



## Modelling Nigeria's Gross Domestic Product – A Modified Vector Autoregressive Model (MVAR) approach

Apantaku F.S<sup>1</sup>, Ogbaola, E.O.<sup>2</sup>, Ogunsola I.A.<sup>3</sup>, Dawodu, G. A.<sup>4</sup> and Agwuegbo S.O.N.<sup>5</sup>

<sup>1,2,3,4,5</sup>Federal University of Agriculture, Abeokuta

\*Correspondence Author's Email: [apantakufs@funaab.edu.ng](mailto:apantakufs@funaab.edu.ng)

Received on: April, 2024

Revised and Accepted on: June, 2024

Published on: July, 2024

### ABSTRACT

On yearly basis, the total value of all goods produced and services rendered in a country is referred to as GDP. It is an indicator of a nation's standard of living and a measure of its economic status. The insecurity level cum poor economy of the country which has grossly affected the standard of living of Nigerians necessitated this study. We adopted a Modified Vector Autoregressive (MVAR) modelling approach based on the absolute values of the mean deviation to capture the relationship between the main effect and interaction effects of Agriculture, Trade and Industry on Nigeria's GDP. Due to the non-stationarity of the process, the random walk transformation technique was considered. The Augmented Dickey-Fuller (ADF) test confirmed that the process is stationary after transformation. The diagnostic tests showed that the data were not highly correlated and no presence of heteroscedasticity. The study confirmed that the MVAR model gave a better fitting of Nigeria's Gross Domestic Product compared to the usual VAR model due to its lower AIC and BIC values.

**Keywords:** VAR, Mean deviation, Random walk, ADF, heteroscedasticity.

### 1.0 INTRODUCTION

The interaction between economic parameters could sometimes be complicated and that was what led to the introduction of the use of systems of equations that are simultaneous in structure. Likewise, determining the endogenous and exogenous variables could also pose a problem during analysis because most economic parameters have a significant influence on one another. As a result of this, some constraints are required sometimes when using simultaneous systems of equations on a structural model to fix the problem of determination (Ivanov, *et al.* 2020). Vector Autoregressive (VAR) model was developed to take care of a complex economic situation that makes use of the simultaneous system of equations (Akkaya, 2021). VAR setup is such that it studies the relationship between the current variable observations and its relationship with its past observations and that of other variables within the system. For instance, VAR could be employed to reveal how GDP affects unemployment rate and how unemployment rate, in turn, affects GDP concurrently.

Currently, the prior knowledge required in carrying out the Vector Autoregressive model would be to use the original values of the variables, which can be thought of to be related to each other over time to capture the interaction between the current occurrence of the endogenous variables with its previous occurrence and previous occurrence of the remaining variables in the system. Since data could be overwhelmingly too large in variation, it is important to create data with smaller variations that can carry out the same relationship among

the variables and also give an overview of the variability of the data. This work seeks to resolve this issue by modifying the conventional Vector Autoregressive model, by simply replacing the original data with the absolute values of the deviation of each observation from the mean of the variable to which it belongs to capture the relationship between the main effect and interaction effects of Agriculture, Trade and Industry on Nigeria's GDP.

Several researchers have put VAR into use in studying economic variables. By applying VAR model, Yuliadi (2020), looked at the factors in the economy that influence the supply of money in the economy. Vector Autoregressions with Dynamic Factor Coefficients were also studied (Doz, 2020). Paolo *et al.* (2021) introduced time-varying coefficient matrices and conditionally heteroskedastic disturbances as a method for analysing vector autoregressive models. Their method was able to capture the dynamism of the model. Their approach was simple to implement and fast to compute. A study was carried out to establish the reliability and robustness of the model and no misspecification of the model was found. They also used experimental illustration where they analysed the relationship among U.S. industrial production, bond spread and inflation with respect to time. Huang, *et al.* (2020) employed Vector Autoregressive model to monitor and predict newly diagnosed cases of HIV/AIDS, gonorrhoea and syphilis in China. Yue, *et al.* (2022), also applied Vector Autoregressive analysis to evaluate the dynamic effects



of mitigation instruments on CO<sub>2</sub> emissions in China's nonferrous metal industry.

