

Forecasting Exchange Rate of Bangladesh – A Time Series Econometric Forecasting Modeling Approach

A. S. M. Borhan and Md. Sharif Hossain

Institute of Statistical Research and Training (ISRT), Dhaka University, Dhaka-100, Bangladesh

Received on 25. 02. 2009. Accepted for Publication on 08. 10. 2010

Abstract

In this paper an attempt has been made to select a model for time series forecasting of Average Exchange Rate (AER) of Bangladesh. Our decision through out this study is mainly concerned with Auto regressive Integrated Moving Average (ARIMA) Model, Holt's Linear Exponential Smoothing Model, Simple Linear Regression Model, Log-Linear Regression Model. This project concerns and analyze on a set of data based on AER during the period July 2003 to June 2007. We try to derive a unique and suitable forecasting model AER. From our study we find that Holt's Linear Exponential Smoothing with $\alpha=0.999$ and $\beta=0.018$ gives less forecasting error than that of others. So we propose that forecasting for the Average Exchange Rate of Bangladesh, one can use the Holt's Linear Exponential Smoothing Model. But before using this model one must verify the validation of the model in different time period, because a forecasting model may lose its validity and suitability as time changes.

I. Introduction

Frequently, there is a time lag between awareness of an impending event and occurrence of that event. If the lead time is long, and the outcome of the final event is conditional on identifiable factors, planning can perform an important role. In such situation forecasting is needed to determine when an event will occur or a need arise, so that appropriate actions can be taken. Forecasting is vital in finance and in financial time series analysis. Opinions on forecasting are probably as diverse as views on any set of scientific methods used by decision makers.

For making valid forecasting, we have to make a choice between a numbers of competing alternative models. The model selection procedure largely depends on the process of searching for a suitable specification. In forecasting point of view, the best model is that one which fits the observed data well and for which the forecast error, on average, is minimum in some sense. For making valid forecasting, we have to make a choice between a numbers of competing alternative models. The model selection procedure largely depends on the process of searching for a suitable specification. In forecasting point of view, the best model is that one which fits the observed data well and for which the forecast error, on average, is minimum in some sense.

Forecasting of exchange rate is very necessary for us. Because Bangladesh is very much depending on export and import and we deal this with international currency such as Dollar, Pound, Euro etc. Here we forecast on the exchange rate of U.S.

Dollar. But unfortunately there is no forecasting model adopted in Bangladesh Bank (BB) to see the future performance of the rate of Dollar.

The price of one country's currency expressed in another country's currency is known as exchange rate. In other words, the rate at which one currency can be exchanged for

another. For example, if the U.S. exchange rate for the Bangladeshi Taka is Tk.60.00, this means that 1 American Dollar can be exchanged for 60.00 Bangladeshi Taka.

This study was conducted on the exchange rate of US Dollar. Our main goal is searching a univariate forecasting model for the average exchange rate of US Dollar, so we had to work on one variable. We took 'Exchange Rate' as our variable. In our data set have 48 observations during the time period 2003-04 to 2006-07.

II. Arima Model

An autoregressive model of order p is conventionally classified as $AR(p)$. A moving average model with q terms is classified as $MA(q)$. A combination model containing p autoregressive terms and q moving average terms is classified as $ARMA(p,q)$ (Gujarati(1995)). If the object series is differenced d times to achieve stationarity, the model is classified as $ARIMA(p,d,q)$, where the symbol "I" signifies "integrated." So the general non-seasonal model is as $ARIMA(p,d,q)$:

AR: p = order of the autoregression part, I: d = degree of differencing involved MA: q = order of the moving average part

The equation for the $ARIMA(p,d,q)$ model is as follows:

$$Y_t = C + f_1 Y_{t-1} + f_2 Y_{t-2} + \dots + f_p Y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_q e_{t-q}$$

Or in backshift notation

$$(1 - f_1 B - f_2 B^2 - \dots - f_p B^p) Y_t = c + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t$$

Where, C = constant term, f_i = i th autoregressive parameter, θ_j = j th moving average parameter, e_t = the error term at time t , B^k = the k th order backward shift operator

The Box-Jenkins (BJ) Methodology

To identify a perfect ARIMA model for a particular data series, Box and Jenkins(Box and Jenkins(1976)) proposed a methodology that consists of three phases

are known as Box-Jenkins methodology, or in short BJ methodology. The total process contains the following phases, namely : Phase I: Identification. Phase II: Estimation of Diagnostic Checking. Phase III: Application.

