ISSN 2581-3463

AJMS
Asian Journal of Mathematical Sciences

RESEARCH ARTICLE

Modeling Unemployment Rates in Nigeria using Time Series Approach

Ogunsola Isqeel Adesegun, Olayiwola Olaniyi Mathew, Sowola Basirat Omotola

Department of Statistics, Federal University of Agriculture, Abeokuta, Nigeria

Received: 01-02-2020; Revised: 10-03-2020; Accepted: 20-04-2020

ABSTRACT

Unemployment is a social-economic situation that has resulted into poor standard of living, increases in crime rate and insecurity in the country. It even possesses greater threats to its development and peaceful co-existence. It is in light of proffering solutions to the problems of unemployment that this study is germane. We examined the trend, model and forecast for the future rate of unemployment in Nigeria. The statistical methodology adopted in this research work is mainly univariate time series analysis using autoregressive integrated moving averages (ARIMA).The time plot showed that rate of unemployment in Nigeria is not stable over the years. ADF test was performed which confirms the non-stationarity of the process ($P > 0.05$). The data was transformed through differencing. KPSS test was used to confirm the stationarity of the process ($P > 0.05$).Various ARIMA (p,d,q) were examined such as ARIMA (1,1,1), ARIMA (2,1,1), ARIMA (2,1,2), ARIMA (2,2,1) and ARIMA (2,2,2) with AICs and Log-likelihood (107.52, −50.76), (109.22, −50.61), (111.08, −50.54), (110.08, −51.04) and (111.36, −50.68) respectively. ARIMA (1,1,1) was selected due to its least AIC and highest log likelihood. Both Shapiro-Wilk test and Box test performed confirm the fitness of the model ($P > 0.05$) for the process. Forecast for 5 years was then made for the process. In conclusion, the model obtained in this research can be used for making inference, monitor and control unemployment rates in Nigeria. Adequate solutions such as industrialization, diversification in the economy, investment in agriculture, modification in education curriculum among others have been identified in this research work. The analysis is completely done using R package.

Key words: Akaike information criteria, box test, log–likelihood, Shapiro-Wilk, trend

INTRODUCTION

Unemployment is one of the critical socio-economic problems faced by Nigerians presently especially the youths. The labor force grows on daily basis sporadically with high number of young graduates from colleges, polytechnics and universities. In 2011, an analysis on the unemployment situation in Nigeria shows that of the total number of unemployed individuals, 43.7% were university graduates, 23.8% were polytechnic graduates and 15.5% were college undergraduates.[1]

The employment growth rate is inadequate to absorb the new labor force entrants. The ministries are filled up yet few new companies are sprouting up. Nigeria has the largest economy in Africa (Salami, 2011). Yet, the average standard of living of the citizens is low. The situation has possess a lot of social-economic, political and moral consequences. The situation is even becoming prominent and out of control. Doğrul and Soytas[2] stated that Unemployment being one of the most important macro-economic issues in every country, it is always of great concern to every nation for its development. It is not easy to measure the rate of unemployment because of the problems of defining who is employed, unemployed or underemployed. Unemployment in Nigeria is defined as the proportion of labor force that was available for work but did not work in the week that precedes the survey period for at least 39 h. Unemployment was initially assumed to be intentionally being idle which was punishable by law. However, it is now recognized

Address for correspondence:
Ogunsola Isqeel Adesegun
E-mail: profisqeel@yahoo.com

as a situation that arises from factors beyond the control of the individual worker. Unemployment may be due to seasonal layoffs (e.g., agricultural jobs), change in technology or technological improvement in industry (particularly changes in modes of production), racial discrimination, lack of adequate skills by the worker, fluctuations in the economy and lots more. Its consequences cannot be over emphasized. It that has resulted into poor standard of living increases in crime rate and insecurity in the country. It even possesses greater threats to its development and peaceful co-existence. It is in light of proffering solutions to the problems of unemployment that this study is germane.

Unemployment has been analyzed by different researchers using regression analysis but little or no attempt has been made on modeling unemployment using ARIMA model. This study tends to examine the trend, model and forecast for the future occurrence of unemployment in Nigeria to sensitize Government about the consequences of unemployment rate in future if they fail to provide adequate measures and solutions.