Other sections are organised as follows. Section 2 discussed the modified VAR. The results were extensively discussed in section 3 and section 4 gave a concise but meaningful concluding remark.

## 2.0 Methodology

This study modified the VAR model in studying the relationship between economic variables and the random-walk transformation approach for stationarity as against the usual differencing method. AIC and BIC and some other tests were employed as model selection criteria and diagnostic tools respectively.

If we were to find the relationship between four different time series variables, given as  $x_{t,1}$ ,  $x_{t,2}$ ,  $x_{t,3}$  and  $x_{t,4}$ . The VAR(1) read as VAR of order 1, is given as:

$$\begin{aligned} x_{t,1} &= \alpha_1 + \phi_{11}x_{t-1,1} + \phi_{12}x_{t-1,2} + \phi_{13}x_{t-1,3} + \phi_{14}x_{t-1,4} + \omega_{t,1} \\ x_{t,2} &= \alpha_2 + \phi_{21}x_{t-1,1} + \phi_{22}x_{t-1,2} + \phi_{23}x_{t-1,3} + \phi_{24}x_{t-1,4} + \omega_{t,2} \\ x_{t,3} &= \alpha_3 + \phi_{31}x_{t-1,1} + \phi_{32}x_{t-1,2} + \phi_{33}x_{t-1,3} + \phi_{34}x_{t-1,4} + \omega_{t,3} \\ x_{t,4} &= \alpha_4 + \phi_{41}x_{t-1,1} + \phi_{42}x_{t-1,2} + \phi_{43}x_{t-1,3} + \phi_{44}x_{t-1,4} + \omega_{t,4} \end{aligned} \quad (1)$$

where  $\omega_{t,1} \sim N(0, \sigma_\mu)$  and  $\alpha_i$  is a  $k \times k$  coefficient matrix.

Suppose we have predictor variables  $x_t = (x_{1,t}, x_{2,t}, \dots, x_{k,t})$ , at time  $t = 0, 1, 2, \dots, T$ , Petrusевич, (2020) made investigation of order selection in VAR model of order  $p$  as:

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \omega_t \quad (2)$$

In matrix form it can be written as:

$$\begin{pmatrix} x_{t,1} \\ x_{t,2} \\ x_{t,3} \\ x_{t,4} \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} & \phi_{13} & \phi_{14} \\ \phi_{21} & \phi_{22} & \phi_{23} & \phi_{24} \\ \phi_{31} & \phi_{32} & \phi_{33} & \phi_{34} \\ \phi_{41} & \phi_{42} & \phi_{43} & \phi_{44} \end{pmatrix} \begin{pmatrix} x_{t-1,1} \\ x_{t-1,2} \\ x_{t-1,3} \\ x_{t-1,4} \end{pmatrix} + \begin{pmatrix} \omega_{t,1} \\ \omega_{t,2} \\ \omega_{t,3} \\ \omega_{t,4} \end{pmatrix}. \quad (3)$$

Generally, for a VAR( $p$ ) model, we would use the first  $p$  lags of each variable in the system as the explanatory variables for each variable.

The model in compact form is of the form

$$x_t = \alpha + \phi x_{t-1} + \omega_t. \quad (4)$$

## 2.2 Modified Vector Autoregressive model VAR( $p$ ) Models

The only modification done to the existing model in this study is that instead of  $x_{t-1}$  being the lag 1 of the exact values of the observation, it is the absolute value of the variation of the mean of the observation from the mean of the variable which is given as:

$$y_{t-1} = |x_{t-1} - \bar{x}_t|. \quad (5)$$

Hence, the modified variable is written as:

$$y_t = \alpha + \phi y_{t-1} + \omega_t. \quad (6)$$

where  $y_{t-1}$  is as given in equation (5)

## 2.1 Vector Autoregressive model VAR( $p$ ) Models

The Vector Autoregressive (VAR) is bi-directional. What this means is that the variables influence each other. This is unlike other Autoregressive models, such as AR, ARMA or ARIMA, which use uni-directional models, in which explanatory variables affect the outcome variable and not vice-versa. VAR models is a generalization of the univariate AR model used for multivariate time series. They are very flexible models. The structure of VAR models is such that each variable is a linear combination of previous lags of other variables and itself.