Holt's Linear Exponential Smoothing Method

In exponential smoothing procedures the weights assigned to observation are exponentially decreased, as the observations get older. Holt (1957)(Makridakis(1998)) extended single exponential smoothing to linear exponential smoothing to allow forecasting of data with trends. The forecast for Holt's linear exponential smoothing is found using two parameter smoothing constants, α and β with values between 0 and 1, and three equations:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \dots \dots \dots (1) \quad b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \dots \dots \dots (2)$$

$$F_{t+m} = L_t + b_t m \dots \dots \dots (3)$$

Where, L_t is the estimate of the level of the series at time t , b_t is the estimate of the slope of the series at time t , Y_t is the observation at time t , α and β are constants between 0 and 1, F_{t+m} is forecasted value at period $t + m$, m is the number of periods ahead to be forecast. To perform a Holt's linear exponential smoothing method we have to go through the following three steps, namely: Step 1: Initialization Step 2: Optimization Step 3: Forecasting

III. Forecasting with Simple Regression

Regression as Statistical Modeling : The simple linear regression model may be defined precisely as: $Y_i = \alpha + \beta X_i + \varepsilon_i$, Where, Y_i and X_i represent the i th observation of the variables Y and X respectively, α and β are fixed (but unknown)parameter and ε_i is a random variable that is normally distributed with mean zero and variance σ^2

Log-linear Regression Model

However, it should be noted that many non-linear functions can be transformed into linear functions. A few simple cases can illustrate the point consider. $W = AB^X \dots \dots \dots (1)$

Equation (1) relates variable W to variable X and a plot of W versus X would be non-linear. Our concern is with the parameters A and B , which appear as a product (therefore not linear) and B is raised to a power other than 1 (therefore not linear). To fit this can be to a set data pairs (W, X) would require an iterative procedure, unless logarithms are taken of both sides: $\log W = \log A + (\log B)X$. Substituting $Y = \log W$, $\alpha = \log A$ and $\beta = \log B$ gives $Y = \alpha + \beta X \dots \dots \dots (2)$ Equation (2) is now a simple linear relationship since the function is linear in α and β (It is also linear in X). Thus we can use simple LS regression or Equation (2), solve α and β and then recover A and B via antilogarithms to get the parameter estimates for equati

IV. Measures of Forecast Error

The model that gives the minimum measures of forecast error will be our expected model for further forecasting.

Suppose, Y_t is the actual observation for time period t and F_t is the one-step forecast for the same period, which is calculated using data Y_1, Y_2, \dots, Y_{t-1} . Then, the one-step forecast error is denoted by e_t and defined as, $e_t = Y_t - F_t$. If there are observations and forecasts for n time periods, then there will be n error terms, and the following standard statistical measures can be defined:

$$(i) \text{Mean Error (ME): } ME = \frac{1}{n} \sum_{i=1}^n e_t \quad (ii) \text{Mean Absolute Error (MAE): } MAE = \frac{1}{n} \sum_{i=1}^n |e_t| \quad (iii) \text{Mean Square Error (MSE): } MSE = \frac{1}{n} \sum_{i=1}^n e_t^2 \quad (iv) \text{Mean Percentage Error (MPE): } MPE = \frac{1}{n} \sum_{i=1}^n PE_t \quad \text{Where, } PE_t = \frac{(Y_t - F_t)}{Y_t} \times 100 \text{ is the relative or percentage error at time.} \quad (v) \text{Mean Absolute Percentage Error (MAPE): } MAPE = \frac{1}{n} \sum_{i=1}^n |PE_t| \quad \text{on } (I).$$

V. Analysis of Data

Our data set have 48 observations during the time period July 2003 to June 2007. We divide the whole set of observations into two segments, namely 'the training segment' and 'the test segment'. The training segment contains first 40 observations and the test set contains the remaining 8 observations. The idea behind this partition is that to select the best forecasting model from Autoregressive Integrated Moving Average (ARIMA) model, Holt's Linear Exponential Smoothing model, Simple Regression model and Log-Linear Regression model using the training segment and then compare among these models by applying them in the test segment.

