

Application of ARIMAX Model on Forecasting Nigeria's GDP

Christogonus Ifeanyichukwu Ugoh, Chinwendu Alice Uzuke, Dominic Obioma Ugoh

Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Nigeria

Email address:

Christogonusugoh2019@gmail.com (C. I. Ugoh)

To cite this article:

Christogonus Ifeanyichukwu Ugoh, Chinwendu Alice Uzuke, Dominic Obioma Ugoh. Application of ARIMAX Model on Forecasting Nigeria's GDP. *American Journal of Theoretical and Applied Statistics*. Vol. 10, No. 5, 2021, pp. 216-225. doi: 10.11648/j.ajtas.20211005.12

Received: July 12, 2021; **Accepted:** July 21, 2021; **Published:** October 29, 2021

Abstract: This paper proposes an appropriate ARIMAX model that is used to forecast the Nigeria's GDP. The data used for the study is sourced from the World Bank for a period of 1990-2019. The ARIMA model is fitted on the residuals using Box-Jenkins approach. The Bayesian Information Criterion (BIC) is adopted to assess the adequacy of the models. The raw data satisfy the assumption of multicollinearity when export is eliminated and the residual series is stationary after the first differencing. This study shows that import is a significant exogenous variable for the GDP dynamics. The ARIMA (0,1,1) with BIC value of 35.253 is considered the appropriate model to be combined with the exogenous variable. The results showed that the ARIMAX (0,1,1) is more ideal and adequate for forecasting Nigeria's GDP based on the Theil's U forecast accuracy measures.

Keywords: GDP, Regression, BIC, ARIMA, ARIMAX, Theil's U Statistic

1. Introduction

Gross Domestic Product (GDP) is the total monetary or market value of all the finished goods and services produced within a country's borders in a specific time period. As a broad measure of overall domestic production, it functions as a comprehensive scorecard of the country's economic health. GDP includes all private and public consumption, government outlays, investments, additions to private inventories, paid-in construction costs, and the foreign balance of trade (exports are added, imports are subtracted). GDP is an important indicator to measure a country's wealth and economic strength. GDP is part of the National income.

Ekhosuehi et al [7] analyzed the link between debt servicing and export earnings of Nigeria using Koyck-kind (KARMAX) model approach, using data extracted from World Bank Database for a period 1970-2018. The result of their findings revealed that the KARMAX model obtained through the maximum likelihood (ML) method is more ideal and inspiring, after comparing the result obtained to the prediction error and the instrumental variable methods. Zheng [20] investigated the methods of least squares in the identification of ARMAX by using 2000 sampled input-output data simulated. The result of the analysis revealed that the BELSX method is considered the best method among the selected methods in identifying

ARMAX, due to the fact that its accuracy is very close to predictor error (PE) of ARMAX. Liu et al [9] proposed a method for locating and quantifying damages in shear structure, based on damage indicator of ARMAX model residual-based KLD using numerical simulation of damage detection on a six-story shear building structure. The results showed that through the evaluation of the models, the proposed CSDF curves of ARMAX model residual can locate clearly the structural damages and the proposed damage indicator of residual-based KLD can locate and quantify the damages in a data-driven way.

Durka and Pastorekova [6] carried out a comparative study between ARIMA and ARIMAX in order to find which of the model is better in the forecasting of macroeconomic time series in Slovakia. They applied Box – Jenkins modelling approach using Gross Domestic Product (GDP) per capita as an output series and unemployment rate as an input series, and the ARIMAX model as fitted explained 92.7% of the variations in the GDP. Musundi et al [12] modeled and forecasted Kenyan GDP using ARIMA model with annual data spanning from 1960 to 2012 which they sourced from the Kenya National Bureau of Statistics. The result of their findings showed that ARIMA (2,2,2) is the best ARIMA model to forecast Kenya's GDP based on minimum AIC.

Ning et al [13] examined Shaanxi GDP, using data which they extracted from Shaanxi Statistical Yearbook from 1952 to 2007.

By applying the Box-Jenkins approach, the result of their findings showed that ARIMA (1, 2, 1) is the best ARIMA model for forecasting Shaanxi's GDP. Ghazo [8] carried out a research on the Jordan GDP and CPI, using data from 1976 to 2019. Applying ARIMA approach and Box-Jenkins approach, the results revealed that ARIMA (3,1,1) is the best model to forecast the Jordan's GDP based on the AIC and SIC criteria. On the other hand, Atanu *et al* [3] examined the Nigeria GDP using data from 1981 to 2019, employed Box-Jenkins approach and obtained ARIMA (1, 2, 1) as the best forecasting model, and Zakai [19] investigates forecasting of Gross Domestic Product (GDP) for Pakistan using quarterly data from 1953 until 2012, using ARIMA (1, 1, 0) model, the findings revealed that Pakistan's GDP will increase for the years 2013- 2025. Maity and Chatterjee [10] examined the GDP growth rate of India data spanning for 60 years, an ARIMA (1, 2, 2) was obtained as the best model for forecasting the GDP which was used to forecast the India's GDP, the result showed that the predicted values intrinsically increases.

Abonazel and Abd-Elftah [1] studied the Egyptian GDP using annual data from World Bank for 1965-2016. ARIMA (1,2,1) was selected as the best model based on minimum AIC, BIC and MSE to forecast the country's GDP for 2017-2026. Their result revealed that Egyptian GDP will steadily rise. Salah and Tanzim [15] predicted Bangladesh's GDP from 2019 to 2025, using ARIMA (1,2,1) model. Their study showed that Bangladesh's GDP trend is steadily improving. Chikumbe and Sikota [5] examined the GDP of Zambia using annual data from 1960 to 2018. Applying the Box-Jenkins approach, ARIMA (5,2,0) model was obtained as the best model to forecast the Zambia's GDP based on the minimum AIC and BIC. The model was used to forecast for next eight (8) years and the result revealed that there will be a decline in the GDP trend for the period of 2020-2022. Touama [16] studied the Jordan GDP series over the period 2003-2013 and applied the ARIMA process. The result revealed that ARIMA (0,1,2) model is the adequate model for the forecast of Jordanian GDP. Wabomba *et al* [18] evaluated Kenyan GDP for 1960-2012 using ARIMA approach. Their study established ARIMA (2,2,2) model as the best model to forecast for future GDP.

