Available online www.jsaer.com

***Journal of Scientific and Engineering Research*, 2016, 3(2):57-66**

Research Article

ISSN: 2394-2630
CODEN(USA): JSERBR

Modeling and Forecasting CPI Inflation in Nigeria: Application of Autoregressive Integrated Moving Average Homoskedastic Model

David Adugh Kuhe, Raymond Chukwuka Egemba

Department of Mathematics/Statistics/Computer Science, University of Agriculture

Abstract This study employs a univariate Autoregressive Integrated Moving Average (ARIMA) homoskedastic model in conjunction with Box and Jenkins modeling procedure to model and forecast annual Consumer Price Index (CPI) data in Nigeria from 1950 to 2014. The annual data on Consumer Price Index is obtained as secondary data from Penn World Table, the National Bureau of Statistics and the Central bank of Nigeria over the period 1950 to 2014. We examine the graphical, statistical, unit root and stationarity properties of the series using time plots, ACF, PACF, Phillips-Perron as well as Dickey-Fuller Generalized Least Squares unit root tests. The results show that the CPI data in Nigeria is non-stationary in level but stationary in logged first difference and thus integrated of order one, I(1). We then applied Box-Jenkins modeling methodology to search for an optimal model and found that ARIMA (3, 1, 0) was the best fitting model to describe CPI data series in Nigeria. The model was validated and found to be adequate and good. Based on this model, we forecast the future annual CPI in Nigeria for a period of 6 years from 2015 to 2020. The forecasts show a steady increase in the annual values of CPI in Nigeria. The study predicts that inflation will increase in Nigeria from 2015 since the confidence intervals of the forecast suggest a consistent increase in annual CPI during the forecasted period of 2015 to 2020.

Keywords Consumer Price Index, Inflation, ARIMA model, Forecasting, Nigeria.

Introduction

Consumer Price Index (CPI) is the most widely used measure of inflation in financial analysis. A price index is a weighted average of the prices of a selected basket of goods and services relative to their prices in some base-year. A consumer price index measures changes in the price level of a market basket of consumer goods and services purchased by households. The CPI is a statistical estimate constructed using the prices of a sample of representative items whose prices are collected periodically. The CPI represents prices paid by consumers (or households). Prices for a basket of goods are compiled for a certain base period. Price data for the same basket of goods is then collected on a monthly basis. This data is used to compare the prices for a particular month with the prices from a different time period. The Bureau of Labor Statistics of the United States Department of Labor defines CPI as: "a measure of the average change over time in the prices paid by urban consumers for a market basket of consumer goods and services." It defines inflation as: "the overall general upward price movement of goods and services in an economy." The difference between the Consumer Price Index (CPI) and inflation is a source of confusion for many. The Consumer Price Index is used to calculate inflation. Thus, their similarities are better understood based on that relationship even if the details of their differences are not.

The CPI is not a perfect measure of inflation. Sources of bias include: (i) Quality adjustments - quality of many goods such as cars, computers, and televisions, etc., goes up every year. Some price increases may reflect quality adjustments that are still counted entirely as inflation. (ii) New goods - new goods may be introduced that will be hard to compare to older substitutes. (iii) Substitution - if the price goes up for one good, consumer may substitute another good that provides similar utility. Although the CPI will go higher due to the price increase in the old commodity, many consumers may not be worse off. Also, when prices go up, consumers may

effectively not pay the higher prices by switching to discount stores. The CPI surveys do not check to see if consumers are substituting discount or outlet stores. Moreover, the macroeconomic data on developing countries can be unreliable due to many reasons: measurement error, imperfect methods of measuring, etc.
The Consumer Price Index, CPI, a proxy for inflation, has been widely used as a leading indicator of economic change. Financial markets continuously assess expectations on the CPI and react to the innovations contained in new published data [1]. Inflation is most likely to affect interest rates, stock prices and exchange rates. Unexpected inflation causes bond prices to drop and yields to rise. An increasing interest rate will negatively affect stock prices. Unexpected inflation also decreases the value of a country's currency in the global market and impacts the exchange rate. Therefore, an effective monetary policy depends largely on the ability of economists and policy makers to develop a reliable model that could help understand the ongoing economic processes and predict future developments. In this regard, this study is important since it is aimed at forecasting CPI, which is a component of inflation in the Nigeria economy.

A Brief Review of Related Works

According to Enders (2004), there are four basic time series models that may describe the behaviour of a dataset [2]. These are Autoregressive (AR), Moving Average (MA), Autoregressive Moving Average (ARMA) and White Noise (WN) models. Wayne (1998) asserts that using vector autoregressive model in forecasting exhibits significant degree of forecast accuracy when compared with other forecasting models [3]. This same conclusion was reached by Meyler *et al.* [4]. They Applied the Bayesian VAR approach in forecasting and found that VAR modeling approach improves forecasting performance. Liu and Han (2007) used ARIMA (2, 2, 1) to model and forecast US CPI for the year 2007. They employed annual CPI data from 1913 to 2006. Their prediction shows an annual increase in the value of CPI for US in 2007 [5].
Adams, *et al.* (2014) fitted a time series model to the quarterly data of consumer price index (CPI) in Nigeria’s Inflation rate between 1980 and 2010 and provided five years forecast for the expected CPI in Nigeria. They applied the Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model and found that the best fitted model is ARIMA (1, 2, 1). The five years forecast showed an average increment of about 2.4% between 2011 and 2015 with the highest CPI being estimated as 279.90 in the 4th quarter of the year 2015 [6]. Kelikume and Salami, (2014) compared Autoregressive Integrated Moving Average model developed by Box and Jenkins and multivariate time series model in the form of Vector Autoregressive model to forecast inflation for Nigeria. They employed changes in monthly consumer price index for the period 2003 to 2012 to predict movements in the general price level. They found that VAR model described inflation situation in Nigeria than ARIMA model. Their forecast shows a gradual increase in inflation over the forecast period [7].
Many studies on inflation forecasting in different part of the world yielded mixed results. Fritzer *et al.* (2002), found that VAR models perform better than ARIMA models in terms of predictive accuracy [8] whereas Bokhari and Feridun (2006), revealed that ARIMA models perform better than the VAR model [9]. Espasa *et al.* (2002), concludes that ARIMA models outperformed the VECM and dynamic factor models [1] while Hubrich (2003), found that VAR models outperformed the autoregressive forecasting models [10]. Alnaa and Ahiakpor, (2005) conducted a study that followed the same pattern as other models proving the VAR modeling approach to be highly efficient in its predictive ability [11].
There are several evidences in the literature supporting the forecasting strength of ARIMA model approach using Box-Jenkins procedure in forecasting [12-15]. Although, recent studies in Nigeria, have shown the VAR modeling and forecasting approach to be highly useful in predicting short run forecast [16-17], there is however a need to revisit the modeling and forecasting ability of ARIMA model in Nigeria using more recent data.