Choosing an ARIMA Model: Identification: Stability in Variance

At first we plot our data containing in training segment, where the variable 'Average Exchange Rate (AER)' is plotted against the period. From the time plot given in Figure 1 we see that over the time period of study AER has been increasing, that is showing an upward trend. This perhaps suggests that AER series is not stationary. Moreover, it seems that the data are non-stationary in the mean level only. So we do not need any transformation of the data to obtain stability in variance.

Testing Stationarity

To test the stationarity we obtain the correlogram of AER. From the correlogram, we observe that the autocorrelations decline very slowly as shown in figure – 2 and most of the spikes fall outside the 2-standard error limits. It indicates that the data set is non-stationary. The ACF is calculated up to 24 lags to see the seasonality. But there is no seasonal

variation (i.e. at lag 12, lag 24), we conclude that there is no seasonality in the data set.

The white noise test

For the autocorrelation values, r_k , in **figure -2**, the Box-

Pierce Q statistic is computed as: $Q = n \sum_{k=1}^h r_k^2 = 40 \sum_{k=1}^{24} r_k^2 = 215.03$. Here we have used $h = 24$ and, since

were not modeled in any way. Hence the degree of freedom is 24. Comparing this value to a chi-square distribution with 24 degrees of freedom, we get the p-value as 0.000 . So, we conclude that the values are significantly different from a null set. Ljung- Box test, which is readily obtained by computer programming, gives the value 294.308 with 24 degrees of freedom. This also indicates that the data set does not follow a white noise series.

Obtaining Stationarity

Since, there is no seasonality in the data we simply take the first difference to make it stationary. The time plot of the first differenced series is given in **figure - 3**, which shows that the values differ around a constant mean. Now the series looks just like a white noise series, with almost no autocorrelations or partial autocorrelations outside the 95% limits. The correlogram of the first differenced data are given in **figure - 4**. The Box-Pierce Q statistic takes the value 16.23 and the Ljung- Box Q^* statistic is equal to 25.421 for these data when the maximum lag being considered, $h = 24$ and the number of parameters in the model, $m = 0$. Compared to a chi-square distribution with 24 degrees of freedom, neither of these is significant. Taking differences has transformed the data into a stationary series which resembles white noise. As we take only one difference to make the original data series to a stationary series so the value of the parameter d in the model would be 1.

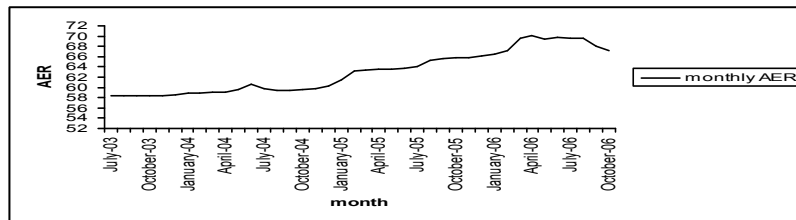


Fig. 1. The time plot of AER

Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	.956	.152					↔*****	*****				39.343	.000
2	.893	.150					↔*****	*****				74.590	.000
3	.818	.148					↔*****	*****				104.958	.000
4	.740	.146					↔*****	*****				130.504	.000
5	.659	.144					↔*****	*****				151.330	.000
6	.577	.142					↔*****	*****				167.779	.000
7	.495	.140					↔*****	*****				180.277	.000
8	.414	.138					↔*****	**				189.272	.000
9	.343	.136					↔*****	**				195.630	.000
10	.273	.134					↔*****					199.817	.000
11	.209	.131					↔*****					202.348	.000
12	.151	.129					↔***					203.717	.000
13	.085	.127					↔**					204.165	.000
14	.012	.124				*						204.174	.000
15	-.065	.122				*	↔					204.459	.000
16	-.136	.120				*	***↔					205.754	.000
17	-.203	.117				*	****↔					208.776	.000
18	-.266	.114				*	*****↔					214.182	.000
19	-.320	.112			**	*	***↔					222.370	.000
20	-.356	.109			*	*	***↔					232.987	.000
21	-.380	.106			*	*	****↔					245.776	.000
22	-.395	.104			*	*	****↔					260.360	.000
23	-.406	.101			*	*	****↔					276.613	.000
24	-.411	.098			*	*	****↔					294.308	.000

Plot Symbols: Autocorrelations * Two Standard Error Limits .