Arneja *et al* [2] analyzed the GDP of India using annual data spanning from 1980 to 2017, and ARIMA (1,1,7) obtained as the best model was used to forecast the Indian GDP. Their result revealed that India GDP will be rising carelessly in the future. Miah *et al* [11] examined the GDP of Bangladesh using data ranging from 1960 to 2017, applying the ARIMA process, the outcome of the result showed that ARIMA (1,2,1) model is the best model for forecasting the Bangladesh GDP. Using the model in forecasting for the next thirteen years, the result showed that the GDP of Bangladesh is expected to improve for the period of forecast. Awel [4] carried out an analysis on the real GDP of Ethiopia using annual data spanning from 1981 to 2014. ARIMA (1,1,1) model was selected as the best forecasting model, which was used to forecast for the period 2015-2017. Rana [14] examined the ARIMA forecast model using monthly data from Mid-July 2016 to Mid-July 2018, applying Box-Jenkins

approach to select ARIMA (0,1,2) as the best model to forecast the GDP of Nepal.

The aim of this study is to fit an appropriate ARIMAX model for Nigeria's Gross Domestic Product (GDP) and other economic variables. This ARIMAX model is similar to a multivariate regression model, but allows to take advantage of autocorrelation that may be present in residuals of the regression which improves the accuracy of the forecast Ulyah *et al* [17]. By using ARIMAX model, it is assumed that the future values of the Nigeria's GDP and other economic variables linearly depend on their past values, as well as on the values of past (stochastic) shocks.

2. Materials and Method

2.1. Functional Variables

This study adopts a dependent variable Y and four predictor variables X_1, X_2, X_3, X_4 , where Y is Gross Domestic Product (GDP), X_1 is the Exchange Rate (Naira/US\$), X_2 is the Real Interest Rate (%), X_3 is the Export (million US\$), and X_4 is the Import (million US\$).

The Gross Domestic Product (GDP) was measured using three methods, the Product Method, Income Method and the Expenditure method. This study adopts the most common method, the Expenditure Method, this method is defined as

$$GDP = C + I + G + A \quad (1)$$

where C is the Consumer Expenditure, I is the Investment, G is the Government Expenditure, and A is the Net Export ($Export - Import$).

The variables adopted in this study all have effect in the determination of GDP by expenditure method.

2.2. ARIMA with Exogenous Variable (ARIMAX)

ARIMAX model is a composition of autoregressive integrated moving average (ARIMA) model and significant exogenous variables (exogenous variable is a covariate X_t , that influence the observed time series values Y_t . In order word, ARIMAX model is a multiple linear regression model with AR (Autoregressive) terms and/or MA (Moving Average) terms. ARIMAX model is denoted by ARIMAX (p, d, q) which is the ARIMA (p, d, q) with significant exogenous variables, expressed as

$$Y_t = \beta X_t + \phi(L)y_t + \theta(L)\varepsilon_t \quad (2)$$

Multiple Regression

Multiple linear regression model is the model that describes the relationship between one Dependent (Response) variable Y and two or more Independent (Predictor) variables (X_1, X_2, \dots, X_p) and shown as

$$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon_i \quad (3)$$

where ε is the error term for $i = 1, 2, \dots, n; j = 1, 2, \dots, p$

Autoregressive (AR) Model

AR model is the regression of the current observations

against one or more past observations. That is the current observation y_t are generated by a weighted averages of past time series data going back p periods, together with a random disturbance in the current period. The process is denoted as $AR(p)$ and it is defined as

$$Y_t = \phi(L)y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} \quad (4)$$

where $\phi_1, \phi_2, \dots, \phi_p$ are the parameters of the AR model; y_t is the current observation, $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ are past observations.

$\phi(L)$ is the Autoregressive polynomial and it is expressed as

$$\phi(L) = 1 - \phi_1 L^1 - \phi_2 L^2 - \phi_3 L^3 - \dots - \phi_p L^p \quad (5)$$

Moving Average (MA) Model

MA model is the regression of current errors and past few forecast errors. MA process of order q each observation y_t is generated by a weighted average of random disturbances going back from q periods backwards. The moving average model of order q is denoted by $MA(q)$ and it is defined as

$$y_t = \theta(L)\varepsilon_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (6)$$

where $\theta_1, \theta_2, \dots, \theta_q$ are the parameters of the MA model; ε_t is the current disturbance, $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ are the past disturbances.

$\theta(L)$ is the Moving Average polynomial and it is expressed as

$$\theta(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \theta_3 L^3 - \dots - \theta_q L^q \quad (7)$$

Autoregressive Integrated Moving Average (ARIMA) Model

ARIMA model is the combination of the autoregressive (AR) model and the moving average (MA) model. The ARIMA model is usually denoted by $ARIMA(p, d, q)$. where p is the order of the AR model, d is the number of times that the actual observations need to be differenced so as to become linearly stationary, and q is the order of the moving average model. ARIMA model is given as

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (8)$$

2.3. ARIMAX Model Fitting

The ARIMAX model follows two stages: viz fitting the multiple linear regression to the response variable and the predictor variable and modelling ARIMA using Box-Jenkins approach on the residuals series.