Time series are sets of observations gathered sequentially in time. It is a stochastic sequence observations made at equal interval of time. One special feature of time series data is that the successive observations are usually known not to be independent, so that the analysis can take into consideration the order of the observations. The applications of time series models are manifold, including sales forecasting, weather forecasting, inventory studies etc. In decisions that involve factor of uncertainty of the future, time series models have been found one of the most effective methods of modeling and forecasting.

METHODOLOGY

The methodology adopted in this research is basically time series analysis using autoregressive integrated moving averages (ARIMA). The data used is an annual time series data obtained from the Central bank of Nigeria (CBN) Statistical Bulletin and national bureau of statistics (NBS).

A general exploratory data analysis (EDA) was also performed on the process. We then examined the trend, model, check for the adequacy of the model and make future forecast for 5 years.

Trend and time plot

Trend is the general direction in which the series or process appears to be moving. This is examined through a time plot. A time plot is the plot of the process against time. We adopt the unit root tests such as Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test and Augmented Dickey fuller test (ADF) for testing the stationarity of the process. The null hypothesis for KPSS is that the process is stationary while the converse is for ADF.

The autocorrelation function (ACF) and partial autocorrelation function (PACF)

The ACF which suggests the order "*q*" of the MA process was also examined and plotted. The autocorrelation is given as:

$$\rho_k = cor(X_t, X_{t-1}) = \frac{\gamma_k}{\gamma_0} \quad (1)$$

Where, ρ_k is a measure of linear dependency between values of the process at lag *k* apart,

$\rho_0 = 1$, $\rho_k = \rho_{-k}$ for all k and $/\rho_k/\leq 1$. The sequence, $\{\rho_k\}$ is called the ACF.

Furthermore, a sharp cut off of the partial autocorrelation functions (PACF) suggests the order "*p*" of the AR. For an arbitrary stationary process, the partial autocorrelations can be calculated using the Durbin-Levinson Algorithm (1947–1960) given below.

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1,} \rho_{K-1}}{1 - \sum_{j=1}^{k-1} \phi_{k-1,} \rho_j} \quad (2)$$

and;

$$\phi_{k,j} = \phi_{k-1,j} - \phi_{kk}, \phi_{k-1,k-j} \quad (3)$$

for j=1,2,…, k−1.

Autoregressive (AR) process

If the current value of the series is linearly dependent upon its previous with some errors, then, the AR (p) can be written as:

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t \quad (4)$$

Which could and be expanded as:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + \varepsilon_t \quad (5)$$

Where ε_t is a white noise time series, that is, ε_t is a sequence of uncorrelated random variable (possibly normally distributed) with mean 0 and variance σ^2 and φ is the AR parameter.
The autocovariance function for AR(p) is giving as:

$$\rho_k = \varphi_1 \rho_{t-1} + \varphi_2 \rho_{t-2} + \ldots + \varphi_p \rho_{\mathrm{t-p}} + \varepsilon_t \quad (6)$$

for k =1,2,…p
The autocorrelation is given as:

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \varphi^k \left(k \geq 0\right) \quad (7)$$

Moving averages (MA)

A moving average process of order q and MA(q) process with mean 0 is a process (X_t) which satisfies the relation:

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \ldots \text{or}$$
$$X_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \quad (8)$$

Using the backward shift, we have;

$$X_t = (1 + \theta_1 \beta + \ldots + \theta_q \beta^q)\varepsilon_t \quad (9)$$

$$\Rightarrow X_t = \theta(\beta)\varepsilon_t \quad (10)$$

Where (ε_t) is a white noise process with mean zero, variance σ^2 and $\theta_q \neq 0$
The auto covariance function for the MA(q) process is:

$$\gamma_k = \begin{cases} \left(1 + \theta_1^2 + \ldots + \theta_q^2\right)\sigma^2 & k = 0 \\ (\theta_k + \theta_1\theta_{k+1} \ldots + \theta_{q-k}\theta_q)\sigma^2 & k = 1, 2, \ldots, q \\ 0 & k > q \end{cases} \quad (11)$$