Fig. 2. The Correlogram of AER

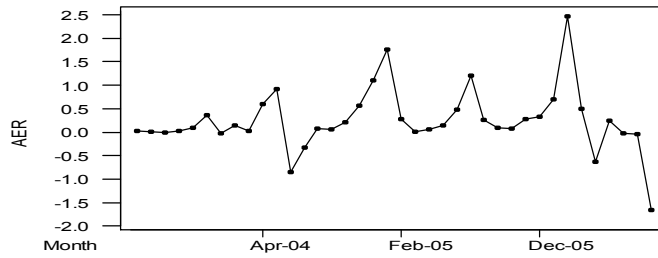


Fig. 3. The Time-Plot First Difference of AER

Auto- Stand.

Lag	Corr.	Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
			□	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓
1	.361	.154					↔	*****	*			5.473	.019
2	-.067	.152					*	↔				5.669	.059
3	-.053	.150					*	↔				5.793	.122
4	-.031	.148					*	↔				5.837	.212
5	-.047	.146					*	↔				5.943	.312
6	-.212	.144					*	****	↔			8.130	.229
7	-.049	.141					*	↔				8.251	.311
8	.099	.139					↔	**				8.759	.363
9	-.037	.137					*	↔				8.832	.453
10	-.047	.135					*	↔				8.953	.537
11	-.025	.132					*	↔				8.989	.623
12	-.013	.130					*					8.999	.703
13	.081	.128					↔	**				9.403	.742
14	.085	.125					↔	**				9.868	.772
15	-.020	.123					*					9.894	.826
16	-.058	.120					*	↔				10.131	.860
17	-.079	.117					**	↔				10.584	.877
18	-.094	.115					**	↔				11.256	.883
19	-.260	.112					*	****	↔			16.665	.613
20	-.310	.109					**	****	↔			24.749	.211
21	.008	.106					*					24.755	.258
22	.072	.103					↔	*				25.239	.286
23	-.043	.100					*	↔				25.421	.329
24	.002	.097					*					25.421	.383

Plot Symbols: Autocorrelations * Two Standard Error Limits .