Materials and Methods

The theoretical model which serves as a basic framework of our analysis is the Autoregressive Integrated Moving Average (ARIMA) model which is a generalization of Autoregressive Moving Average (ARMA) model. The ARMA (p, q) is given by:

$$CPI_t = \alpha_0 + \alpha_1 CPI_{t-1} + \cdots + \alpha_p CPI_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \quad (1)$$

Where CPI is the consumer price index at levels, $\alpha_i's$ are coefficients of AR (p) process while $\theta_i's$ are coefficients of MA (q) process, ε_t is the error term. AR (p) is the autoregressive process of order p whereas MA (q) is the moving average process of order q. Equation (1) can also be written as:

$$\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right) CPI_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i\right) \varepsilon_t \quad (2)$$

Where L is the lag operator, define by $LCPI_t = CPI_{t-1}$, the α_i are the parameters of the autoregressive part of the model, the θ_i are the parameters of the moving average part and ε_t are error terms which are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean.
Assuming that the polynomial $\left(1 - \sum_{i=1}^{p} \alpha_i L^i\right)$ has a unitary root of multiplicity d, then it can be written as:

$$\left(1-\sum_{i=1}^{p}\alpha_i L^i\right)=\left(1+\sum_{i=1}^{p-d}\theta_i L^i\right)(1-L)^d \quad (3)$$

An ARIMA (p,d,q) process expresses this polynomial factorization property, and is given by:

$$\left(1-\sum_{i=1}^{p}\Phi_i L^i\right)(1-L)^d CPI_t=\left(1+\sum_{i=1}^{q}\theta_i L^i\right)\varepsilon_t \quad (4)$$

Autoregressive Model

The Autoregressive model is a dynamic model in which the independent variables include lagged values of the dependent variable. The AR (p) model specifies the dependent variable as a function of "p" past values of itself. It can be expressed as

$$CPI_t=\alpha_0+\alpha_1 CPI_{t-1}+\alpha_2 CPI_{t-2}+3CPI_{t-3}+\cdots+\alpha_p CPI_{t-p}+\varepsilon_t \quad (5)$$

where ε_t is identically and independently distributed and E$[\varepsilon_t]=0$ for all t, $E[(\varepsilon_t),(\varepsilon_{t-j})]=0$, for all $t\neq j$ and var$[\varepsilon_t]$= σ^2 is a constant. An AR (p) process is stationary if $\sum\alpha_i<1$.

Moving Average Model

A Moving Average process is a process in which the value of a variable in the current period is a function of the value of shocks (innovations) from one or more past periods. The MA (q) model specifies the dependent variable as a function of "q" past shocks. It can be expressed as

$$CPI_t=\theta_0-\theta_1\varepsilon_{t-1}-\theta_2\varepsilon_{t-2}-\theta_3\varepsilon_{t-3}-\cdots-\theta_q\varepsilon_{t-q}+\varepsilon_t \quad (6)$$

where ε_t is identically and independently distributed and E$[\varepsilon_t]=0$ for all t, $E[(\varepsilon_t),(\varepsilon_{t-j})]=0$, for all $t\neq j$ and var$[\varepsilon_t]$= σ^2 is a constant. An MA process is always stationary.

Stationarity

Stationarity of Order M: A time series $\{Y_t\}$ is stationary of order M if for any admissible set $\{t_1,t_2,\dots,t_m\}$ and for any k, the joint moments of $\{Y_{t1},Y_{t2},\dots,Y_{tm}\}$ up to order M exists, and are equal to the joint moments of $\{Y_{t1+k},Y_{t2+k},\dots,Y_{tm+k}\}$ up to order M. That is $E\{(Y_{t1})^\alpha(Y_{t2})^\beta\dots(Y_{tm})^\gamma\}=E\{(Y_{t1+k})^\alpha(Y_{t2+k})^\beta\dots(Y_{tm+k})^\gamma\}$ for all $\alpha,\beta,\dots,\gamma$ such that $\alpha+\beta+\dots+\gamma\leq M$.

Weakly or Covariance Stationary: A time series $\{Y_t\}$ is said to be weakly or covariance stationary if its mean and variance are constant over time and its covariance function depends only on the time lag. A covariance stationary series satisfies the following conditions:

(i) E$(Y_t)=\mu$, where μ is a constant (ii) E$(Y_t-\mu)^2=$ Var$(Y_t)=\sigma^2$, where σ^2 is a constant and (iii) $E(Y_t,Y_s)=E(Y_t,Y_{t+k})$ is a function of $s-t=k$ only where k is the lag.

A series which becomes stationary after first differencing is said to be integrated of order one, denoted I(1).