Under Stage One: Multiple Linear Regression Fitting

Here, the first step is to consider the diagnosis of the data, making sure that all the assumptions under the multiple regression are met. Secondly, the data is modelled using multiple linear regression, and the final step is testing the parameters for significance, since ARIMAX deals with significant exogenous variables, the Residual series was obtained which will help in ARIMA modelling.

Under Stage Two: ARIMA Modelling

This stage is also divided into three major steps under the application of Box-Jenkins. But firstly, we have to check if the residual series is stationary that is it has constant mean, constant variance, and there is existence of serial correlation. If the series is not stationary, then there is need for first or two differencing, so as to bring the series to stationarity. The differencing of the actual observations y_t is given as $\Delta^d y_t$, where d is the number of differencing to attain stationarity. The first and second difference are given as:

First differencing of y_t is $\Delta^1 y_t = y'_t$ and it is defined as

$$y'_t = y_t - y_{t-1} \quad (9)$$

Second difference is given as $\Delta^2 y_t = y''_t$, which defined as

$$y''_t = y'_t - y'_{t-1} = (y_t - y_{t-1}) - (y_{t-1} - y_{t-2}) \quad (10)$$

First step deals with the model identification and it involves the process of selecting the appropriate ARIMA by selecting the appropriate orders of AR and/or MA. In selecting the ARIMA model, the Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) play a very important role. The ACF measures the amount of linear dependence between observations in a time series that are separated by lag k , while the PACF plot shows the number of AR terms.

Second step deals with the estimation of the parameters, Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF).

Third step deals with diagnostic check for model adequacy. The Akaike Information Criteria (AIC) and/or Bayesian Information Criteria (BIC) is/are used to check for model adequacy. The AIC is expressed as

$$AIC = n \log(\hat{\sigma}^2) + 2k \quad (11)$$

k is the number of model parameters; $\hat{\sigma}^2$ is the residual sum of squares, and n is the sample size

Bayesian Information Criterion (BIC) is expressed as

$$BIC = n \log(\hat{\sigma}^2) + k \log(n) \quad (12)$$

The model with the lowest Akaike' Information Criterion (AIC) or/and Bayesian Information Criterion (BIC) is considered the best model among others.

Autoregressive (AR) model with significant exogenous variables (ARX)

$$Y_t = \phi(L)Y_t + \beta X_t + \varepsilon_t \quad (13)$$

Moving Average (MA) model with significant exogenous variables (MAX)

$$Y_t = \beta X_t + \theta(L)\varepsilon_t \quad (14)$$

Autoregressive Moving Average with exogenous variable

$$\phi(L)Y_t = \beta X_t + \theta(L)\varepsilon_t \quad (15)$$

where X_t is the exogenous variable at time t , β is the coefficient of the exogenous variable.

2.4. Measure of Forecast Accuracy

The measure of forecast accuracy adopted in this study is Theil's U Forecast Accuracy. The Theil's U shows how the forecast conforms to the values of the future periods. It is defined as

$$U = U_1 = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n Y_t^2 + \frac{1}{n} \sum_{t=1}^n \hat{Y}_t^2}} \quad (16)$$

where Y_t is the actual value of a point for a given time period t , \hat{Y}_t is the forecast value, n is the number of the data points. U_2 is the measure of forecast quality and shows how adequate the ARIMAX model is. It is defined as

$$U = U_2 = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{Y}_t - Y_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n Y_t^2}} \quad (17)$$

when $U = 0$, it means the ARIMAX model obtained forecasts perfectly; $U > 1$, it means the ARIMAX model obtained does not forecast as well as the naïve model; $U < 1$, it means the ARIMAX model obtained forecasts better than the naïve model; and $U = 1$, it means the ARIMAX model obtained forecasts as well as naïve model.

3. Results/Findings

Figure 1 shows the timeplot of GDP, Exchange Rate, Interest Rate, Export, and Import from 1990 to 2019.

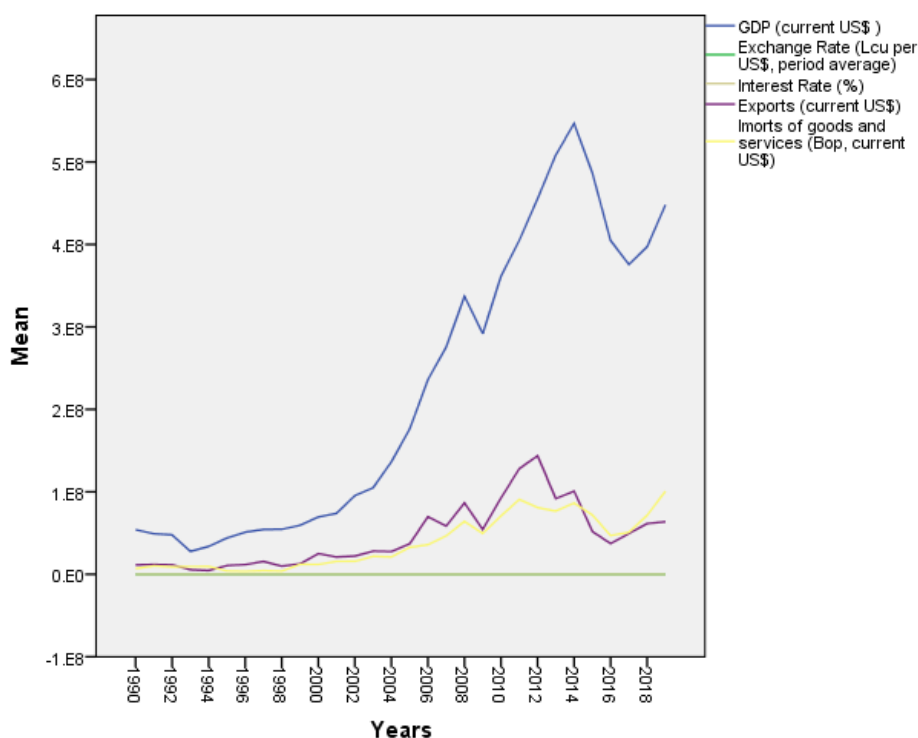


Figure 1. Timeplot for GDP, Exchange Rate, Real Interest Rate, Exports, and Import.

