Investment in Malaysia: Forecasting Fixed Deposit Using Time Series and Regression Analysis

Nuzlinda Abdul Rahman[1], Kevin Pang[1,2,3], Chen Yi Kit[1], Fauhatuz Zahroh Shaik Abdullah[1(✉)], Nur Sofiah Izani[2], and Wan Muhamad Amir W. Ahmad[3]

[1] School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia
fauhatuz@usm.my
[2] Gelugor, 11700 Pulau Pinang, Malaysia
[3] School of Dental Sciences, Universiti Sains Malaysia, 11800 Minden, Pulau Pinang, Malaysia

Abstract. This paper studies Malaysian banking fixed deposit (FD) rates from 1997 to 2018 using time series and regression analysis. The FD rates is based on rates set by Bank Negara Malaysia. Multiple Linear Regression (MLR) is used to study the linear relationship between FD Rates and certain economic and financial indicators. The findings suggest FD Rates is heavily affected by Base Lending Rate (BLR), Consumer Price Index (CPI) and Real Effective Exchange Rate (REER). However, autocorrelation occurs and ARDL method is employed by adding variable lags. The new model adds in the lag of FD Rates, BLR and REER to fulfill the independent assumptions. Subsequently, the time series behavior of the three variables is investigated using ARIMA model approach. Forecasts of three explanatory variables for the next three years is made in order to predict the next three years FD Rates using the regression equation. The computed rates are then converted to 12-month period FD Rates. Results indicates that fixed deposit give a low but consistent return.

Keywords: ARIMA · Regression

1 Introduction

A Fixed Deposit (FD) rates is a special type of bank savings account where a higher rate of interest is earned provided the deposit, a fixed amount, is not withdrawn over a fixed period. Typical periods are one month, three months, six months and a year. The interest is paid by the bank at the end of the stipulated period. FD are popular in Malaysia because it is very safe and can earn better returns than an ordinary savings account. FD rates are usually reference to a certain rates determined by Bank Negara Malaysia.

It would be highly beneficial if one can forecast the performance of investments in FD with a reasonable degree of accuracy. There are several time series modeling techniques to obtain reliable forecasts of investments. An important forecasting technique

N. Wahi et al. (Eds.): ICMSS 2022, ACSR 98, pp. 34–46, 2023.
https://doi.org/10.2991/978-94-6463-014-5_5

introduced in 1970 called Box and Jenkins is the Autoregressive Integrated Moving Average (ARIMA) model [1]. The ARIMA model is ideal for short-term prediction as the predicted result shown in their analysis was quite similar to the actual values [2]. Other techniques are the ARCH model introduced by Robert Engle in 1982, and an extension of the conditional variance function proposed by Bollerslev, which named as GARCH in 1986. Forecasting the fixed deposit rate, will be more accurate if takes into consideration several variables that give a significant impact to rate changes. In such a situation, multiple linear regression can be used to identify the relationship between fixed deposit rate and the independent variables. It is a statistical method that makes use of several explanatory variables to predict the outcome of the response variable. It helps in understanding how the dependent variable changes when the independent variables changes.

The objectives of this study are to determine the linear relationship between fixed deposit rate and certain financial or economic indicator, to predict the fixed deposit rate of Bank Negara Malaysia based on the fitted model and to forecast the profit of investing in fixed deposit for the coming three years.

All the data set are obtained from the official website of Bank Negara Malaysia. The study period starts from January 1997 until December 2018 with a total of 22 years. Since the data is collected based on monthly basis, there are 264 observations available for each variable for analysis.

The time series analysis of Box and Jenkins model is used in this research to identify the pattern of fixed deposit rate of return to the investors. Multiple linear regression is implemented to determine which factor has significant impact on the fixed deposit rate. Five different variables were investigated in this study.

2 Literature

Lairi and Lee (2016) studied the effect of changes in GDP and interest rate on savings using panel data evidence of 6 ASEAN countries. The main purpose of this study is to investigate the relationship between GDP, interest rate and saving across the ASEAN-6 countries by using Panel Data Econometrics [3]. The data used is the secondary data from World Bank and international finance statistics publication. Panel data regression is employed with saving as the dependent variable. It was found that, there is a positive relationship between GDP and interest rate with savings [3].