## 2.3 Random Walks

A simple random walk describes a process of a path consisting of a succession of random steps. It can also be defined as the sum of an order of random variables, for example, the movement of a drunkard as he takes one step forward and one step backwards in a random manner or the motion of a particle in space. The directions for the different steps do not depend on each other. The defining condition of the random walk is

$$S_n = X_1 + \dots + X_n, \quad n = 1, 2, \dots$$



it is a stochastic sequence  $\{S_n\}$ , with  $S_0 = 0$ , defined by  $S_n = \sum_{t=1}^n X_t$

where  $\{X_t\}$  are independent and identically distributed random variables (i.i.d).

The main interest is to determine the distribution of the random walk,  $S$  for finite  $n$  mentioning the assumptions for the distribution of the  $X_t$ . The sequence of values  $X_t$  can take only  $\pm 1$  for a binomial process with

$$X_t = \begin{cases} +1 & \text{with probability } p \\ -1 & \text{with probability } q \end{cases}$$

Such random variable with more generally  $P(X_t = 1) = p$  for arbitrary  $p \in (0,1)$  are called Bernoulli variables. The sequence  $X_t$  can also be considered as a Gaussian random walk, where  $X_t$  are jointly normally distributed.

Binomial and trinomial processes are common examples of random walks. The increments in this random variable have a distribution of a real-valued random variable, and it can take a finite or countable infinite number of values. However, the possibility for to take values out of a continuous set also exists. We can construct such random walks by assuming that the increments are identical, independent, and normally distributed

## 2.4 Augmented Dickey Fuller (ADF) Test

The ADF test is an advanced or expanded form of Dickey-Fuller test. It added high order regressive process in the model given as:

$$y_t = c + \beta_t + \alpha y_{t-1} + \phi_1 \Delta Y_{t-1} + \phi_2 \Delta Y_{t-2} + \dots + \phi_p \Delta Y_p + e_t \quad (7)$$

This makes the test to be more thorough but the null hypothesis remains unchanged with that of Dickey-Fuller test.

## 2.5 VAR Model Selection Criterion

In order to overcome the issue of producing too large models, we chose the lag order that minimized at least one of the following two information criteria:

- (i) The Bayesian information criterion (BIC):

$$BIC(p) = \log \left( \frac{SSR(p)}{T} \right) + (p + 1) \frac{\log(T)}{T}$$

- (ii) The Akaike information criterion (AIC):

$$AIC(p) = \log \left( \frac{SSR(p)}{T} \right) + (p + 1) \frac{2}{T}$$

The two criteria estimate the optimal length of the shift in time. Basically, the SSR decreases for the two criteria as additional lags are added to the model in a way that the first term decreases when the second increases as the lag order increases. In this study, both estimators were used to decide the lag length of the models that will produce optimal results.

## 3.0 RESULTS AND DISCUSSION

The yearly data used in this study was secondary data of the Gross Domestic Product (GDP) of Nigeria for the period of (1981 – 2018), obtained from Central Bank of

Nigeria ([www.cbn.gov.ng](http://www.cbn.gov.ng)). The four (4) factors considered were GDP, Agriculture, Trade, and Industry.

### 3.1 Continuous Sample Paths

The study modelled the change from one state to the other of the fluctuating prices of the GDP of Nigeria and arrived at the likelihood of each data occurring. The distribution was evaluated using the binomial process as indicated in Table 1 which gives the actual random walk paths of the various factors in the study. The transformed data was tested for stationarity using the Augmented Dickey-Fuller test and was confirmed stationary.



