rate = {}
    date = df['Date']

    # Input given tenors
    for t in df.index.drop('Date'):
        interp[to_year(t)] = df[t]
    N = max(interp.keys()) + dt

    for t in np.arange(dt,N,dt):
        # Interpolated Rates
        if t not in interp:
            for T in np.arange(t,N,dt):
                if T in interp: break
            interp[t] = np.interp(maturity(date, t), [maturity(date, t-dt), maturity(date, T)], [interp[t-dt], interp[T]])

        # Discount Factor
        if t <= 1:
            discount[t] = 1 / (1 + interp[t] / 100 * maturity(date, t))
        else:
            discount[t] = 1
            for tt in np.arange(dt,t,dt):
                discount[t] -= interp[t] / 100 * discount[tt] * maturity(date, tt, tt-dt)
            discount[t] /= (1 + interp[t] / 100 * maturity(date, t, t-dt))

        # Zero-Rate
        zero_rate[t] = - 100 * np.log(discount[t]) / maturity(date, t)

        # Modify entry
        if to_string(t) in df.index: df[to_string(t)] = zero_rate[t]

    # Convert the rest of the tenors to continuous
    for t in df.index.drop('Date'):
        if to_year(t) < dt:
            df[t] = 100 * np.log(1 + df[t] / 100 * maturity(date, to_year(t))) / maturity(date, to_year(t))

    return df
```

Listing 2: Bootstrapping Python Code