Ghulam et al. (2018) states that regression performed a vital role in construction of time series econometrics model. However, due to time series properties of the variables, particularly non-stationary series, autocorrelation is sure to be present and leads to a spurious regression [4]. The consequences might lead the model to be inappropriate, as the estimator properties of the model is no more reliable [4]. Since autocorrelation suggest the variables is correlated by its lag, the paper claim that it is reasonable to include the lag variables into the regression model, which leads to the ARDL method. In addition, Monte Carlo simulations further provide justification that ARDL model can be used as an alternative tool to avoid the autocorrelation problem that exist in the model.

Aliu, et al. (2016) investigated the factors affecting interest rate risk using the case of Kosovo [5]. The objective of this study is to capture the short-term and long-term effects of the factors affecting the interest rate in Kosovo. Several factors were included and the data were collected from different resources [5]. The authors listed down all the expected impact based on different factors. Statistically, the results in the short-term basis showed that the impacts are contrary to the theoretical expectations [5]. On the other hand, for long-run basis, most results were as expected where net profit (NP) and overall expenses (EXP) have positive effect while GDP and inflation appeared to have negative effect. However, there is one factor, which is interest income (INI) showed unexpected impact where it affects interest rate negatively.

Several other studies on ARIMA and forecasting are [6, 7] and [8]. Based on representative past studies, ARIMA model is selected as the forecasting method in this research as this model is better in predicting for short-term period. Multiple linear regression also was proven to be a good approach to select the independent variables which gives significant impact to the fixed deposit interest rate. It is important to determine which are the influential factors before forecasting the interest rate.

3 Methodology

Time series data is normally used to monitor industrial process and track business metrics. The time series means ordered sequence of values of a variable at equally spaced time intervals. The main purpose of time series modelling is to study the past observations in order to generate an appropriate model describing the structure of the series. In this particular study, time series statistical technique is applied to obtain the most adequate model for forecasting purpose. All the analysis of time series will be conducted on SPSS, EViews and Minitab software.

There are various methods of time series model fitting. These includes Box-Jenkins ARIMA models, Box-Jenkins Multivariate Models and Holt-Winters Exponential Smoothing (single, double, triple). In this study we use Box-Jenkins ARIMA models to fit the time series model. Box-Jenkins ARIMA model or also known as “univariate time series” refers to a time series that consists of single (scalar) observations recorded sequentially over equal time increments. The model developed known as ARMA model is defined as:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \theta_q \varepsilon_{t-q}$$

$$\left(1 + \phi_1 B + \phi_2 B^2 + \ldots \phi_p B^p\right) Y_t = \left(1 - \theta_1 B - \theta_2 B^2 - \theta_q B^q\right) \varepsilon_t \tag{1}$$

$$\phi_p(B) Y_t = \theta_q(B) \varepsilon_t$$

However, for the case of non-stationary ARMA model, Box and Jenkins proposed the Autoregressive Integrated Moving Average, ARIMA (p, d, q) model to counter the time correlated modelling. The term (I) is integration referring as the differencing procedure with the notation d as the degree of differencing. It is introduced for a non-stationary ARMA model to solve the stationary problem of the series. The ARIMA model can be defined as:

$$\phi_p(B)(1 - B)^d Y_t = \theta_q(B) \varepsilon_t \tag{2}$$

Basically, there are four major steps involved in Box and Jenkins method which are model identification, parameter estimation, model validation with diagnostic checking and lastly forecasting.

Model identifications involves data transformation, Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF), differencing and model selection which is based on the characteristic of the ACF and PACF.

The parameter is estimated using conditional least squares whereas the estimators using this method will make the sum of squared residuals of the model become minimum. The Box and Jenkins model is fitted for parameter estimation. Next is model validation with diagnostic checking whereby this phase aimed to identify whether the estimated model is statistically adequate. The diagnostic checking is implemented based on residuals, which are defined by:

$$Residual = Observation - Fitted\ value \tag{3}$$

Diagnostic checking involves residual analysis, ACF and PACF of the residuals, Breusch-Godfrey Lagrange-Multiplier test, heteroskedasticity test, overfitting of the model and Bayesian Information Criterion (BIC). If heteroskedasticity effect exists, GARCH (Generalized ARCH) model, Exponential GARCH (EGARCH) model, GARCH-in-mean (GARCH-m) model) is proposed to model the volatility of the data. Once the most adequate model is identified, then only it can be used to generate accurate forecasting.