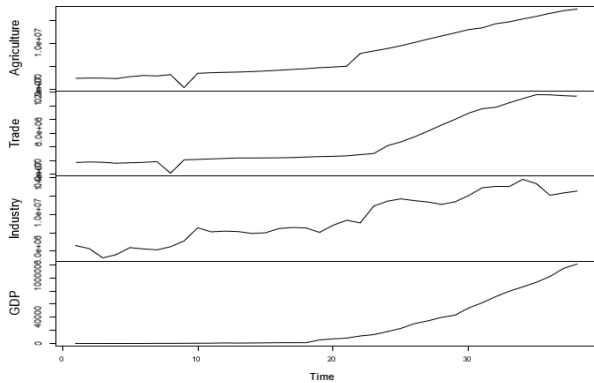


Fig 1: GDP of Nigeria (1981 – 2018) Time Plot

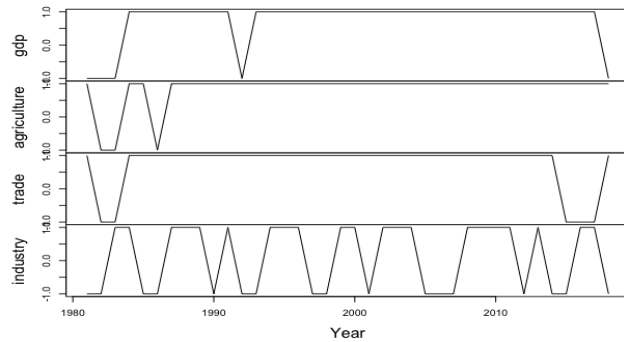


Fig 2: Nigeria's GDP Time-plot of the Random-walk of adjusted data

### 3.4 Stationarity Test

Augmented Dickey-Fuller (ADF) test result as presented in Tables 5 and 6 was adopted to confirm stationarity status of the various data sets.

#### 3.4.1 Augmented Dickey-Fuller Test

The p-values of the Augmented Dickey-Fuller test in Table 5 showed that they were all non-stationary variables before differencing but after differencing only GDP was still not stationary. In Table 6, the p-values of the Augmented Dickey-Fuller test of the random walk model showed that only Trade was not stationary before differencing but after differencing they were all stationary.

Table 5: ADF test for Stationarity of the data

Variable	p-value at 95% C.I. (before differencing)	Stationarity status	p-value at 95% C.I. (after differencing)	Stationarity status
GDP	0.8045	Non-stationary	0.6532	Non-Stationary
Agriculture	0.9841	Non-stationary	0.2656	Stationary
Trade	0.7985	Non-stationary	0.4430	Stationary
Industry	0.6201	Non-stationary	0.0750	Stationary

Table 6 ADF test for Stationarity of random walk data

Variable	p-value at 95% C.I. (before differencing)	Stationarity status	p-value at 95% C.I. (after differencing)	Stationarity status
GDP	0.2193	Stationary	0.0618	Stationary
Agriculture	0.0100	Stationary	0.0404	Stationary
Trade	0.5802	Non-stationary	0.0252	Stationary
Industry	0.0135	Stationary	0.0100	Stationary

### 3.5 Model Identification

The ACF and PACF plots in Figures 3 and 4 were also used to check the stationarity of the data and to identify the order of the model.

#### 3.5.1 ACF and PACF of Nigeria's GDP (1981 – 2018)

The correlogram of the original data in Figure 3 showed that the data was not stationary. In Figure 4, the examination of the autocorrelation function (ACF) and the partial auto-correlation functions (PACF) of the random data indicate an autoregressive model of order 1.

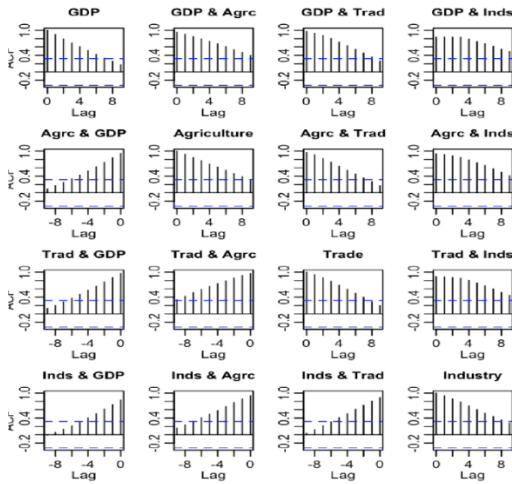


Figure 3: Autocorrelation Function (ACF) of the GDP of Nigeria (1981 – 2018)