The estimated coefficients of the predictor variables (Exchange Rate, Interest Rate, Export and Import) and the variance inflation factor (VIF) are shown in Table 1.

Table 1. Multicollinearity Check for Exchange Rate, Interest Rate, Export and Import.

Coefficients ^a							
Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	5136080.601	17543132.261		.293	.772	
	Exchange Rate	292490.673	186710.854	.146	1.567	.130	.326
	Interest Rate	937678.271	962922.271	.056	.974	.339	.855
	Export	.367	.611	.078	.601	.553	.167
	Import	4.274	.961	.757	4.446	.000	.098
a. Dependent Variable: Gross Domestic Product (US\$)							

a. Dependent Variable: Gross Domestic Product (US\$)

Table 1 showed that the VIF for Exchange Rate and Interest Rate is less than 4, while the VIF of Export and Import is greater than 4, this implies that there is an existence of multicollinearity in the predictor variables. However, from

the correlation matrix of Table 2, Export and Import are highly correlated, and Export is the cause of the multicollinearity and it is discarded from the predictor variable list.

Table 2. Correlation Matrix for Exchange Rate, Interest Rate, Export and Import.

Coefficient Correlations ^a						
Model		Import	Interest Rate	Exchange Rate	Export	
1	Correlations	Import	1.000	-.070	-.723	-.872
		Interest Rate	-.070	1.000	-.103	-.017
		Exchange Rate	-.723	-.103	1.000	.470
		Export	-.872	-.017	.470	1.000
	Covariances	Import	.924	-64891.658	-129701.062	-.512
		Interest Rate	-64891.658	927219300780.261	-18465910304.089	-10001.814
		Exchange Rate	-129701.062	-18465910304.089	34860942819.884	53579.124
		Export	-.512	-10001.814	53579.124	.373
a. Dependent Variable: Gross Domestic Product (US\$)						

a. Dependent Variable: Gross Domestic Product (US\$)

The estimated coefficients of the predictor variables (Exchange Rate, Interest Rate, and Import) and the variance inflation factor (VIF) are shown in Table 3.

Table 3. Multicollinearity Check and Estimated Coefficients of Exchange Rate, Interest Rate and Import.

Coefficients ^a									
Model		Unstandardized Coefficients		Standardized Coefficients		T	Sig.	Collinearity Statistics	
		B	Std. Error	Beta				Tolerance	VIF
1	(Constant)	9032802.234	16098726.524			.561	.580		
	Exchange rate	239791.310	162788.383	.120		1.473	.153	.418	2.390
	Interest rate	947515.857	950878.295	.057		.996	.328	.855	1.169
	Import	4.778	.464	.846		10.294	.000	.411	2.436
a. Dependent Variable: GDP									

a. Dependent Variable: GDP

Table 3 showed that the VIF of Exchange Rate, Interest Rate and Import is less than 4, which implies that there is no existence of multicollinearity. Thus, multiple linear regression can then be modelled to the data of interest.

Again, inferences on the parameters showed that Exchange

Rate and Interest Rate are not significant, while the Import is significant with a p-value less than 0.05. Then the regression model is given as

$$\hat{Y} = 9032802.234 + 4.778X_4 \quad (18)$$

Table 4. Estimated ACF and PACF for the Residuals.

Series: Residuals					
Lag	Autocorrelation	Partial Autocorrelation	Box-Ljung Statistic		
			Value	Df	Sig. ^b
1	.700	.700	16.218	1	.000
2	.384	-.207	21.283	2	.000
3	.194	.028	22.617	3	.000
4	.037	-.120	22.668	4	.000
5	-.127	-.155	23.287	5	.000
6	-.085	.247	23.577	6	.001
7	.064	.129	23.746	7	.001
8	.179	.083	25.142	8	.001
9	.118	-.234	25.776	9	.002
10	.053	-.035	25.911	10	.004
11	-.006	-.016	25.912	11	.007
12	-.110	-.051	26.554	12	.009
13	-.208	-.003	28.986	13	.007
14	-.184	-.017	31.016	14	.006
15	-.181	-.192	33.106	15	.005
16	-.114	.135	33.997	16	.005

a. The underlying process assumed is independence (white noise).

b. Based on the asymptotic chi-square approximation.

Table 4 showed that the Q-statistic of the 16th lag has a p-value 0.005, which is less than 0.05, implying that the residual series do not have constant mean and constant variance. Thus, Residual Series is not stationary, however,

requires to be differenced. The estimated ACF and PACF for the first difference Residuals and the Box-Ljung statistic (Q-statistic) are shown in Table 5.

Table 5. Estimated ACF and PACF for the First Difference.

Series: Residuals					
Lag	Autocorrelation	Partial Autocorrelation	Box-Ljung Statistic		
			Value	Df	Sig. ^b
1	.167	.167	.895	1	.344
2	-.102	-.134	1.241	2	.538
3	.012	.056	1.246	3	.742
4	-.038	-.067	1.298	4	.862
5	-.384	-.376	6.818	5	.235
6	-.180	-.066	8.078	6	.232
7	-.056	-.128	8.208	7	.315
8	.171	.218	9.463	8	.305
9	.076	-.016	9.724	9	.373
10	.048	-.070	9.831	10	.455
11	.039	-.061	9.906	11	.539
12	.035	-.060	9.971	12	.619
13	-.217	-.129	12.622	13	.477
14	-.020	.120	12.646	14	.555
15	-.101	-.174	13.306	15	.579
16	-.023	.019	13.342	16	.648

a. The underlying process assumed is independence (white noise).
b. Based on the asymptotic chi-square approximation.