The Phillips-Perron (PP) Unit Root Test

Phillips and Perron (1988) propose a nonparametric method of controlling for serial correlation when testing for a unit root. The PP method estimates the non-augmented Dickey-Fuller test equation [18]:

$$\Delta Y_t=\alpha Y_{t-1}+x_t'\delta+\varepsilon_t \quad (7)$$

Where $\alpha=\rho-1$ and modifies the t-ratio of the α coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic. The PP test is based on the statistic:

$$\tilde{t}_\alpha=t_\alpha\sqrt{\left(\frac{\gamma_0}{f_0}\right)}-\frac{T(f_0-\gamma_0)(se(\tilde{\alpha}))}{2\sqrt{f_0}\,s} \quad (8)$$

Where $\tilde{\alpha}$ is the estimate, and t_α the t-ratio of $\alpha, se(\tilde{\alpha})$ is coefficient standard error, and s is the standard error of the test regression. In addition, γ_0 is a consistent estimate of the error variance in (7) (calculated as $(T-k)s^2/T$ where k is the number of regressors).The remaining f_0 term, is an estimator of the residual spectrum at frequency zero. The PP test evaluate the pair of hypothesis $H_0:\alpha=0$ against $H_1:\alpha<0$

Dickey-Fuller Generalized Least Squares Unit Root Test

The DFGLS test involves estimating the standard ADF test equation

$$\Delta Y_t=\alpha Y_{t-1}+x_t'\delta+\beta_1\Delta Y_{t-1}+\beta_2\Delta Y_{t-2}+\cdots+\beta_p\Delta Y_{t-p}+v_t \quad (9)$$

after substituting the GLS detrended Y_t^d for the original Y_t:

$$\Delta Y_t^d=\alpha Y_{t-1}^d+\beta_1\Delta Y_{t-1}^d+\cdots+\beta_p\Delta Y_{t-p}^d+v_t \quad (10)$$

Since the Y_t^d are detrended, we do not include the x_t in the DFGLS test equation. As with the ADF test, we consider the t-ratio for $\tilde{\alpha}$ from this test equation. ERS (1996) define GLS detrended data Y_t^d using the estimates associated with $\bar{a}$:

$$Y_t^d\equiv Y_t-x_t'(\bar{a}) \quad (11)$$

where $\bar{a} = \begin{cases} 1 - 7/T, & \text{if } x_t = \{1\} \\ 1 - 13.5/T, & \text{if } x_t = \{1, t\} \end{cases}$ (12)

Information Criteria for Model Order Selection

The most common approach for model order selection involves selecting a model order that minimizes one or more information criteria evaluated over a range of model orders. The information criteria used in this work for model order selection are Akaike Information criterion (AIC) [19], Schwarz information Criterion (SIC), and Hannan-Quinn Criterion (HQC), [20]. Each criterion is a sum of two terms, one that characterizes the entropy rate or prediction error of the model, and a second term that characterizes the number of freely estimated parameters in the model (which increases with increasing model order). By minimizing both terms, we seek to identify a model that does not over-fit the data with too many parameters while also accurately modeling the data. The information criteria are given below:

$$AIC = log\left(\frac{RSS}{n}\right) + \left(2 \times \frac{k}{n}\right) \quad (13)$$

$$SIC = \log\left(\frac{RSS}{n}\right) + \left(\log(n) \times \frac{k}{n}\right) \quad (14)$$

$$HQC = \log\left(\frac{RSS}{n}\right) + \left(2 \times \log(\log(n)) \times \frac{k}{n}\right) \text{ and} \quad (15)$$

Where n is the number of observations; k is the number of free parameters to be estimated, RSS is the residual sum of squares.

Forecast Evaluation

Suppose the forecast sample is $j = T+1, T+2, \dots, T+h$, and denote the actual and forecasted value in period t as y_t and $\hat{y}_t$, respectively. The reported forecast error statistics are computed as follows:

$$\text{Root Mean Square Error (RMSE)} = \sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}$$

$$\text{Mean Absolute Error(MAE)} = \sum_{t=T+1}^{T+h} |\hat{y}_t - y_t| / h$$

$$\text{Mean Absolute Percentage Error(MAPE)} = 100 \times \sum_{t=T+1}^{T+h} \left|\frac{\hat{y}_t - y_t}{y_t}\right| / h$$

$$\text{Theil Inequality Coefficient(TIC)} = \frac{\sqrt{\sum_{t=T+1}^{T+h} (\hat{y}_t - y_t)^2 / h}}{\sqrt{\sum_{t=T+1}^{T+h} \hat{y}_t^{\,2} / h} + \sqrt{\sum_{t=T+1}^{T+h} y_t^{\,2} / h}}$$

The smaller the error, the better the forecasting ability of the model according the criterion. The Theil Inequality Coefficient always lies between zero and one, the zero value indicates a perfect fit.

Results and Discussion

Graphical and Statistical Properties of the Series

In this study we employ the Box-Jenkins modeling procedure to fit ARIMA model to annual time series CPI data in Nigeria from 1950-2014. The time plot of the original series is graphed in Figure 1.