Autoregressive moving averages (ARMA)

An ARMA (p,q) process of order p,q is a stationary process X_t that satisfies the relation

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q} \quad (12)$$

Where $\{\varepsilon_t\}$ is a white noise.
In lag form, the equation becomes:

$$X_t - \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \ldots + \varphi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \ldots + \theta_q \varepsilon_{t-q}$$
$$\varphi(B) X_t = \theta(B)\varepsilon_t$$
$$X_t = \theta(B)\varepsilon_t / \varphi(B)$$
$$X_t = \theta(B)\varepsilon_t \varphi^{-1}(B) \quad (13)$$

Also,

$$\varepsilon_t = \varphi(B) X_t / \theta(B)$$
$$\varepsilon_t = \varphi(B) X_t \theta^{-1}(B) \quad (14)$$

Where,
$\varphi(B)$ is the AR characteristics polynomial and $\theta(B)$ is the MA characteristics polynomial.

ARIMA process

A process X_t is said to be an AR integrated moving average process of order (p,d,q) ARIMA (p,d,q) if its d[th] difference is an ARMA (p,q) process. An ARIMA (p,d,q) model can be defined by:

$$\varphi\left(\beta\right)\nabla^d X_t = \theta\left(\beta\right)\varepsilon_t \quad (15)$$

Where p,d,q are non-negative integers.

Data used

The data used for this research were a secondary data obtained from NBS bulletin, CBN publications and International Monetary Fund, and World Economic Outlook Database. The data cover unemployment rates in Nigeria from the period of year 2000 to 2018.

ANALYSIS AND RESULT

Here, we present the result of the analysis obtained from Nigeria unemployment rates.

DISCUSSION

Table 1 and Figure 1 show the summary of statistics of unemployment rates and the histogram, respectively. Both show that unemployment rates are not normally distributed over the years in Nigeria. A great variation is also noticed among the rates due to the high variance (22.92) and standard error (4.79) values. Figures 2 and 3 also show the time plots for unemployment rates in Nigeria before and after differencing, respectively. It is evident from the plots that the process is not stable over the years indicating some assignable cause(s) of variation before differencing and stable after first difference.[3,4,5,6] The ADF unit root test also confirmed the non-stationarity of the process and KPSS test confirmed the stationarity of the differenced process after first differencing. The results are shown in Table 2.

Figures 4 and 5 show the ACF and PACF of the process. The ACF cuts off at lag 9 while the PACF tails to zero at lag 11. None of the points exceed the significant bounds. Table 3 shows the summary of ARIMA (*p,d,q*) models examined with their standard errors, AICs, and Log-likelihoods. The result showed that ARIMA (1,1,1) has the least AIC (107.52) and the highest Log-likelihood (−50.76). Therefore, ARIMA (1,1,1) is selected for the process. The model selected is written as:

$$X_t = 0.1632X_{t-1} + \varepsilon_t + \varepsilon_{t-1}$$

Again, Figure 6 shows the diagnostic plot for the model. The results show that the model fits the

Table 1: Descriptive statistics

Summary of statistics	
Mean	14.30
Variance	22.92
Standard deviation	4.79
Median	13.3
Mode	13.3
Minimum	7.3
Maximum	23.9

Table 2: Summary of unit root tests

Test	Statistic	Lag order	*P*-value	Conclusion
ADF	−2.9754	2	0.2008 (*P*>0.05)	Not stationary before differencing
KPSS	0.072448	2	0.1 (*P*>0.05)	Stationary after differencing