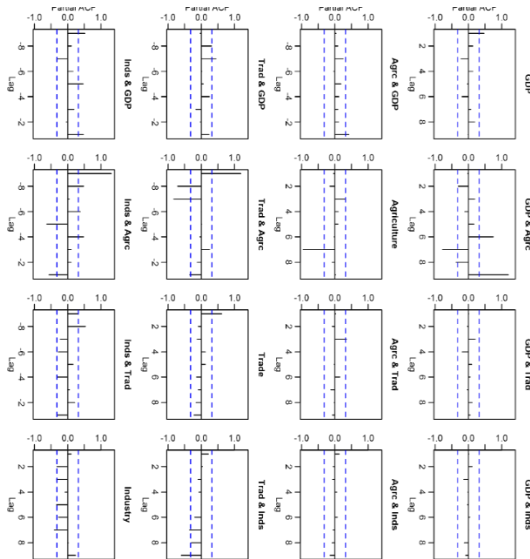


Figure 4: Plot of the Autocorrelation Function (ACF) of the Random walk

### 3.6 Estimation of Model Parameters

Table 7 displays the different models with the standard errors of the exogenous variables written below the coefficients of the factors at lag 1. The result for the original data showed that at lag 1, with GDP as an endogenous variable, only GDP and Trade were significant at 0.01 and 0.05 respectively. With Agriculture as an endogenous variable, Trade was significant at 0.01 and Industry significant at 0.1. When Trade is an endogenous variable, no variable was significant. When Industry is the endogenous variable only Agriculture and Industry are significant at 0.1 and 0.05 respectively. Whereas the results for the random walk model in Table 8 showed that at lag 1, with GDP as an endogenous variable none of the variables is significant. With Agriculture as endogenous variable Agriculture and Industry were significant at 0.1 and 0.05 respectively. When Trade is an endogenous variable none of the variables is significant. When Industry is the endogenous variable only Industry was significant at 0.1. Using the R software to find the AIC and (BIC) of the model, the result came out as follows.

For the existing VAR model:

$$AIC = 702.0677 \text{ and } BIC = 716.8060$$

Whereas for the modified model:

$$AIC = 58.1558 \text{ and } BIC = 71.0432.$$

Table 7: Parameter estimation of the adjusted data model  
Dependent variable (y):

Variables	(GDP)	(Agriculture)	(Trade)	(Industry)
GDP(t-1)	**0.674	-7.972	31.695	**47.193
	0.111	30.075	27.843	22.594
Agriculture(t-1)	0.001	-0.111	0.089	***-0.354
	0.001	0.139	0.128	0.104
Trade(t-1)	*0.002	***1.026	0.007	-0.099
	0.001	0.214	0.198	0.161
Industry(t-1)	-0.001	*0.356	0.050	**0.321
	0.001	0.201	0.186	0.151
const	318.044	53685.440	27282.480	-28311.550
	430.518	116654.000	107995.800	87634.870
Observations	37	37	37	37



<b>R2</b>	0.742	0.537	0.101	0.317
<b>Adjusted R2</b>	0.710	0.480	-0.011	0.231
<b>Residual Std. Error (df = 32)</b>	2422.525	656412.100	607692.800	493121.600
<b>F Statistic (df = 4; 32)</b>	***23.026	***9.293	0.900	**3.710

Note: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01

The original VAR model for order 1 was:

$$\begin{aligned}
 GDP_t &= +00318.044 + 00.674GDP_{t-1} + 0.001Agriculture_{t-1} + 0.002Trade_{t-1} - 0.001Industry_{t-1} \\
 &\quad + 002422.525 \\
 Agriculture_t &= +53685.440 - 07.972GDP_{t-1} - 0.111Agriculture_{t-1} + 1.026Trade_{t-1} + 0.356Industry_{t-1} \\
 &\quad + 656412.100 \\
 Trade_t &= +27282.480 + 31.695GDP_{t-1} + 0.089Agriculture_{t-1} + 0.007Trade_{t-1} + 0.050Industry_{t-1} \\
 &\quad + 607692.800 \\
 Industry_t &= -28311.550 + 47.193GDP_{t-1} - 0.354Agriculture_{t-1} - 0.099Trade_{t-1} + 0.321Industry_{t-1} \\
 &\quad + 493121.600
 \end{aligned}$$