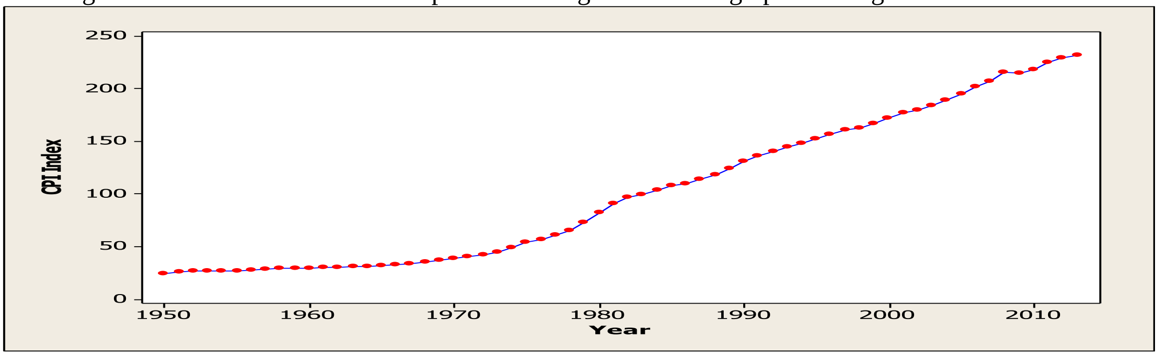

Figure 1: Plot of CPI Inflation in Nigeria in Level

We linearize the exponential growth in the original series and also stabilize the variance of the annual changes in the CPI by applying a natural log transformation. The time plot of natural log transform is represented in Figure 2.

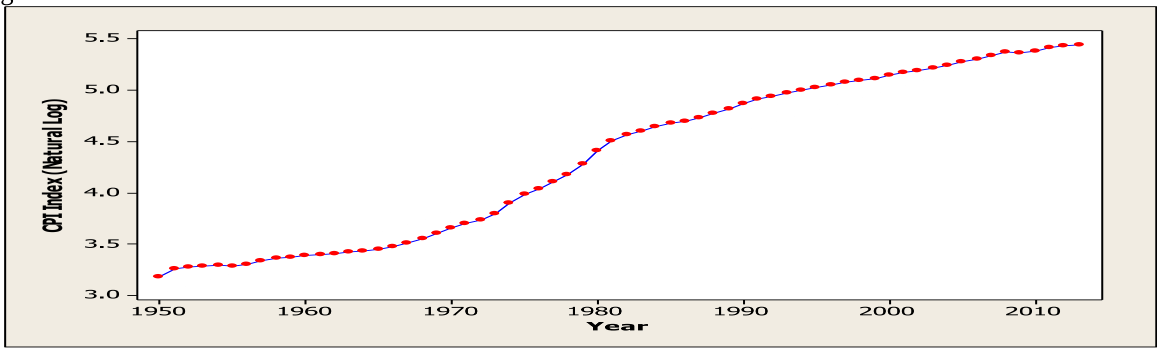

Figure 2: Plot of CPI Inflation in Nigeria (Natural Log)

From Figure 2, we observe that the variance and mean of the series seemed to be more stable in log units than in the original level form. However, the mean and variance of the series appeared to be changing with time. This suggests that the CPI series is not covariance stationary. This call for an alternative way of stationarizing the series and we therefore applied the first difference operator on the series. The time plot of the first difference of the natural log is presented in Figure 3.

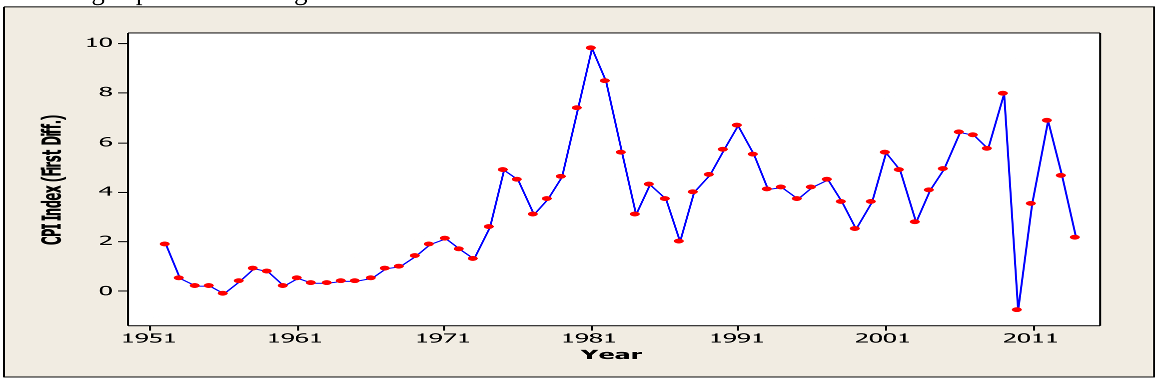

Figure 3: Plot of CPI Inflation in Nigeria (First Difference of Natural Log)

In Figure 3, the difference-logged series appeared to be more dynamically stable. This suggests a mean reverting series and homoskedasticity of the variance. This also indicates that the first difference of the CPI series is covariance stationary. We check whether the difference-logged values of the series are statistically independent by plots of ACF and PACF. The plots are presented in Figure 4.