Fig. 4. The Correlogram First Difference of AER

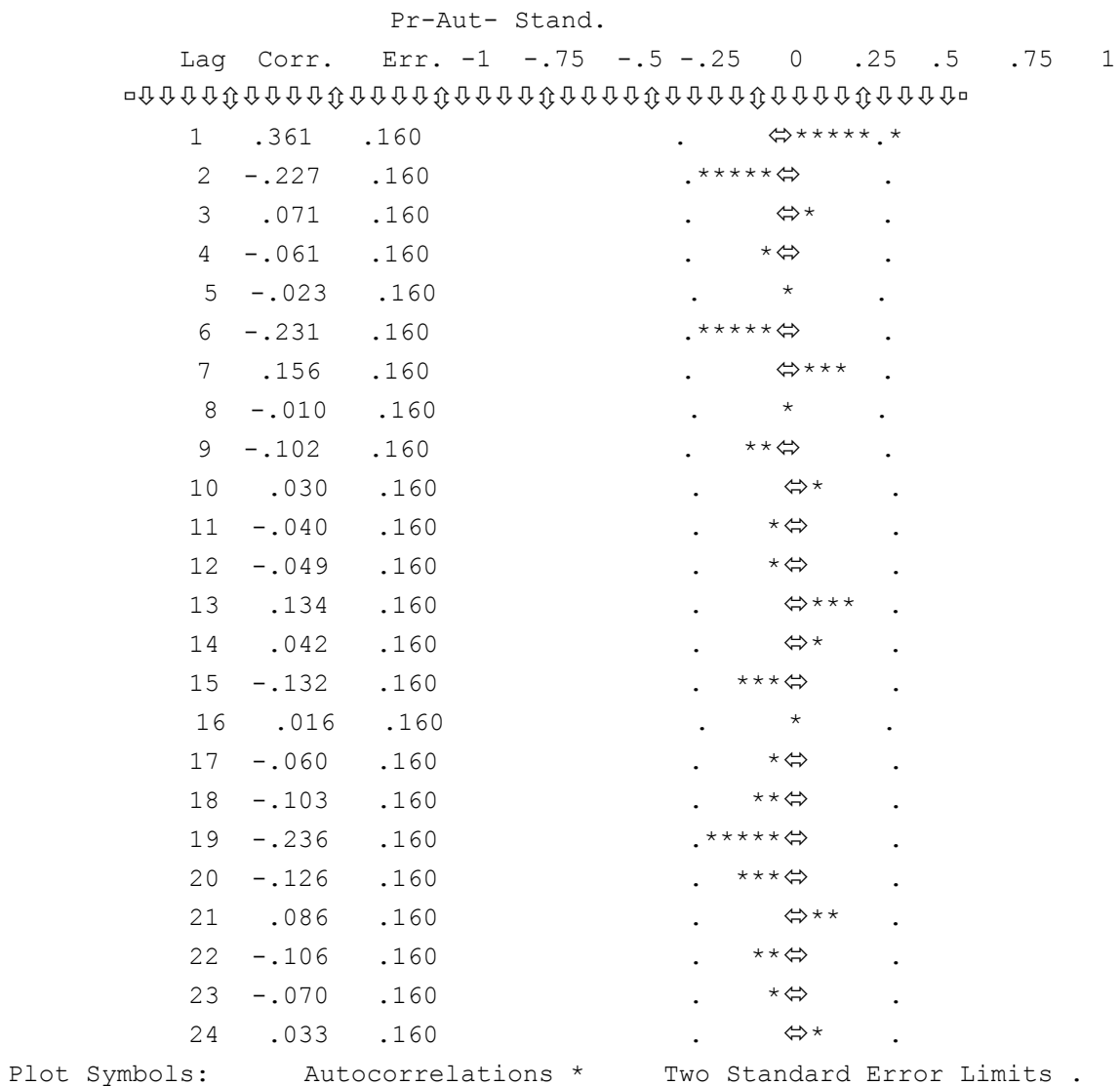


Fig. 5. The Partial Correlogram First Difference of AER

Model Selection

As mentioned earlier, the plot of ACF and PACF can give a primary guess about the order of the parameters, p and q for ARIMA model. We use the Akaike Information Criterion (AIC) to choose the best model among the class of plausible models. The model which has the minimum AIC value is our model of interest. For different values of p and q we find the AIC value using computer program, namely, SAS (Statistical Assessment Software) or R. It gives the minimum AIC value using the maximum likelihood estimation. The AIC values for different p and q values for ARIMA (p, l, q) are given in Table 1.

Estimation and Diagnostic Checking

From the Table 1, we found that the AIC value for the model ARIMA ($0, l, l$) is minimum. This model includes no AR coefficients and one MA coefficient and takes the form: $Y_t^l = e_t - \theta_l e_{t-l}$, Or, $Y_t - Y_{t-l} = e_t - \theta_l e_{t-l}$, Or, $Y_t = Y_{t-l} + e_t - \theta_l e_{t-l}$

Where, θ_l is the l th MA coefficient, e_t is the error term.

Now we have to test the significance of the parameters. The coefficients with their estimated

value and corresponding value of the z -statistic are given in the following Table 2:

Table 2. The significance test of the parameters of ARIMA (0,1,1)

Coefficients	Parameters	Standard error	z-value	Decision
θ_1	0.5214	0.1534	3.40	Significant

From the above table we see that the p-values corresponding to the coefficients θ_1 less than 0.05, which leads to the conclusion that these parameter are significant. So, the revised model becomes, $Y_t = Y_{t-1} + e_t - 0.5215e_{t-1}$
 And the estimated model is , $Y_t = Y_{t-1} - 0.5214e_{t-1}$
 Hence, this is our ARIMA model that we select for forecasting the average exchange rate of Dollar.

For diagnostic checking, we have to look at the **figure 6**, which shows the behavior of the residuals left over after fitting the ARIMA (0,1,1) model. The plot of the standardized residuals shows that most of the standardized residuals are within the 95% limits. The plot of ACF of residuals is given. In both cases, all the spikes are in the 95% limits and near to zero. The plot of p-values of the Ljung-Box statistic indicates that the residuals left over after fitting the model is white noise. All these diagnostic check support that our selected model is not only have the smallest AIC value but also the better behaved residuals. As we obtain an appropriate ARIMA model for forecasting, we use it to predict the future values in the test set.