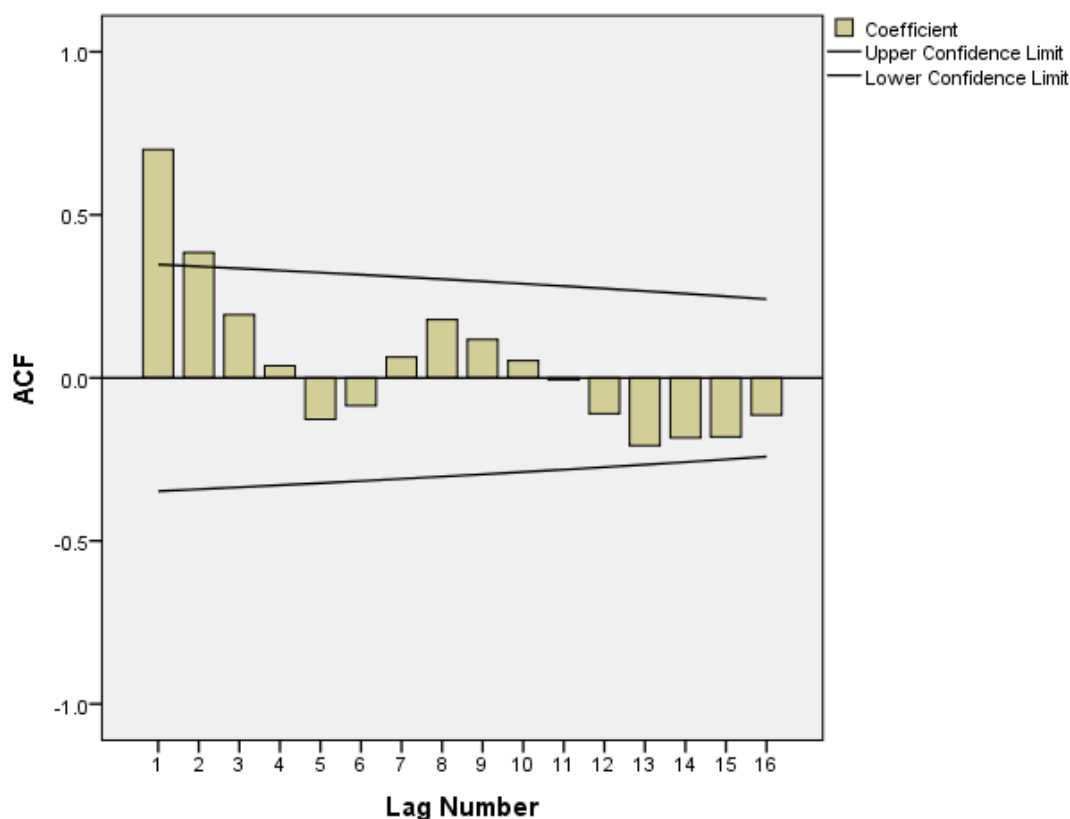


Figure 2. ACF Plot for Residuals.

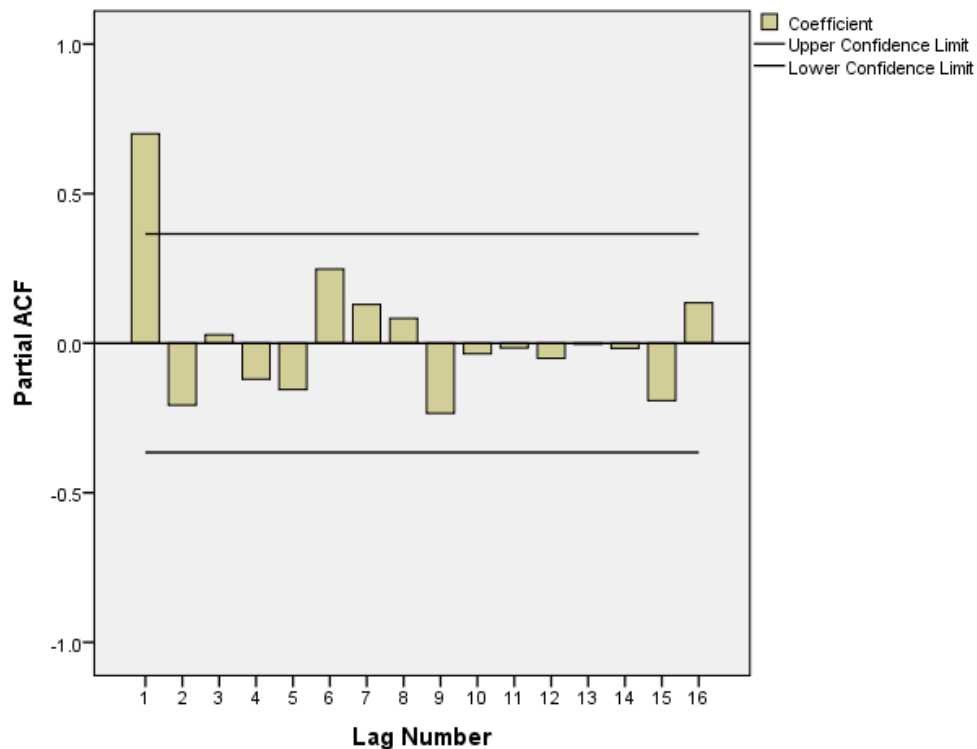


Figure 3. PACF Plot for Residuals.

The Q-statistic of the 16th lag of the first differenced Residual Series has a p-value 0.005 which is less than 0.05, denoting constant mean and variance as shown in Table 5, implying that the first difference Residual series is stationary. Hence there is need to estimate ACF and PACF.