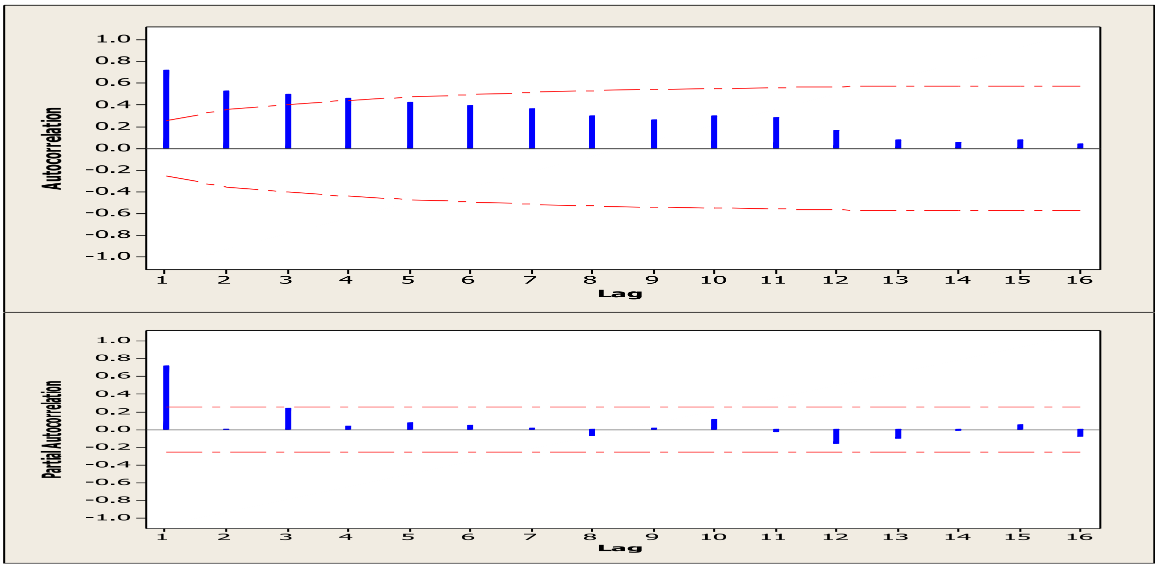

Figure 4: Plot of ACF and PACF of CPI Inflation in Nigeria (First Difference)

The plots of ACF and PACF in Figure 4 show that the difference-logged values of the series are statistically independent meaning that a CPI value in a present year is independent of the previous year. This is justified by almost all the sample autocorrelation coefficients of the series falling within the confidence bounds. Figure 4 also shows that the series is differenced stationary and that the residuals are purely random process.
We further investigate the unit root and stationarity properties as well as the order of integration of the series by applying unit root tests on both the natural log and log difference of the series. Here, we employ Phillips-Perron unit root test as well as Dickey-Fuller Generalized Least Squares unit root test already discussed in methodology. The results of the tests are reported in Table 1.

Table 1: Phillips-Perron & Dickey-Fuller GLS Unit Root Tests Result

Variable	Option	PP Adjusted t-statistic	DF-GLS Test Statistic
Y	Intercept only	0.09588[0.9639]	1.2789
	Intercept & Trend	-1.5663[0.7993]	-1.7410
ΔY	Intercept only	-4.5410[0.0003]***	-3.8416***
	Intercept & Trend	-3.4977[0.0028]***	-4.1363***

Note: *** denotes the significant of the test statistics at 1%, 5% and 10% levels. Numbers in [ ] are p-values. ΔY denotes first difference of Y.

Phillips-Perron and Dickey-Fuller GLS unit root tests are applied on both the natural log as well as the logged difference series. The Phillips-Perron and Dickey-Fuller GLS unit root results of the natural log series with intercept only and with intercept and linear time trend both indicate non-stationarity of the series meaning that the natural log of CPI data in Nigeria contains a unit root. However, the Phillips-Perron and Dickey-Fuller GLS unit root test results of the logged difference of the series with intercept only and with intercept and time trend are both sufficiently negative, thereby rejecting the presence of a unit root in the series. This means that the first logged difference of CPI data in Nigeria is stationary. This shows that the order of integration of the series is one. At this point, we proceed to model identification of the stationary series.

Model Identification

Looking carefully at the ACF and PACF plot of Figure 4, we observe that the spikes of ACF decay gradually towards zero while the spikes of PACF decay quickly to zero. This suggests an Autoregressive (AR) process. From Table 1, we noted that the order of integration, $d = 1$. Therefore we shall fit an ARIMA (p, d, q) in which the MA (q) component is zero.

A Search for an Optimal Model

The most common approach of searching for an optimal model involves selecting a model order that minimizes one or more information criteria evaluated over a range of model orders. The information criteria employed in this paper are Akaike Information criterion (AIC), Schwarz information Criterion (SIC) and Hannan-Quinn Criterion (HQC). The execution of the model is repeated for different number of lags following this procedure; we choose using parsimony the model with the least information criteria. The result is presented in Table 2.

Table 2: ARIMA Model Order Selection Using Information Criteria

Model	AIC	SC	HQC	R^2	Adj. R^2	DW
ARIMA (1,1,0)	3.6731	3.7355	3.6943	63.78	63.40	2.03
ARIMA (2,1,0)	3.7022	3.7813	3.7342	63.57	62.80	1.99
ARIMA (3,1,0)**	**3.6459**	**3.7221**	**3.6888**	66.51	65.24	2.05
ARIMA (4,1,0)	3.6630	3.7966	3.7170	67.00	65.55	2.03
ARIMA (5,1,0)	3.6663	3.8276	3.7315	67.95	66.15	1.99
ARIMA (6,1,0)	3.6954	3.8848	3.7719	68.08	65.88	1.99
ARIMA (7,1,0)	3.7178	3.9356	3.8057	68.43	65.83	1.82
ARIMA (8,1,0)	3.6564	3.8827	3.7356	70.96	68.16	1.94
ARIMA (9,1,0)	3.6567	3.9219	3.7573	71.08	67.87	2.01
ARIMA (10,1,0)	3.6596	3.9651	3.7828	71.49	67.89	1.96
ARIMA (11,1,0)	3.6918	4.0274	3.8271	71.32	67.23	1.99
ARIMA (12,1,0)	3.6993	4.0653	3.8467	71.96	67.47	1.99

Note: ** denotes ARIMA model with the least information criteria

From the result of Table 2, ARIMA (3,1,0) appears to provide statistically adequate representation of the given data because it has the least information criteria. Having chosen the best model, we now estimate the parameters of the model. The result of the parameter estimates of the optimal model is presented in Table 3.