**Table 8 Parameter estimation of the Random Walk model**  
Dependent variable (y):

Variables	(GDP)	(Agriculture)	(Trade)	(Industry)
<b>GDP(t-1)</b>	-0.258	0.255	0.173	0.266
	0.181	0.160	0.168	0.343
<b>Agriculture(t-1)</b>	0.104	*-0.329	0.043	-0.430
	0.207	0.184	0.193	0.393
<b>Trade(t-1)</b>	-0.061	-0.047	-0.239	-0.614
	0.202	0.179	0.187	0.382
<b>Industry(t-1)</b>	-0.036	**0.154	0.132	*-0.318
	0.084	0.074	0.078	0.159
<b>const</b>	0.000	-0.000	0.000	-0.000
	0.112	0.099	0.104	0.212
<b>Observations</b>	37	37	37	37
<b>R2</b>	0.074	0.273	0.202	0.262
<b>Adjusted R2</b>	-0.042	0.182	0.102	0.170
<b>Residual Std. Error (df = 32)</b>	0.681	0.603	0.632	1.288
<b>F Statistic (df = 4; 32)</b>	0.635	**3.008	2.023	**2.844

Note: \*p<0.1, \*\*p<0.05, \*\*\*p<0.01

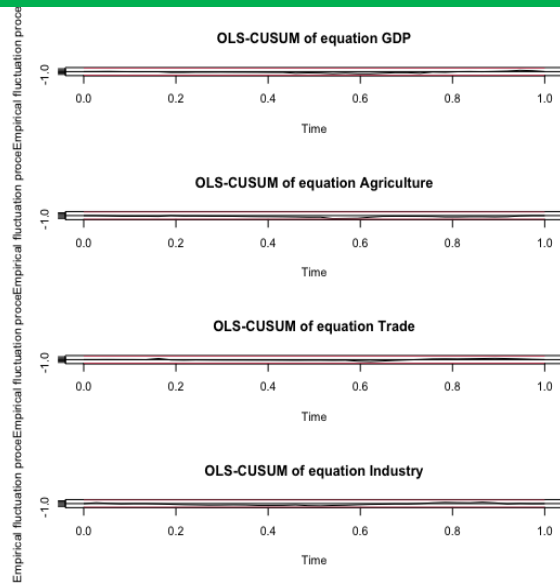
The modified VAR model for order 1 is:

$$\begin{aligned}
 GDP_t &= -0.258GDP_{t-1} + 0.104Agriculture_{t-1} - 0.061Trade_{t-1} - 0.036Industry_{t-1} + 0.681 \\
 Agriculture_t &= +0.255GDP_{t-1} - 0.329Agriculture_{t-1} - 0.047Trade_{t-1} + 0.154Industry_{t-1} + 0.603 \\
 Trade_t &= +0.173GDP_{t-1} + 0.043Agriculture_{t-1} - 0.239Trade_{t-1} + 0.132Industry_{t-1} + 0.632 \\
 Industry_t &= +0.266GDP_{t-1} - 0.430Agriculture_{t-1} - 0.614Trade_{t-1} - 0.318Industry_{t-1} + 1.288
 \end{aligned}$$

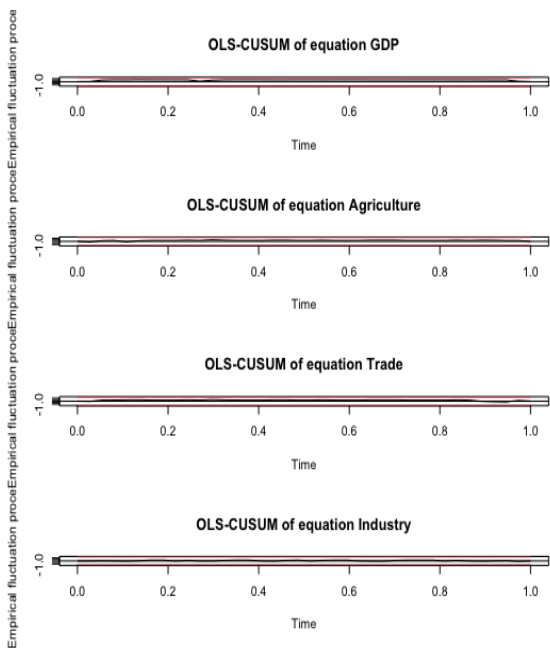
### 3.7Diagnostic Tests

It was observed in Figures 5 and 6 that the variable lines in both original and random walk data do not cross the red line. Hence, there are no structural breaks in the models. That confirms that the models are stable.