The correlogram of ACF and PACF shown in Figure 2 and 3 respectively indicates that the actual residual series is not

stationary based on the lags which do not show a rapid fall, as indicated by Table 5.

Figures 4 and 5 shows the correlogram of the first difference ACF and PACF respectively. The first difference ACF shows a rapid fall of the lags indicating that the first difference residual series is stationary. PACF plot is the basis for the model selection.

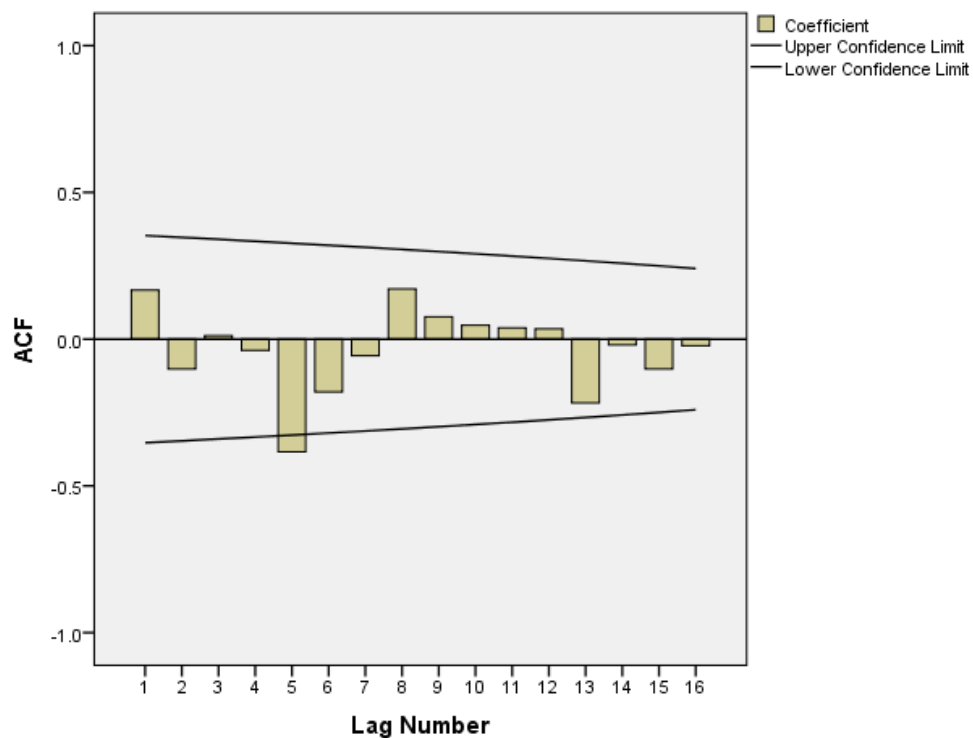


Figure 4. ACF Plot for the First Differenced Residuals.

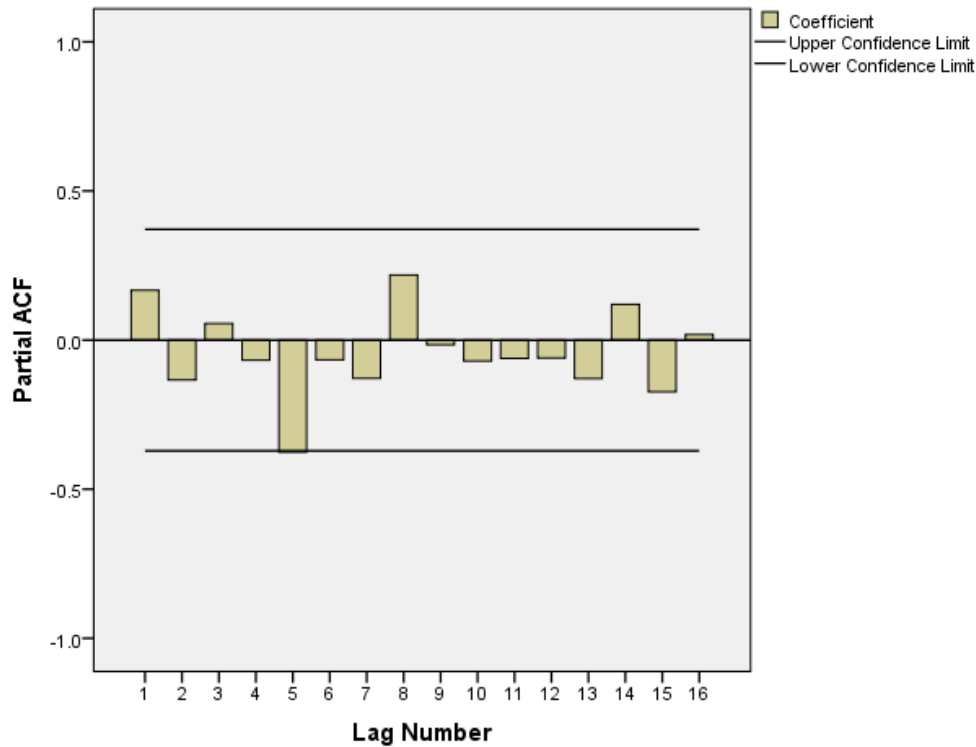


Figure 5. PACF Plot for the First Differenced Residuals.

Table 6 shows the various ARIMA models selected from Table 5 under the range of exploration, with Bayesian Information Criteria (BIC), and Q-statistic.

Table 6. Model Comparison within the Range of Exploration using BIC.

