Modelling of Short Rate ($\beta_2$) Parameter Diebold-Li Model Using Vasicek Stochastic Differential Equations

Meylita Sari, Endah R.M Putri, and Nuri Wahyuningsih
Department of Mathematics, Institut Teknologi Sepuluh Nopember, Surabaya, 60111
E-mail: endahrmp@matematika.its.ac.id

Abstract. Investment is the activity of investing or allocating money to earn some profit. There are several type of investments, one of which is bond. The interest rate of the bond is called a yield. A yield is considered as a representation of market expectations depends on interest rate movements according to market price at a certain time. The yield’s value will change following a stochastic process. Estimation of a short rate parameter ($\beta_2$) of a Diebold-Li model is conducted by applying a least square method. Then, the yield modeling is developed based on the estimated short rate parameter Diebold-li model ($\beta_2$) using Vasicek Stochastic Differential Equation. The final result is that modelling of short rate ($\beta_2$) parameter Diebold-Li model using Vasicek Stochastic Differential Equation generates prediction values that have a high level of accuracy by MAPE prediction value of 8.65%.

1. Introduction
Investing has become one of the people’s choices in doing business. Investment is an activity carried out by investors, both foreign and domestic in various business fields to obtain profits. There are various types of investment instruments, one type of investment is to buy financial products in the form of securities, namely bonds. The interesting benefit about bond investment is to provide fixed income in the form of coupons, and this is the main characteristic of bonds where bondholders will get a regular interest income until the term of the bond is mature. In addition, one of important factors that investors considers before investing in bonds is the yield [1].

There are various models used to model yield. One of the models used is a Diebold-Li model which has parameters ($\beta_1, \beta_2, \beta_3$) to calculate long term, short term, and medium term [2]. The short term factor ($\beta_2$) has negative values which leads to the use of a Vasicek stochastic differential equation. Therefore, this study discusses the short rate ($\beta_2$) parameter modeling on the Diebold-Li model using the stochastic differential equation of the Vasicek model. And at the same time to analyze the results of the short rate parameter modeling ($\beta_2$) in the Diebold Li model which then becomes the output of this research.

2. Model and Preliminaries
In this section, some theoretical backgrounds used in this study are discussed to give an insight to the object of study. The theoretical backgrounds includes are Diebold-Li model,
return of yield, normality test, a stochastic processes, a Vasicek stochastic differential equation, parameter estimations, MAPE, dan confident interval estimations.

2.1. Diebold-Li Model

In this section, the yield model used is Diebold-Li model [3]. This model is an extension of Nelson-Siegel model and can be considered as a dynamic version of it. The model describes a yield bond dynamic with a maturity $\tau$ and three parameters ($\beta_1, \beta_2, \beta_3$) as the long term, short term, and medium term. The first parameter is constant for different maturities, the second parameters has more impact on short maturities, and the third parameters will result to a peak and a decay to zero as the maturities increase. The decay factor $\lambda$ is considered as a non-linear factor in the Diebold-Li model. The model is written as follows:

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

where

- $y_t(\tau)$ : yield at the time of $t$ with maturity $\tau$
- $\beta_{1t}, \beta_{2t}, \beta_{3t}$: parameter of long term, short term, medium term at time $t$.
- $\lambda$ : decay exponential rate
- $\tau$ : time of maturity yield

Parameter $\beta_{2t}$ will be modelled as a Vasicek differential equation considering the fact that the values based on the market data is negative and stochastic.

2.2. Yield Return

In order to apply the Vasicek stochastic differential equation, it is necessary to calculate the return of yield obtained from the bond. To calculate return of yield bond, the formula is as follows:

$$R_t = \left(\frac{r(t+1) - r(t)}{r(t)}\right)$$

where

- $R_t$ : return yield of bond
- $r(t+1)$ : yield at $t + 1$ time
- $r(t)$ : yield at $t$ time

2.3. Normality test

As the Vasicek stochastic differential equation follows a Wiener process, a normality condition of the yield return should be satisfied. The normality test used is written as follows:

- Hypothesis :
  - $H_0 : F(x) = F_0(X)$ To sample data normally distribution.
  - $H_1 : F(x) \neq F_0(X)$ To sample data do not normally distribution.