Table 3: OLS Parameter Estimates of ARIMA (3,1,0) Model

		Dependent Variable: ΔCPI		
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**P-value**
C	2.547561	1.080430	2.357914	0.0205
AR(1)	0.752070	0.100340	7.495238	0.0000
AR(2)	-0.190040	0.125888	-1.509598	0.0503
AR(3)	0.298174	0.102271	2.915532	0.0044
R-squared				0.665071
Adjusted R-squared				0.654267
Durbin Watson Statistic				2.045956
F-statistic	61.55698		Probability [F-statistic]	0.000000
LLIK= -172.8267	AIC= 3.645912		SC= 3.722086	HQC= 3.688843

The result of the parameter estimates of Table 3 shows that our data fits an ARIMA (3,1,0) model which is presented below:

$$CPI_t = 2.547561 + 0.752070CPI_{t-1} - 0.190040CPI_{t-2} + 0.298174CPI_{t-3} + \varepsilon_t \qquad (16)$$

where CPI_t = consumer price index response variable at time t;

$y_{t-1}, y_{t-2}, y_{t-3}$ = CPI response variables at time $t-1, t-2, t-3$ respectively;

ε_t = Error term at time t;

The result of Table 3 shows that the intercept (C) is positively related with CPI and statistically significant indicating that the predicted value of CPI in Nigeria will be 254.76 if all the explanatory variables are held constant. All the coefficients of the model are significant at 5 percent levels. The coefficient of determination (R^2) of the regression model is 0.665071 indicating that about 66.51% of the total variations in CPI have been explained by the model while the remaining 33.49% unexplained variations is being accounted for by the error term or by factors not included in the model. The F-statistic is a goodness of fit test which measures the overall significance of the model parameters. F=61.55698 with a p-value of 0.000000 indicates that the model is a good fit. The Durbin Watson statistic value of 2.045956 which is higher than R^2 and R^2 adjusted means that the model is non-spurious. The estimated model have also satisfied the stationarity condition because $\alpha_1 + \alpha_2 + \alpha_3 = 0.752070 + (-0.190040) + 0.298174 = 0.860204 < 1$. This shows that the estimated ARIMA (3,1,0) is stationary. The following subsection contains residual diagnosis of the estimated ARIMA (3,1,0) model.

Model Diagnosis

We check the fitted model for adequacy and examine the goodness of fit by means of plotting the ACF and PACF of residuals of the fitted model. If all the sample autocorrelation coefficients of the residuals are within the limits $\pm 1.96/\sqrt{T}$ where T is the number of observations upon which the model is based, then the residuals are white noise indicating that the model is a good fit (See Figure 5).

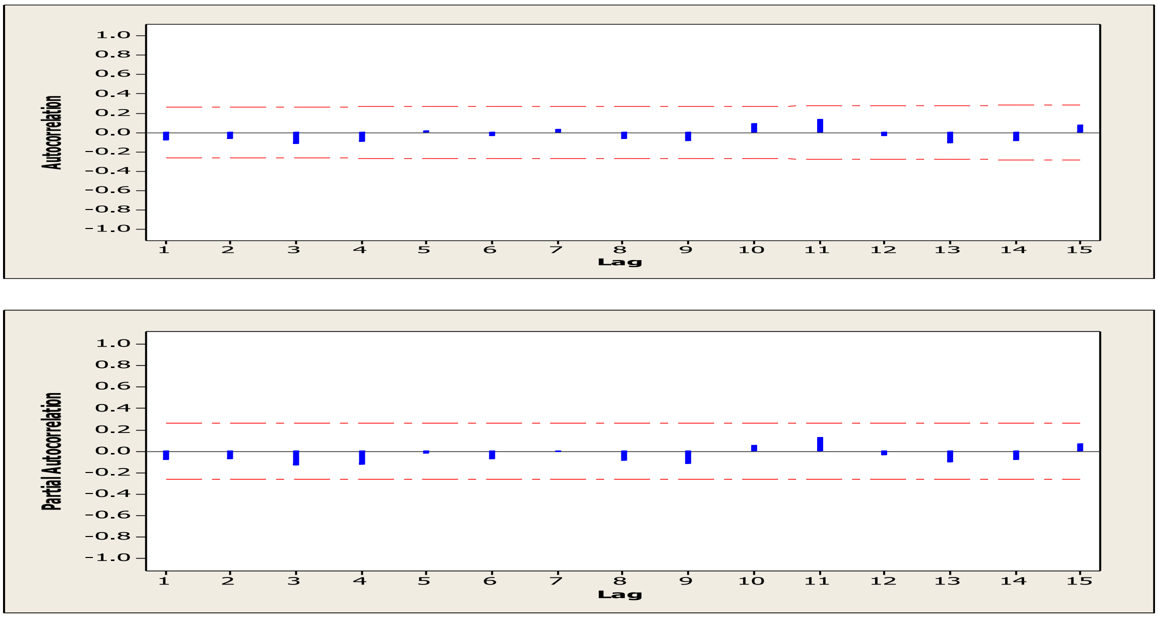

Figure 5: ACF and PACF Plot of Residuals of the Fitted Model

From Figure 5, we observe that all the sample autocorrelation coefficients of the residuals are within the confidence limits. This shows that the residuals are white noise and the fitted model is stable and stationary. This is also justified by the time plot of residuals against time which is covariance stationary as represented in Figure 6.