### 3.7.1Test for Structural breaks



**Figure 5** Stability test for structural breaks in the residuals of original data



**Figure 6** Stability test for structural breaks in the residuals of random walk data

### 3.7.2 Testing for serial correlation

The result from the serial test carried for both the original data and random walk model are given in table 9. The p-values are given respectively as 0.9864 and 0.9998. The result showed that there is no serial correlation in the two models because the p-values are greater than 0.05 level of significance.

**Table 9** Result of Serial correlation test

Model	p-value at 95% C.I
Original data model	0.9864
Random walk model	0.9998

### 3.7.3 Testing for the presence of Heteroscedasticity

Table 10 shows the result of the Arch test for both the original data model and random walk model. The p-values obtained for the models are given as 1 and 0.9998 for original data and random walk transformed data respectively. Since both the p-values are greater than 0.05 level of significance, it confirms that the data sets do not suffer from heteroscedasticity i.e. they are homogenous in nature.

**Table 10** Result of Arch test

Model	p-value at 95% C.I
Original data model	1.0000
Random walk model	0.9998

## 4.0 CONCLUSION

This study builds a Modified Vector Autoregressive (MVAR) model for the analysis of Nigeria's Gross Domestic Product (GDP) using data in Agriculture, Trade and Industry. This was based on the data transformation which utilized the absolute values of the deviation of each observation from the mean of the variables. The data was made stationary by a random walk process which was confirmed by the Augmented Dickey-Fuller test. Compared with the conventional Vector Autoregressive (VAR), it can be concluded that the MVAR has a better fit due to its least AIC and BIC. Users can employ the use of MVAR to improve their practices and policies. More research on the statistical techniques that can help to reduce the variability in the observations of multivariate time series data could also be studied in future.

## REFERENCES

Akkaya, M. (2021). Vector Autoregressive Model and Analysis. *Handbook of Research on Emerging Theories, Models, and Applications of Financial*

[https://doi.org/10.1007/978-3-030-54108-8\\_8](https://doi.org/10.1007/978-3-030-54108-8_8).

- Doz, Y. (2020). Fostering strategic agility: How individual executives and human resource practices contribute. *Human Resource Management Review*, 30(1). 100693.
- Huang R, Luo G, & Duan Q, (2019). Using Baidu search index to monitor and predict newly diagnosed cases of HIV/AIDS, syphilis and gonorrhoea in China. *Estimates from a vector autoregressive (VAR) model. BMJ Open* 2020;10: e036098. doi: 10.1136/bmjopen-2019-036098.
- Ivanov, D., Sokolov, B., Chen, W., Dolgui, A., Werner, F., and Potryasaev, S. (2021). A control approach to *scheduling* flexibly configurable jobs with dynamic structural-logical constraints. *IISE transactions*, 53(1), 21-38.
- Paolo, G., Siem, J. K. and Julia, S. (2021). Vector Autoregressions with Dynamic Factor Coefficients and Conditionally Heteroscedastic Errors. *Tinbergen Institute Discussion Paper 2021-056/III*.
- Petrusevich, D. (2020). Investigation of order selection in the Vector Autoregression Model. *International Multidisciplinary Scientific GeoConference: SGEM; Sofia, Vol. 20, Iss. 2.1, (2020)*. DOI: 10.5593/sgem2020/2T/s07.025.
- Yue Cao, Lingling Guo, Ying Qu, (2022). Evaluating the dynamic effects of mitigation instruments on CO2 emissions in China's nonferrous metal industry: A vector autoregression analysis, *Science of The Total Environment, Volume 853, 2022, 158409, ISSN 0048-9697*, <https://doi.org/10.1016/j.scitotenv.2022.158409>.
- Yuliadi, I. (2020). An Analysis of Money Supply in Indonesia: Vector Autoregressive (VAR) Approach. *The Journal of Asian Finance, Economics and Business*.