- Statistics Test :
  $$D_{calculate} = max|F_0(X) - S_N(X)|$$

where
\[ \text{max}|F_0(X) - S_N(X)| : \text{maximum value to } F_0(X) \text{ minus } S_N(X) \]

\[ D_{\text{Calculate}} \text{ maximum deviation} \]

\[ F_0(X) : \text{distribution functions that are hypothesized are normally distribution} \]

\[ S_N(X) : \text{cumulative distribution functions of sample data.} \]

### 2.4. A Vasicek Stochastic Differential Equation

The Vasicek is a method of modeling interest rate movement that describes the movement of an interest rate as a factor of market risk, time and equilibrium value that the rate tends to revert towards. Essentially, it predicts where interest rates will end up at the end of a given period of time given current market volatility, the long-run mean interest rate value, and a given market risk factor. It is important to note that the equation can only test one market risk factor at a time. This stochastic model is often used in the valuation of interest rate futures.

\[
\text{dr}(t) = \eta(\theta - r(t))dt + \sigma dZ
\]

where :

\[ \text{dr}(t) : \text{interest rate change.} \]

\[ \eta : \text{parameter of mean reversion} \]

\[ \theta : \text{parameter of drift.} \]

\[ \sigma : \text{parameter of volatilitas} \]

\[ Z : \text{Wiener Process or Brownian Motion} \]

The drift part of Vasicek model may consist of negative value in a mean reversion environment.

### 2.5. Parameter Estimation

The Vasicek stochastic differential equation has some key parameters which are required to be estimated based on the market data. One of most common estimation method is Ordinary Least Square methods based on a linear regression (simple or multiple depending on the number of explanatory variables). The ordinary least square method corresponds to minimizing the sum of square differences between the observed and predicted values. The model is written as follows:

\[
r_t = \gamma_0 + \gamma_1 r_{t-1} + \epsilon_t, t = 1, 2, ..., n
\]

where :

\[ \gamma_0 : \text{regression coefficient} \]

\[ \gamma_1 : \text{regression coefficient} \]

\[ \epsilon_t : \text{error value} \]

\[ r_t : \text{dependent variable} \]

\[ r_{t-1} : \text{independent variable}. \]

It assumed expectation value of \( \epsilon_t \) is 0. \( E[\epsilon_t] = 0 \). And Estimator of ordinary least square to koefisien \( \gamma \) in this equation is:

\[
\hat{\gamma}_1 = \frac{\sum_{t=2}^{n} Y_{t-1} Y_t}{\sum_{t=2}^{n} Y_{t-1}^2}
\]
2.6. MAPE

In this section, Mean Absolute Percentage Error (MAPE) is used to know the accuracy of prediction result of the modelling of short rate ($\beta_2$) parameter Diebold-Li. MAPE

$$MAPE = \frac{\sum_{k=1}^{N} \left| \frac{r_t - P_t}{r_t} \right| \times 100}{N},$$

where:

- $r_t$: actual short rate $\beta_2$ parameter value at $t$ times
- $F_t$: prediction short rate $\beta_2$ parameter value at $t + 1$ times
- $N$: sum of parameter short rate $\beta_2$ prediction data.

3. Result and Discussion

In this section, implementation of the Vasicek stochastic differential equation to model parameter $\beta_2$ using yield data is conducted. The steps in implementing the model are started by parameter estimation using Ordinary Least Square method. Subsequently, simulations of modelling the short-rate ($\beta_2$) parameter model Diebold-Li with MATLAB are applied.