Next, multiple linear regression is implemented to determine the factors that has significant impact on the FD rate. Multiple Linear Regression (MLR) is a statistical technique that model the linear relationship between a response variable with two or more independent variables. This method is chosen because MLR can easily provide understanding of the rate of change of response variable due to the change of regressors. This could then use to predict future values. Multiple linear regression model is defined as:

$$Y_t = \beta_0 + \beta_1 x_{t1} + \cdots + \beta_k x_t + \varepsilon_t,\ \ t = 1, 2, \ldots, n \tag{4}$$

where Y_t is the value of the response variable at t^{th} observations, and $x_t, \ldots, x_{tk}$ denote the value of t^{th} sample unit of k explanatory variables. The model is a linear relationship with coefficients $\beta_0, \ldots, \beta_k$. ε_t is the random error term that contains factors affecting y apart from the independent variables. F-test is used to test the significance of the model. For k regressors, the hypothesis test is:

$$H_0{:}\beta_1 = \beta_2 = \cdots = \beta_k = 0\ vs\ H_a{:}\,at\ least\ one\ \beta_j \neq 0 for\, 0 < j < k \tag{5}$$

The F-statistic is calculated by:

$$F = \frac{SSR/k}{SSE/(n-k-1)} = \frac{MSR}{MSE} \tag{6}$$

The null hypothesis is rejected when $F > F_{\alpha,k,n-k-1}$, and this indicates that at least one of the explanatory variables contributes significantly to the model.

The coefficient of multiple determination R^2 provides a description for the goodness of fit of the regression model. The R^2 could be interpreted as the total percentage of variation in dependent variable that can be explained by the independent variables.

After the model is determined, diagnostic checking is conducted to ensure the current model is statistically adequate. The presence of multicollinearity, autocorrelation and heteroskedasticity should be detected to avoid an inappropriate model.

When at least one of the independent variables is highly correlated to another independent variables, there exist multicollinearity. The consequence is that the variance of the least squares estimates of the regression coefficient will be very large. To detect multicollinearity, one of the methods is by the variance inflation factor (VIF). VIF measures the tendency of increment in variances and covariances of coefficients, which is:

$$VIF = \frac{1}{1 - R_j^2} \tag{7}$$

where R_j^2 is the coefficient of determination obtained when x_j is regressed on the remaining variables, which $0 < R_j^2 < 1$. The more correlated x_j is with other variables, the closer R_j^2 is to 1, resulting a larger VIF. Values of VIF that exceeds 10 shows the presence of multicollinearity. To deal with multicollinearity, one direct method is to drop the variables that are suspected to have high multicollinearity with the other variables, in other words, removing the variable with highest VIF from the model.

Autocorrelation is defined as the correlation between the observation of the series and thus violates the independent assumption. However, time series data usually exhibit this property, as the observations are ordered in chronological order. Furthermore, in time series regression, the response variable at current period might not only be affected by the regressors at same period but also the regressors at previous period. The consequences of autocorrelation are R^2 is likely to be overestimate and would give misleading conclusions about independent variables that affects the response variable. To detect autocorrelation in regression, Durbin-Watson d statistic is checked, which is defined as:

$$d = \frac{\sum_{t=2}^{t=n} \left(\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1}\right)^2}{\sum_{t=2}^{t=n} \hat{\varepsilon}_t} \tag{8}$$

where $\hat{\varepsilon}_t$ is the residuals at time t. The d statistic is simply the ratio of the sum of squared differences in successive residuals to the sum of squares of residual. An advantage of Durbin-Watson d statistic is that it is based on residuals only, which can be easily compute in regression analysis. The decision rules are summarized in Table 1. After detecting the presence of autocorrelation, one of the remedial procedures is by applying Autoregressive Distributed Lag (ARDL) model as shown in Eq. 9 adapted from [9]. This method solves the autocorrelation by adding lag variables to the current regression model, whereas the length of autoregressive is freely determined using trial and error method.