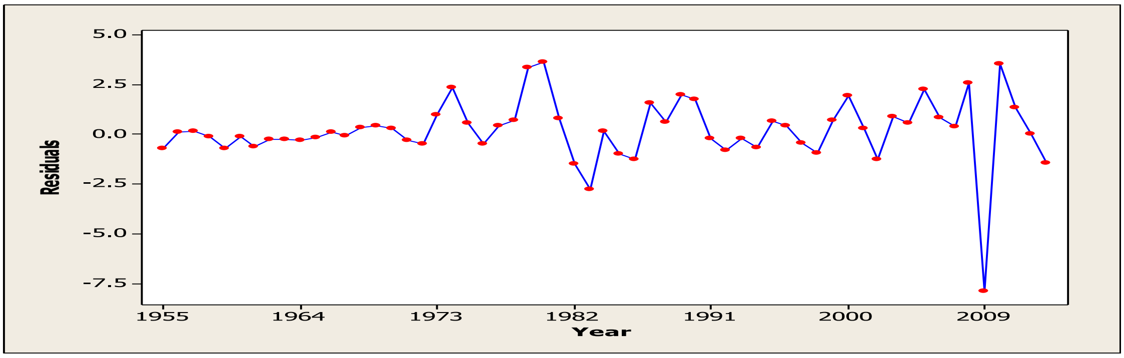

Figure 6: Plot of Residuals of the Fitted ARIMA (3,1,0) Model

Another evidence to show that the estimated model is dynamically stable is that the inverse roots of AR polynomials are within a unit circle. This is represented in Figure 7.

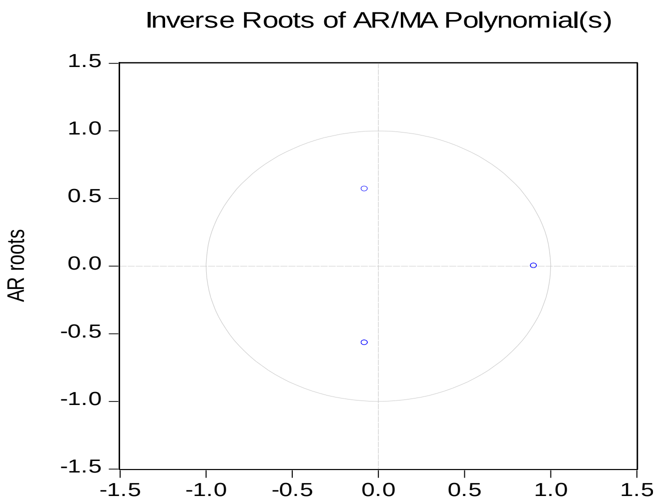

Figure 7: Inverse Roots of AR Polynomials

From the roots of AR polynomials of the fitted model, we estimate that $\tan\theta = y/x = 0.57/0.08 = 7.125$ and $\theta = 82.01°$. Thus, the cycle is $360°/82.01° = 4.4$ years and we say that CPI inflation in Nigeria has a long cycle of 4.4 years.

Residual Tests of Arima (3,1,0)

We also conducted some tests on the residuals of the fitted model. From the result of the test presented in Table 4, the residuals of the estimated model have satisfied the Bruesch-Godfrey serial correlation Lagrange Multiplier (LM) test because the p-values of F-statistic and nR^2 are 0.1940 and 0.1796. The null Hypothesis of no serial correlation in the residuals at all lags is accepted since the p-values are all greater than 0.05.

Table 4 also shows the result of Ramsey Regression Equation Specification Error Test (RESET). This test tests whether an estimated equation is mis-specified. Our fitted model has passed the Ramsey RESET test since the p-values of t-statistic, F-statistic and Likelihood ratio are all greater than 0.05. We thus conclude that the parameters of our model equation are not mis-specified.

Table 4: Ramsey RESET Test & Serial Correlation LM Test for the Fitted Model

Test	Test value	P-value
t-statistic	1.147617	0.2541
F-statistic	1.317026	0.2541
Likelihood ratio	1.378758	0.2403
Breusch-Godfrey Serial Correlation LM Test		
F-statistic	1.669718	0.1940
nR^2	3.433615	0.1796

Since the residuals of our model have passed the diagnostic tests, we validate it as being an adequate and good model. An adequate, valid and good model should be able to forecast future values of the relevant series. In the following subsection, we will consider the ability of the series to forecast future values.

Model Forecast Evaluation

We want to see the ability of the fitted model to forecast future time series. The ability to do so will further testify the validity of this model. We employ four accuracy measures to evaluate this forecast ability.

Table 5: Result of Forecast Comparison of ARIMA (3,1,0) Model Using Accuracy Measures

Mode of forecast	RMSE	MAE	MAPE	TIC
In-sample	2.382958	2.043889	300.1816	0.413742
Out-of sample	1.437320	0.906917	89.91606	0.224006

We consider four measures of accuracy, namely: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Theil Inequality Coefficient (TIC) to compare the performance ability of the In-sample and Out –of sample forecasts of the estimated ARIMA (3,1,0) model and to decide on which mode of prediction is suitable for the series. The result of Table 5 shows that the RMSE, MAE and the MAPE of the out –of sample forecasts are smaller than those of the In-sample forecasts, and the decision is that the smaller the forecast errors, the better the forecasting performance of that model, according to the criterion, our model is good for post-sample forecast. The Theil Inequality Coefficient always lies between 0 and 1, the 0 value indicates a perfect fit. Comparing our In-sample and post-sample forecasts using the theil inequality coefficient, the post-sample forecast fits more perfectly than that of the In-sample forecast. We therefore conclude that the post-sample forecast is the best forecast mode for this model.

Forecast of CPI Inflation in Nigeria

Having voted the post sample forecast approach for the series, we use the estimated ARIMA (3,1,0) model to forecast future values of CPI in Nigeria for the period of 6 years starting from 2015 to the year 2020. The result of the forecast is presented in Table 6.