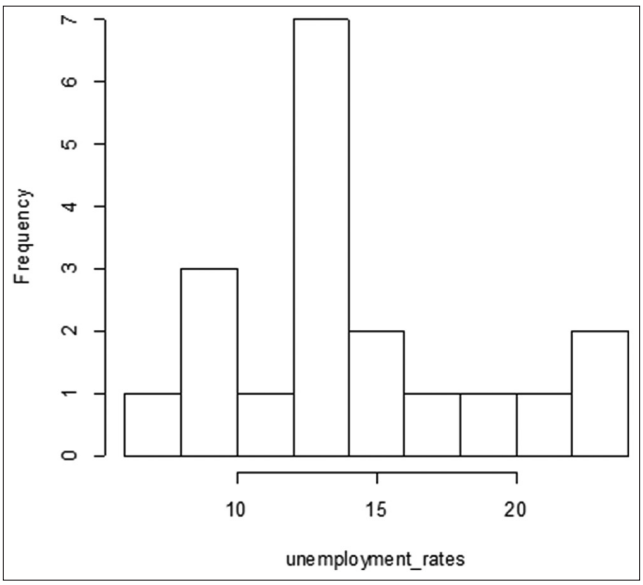

Figure 1: Histogram of unemployment rates

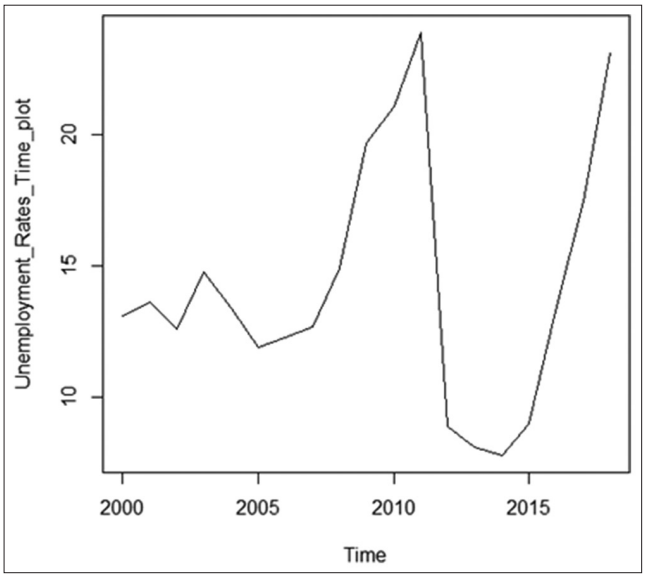

Figure 2: Time plot showing the trend of unemployment rates

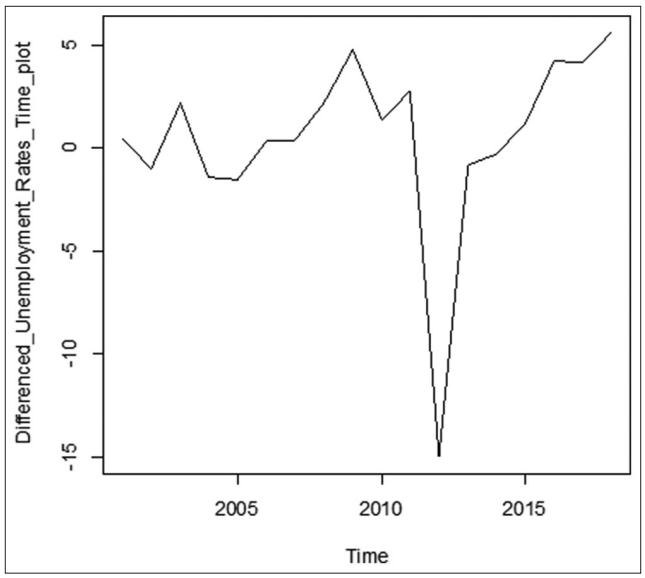

Figure 3: Time plot of the differenced unemployment rates

Table 3: Summary of ARIMA models examined

S/N	ARIMA							SE	AIC	Log likelihood
	P	D	Q	AR(1)	AR(2)	MA(1)	MA(2)			
1	1	1	1	0.1632	-	−1.0	-	4.4	107.52	−50.76
2	2	1	1	0.1663	0.128	−1.0	-	4.4	109.22	−50.61
3	2	1	2	−0.1159	0.204	−0.7167	−0.2833	4.4	111.08	−50.54
4	2	2	1	−0.4731	−0.467	−1.0	-	5.2	110.08	−51.04
5	2	2	2	0.3077	0.235	−1.9751	1.00	4.7	111.36	−50.68