$$\Delta I_t = \alpha_0 + \theta_1 I_{t-1} + \theta_2 S_{t-1} + \sum_{i=1}^{P} \overline{\omega}_i \Delta I_{t-i} + \sum_{j=0}^{P} \beta_j \Delta S_{t-j} + \varepsilon \tag{9}$$

Table 1. Durbin-Watson d statistics: Decision Rules

Condition	Results
$0 < d < d_L$	Positive autocorrelation occurs
$d_L \leq d \leq d_U$	No decision
$d_U < d < 4 - d_U$	No autocorrelation
$4 - d_U \leq d \leq 4 - d_L$	No decision
$4 - d_L < d < 4$	Negative auto correlation occurs

At this stage, Durbin-Watson d statistic is no longer applicable as the regression model is converted to an autoregressive model, where LM test works well in testing the autocorrelation. The added lag variables will subsequently lead the model to evade autocorrelation because the new model already accounts for the relationship between response variable and the lag of independent or even the dependent variables itself. Thus, the modified model will provide more precise explanation of the response variable.

The homoskedasticity assumption can be identified through graphical method, by plotting the standardized residuals against the fitted value from the regression model. The points should be scattered randomly within the horizontal band from -2 to 2 (assume 5% significance level). Heteroskedasticity is suspected when there is obvious pattern, or the points lies outside the band.

Lastly the predicted mean response and its corresponding confidence interval is generated through statistical software. The estimated value is applied on validation to ensure the model explains the variables well and also projection to estimate the value of response at certain values of regressors.

4 Results

In this study FD are studied to forecast its return from 2019 to 2021. Monthly data selected ranges from January 1997 to December 2018. The data fitted to an appropriate model and then forecast the future values of the series in order to compute and determine its feasibility. Five variables suspected to have high influence on the FD rate which are Base Lending Rate (BLR), Consumer Price Index (CPI), Money Supply (M2), KLCI and Real Effective Exchange Rate (REER) were selected to investigate their relationship with fixed deposit rate (FD Rates) of Malaysia.

As mentioned previously, FD is a low risk investment with a guaranteed return. However, the Malaysia bank FD rates may change over time and is determined by several economic and financial indicators. The initial model suggests that FD rates is influenced by explanatory variables: BLP, CPI, M2, KLCI and REER. Data from year 1997 to 2016 is selected for the study and linear regression is conducted to identify the relationship between the dependent and independent variables. Next, the time series model of the significant variables in the regression model is determined and the future three years value is forecasted for regression.

From Table 2, the FD Rates has a mean of 3.989% ranging from 2.5% to 10.28%, with a standard deviation of 1.557. Clearly, the mean value is nearer to the minimum value rather than the maximum value, indicates in some short period, high rates were given by the bank to encourage people to save more by fixed deposit.

Based on Table 3, the correlation coefficient of FD Rates and BLR indicates there is a strong linear relationship between them. Subsequently, the correlation of CPI, M2 and KLCI with FD Rates has coefficient of −0.622, −0.543, and −0.494, showing there might be a negative linear relationship between them. In addition, the high correlation of CPI and M2, CPI and KLCI, M2 and KLCI suggest multicollinearity possibly exist in the model. Lastly, REER with a correlation coefficient of 0.384 with FD Rates implies there is a weak relationship between them. From Table 4, the result of stepwise regression yields a new model, which BLR, CPI and REER are selected to be the variables in the model. The model is tentatively the best model before diagnostic checking and given by:

$$FD\,Rates = 0.708 + 1.064BLR - 0.036\,CPI - 0.006REER \tag{10}$$

To ensure the current model is statistically adequate, diagnostic checking is conducted by detecting the presence of multicollinearity, autocorrelation and heteroskedasticity. Since the Variance Inflation Factor (VIF) value for BLR, CPI and REER are 1.239, 1.408 and 1.316 respectively which is greatly less than 10, model does not suffer from multicollinearity problems.

The presence of autocorrelation is detected using Durbin-Watson d statistics. The test indicates that positive autocorrelation occurs in the model. This implies certain variables have positive relationship with its own lags. This is not surprising because the FD Rates at time ‘t’ does not only depends on explanatory variables at time ‘t’, but might also related on these variables of previous time period. To cope with the issue, ARDL method is employed to manage the autocorrelation by adding lag variables to modify the regression model as shown in Table 5.