3.1. Implementation of yield SBN Bonds Data on Diebold-Li Models

The bond data is obtained based on the yield data of Surat Berharga Negara Republik Indonesia (SBN-RI) available on website of Ministry of Finance Republik Indonesia. The implementation yield SBN Bonds Data on Diebold-Li model is to get $\beta_2$ value. Based on equation (1), so the equation can be formed in the matrix form for some maturity $\tau$ values as follows,

$$\begin{bmatrix}
y_t(\tau_1) \\
y_t(\tau_2) \\
\vdots \\
y_t(\tau_n)
\end{bmatrix} = 
\begin{bmatrix}
1 & 1 - e^{-\lambda_1 \tau_1} & 1 - e^{-\lambda_1 \tau_1} \\
1 & 1 - e^{-\lambda_2 \tau_2} & 1 - e^{-\lambda_2 \tau_2} \\
\vdots & \vdots & \vdots \\
1 & 1 - e^{-\lambda_n \tau_n} & 1 - e^{-\lambda_n \tau_n}
\end{bmatrix} 
\begin{bmatrix}
\beta_{1t} \\
\beta_{2t} \\
\beta_{3t}
\end{bmatrix}$$

The equation is used to calculate the yield on certain time $t$ for some maturity values $\tau$. To get $\beta_t$ value using the ordinary least square method is given as follows:

$$\hat{\beta}_t = (Z^T Z)^{-1} Z^T y_t(\tau_n)$$

(7)

Based on the data collected from Ministry of Finance’s website, there are 99 yield data and using Diebold-Li equation , $\beta_2$ values are obtained for each maturity. Subsequently, $\beta_2$ data is divided into 2 parts, 93 in-sample-data and 6 out-sample-data. In-sample-data is used to get parameter estimation of Diebold-Li model and the out-sample-data is used to validate the model.

3.2. Return of $\beta_2$ and its normality test

Subsequently, based on $\beta_2$ data obtained in previous subsection, the return is calculated using Equation (2). A normality test is applied on $\beta_2$ data and the result can be seen in Figure 1-2.
Figure 1. Histogram of normality test

Figure 2. Q-Q Plot of $\beta_2$

Figure 1 shows that the distribution of return of $\beta_2$ data in histogram has followed the normal distribution. Moreover a Q-Q plot of the return data is shown in Figure 2 supports the evidence that the return of $\beta_2$ follows normal distribution significantly in statistics by the p-value greater that the level of significance 5%. Now, the $\beta_2$ data is fitted to be implemented in Vasicek stochastic differential equation.

3.3. Solution of Vasicek Stochastic Differential Equations

In the section, the solution of Vasicek stochastic differential equation is obtained based on Equation (4). First, the equation is defined as follows,

$$Y(t) = f(t, r(t)) = e^{\eta t}r(t)$$

(8)

Because $r(t)$ is a stochastic process, so the function $f(t, r(t))$ that contains $r(t)$ is also a stochastic process. For a small change of $Y(t)$ value with respect to small change of time $t$, the change can be expressed as follows:

$$dY(t) = d(e^{\eta t}r(t))dt$$

$$= \eta e^{\eta t}r(t)dt + e^{\eta t}dr(t)$$

$$\Rightarrow d(e^{\eta t}r(t)) = \eta r(t)dt + e^{\eta t}(\eta(\theta - r(t))dt + \sigma dZ)$$

$$d(e^{\eta t}r(t)) = \eta \theta e^{\eta t}dt + \sigma e^{\eta t}dZ$$

(9)

Integration to Equation (9) from $u$ to $t$, $u \leq t$ results in:

$$\int_u^t d(e^{\eta t}r(t)) = \int_u^t \eta \theta e^{\eta t}dt + \int_u^t \sigma e^{\eta t}dZ$$