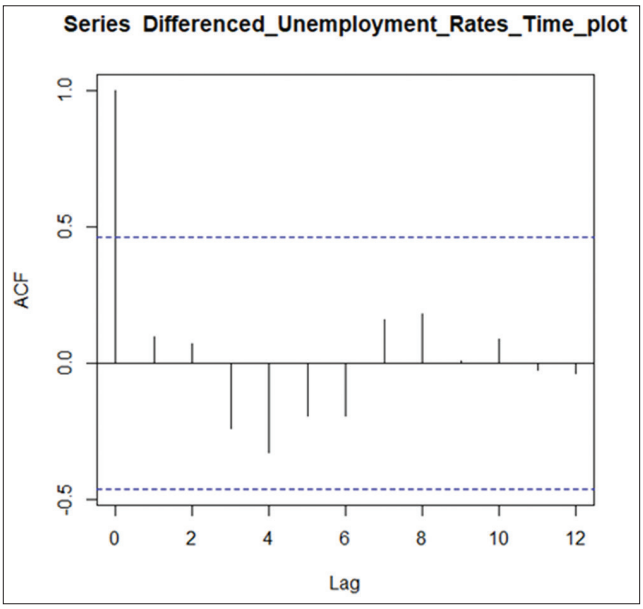

Figure 4: Autocorrelation function plot

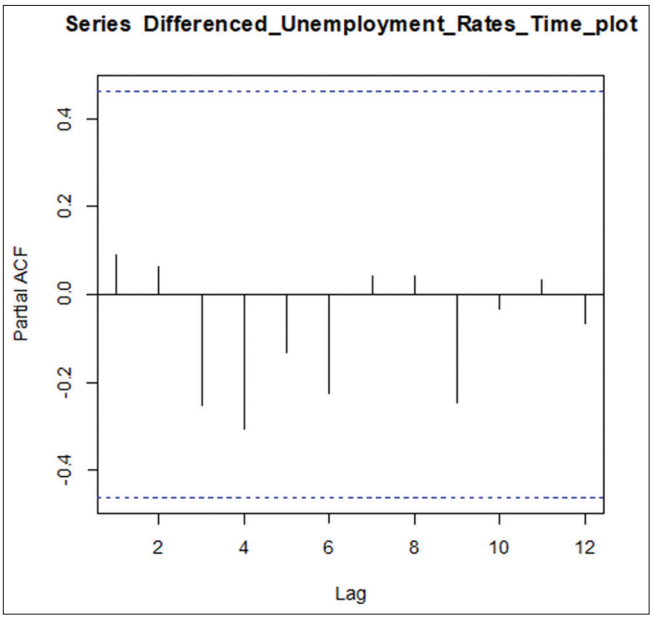

Figure 5: Partial autocorrelation function plot

process as the $P > 0.5$; thereby we fail to reject the null hypothesis and conclude that the model fit the process. Table 4 shows Shapiro-Wilk and Box test and the two tests confirmed the fitness of the model to the process. Finally, Table 5 shows 5 years forecast for the process to show reasons for quick intervention on curbing high and persistence rate of future unemployment in the country.

Table 4: Shapiro-Wilk test and box test results

Test	W	*P*-value	Conclusion
Shapiro-Wilk	0.6708	0.37 (*P*>0.05)	The error terms are normally distributed
Test	**X-squared**	**Df**	***P*-value** **Conclusion**
Box test	0.07982	1	0.7775 (*P*>0.05) Residuals are uncorrelated

Table 5: Five years point and interval forecast for unemployment rates (2019–2023)

Year	Point forecast	80% Interval forecast		90% Interval forecast	
		Lo 80	Hi 80	Lo 95	Hi 95
2019	23.09944	15.645029	30.55384	11.6989042	34.49997
2020	23.09944	12.392052	33.80682	6.7239050	39.47497
2021	23.09944	9.777521	36.42135	2.7253252	43.47355
2022	23.09944	7.469923	38.72895	−0.8038407	47.00271
2023	23.09944	5.342035	40.85684	−4.0581658	50.25704