Based on Table 5, applying ARDL method by adding lag variables to the original model successfully solve the autocorrelation. Based on the modified regression model, the regression equation is:

$$\begin{aligned} FD\,Rates =\, & 0.063 + 1.005\,BLR - 0.03\,CPI - 0.07\,REER \\ & + 0.906\,FD\,Rates\,(lag\ 1) - 0.926\,BLR\,(lag\ 1) \\ & + 0.028\,REER\,(lag\ 1) \end{aligned} \tag{11}$$

Compared to the previous regression equation, the relationship between FD Rates and BLR, CPI and REER remain the same, and did not change much in the magnitude. However, the FD Rates is found to have significant positive relationship with its own at one lag. These suggest that FD Rates in current period is not only affected by the previous determined variable, but also depend on BLR, REER and FD Rates of previous period.

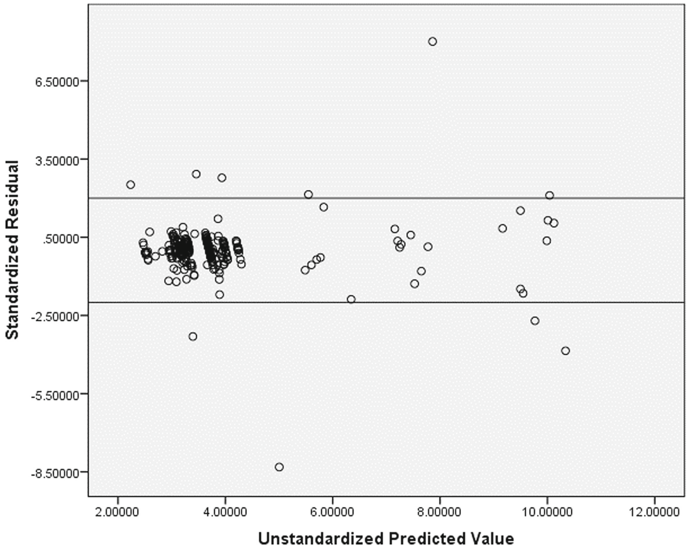

Fig. 1. Standardized residuals versus fitted value

In order to check for homoskedasticity, standardized residuals are plotted against the fitted value. Based on Fig. 1, around 10 points are outside the horizontal band whereas the other points all lie between the horizontal band from −2 to 2. The model has a constant variance since no clear pattern seen. However, several outliers might present. Around 95% of the residuals lies within the band makes the situation acceptable. Furthermore, existence of financial crisis in the sample period will yield some extreme data points which contribute to these outliers.

According to the Box and Jenkins model, it is only applicable to a stationary series. The homogenous non-stationary series could be reduced to a stationary series by carrying out suitable degree (order) of differencing, d on the series. To determine the non-stationary of the series, one could examine the data graphically or inspect the estimated ACF of the series. Whenever the ACF displays pattern of tailing off extremely slowly at all lags, it is believed that the series is non-stationary. Hence, the differencing method is employed to all variables to obtain a stationary series. The first differencing has stabilized the mean and the characteristic of non-stationarity has been removed. The differenced series is now stationary in mean.

Box-Jenkins modelling procedure was done to all the stationary variables including BLR, CPI and REER. Subsequently, out-sample forecast is performed to obtain the next 3 years value of BLR, CPI and REER. Based on the out-sample forecasting on the period from January 2019 to December 2021, the predicted value of BLR, CPI and REER remain constant.