Estimation of Holt’s Linear Smoothing Parameter

In order to select a Holt’s linear method for forecasting, we take the first value of the original series as initial value i.e. $L_1 = Y_1$ and the difference between the first and second observation of the original series as an estimate of the slope of the series. i.e. $b_1 = Y_2 - Y_1$

We consider different values for the smoothing parameters α and β ranging from 0 to 1 and calculate the forecasted series for each value of α and β . Now through computer programming, we obtain the set of value α and β as smoothing parameter which gives the minimum Mean Square Error (MSE). Following this procedure we obtain the value of $\alpha = 0.999$ and $\beta = 0.018$ which minimizes the MSE. Thus, our Holt’s linear model becomes,

$$L_t = 0.999Y_t + (1 - 0.999)(L_{t-1} + b_{t-1}) = 0.999Y_t + 0.001(L_{t-1} + b_{t-1})$$

$$b_t = 0.018(L_t - L_{t-1}) + (1 - 0.018)b_{t-1} = 0.018(L_t - L_{t-1}) + 0.982b_{t-1}$$

and, $F_{t+m} = L_t + b_t m \quad m = 1, 2, 3, \dots, \dots, \dots$

Using $m = 1$. We get $F_{t+m} = L_t + b_t$. This model is used to forecast 8 values in the test set.

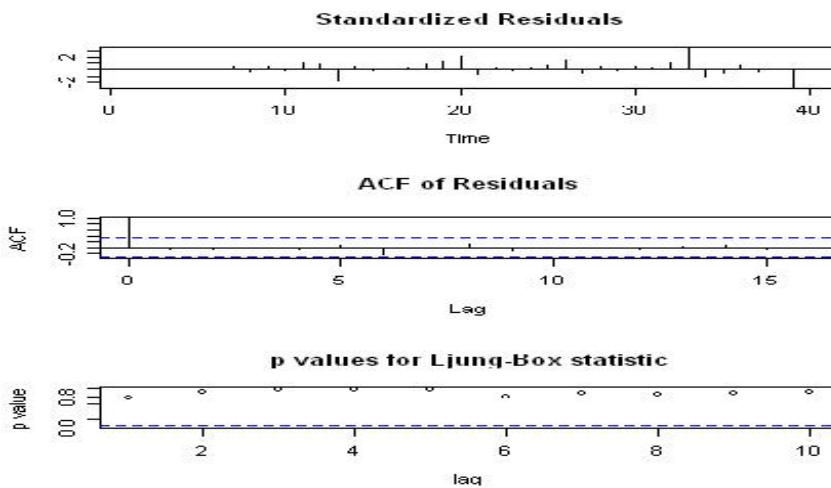


Fig. 6. The diagnostic checking of ARIMA (0,1,1) model of AER

Fitting of Simple Regression Model

We have assumed that a linear relationship exists between the dependent variable Average Exchange Rate (Y) and explanatory variable time period (t). We have a sample of 40 observations. We can write

$$Y_i = \alpha + \beta t_i + \varepsilon_i \quad i = 1, 2, 3, \dots, \dots, \dots, 40$$

The coefficients are unknown and our problem is to obtain estimates of these unknown. From the analysis, we get

ANOVA table for Simple Linear Regression b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	595.768	1	595.768	407.931	.000 ^a
	Residual	55.498	38	1.460		
	Total	651.265	39			

a. Predictors: (Constant), PERIOD

b. Dependent Variable: AER

Coefficients a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	56.218	.389		144.355	.000
	PERIOD	.334	.017	.956	20.197	.000

a. Dependent Variable: AER

Hence the fitted model is, $\hat{Y} = 56.218 + 0.334 t_i$
Fitting of Log-linear Regression Model

Let us consider a non-linear function of parameters as: $Y_i = e^{\alpha + \beta t_i}$

Taking logarithms to base e of both side yields $Log Y_i = (\alpha + \beta t_i) Log e = \alpha + \beta t_i$

Which is now a linear form, so that α and β can be estimated directly. From the analysis we get

ANOVA table for Log-linear Regression b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	.148	1	.148	442.967	.000 ^a
	Residual	.013	38	.000		
	Total	.161	39			

a. Predictors: (Constant), PERIOD

b. Dependent Variable: LOGAER

Coefficients a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	4.034	.006		684.627	.000
	PERIOD	.005	.000	.960	21.047	.000

a. Dependent Variable: LOGAER

So, the fitted model is, $log \hat{Y}_i = 4.034 + 0.005 t_i$, $\hat{Y}_i = e^{4.034 + 0.005 t_i}$