ARIMA (p, d, q)	BIC	Ljung-Box (Q)	Sig.
ARIMA (1, 1, 0)	35.264	12.559	0.765
ARIMA (2, 1, 0)	35.401	13.871	0.608
ARIMA (3, 1, 0)	35.552	14.581	0.482
ARIMA (4, 1, 0)	35.699	15.278	0.359
ARIMA (5, 1, 0)	35.592	10.132	0.683
ARIMA (0, 1, 1)	35.253	12.904	0.743
ARIMA (0, 1, 2)	35.401	13.989	0.600
ARIMA (0, 1, 3)	35.383	15.535	0.414
ARIMA (0, 1, 4)	35.535	15.693	0.332
ARIMA (0, 1, 5)	35.544	10.544	0.649

ARIMA (0, 1, 1) has the lowest BIC value 35.253, implying that the best ARIMA model is the ARIMA (0, 1, 1) as shown in Table. The selected model ARIMA (0, 1, 1) has a Q-statistic of 12.904 and a p-value of 0.743, which is greater than 0.05, implying that the model is adequate and that an exogenous variable can be added to it. Hence, the best ARIMAX model is ARIMAX (0, 1, 1). Table 7 shows the estimated parameters of ARIMAX (0,1,1).

Table 7. Estimated ARIMA (0, 1, 1) Parameters.

	Estimate	T	Sig.
Difference	1		
MA Lag1	-0.244	-1.249	0.222

The moving average model has an estimated parameter is -0.244, with a p-value of 0.222, which shows that the estimated parameter of MA is significant; the ARIMAX (0, 1,

1) is given as

$$\hat{Y}'_t = \hat{Y}_t(ARIMAX) = 4.778X_{t-1} - .244\varepsilon_{t-1} \quad (19)$$

The estimated ARIMAX is given in Appendix 1. The Theil's U Forecast Accuracy is given as

$$U = U_1 = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n y_t^2 + \frac{1}{n} \sum_{t=1}^n \hat{y}_t^2}} \quad (20)$$

$$U = U_1 = \frac{91084460.06}{282509083.5 + 206115229.5} = 0.18641$$

The result of Theil's U_1 forecast accuracy in equation (20) which is 0.1861 is less than 1, thus, it implies that the proposed ARIMAX (0,1,1) is a good model and the quality of the forecast 0.05886 given by Theil's U_2 shows that the ARIMAX (0,1,1) model obtained is adequate.

$$U = U_2 = \frac{\sqrt{\frac{1}{n} \sum_{t=1}^n (\hat{y}_t - y_t)^2}}{\sqrt{\frac{1}{n} \sum_{t=1}^n y_t^2}} = \frac{91084460.06}{1547365977} = 0.05886 \quad (21)$$

4. Conclusion

The aim of this paper is to identify an adequate ARIMAX model that will be used to forecast Nigeria's GDP. Multiple linear regression was fitted to the data under the conditions that the assumptions were met, and the stages of ARIMA model fitting on the residuals were reached and the explored models were compared based on the Bayesian Information Criteria (BIC), and ARIMA (0, 1, 1) was selected. The ARIMA model and the exogenous variable obtained were

combined together. The Theil's U statistic showed that ARIMAX (0,1,1) is appropriate for forecast the Nigeria's GDP.

The fitted ARIMAX (0,1,1) model proposed in this study can play an important role in Monetary Policy. The Central Bank of Nigeria can use the proposed model to predict

future short-term macroeconomic conditions and decide about their future operations. The ARIMAX (0,1,1) will help to know whether it is justified to continue or withdraw the Monetary Policy, and it will act as a useful tool in the decision making process of Monetary authorities so as to optimize resources.

Appendix

Table 8. Computation of Theil's U Forecast Accuracy.

Year	Actual GDP (y_t)	$\hat{Y}_t(ARIMAX)$	$(y_t - \hat{Y}_t)^2$	y_t^2	\hat{Y}_t^2
1990	54035795.39	0	2.91987E+15	2.91987E+15	0
1991	49118433.05	46103135.04	3.62318E+14	2.41262E+15	9.05034E+14
1992	47794925.81	45138953.93	1.16743E+13	2.28435E+15	2.62264E+15
1993	27752204.32	45877306.03	2.63346E+14	7.70185E+14	1.93425E+15
1994	33833042.99	51854473.27	3.02676E+14	1.14467E+15	2.62458E+15
1995	44062465.80	24136872.69	4.1738E+13	1.9415E+15	2.55257E+15
1996	51075815.09	12976250.87	1.28688E+15	2.60874E+15	2.31122E+14
1997	54457835.19	14571321.50	1.91285E+15	2.96566E+15	1.14956E+14
1998	54604050.17	12183067.20	1.5935E+15	2.9816E+15	2.15661E+14
1999	59372613.49	50965461.17	2.28826E+15	3.52511E+15	1.33098E+14
2000	69448756.93	59201130.54	1.00525E+14	4.82313E+15	3.53104E+15
2001	74030364.47	74455757.17	3.00739E+14	5.48049E+15	3.21359E+15
2002	95385819.32	77966427.27	3.13742E+14	9.09845E+15	6.0331E+15
2003	104911947.83	101825128.60	1.02943E+15	1.10065E+16	5.30381E+15
2004	136385979.32	102351449.20	8.88621E+14	1.86011E+16	1.13585E+16
2005	176134087.15	149277405.10	6.80512E+15	3.10232E+16	8.76863E+15
2006	236103982.43	168848974.80	6.88067E+15	5.57451E+16	2.34562E+16
2007	275625684.97	209328362.20	1.3825E+16	7.59695E+16	2.49785E+16
2008	337035512.68	296083151.50	1.55844E+16	1.13593E+17	4.5028E+16
2009	291880204.33	231243351.50	9.38365E+13	8.51941E+16	9.09427E+16
2010	361456622.22	327209423.50	1.85925E+16	1.30651E+17	5.0671E+16
2011	404993594.13	430423267.20	4.87856E+15	1.6402E+17	1.12323E+17
2012	455501524.58	395652780.30	1.55098E+14	2.07482E+17	1.96291E+17
2013	508692961.94	351599882.10	1.87494E+16	2.58769E+17	1.38209E+17
2014	546676374.57	380495017.40	4.53585E+16	2.98855E+17	1.11356E+17
2015	486803295.10	313363203.10	1.0849E+16	2.36977E+17	1.46417E+17
2016	404650006.43	191860317.60	8.75768E+15	1.63742E+17	9.6763E+16
2017	375746469.54	201624696.60	3.72209E+16	1.41185E+17	3.34228E+16
2018	397190484.46	312185807.10	3.38283E+16	1.5776E+17	4.54823E+16
2019	448120428.86	470525571.30	1.36961E+16	2.00812E+17	1.09621E+17
Total			2.49E+17	2.39E+18	1.2745E+18