Table 2. Descriptive Statistics

Variables	Mean	Variance	Std. Deviation	Min	Max
FD Rates	3.989	2.424	1.557	2.500	10.280
BLR	6.836	1.505	1.227	5.510	2.270
CPI	93.305	156.800	12.522	72.700	116.800
M2	806811.470	2.114 × 1011	459762.340	246200	1647269.480
KLCI	1140.1432	181921.811	426.523	302.910	1882.710
REER	98.187	54.120	7.357	82.970	132.600

Table 3. Correlation Coefficient between Variables

Variables	FD Rates	BLR	CPI	M2	KLCI
BLR	0.948				
CPI	−0.622	−0.410			
M2	−0.543	−0.319	0.988		
KLCI	−0.494	−0.277	0.896	0.929	
REER	0.384	0.331	−0.465	−0.407	−0.167

Since the appropriate regression model and the desired forecast values are obtained, the values of FD Rates for the next three years can be estimated by using the regression Eq. (10). However before forecasting the FD Rates, the comparison of actual values and predicted values of regression model should first be validated by the reserved observation from the last two years, year 2017 and 2018 of the data as shown in Table 6. The results show that all the actual values throughout the last two years lies inside the prediction interval. This indicates that the model is considered sufficient to explain the behavior of FD Rates. The model is appropriate to estimate the FD Rates from year 2019 to 2021. The forecast of FD rates from 2019 to 2021 shown in Table 7.

Based on Table 7, the average values of fixed deposit rates in each year is considered to represent the fixed deposit rates for the respective year. The return is calculated by the formula:

$$\begin{gathered} 1 \times \left(1 + \tfrac{3.165}{100}\right) \times \left(1 + \tfrac{2.989}{100}\right) \times \left(1 + \tfrac{2.890}{100}\right) = 1.093 \\ 1.093 \times 100 = 109.3\% \end{gathered} \tag{12}$$

The investment for three years of the one-year period fixed deposit provides a return of 9.3% to the investors that invest in fixed deposit for the coming three years. These phenomena indicate that the fixed deposit gives a consistent return.

Table 4. Coefficient Statistics

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	0.708	0.401	1.763	0.079
BLR	1.064	0.017	62.535	<0.0001
CPI	−0.036	0.002	−20.386	<0.0001
REER	−0.006	0.003	−2.115	0.035
R-Square		0.966 (96.6%)		

Table 5. Coefficient Statistics

Model	Unstandardized Coefficients		t	Sig.
	B	Std. Error		
(Constant)	0.063	0.161	0.389	0.698
BLR	1.005	0.036	28.169	<0.0001
CPI	−0.03	0.001	−2.774	0.006
REER	−0.027	0.004	−6.847	<0.0001
FD Rates (lag1)	0.906	0.027	34.080	<0.0001
BLR (lag1)	−0.926	0.038	−24.633	< 0.0001
REER (lag1)	0.028	0.004	7.145	< 0.0001
R-Square		0.994		
F-test		<0.0001		

Table 6. Comparison of Actual Values and Predicted Values of Regression Model

Date	Actual Values	Predicted Values	Lower Prediction Interval	Upper Prediction Interval
Jan 2017	3.09	3.029	2.804	3.255
Feb 2017	3.09	3.044	2.818	3.270
Mar 2017	3.07	3.063	2.838	3.289
Apr 2017	3.08	3.052	2.827	3.278
May 2017	3.08	3.017	2.791	3.242
Jun 2017	3.10	3.058	2.832	3.283
Jul 2017	3.10	3.096	2.870	3.321
Aug 2017	3.10	3.104	2.879	3.330

(*continued*)

Table 6. (*continued*)

Date	Actual Values	Predicted Values	Lower Prediction Interval	Upper Prediction Interval
Sep 2017	3.10	3.049	2.824	3.275
Oct 2017	3.10	3.071	2.845	3.296
Nov 2017	3.10	3.039	2.813	3.265
Dec 2017	3.10	3.029	2.803	3.255
Jan 2018	3.16	3.082	2.857	3.308
Feb 2018	3.33	3.299	3.073	3.525
Mar 2018	3.33	3.295	3.069	3.520
Apr 2018	3.33	3.283	3.058	3.509
May 2018	3.33	3.308	3.082	3.534
Jun 2018	3.33	3.331	3.106	3.557
Jul 2018	3.33	3.300	3.075	3.526
Aug 2018	3.33	3.315	3.090	3.541
Sep 2018	3.33	3.314	3.089	3.540
Oct 2018	3.33	3.292	3.067	3.518
Nov 2018	3.33	3.316	3.090	3.541
Dec 2018	3.33	3.294	3.069	3.520