Table 6: Forecasts of CPI in Nigeria from Fitted ARIMA (3,1,0) Model

Year	LCL	Forecast	UCL
2014	---	231.74	---
2015	220.24	235.69	252.22
2016	210.10	241.81	278.31
2017	203.31	248.89	304.68
2018	198.89	256.32	330.33
2019	195.82	264.09	356.17
2020	193.65	272.25	382.76

Note: Forecast begins in 2015. For 95% confidence intervals, $Z_{0.025} = 1.96$

We forecast CPI in Nigeria from 2015 to 2020. The forecast value for the year 2015 is 235.69 with a 95% confidence interval of [220.24, 252.22]. By this we are 95% confident that the outcome for the next period will fall within this interval. Comparing with the annual CPI in 2014 (231.74), we predict that in 2015 CPI will consistently increase from the current year. This interval implies that the annual CPI increase may lie between 220.24 and 252.22 (i.e. it may increase at least by 15.45 or at most by 16.53) in 2015. The forecast for the following years shows a significant and gradual increase in CPI in Nigeria over the forecasted period with the highest CPI been predicted to occur in the year 2020. The confidence intervals of the forecast suggest a consistent increase in annual CPI during the forecasted period of 2015 to 2020 implying that inflation will also be on the increase within this period.

Concluding Remarks

In this study, we attempted to search for an optimal ARIMA model that will best forecast annual CPI data in Nigeria from 1950 to 2014. We employed time plots, ACF, PACF, Phillips-Perron and well as Dickey-Fuller Generalized Least Squares to investigate the graphical, statistical and unit root as well as stationary properties of the series. The results show that CPI data in Nigeria is non-stationary in level but stationary in logged first difference and thus integrated of order one, I(1). We then applied Box-Jenkins modeling methodology to search for an optimal model and found that ARIMA (3, 1, 0) was the best fitted model to describe CPI data series in Nigeria. The model was validated and found to be adequate and good. Based on this model, we forecast the future annual CPI in Nigeria for the period of 6 years from 2015 to 2020. The forecasts show a steady increase in the annual values of CPI in Nigeria. The study predicts that inflation will increase in Nigeria from 2015 since the confidence intervals of the forecast suggest a consistent increase in annual CPI during the forecasted period of 2015 to 2020.

This finding will provide useful information for the Central Bank of Nigeria, other concern authorities, financial and economic analysts who are concerned about the economy. Judging from the inferences of this study, we recommend that the Federal government of Nigeria should take necessary actions to contract the economy in this contending period.

References

[1]. Espasa, A., Poncela, P. and Senra, E. (2002). Forecasting Monthly US Consumer Price Indexes through a Disaggregated I(2) analysis. Working Paper. Universidad Carlos III De Madrid.

[2]. Enders, W. (2004). Applied Econometric Time Series. John Wiley & Sons, Inc.

[3]. Wayne, R. (1998). Forecasting Inflation Using VAR Analysis. *Bank of Jamaica.*

[4]. Meyler A., Kenny, G. and Quinn, T. (1998). Bayesian VAR Models for Forecasting Irish Inflation. *Central Bank of Ireland Technical Paper, 4*

[5]. Liu, S. X., and Han, Y. S. (2007). Forecasting consumer price index of USA. International Academy of Business and Economics. 7(1): 76-91.

[6]. Adams, S. O., Awujola, A., & Alumgudu, A. I. (2014). Modeling Nigeria's Consumer Price Index Using ARIMA Model. International Journal of Development and Economic Sustainability, 2(2): 37-47.

[7]. Kelikume. I., and Salami, A. (2014). Time Series Modeling and Forecasting Inflation: Evidence From Nigeria. *The International Journal of Business and Finance Research*, 8(2); 41-51

[8]. Fritzer, R., Moser, G., and Scharler, J. (2004.). Forecasting Austrian Inflation. Working Papers 91, Oesterreichische National bank (Austrian Central Bank).

[9]. Bokhari, S. M. and Feridun, M. (2006). Forecasting Inflation through Econometric Models: An Empirical Study on Pakistani Data. *Dogus universitesi*, Vol. 7(1), p. 39-47

[10]. Hubrich, K. (2003). Forecasting Euro Area inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy? *International Journal of Forecasting,* Vol. 21(1), p. 119-136

[11]. Alnaa, S. E. and Ahiakpor, F. (2011). ARIMA Approach to Predicting Inflation in Ghana,'' *Journal of Economics and International Finance,* Vol. 3(5), p. 328-336

[12]. Poulos, L., Kvanli. A. and Pavur, R (1987). A comparison of the accuracy of the Box-Jenkins method with that of automated forecasting model. *International Journal of forecasting*, Vol. 3, p. 261-267

[13]. Libert, G. (1983). The M-competition with a Fully Automatic Box- Jenkins Procedure," *International Journal of Forecasting*, Vol. 2, p. 325-328

[14]. Texter, P. A and Ord, J. K.(1989). Forecasting Using Automatic Identification Procedures: A comparative Analysis. *International Journal of Forecasting*, Vol. 5, p. 209-215

[15]. Hill, G., and Fildes, R. (1984). The Accuracy of Extrapolation Methods: An Automatic Box-Jenkins package (SIFT). *Journal of Forecasting,* vol. 3, p. 319-323

[16]. Adebiyi, M. A., Adenuga, A. O., Abeng, M. O., Omanukwe, P.N. and Ononugbo, M.C. (2010). Inflation Forecasting Models for Nigeria. Research Department, Central Bank of Nigeria Occasional Paper

[17]. Uko, A. K., and Nkoro, E. (2012). Inflation Forecasts with ARIMA, Vector Autoregressive and Error Correction Models in Nigeria. *European Journal of Economics, Finance and Administrative Sciences* issue 50

[18]. Dickey, D. and W. Fuller. 1979. Distribution of the Estimates for Autoregressive Time Series with a Unit Root. Journal of the American Statistical Association 74: 427-31.

[19]. Akaike, H. (1974). A New Look at Statistical Model Identification, IEEE Transactions on Automatic Control, AC-19, pp. 716-723

[20]. Hannan, E. (1980). The Estimation of the Order of ARMA Process. Annals of Statistics, Vol.8, 1071-1081