References

- [1] ABONAZEL, M. R., ABD-ELFTAH, A. J., Forecasting Egyptian GDP using ARIMA Models, Reports on Economics and Finance, 2019, Vol. 5, No. 1, 35-47.
- [2] ARNEJA, R. S., KAUR N., SAIHHJPAL, Analyses and Forecasting Evaluation of GDP of India using ARIMA Model, International Journal of Advanced Science and Technology, 2020, vol. 29, No. 11s, 1102-1107.
- [3] ATANU ENEBI Y., ETTE HARRISON E., NWUJU KINGDOM, NWAHOA WILLIAN C., ARIMA Model for Gross Domestic Product (GDP): Evidence from Nigeria, Archives of Current Research International, 2020, 20 (7), 49-61.
- [4] AWEL, Y. M., Forecasting GDP Growth: Application of Autoregressive Integrated Moving Average Model, Empirical Economic Review, 2018, Vol. 1, No. 2, 1-16.
- [5] CHIKUMBE, E. S., SIKOTA, S., Forecasting Zambia's Gross Domestic Product using Time Series Autoregressive Integrated Moving Average (ARIMA) Model, International Journal of Innovation Science and Research Technology, 2020, Vol. 5, Issue 9.
- [6] DURKA, P., PASTOREKOVA, S., ARIMA vs ARIMAX – which approach is better to analyze and forecast macroeconomic time series? Proceedings of 30th international conference on mathematical methods in economics, 2012, 136-140.
- [7] EKHOSUEHI, V. U., OMOREGIE, D. E., Inspecting debt servicing mechanism in Nigeria using ARMAX model of the koyck-kind, Operations Research and Decisions, 2021, 31 (1), 5-20. DOI: 10.37190/ord210101.
- [8] GHAZO ABDULLAH, Applying the ARIMA Model to the Process of Forecasting GDP and CPI in the Jordanian Economy, International Journal of Financial Research, 2020, Vol 12, No 3.

- [9] LIU MEI, LI H., ZHOU Y., WANG W., XING F., Substructural damage detection in shear structures via ARMAX model and optimal subpattern assignment distance, *Engineering Structures*, 2019, 191, 625-639.
- [10] MAITY, B., CHATTERJEE, B., Forecasting GDP growth rates of India: An empirical study, *International Journal of Economics and Management Sciences*, 2012, 1 (9), 52-58.
- [11] MIAH, M. M., TABASSUM, M., RANA, M. S., Modeling and Forecasting of GDP in Bangladesh: An ARIMA Approach, *Journal of Mechanics of Continua and Mathematical Science*, 2019, 14 (3), 150-166.
- [12] MUSUNDI SAMMY WABOMBA, M'MUKIIRA PETER MUTWIRI, MUNGAI FREDRICK, Modeling and forecasting Kenyan GDP using autoregressive integrated moving average (ARIMA) models, *Science Journal of Applied Mathematics and Statistics*, 2016, 4 (2), 64-73.
- [13] NING, WEI, KUAN-JIANG, BIAN, ZHI-FA, YUAN, Analysis and Forecast of Shaanxi GDP Based on the ARIMA Model, *Asian Agricultural Research, USA-China Science and Culture Media Corporation*, 2010, 2 (01), 1-4.
- [14] RANA, S. B., Forecasting GDP Movements in Nepal using Autoregressive Integrated Moving Average (ARIMA) Modelling Process, *Journal of Business and Social Sciences Research*, 2019, Vol. iv, No. 2.
- [15] SALAH UDDIN, K. M, TANZIM, N., Forecasting GDP of Bangladesh using ARIMA Model, *International Journal of Business and Management*, 2021, Vol. 16, No. 6.
- [16] TOUMA, H. Y., Application of the Statistical Analysis for Prediction of the Jordan GDP using ARIMA Time Series and Holt's Linear Trend Models for the Period 2003-2013, *Mathematical Theory and Modelling*, 2014, 4 (14), 19-26.
- [17] Ulyah, S. M., Mardianto, M. F., Sediono, F. (2019) Comparing the Performance of Seasonal ARIMAX Model and Nonparametric Regression Model in Predicting Claim Reserve of Education Insurance, *Journal of Physics: Conference Series* 1397 (2019) 012074 IOP Publishing doi: 10.1088/1742-6596/1397/1/012074.
- [18] Wabomba, M. S., Mutwiri, M. P., Frederick, M., Modelling and Forecasting Kenyan GDP using ARIMA models, *Science Journal of Applied Mathematics and Statistics*, 2016, 4 (2), 64-73.
- [19] ZAKAI, M., A time series modeling on GDP of Pakistan, *Journal of Contemporary Issues in Business Research*, 2014, 3 (4), 200-210.
- [20] ZHENG W. X., On least-squares identification of ARMAX models, 15th Triennial World Congress, Barcelona, Spain, 2002, 391-296.