Comparison among ARIMA, Holt’s Linear, Simple Linear Regression, Log-linear Regression models

We calculate different measures of error and the summary measures are listed in the table below:

Table. 3. Comparison among ARIMA, Holt’s Linear, Simple Linear Regression, Log-linear Regression models

Measures of error	ARIMA model	Holt’s linear model	Simple Linear Regression model	Log-linear Regression model
ME	-0.002244	-0.010220	0.0063995	0.383711
MAE	0.402896	0.429813	0.906627	0.975295
MSE	0.411922	0.453091	1.387493	1.513052
MPE	-0.006641	-0.018089	-0.021802	0.556585
MAPE	0.626526	0.668101	1.436760	1.521674

From the above table we see that for all measure of errors, Holt’s Linear method with $\alpha = 0.999$ and $\beta = 0.018$ gives the best results over ARIMA (0,1,1) model, Simple Linear Regression model and Log-linear Regression model.

Forecasting for the period November-2006 to June-2008 (Out of sample)

Now to see the performance of these models in out-of-sample forecasting, we derive the forecasted value of Average Exchange Rate using these models for the

period November 2006 – June 2008 and compare with the actual value of Average Exchange Rate during the period November 2006 to June 2007. The following table gives four forecasted series obtained by the four models as well as the actual data set.

Table 4. The forecasted and actual value of average exchange rate during the period November 2006- June 2008

Period	Average Exchange Rate (AER)	Forecasted AER by ARIMA (0,1,1)	Forecasted AER by Holt's Linear with $\alpha = 0.999$ and $\beta = 0.018$	Forecasted AER by Simple Linear Regression	Forecasted AER by Log-Linear Regression
November-06	69.912115	67.21118	67.31038	69.912	69.338
December-06	69.458654	70.45731	70.17317	70.246	69.686
January-07	69.694200	69.56222	69.71018	70.580	70.035
February-07	69.017500	69.93273	69.94474	70.914	70.386
March-07	68.940563	69.13102	69.25228	71.248	70.739
April-07	68.935288	69.14060	69.16912	71.582	71.094
May-07	69.108148	69.13356	69.15956	71.916	71.540
June-07	68.938942	69.32789	69.33132	72.250	71.8080
July-07	69.11529	69.15539	72.584	72.168

From the above table, we see that the forecasted value for each month in the time period can be obtained by ARIMA (0,1,1) model, Holt's Linear method with $\alpha = 0.999$ and $\beta = 0.018$, Simple Linear Regression model and Log-linear Regression model. Comparing this out-of-sample forecast, Holt's Linear method with $\alpha = 0.999$ and $\beta = 0.018$ gives the best result compared to the others.

VI. Conclusion

The basic aim of this paper was to select a model for forecasting Average Exchange Rate of Dollar in Bangladesh. In this context we took our interest on ARIMA, Holt's Linear method, Simple Linear Regression Model and Log-linear Regression Model. We obtained the forecast error for all the models and found various measures of forecast error, which are listed in Table 3. From this table,

we found that Holt's Linear method with $\alpha = 0.999$

and $\beta = 0.018$ gave less average forecasting error than others. The comparison shows that the forecasting performance of Holt's Linear method

with $\alpha = 0.999$ and $\beta = 0.018$ is better than that of ARIMA model with order (0,1,1), Simple Linear Regression model and Log-linear Regression model.

On the basis of the above comparison, we can conclude that, to forecast monthly average exchange rate of Dollar in Bangladesh one can easily use the Holt's Linear method with $\alpha = 0.999$ and $\beta = 0.018$.

It should also be borne in mind that a good forecasting technique for a situation may become bad technique for a different situation. The validation of a particular model must be examined as time changes.

-
1. Box, G.E.P. and Jenkins G. M.(1976), Time Series Analysis Forecasting and Control, Holden-Day.
 2. Makridakis S.G., S.C.Wheelwright and R. J.Hyndman (1998), Forecasting: Methods and Application, John Wiley & Sons, Inc., New York.
 3. Gujarati D. N. (1995). Basic Econometrics, McGraw-Hill Book Co, New York.
 4. www. Bangladesh-bank. org. bd