Table 7. Forecast of Fixed Deposit Rates

Fixed Deposit Rates	
Year 2019	3.165
Year 2020	2.989
Year 2021	2.890

5 Discussion and Conclusion

FD rates does not solely depend on its past values. FD Rates is declared upon considering the economic and financial indicators. Therefore, Base Lending Rate, Consumer Price Index, Money Supply M2, KLCI, and Real Effective Exchange Rate are selected as they are believed to provide significant impact on the FD Rates. Evidence such as Pearson correlation coefficient do support that some of these variables have strong relationship with FD rates. Stepwise regression detects that BLR, CPI and REER are significant in affecting the values of FD Rates. R-Square value of 96.6% indicates a very large percent of variation in FD Rates can be explained by these three variables. The model is said to be very efficient in explaining the FD Rates using BLR, CPI and REER. To solve autocorrelation problem, ARDL method is used as it makes less changes to the original model, resulting a simpler model which are statistically adequate to estimate the future values. Since these three variables are significant, time series analysis is conducted to investigate the time series pattern of them. The validation of the time series model is done by forecasting the series from year 2016 to year 2018. The model successfully provides reliable forecast as all observation from year 2016 to 2018 falls inside the forecast interval. The actual values of FD Rates from year 2016 to 2018 did lie between the confidence interval of the fitted value, implying the model is accurate in estimating the FD Rates. Therefore, prediction as well as forecast of the future three years of FD rates can be acquired.

The annual FD Rates is calculated by averaging the FD Rates of 12 months in a year to obtain a 12-month period FD Rates for three years. The findings reported FD Rates will be 3.165% in the first year, 2.989% in the second year, 2.890% in the third year and provide a total of 9.30% return for investing in fixed deposit for future three years' time. For a risk averse, fixed deposit gives not large but a consistent return therefore it will be appropriate to select fixed deposit as its rate of return are guaranteed.

In conclusion, time series analysis and regression analysis can be used although they have their limitations, in investigating the properties of the series and provide useful information for investors to make decision in their investments.

Acknowledgments. Acknowledgement to Ministry of Higher Education Malaysia (203.PPSG .6711666) for Fundamental Research Grant Scheme with Project Code: FRGS/1/2018/STG 06/USM/02/12 and the fourth author gratefully acknowledges support from USM Research University Grant (1001.PMATHS.8011101).

References

1. Box, G., Jenkins, G.: Time Series Analysis: Forecasting and Control. Holden-Day, San Francisco (1970)
2. Adebiyi, A., Adewumi, A., Ayo, C.: Stock price prediction using the ARIMA model. In: UKSim-AMSS 16th International Conference on Computer Modelling and Simulation, 26–28 March 2014, Cambridge, United Kingdom (2014)
3. Lairi, N.O., Lee, C.W.: The effect of changes in GDP and interest rate on savings: panel data evidence on 6 ASEAN countries. Bus. Manag. Stud. **2**(3) (2016). http://redfame.com/journal/index.php/bms/article/view/1699. Accessed 12 Mar 2019
4. Ghulam, G.: ARDL model as a remedy for spurios regression: problems, performance and prospectus. https://mpra.ub.uni-muenchen.de/83973/. Accessed 10 Jan 2019
5. Aliu, F., Knapkova, A., Aliyev, K., Nadirov, O.: Factors affecting interest rate risk: the case of kosovo. In: The 6th EACO International Scientific Conference, 8–9 December 2016, Krakow, Poland (2016)
6. Gupta, B.K., Mallick, M.K., Hota, S.: Survey on stock price forecasting using regression analysis. In: Mishra, D., Buyya, R., Mohapatra, P., Patnaik, S. (eds.) Intelligent and Cloud Computing. SIST, vol. 153, pp. 147–156. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-6202-0_16
7. Pillay, S.: Determining the optimal arima model for forecasting the share price index of the Johannesburg stock exchange. J. Manag. Inf. Decis. Sci. **23**(5), 527–538 (2020)

8. Sun, Z.: Comparison of trend forecast using ARIMA and ETS models for S&P500 close price. In: 4th International Conference on E-Business and Internet, ICEBI, Singapore (2020)
9. Narayan, P.K.: The saving and investment nexus for China: evidence from cointegration tests. Appl. Econ. **37**(17), 1979–1990 (2005)