(10)
and by using a dummy variable to simplify the integration process, Equation (10) can be written as follows,

\[
\int_u^t dY(s) = \int_u^t \eta \theta e^{\eta s} ds + \int_u^t \sigma e^{\eta s} dZ_s
\]

\[
Y(t) = Y(u) + \int_u^t \eta \theta e^{\eta s} ds + \int_u^t \sigma e^{\eta s} dZ_s \tag{11}
\]

Then the two sections in Equation (11) are multiplied by \(e^{-\eta t}\). Because \(Y(t) = r(t)e^{\eta t}\) then \(r(t) = Y(t)e^{-\eta t}\). So Equation (11) becomes:

\[
r(t) = r(u)e^{-\eta(t-u)} + \frac{\eta \theta}{\eta} e^{-\eta t} \left[ e^{\eta(t-u)} \right] + e^{-\eta t} \int_u^t \sigma e^{\eta s} dZ_s \tag{12}
\]

Equation (13) is the final form of Equation (12). The solution of \(r(t)\) is in the form of the value of one step of \(r(u)\). Given \(\Delta t = t - u\), \(r(t) = r_t\) and \(r(u) = r_{t-1}\), then equation (13) can be written as follows:

\[
r_t = r_{t-1}e^{-\eta \Delta t} + \theta(1 - e^{-\eta \Delta t}) + \int_u^t e^{-\eta(t-s)} \sigma dZ_s \tag{13}
\]

After the solution from Vasicek’s stochastic differential equation is obtained, then the mean and variance can be found. So, the expected value and variance of Vasicek stochastic differential equation are written as follows: The expectation value of \(r_t\) is:

\[
E(r_t) = r_{t-1}e^{-\eta \Delta t} + \theta(1 - e^{-\eta \Delta t}) \tag{14}
\]

The variance value of \(r_t\) is:

\[
Var \ (r_t) = \frac{\sigma^2}{2\eta} (1 - e^{-2\eta(t-u)}) \tag{15}
\]

3.4. Estimation of the parameter Vasicek Stochastic Differential Equation

In this section, an ordinary least square to estimate the parameter \(\eta, \theta\) and \(\sigma\) in the following Vasicek stochastic differential equation is used.

\[
r_t = r_{t-1}e^{-\eta \Delta t} + \theta(1 - e^{-\eta \Delta t}) + \int_u^t e^{-\eta(t-s)} \sigma dZ_s
\]

Equation (13) can be expressed as an AR order 1 model and the equation is written as follows:

\[
r_t = \gamma_0 + \gamma_1 r_{t-1} + \epsilon_t \tag{16}
\]
so that the estimator $\eta$ parameter is obtained as follows
\[
\hat{\gamma}_1 = e^{(-\eta \Delta t)} \\
\hat{\eta} = \frac{-1}{\Delta t} \ln \left[ \frac{n \sum_{i=1}^{n} r_t r_{t-1} - \sum_{i=1}^{n} r_t \sum_{i=1}^{n} r_{t-1}}{n \sum_{i=1}^{n} (r_{t-1})^2 - \sum_{i=1}^{n} r_{t-1}\sum_{i=1}^{n} r_{t-1}} \right]
\] (17)

and estimator $\theta$ parameter is obtained as follows
\[
\hat{\gamma}_0 = \hat{\theta}(1 - e^{-\eta \Delta t}) \\
\hat{\theta} = \frac{\hat{\gamma}_0}{(1 - e^{-\eta \Delta t})} \\
\hat{\theta} = \frac{\sum_{i=1}^{n} r_t - e^{(-\hat{\eta} \Delta t)} \sum_{i=1}^{n} r_{t-1}}{n(1 - e^{-\eta \Delta t})}
\] (18)

Last, the estimator $\sigma$ parameter is obtained as follows,
\[
\text{st.d}(\epsilon) = \hat{\sigma} \sqrt{\frac{1 - e^{-2\eta \Delta t}}{2\eta}} \\
\hat{\sigma} = \text{st.d}(\epsilon) \sqrt{\frac{-2 \ln \hat{\gamma}_1}{\Delta t(1 - \hat{\gamma}_1^2)}}
\] (19)

All the estimators for Vasicek model are obtained using an ordinary least square method explicitly. The estimators then are applied to the data to get the values.