Recommendation

Since unemployment results in a lot of social disorganization, it therefore requires a multiplicity of policies to reduce it.[7-11] Creation of more employment opportunities, improvement on infrastructural facilities in the rural area to reduce the rate of rural-urban migration which brings lack of labor in rural areas and congested labor in urban areas, promotion of technical and self-employment education, diversification in the economy and proper access to information, and subsidization of education cost by the Government are hereby recommended as palliative measures to lingering phenomenon. It is hoped that effective implementation of the above recommendations will help in solving the unemployment situation and reduce its rate which will in turn lead to a better economic and social situation in Nigeria.[12-15]

CONCLUSION

Unemployment is one of the factors that affect the national economy, so its control must be put into consideration. Hence, the model obtained from

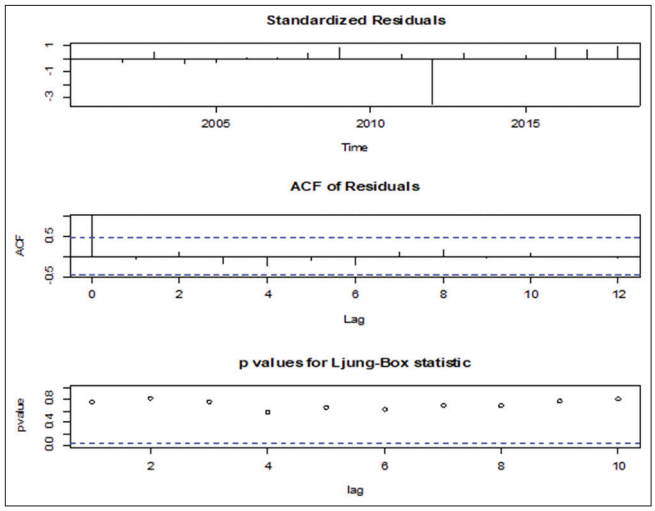

Figure 6: Diagnostic plot

this study can be used to monitor and forecast the future rate of unemployment in Nigeria to achieve a better economy and thus brings about national development in Nigeria.

ACKNOWLEDGMENT

The authors of this paper want to acknowledge Dr. (Mrs.). A. A. of the Department of Statistics, Federal University of Agriculture, Abeokuta, for the experience shared and guide provided during the cause of writing this research work.

REFERENCES

1. Coghlan A. A Little Book of R for Time Series. United States: Calvin University; 2015.
2. Card D. Origins of the unemployment rate: The lasting legacy of measurement without theory. Am Econ Rev 2011;101:552-57.
3. Akintunde AA. Time Series Analysis Lecture Note. Abeokuta: Department of Statistics, College of Physical Sciences, Federal University of Agriculture; 2015.
4. Ande CE. Essential Economics for Senior Secondary Schools. Ogun: Tonad Publisher; 2003.
5. CBN. Statistical Bulletin. Abuja: Central Bank of Nigeria; 2012.
6. Wittenburg DC, Fishman M, Stapleton D, Crivner S, Tucker A. Literature Review and Empirical Analysis of Unemployment Insurance, Recipiency Ratios; 1999. p. 12.
7. Doğrul HG, Soytas U. Relationship between oil prices, interest rate, and unemployment: Evidence from an emerging market. Energy Econ 2010;32:1523-8.
8. Thomas DR, Hodges ID. Designing and Planning Your Research; 2010.
9. Shaddick G. Using R with Application with time Series; 2004.
10. Available from: http://www.localhistories.org/unemployment.html. [Last accessed on 2019 Dec 06].
11. Available from: http://www.doublegist.com/effectsunemployment-society. [Last accessed on 2019 Dec 19].
12. McLeod I, Yu H, Mahdi E. Time Series Analysis with R. London, Ontario, Canada: The University of Western Ontario, Department of Statistical and Actuarial Sciences; 2012.
13. Aiyedogbon JO, Ohwofasa BO. Poverty and youth unemployment in Nigeria, 1987-2011. Int J Bus Soc Sci 2012;3:269-79.
14. Crawley MJ. The R Book. 2nd ed. Chichester: John Wiley and Sons; 2013.
15. Ahmad MS, Ibrahim I, Oye ND. Unemployment in Nigeria: Implication on the gross domestic product (GDP) over the years. Int J Econ Res 2011;2:66-71.