3.5. Analysis modeling of short rate ($\beta_2$) parameter model Diebold-Li

1. Calculate Parameter value of Vasicek stochastic differential equation.

The results of estimating Vasicek stochastic differential equation parameters using ordinary least square based Equation (17), (18), (19) are in Table 1.

| $\eta$ | $\theta$ | $\sigma$ |
|---|---|---|
| 0.1571 | -2.7639 | 0.4562 |

2. Form the modeling of short rate ($\beta_2$) parameter model Diebold-Li

Based on Equation (13), the model of a short rate ($\beta_2$) parameter model Diebold-Li is obtained. Given $r_t = F_t$ and $r_{t-1} = F_{t-1}$, so the equation can be written as follows:
\[
F_t = F_{t-1} e^{-\eta(t-u)} + \theta(1 - e^{-\eta(t-u)}) + \sigma e^{-\eta t} \int_{t}^{\infty} e^{\eta s} dZ_s
\] (20)

Then, the solution of stochastic integration $\int_{t}^{\infty} e^{-\eta(t-s)} \sigma dZ_s$ is obtained using a Riemann sum approach. So, it can be written as follows:
\[
I(t) = \int_{0}^{t} e^{\eta s} dZ_s = \sum_{j=0}^{k+1} e^{\eta s_j} |Z(s_{j+1}) - Z(s_j)|,
\] (21)
Replacing $I(t) = \int_{u}^{t} e^{\eta s} d(Z_s)$ in Equation (20), we get modeling of short rate ($\beta_2$) parameter model Diebold-Li and the equation can be written as follows:

$$F_t = F_{t-1} e^{-\eta(t-u)} + \theta(1 - e^{-\eta(t-u)}) + \sigma e^{-\eta t} \sum_{s=u}^{t} e^{\eta s} \Delta Z.$$  \hspace{1cm} (22)

where

$F_{t-1}$ : parameter short rate $\beta_2$ model Diebold-Li at $t - 1$ time
$F_t$ : prediction of short rate $\beta_2$ parameter Diebold-Li model value at time $t$
$\eta$ : rate of reversion
$\theta$ : drift
$\sigma$ : volatility
$Z$ : Wiener processes.

3. Validation the modeling of short rate ($\beta_2$) parameter model Diebold-Li

In this section after we get the form modeling of short rate ($\beta_2$) parameter model Diebold-Li. Then, with Equation (22) and $\eta$, $\theta$, $\sigma$ value in Table 1 which has been obtained in the previous section and the modelling of short rate $\beta_2$ parameter on Diebold-Li model as follows:

$$F_t = F_{t-1} e^{-0.1571(t-u)} + (-2.7639)(1 - e^{-0.1571(t-u)}) + (0.4562)e^{-0.1571 t} \sum_{s=u}^{t} e^{0.1571 s} \Delta Z.$$  \hspace{1cm} (23)

Then, by using 6 out-sample data and Equation (23), validation is conducted using data from October 2017 until March 2018. The result of prediction to validate modelling of short rate $\beta_2$ parameter on Diebold-Li model using Vasicek stochastic differential equation is shown by Table 2:

| No | t  | Date       | Actual Data $\beta_2$ | Prediction result $\beta_2$ |
|----|----|------------|------------------------|-----------------------------|
| 1  | 94 | October 2017 | -2.24947               | -2.27252                    |
| 2  | 95 | November 2017| -2.33533               | -2.75145                    |
| 3  | 96 | December 2017| -2.54590               | -2.68071                    |
| 4  | 97 | January 2018 | -2.63515               | -2.69322                    |
| 5  | 98 | February 2018| -2.71215               | -3.22366                    |
| 6  | 99 | March 2018   | -2.63823               | -2.81616                    |
From the result prediction in Table 2, we can calculate the MAPE to know accuracy of the prediction (in %). Based on Equation (6) and result modelling on Table 2, we can get value of MAPE=8.6577%. Based on Table 2, the MAPE value is 8.6577 % means that the result of modelling short rate ($\beta_2$) parameter have a high accuracy prediction. Each realization of the path is 6 points is using Equation (23). Suppose if $i$ is any realization of a possible path from modeling.

![Figure 3. With $i = 10$ trajectories](image3)

![Figure 4. With $i = 100$ trajectories](image4)

![Figure 5. With $i = 1000$ trajectories](image5)

Iteration is conducted up to $i = 1000$. Each $i$ will form a green line that means realization of the prediction from the results of the $\beta_2$ parameter modeling on the Diebold-Li model that might occur. The result of realization with 10, 100 and 1000 trajectories can be seen in Figure 3-5.
3.6. Prediction of short rate ($\beta_2$) parameter model diebold-Li using Vasicek Stochastic Differential Equation

In this section based on Equation (23), we can get the prediction of short rate ($\beta_2$) parameter model Diebold-Li value on April until December 2018 which is shown on Table 3.

Table 3. Confident Interval of prediction short rate ($\beta_2$) parameter model Diebold-Li

| No | Date    | Lower  | Prediction | Upper  |
|----|---------|--------|------------|--------|
| 1  | April 2018 | -3.12080 | -2.63824 | -2.19221 |
| 2  | Mei 2018   | -3.12080 | -2.48687 | -2.19221 |
| 3  | Juni 2018  | -2.99144 | -2.53691 | -2.06285 |
| 4  | Juli 2018  | -3.03419 | -2.08842 | -2.10561 |
| 5  | Agst 2018  | -2.65091 | -2.63908 | -1.72233 |
| 6  | Sept 2018  | -3.12151 | -2.42795 | -2.19293 |
| 7  | Okt 2018   | -2.94108 | -2.51845 | -2.01249 |
| 8  | Nop 2018   | -3.01842 | -2.83355 | -2.08984 |
| 9  | Des 2018   | -3.28771 | -2.94063 | -2.35912 |

4. Conclusion

Based on the study conducted, the are two main concluding remarks can be drawn. Firstly, the method has been obtained to get the modelling short rate ($\beta_2$) parameter on the Diebold-Li model using Vasicek stochastic differential equation. The modelling formed to get the $\beta_2$ parameter value. Secondly, the results of analysis modeling short rate parameter on the Diebold-Li model using the Vasicek stochastic differential equation get the parameter value $\eta = 0.1571$, $\theta = -2.7639$, and $\sigma = 0.4562$. The modeling results in MAPE value of 8.65 % which means that the modeling results have a high prediction accuracy.

Bibliography

[1] Setiadi. (1996). Obligasi dalam Perspektif Hukum Indonesia. Jakarta. Citra Aditya Bakti.
[2] Diebold, F.X., Li. (2006). Forecasting the Term Structure of Government Bond Yields. Journal of Econometrics, 130, 337-364.
[3] Diebold, F.X., dkk. (2006). The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. Journal of Econometrics, 131, 309-338.
[4] Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of Financial Economics 5, 177-188.
[5] Willian, W.S.W. (1990). Time Series Analysis: Univariate and Multivariate Methods. Addison-Wesley Publishing Company. United States of America.
[6] Walpole, R.E.. (1995). Pengantar Statistika. Edisi ke-3, PT. Gramedia Pustaka Utama, Jakarta.
[7] Smith, W.. (2010), On The Simulation and Estimation of The Mean-Reverting Ornstein-Uhlenbeck Process, Verson1.01
[8] Higham, D.J. (2001). An Algorithmic Introduction to numerical Simulation of Stochastic Differential Equations”. Society for Industrial and Applied Mathematics Vol. 43, No. 3, pp. 525546.