THE NEXUS BETWEEN FOREIGN AID AND FOREIGN DIRECT INVESTMENT IN NIGERIA: SIMULTANEOUS EQUATIONS APPROACH

Mohammed Kabir Garba[1*], Saheed Busayo Akanni[2], Alabi Waheed Banjoko[3] and Toluwa Celestine Oladele[4]

[1*, 2, 3] *Department of Statistics, University of Ilorin, Nigeria*

[4]*Department of Finance, University of Ilorin, Nigeria*

[1*]garba.mk@unilorin.edu.ng, [2]akannisb2018@gmail.com, [3]banjoko.aw@unilorin.edu.ng, [4]ct.oladele@mail1.ui.edu.ng

ABSTRACT

Undoubtedly, the choice of models for analyzing the relationship between level stationary Foreign Aid (FA) and Foreign Direct Investment (FDI) variables using single-equation modeling techniques such as Multiple Linear Regression (MLR) model produces non-spurious results in the sense that the coefficient of determination (R^2) is always less than the Durbin-Watson (DW) statistic. However, non-spurious MLR in this case might not adequately fit the FA-FDI relationship because macroeconomic variables are usually prone to problems of simultaneity, serial correlation as well as autocorrelation. This study therefore shed light on the FA-FDI relationship in Nigeria using a system of simultaneous equation modelling techniques. The FA and FDI variables are proxied as endogenous variables while the Inflation (INF), Population Growth (PG), Trade (TR), and Real Interest Rate (RI) are proxied as instruments. Pre-tests analysis of these time series datasets extracted from the repository of World Governance Index did not only reveal that the six series are level stationary series I(0) but also shown that there was a two-way causation between FA and FDI variables. Results from the estimation techniques further showed that the Three-Stage Least Squares (3SLS) and Seemingly Unrelated Regression (SUR) outperformed the Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS) estimators at estimating the structural parameters of the exactly identified and overidentified equations embedded in the model. Findings from the SUR estimates of the overidentified equation established that Nigeria's FA is projected to increase by 0.217188 units on average for each one-unit increase in the FDI.

Keywords: *Foreign Aid, Foreign Direct Investment, Instruments, Level Stationary Series, Simultaneous Equation, Nigeria.*

Received for review: 12-07-2023; Accepted: 29-09-2023; Published: 10-10-2023

DOI: 10.24191/mjoc.v8i2.16635

1. Introduction

The relationship between Foreign Aid (FA) and Foreign Direct Investment (FDI) is contentious and research results on it remain inconclusive (Selaya & Sunesen, 2012). Consequently, proper understanding of FA and FDI relationship is fundamental because, FDI provides an important source of capital inflow force (Aluko *et al.*, 2021) while FA promotes growth in a good policy environment (Easterly, 2003). Statistics show that African economies are growing fast with

Nigeria leading the continent's economy with an estimated Gross Domestic Product (GDP) of US$490 billion as reported by Visafrican (2021). Nigeria is also one of the major faces in Sub-Saharan Africa. One of the best ways to further reinforce the country's economy is through the formulation of policies that will promote the judicious use of foreign direct investment and foreign aid variables. Empirical evidence has established that FDI and FA are some of the major determinants of a country's economic growth (see Okafor *et al.*, 2016).

In practice, two or more related level stationary macroeconomic variables or series {i.e., I(0)s} are scarcely found in most econometric studies. This is because, many macroeconomic series are either integrated or non-stationary (Yaffee & McGee, 2000). Noticeably, time series variables under any econometric studies are either difference stationary series of order one {i.e., I(1)s}or mixture of I(0)s and I(1)s. For instance, analytical experience has shown that macroeconomic variables such as Gross Domestic Product (GDP) Per Capita and Foreign Aid are frequently I(1) and I(0) respectively. Sometimes, the Foreign Direct Investment (FDI) can either be I(1) or I(0) depending on how the variable is measured. When the FDI is measured as net bilateral aid flows from Development Assistance Committee (DAC) donors, Germany (in current US$) for example, it is usually I(1) whereas when it is measured as the percentage of GDP it is usually I(0) (World Governance Index, 2019). Furthermore, Foreign Aid (measured as % of Gross National Income), Inflation (measured as annual %), Population Growth (measured as annual %), Trade (measured as % of GDP), and Real Interest Rate (measured as %) are usually I(0)s.

In the econometrics context, regression of level stationary series on another level stationary series yields a non-spurious regression model. However, for MLR to fit time series datasets perfectly, the cause-and-effect relationship (if any) in such a model should run from the explanatory variables to the regressor. Apart from the stationarity assumptions, time series variables are expected to satisfy the assumption of multicollinearity for the usual MLR to adequately capture the dynamics of the series (see Bayo *et al*., 2021; Garba and Sikiru, 2022). If the assumption of multicollinearity is violated, then alternative methods that are robust for existence of multicollinearity like ridge regression or LASSO among others is desirable.

But when the dependent variable or regressor is determined by the explanatory variables and some of the explanatory variables are in turn determined by the regressor, then a system of Simultaneous Equations Model (SEM) is desirable (Gujarati & Porter, 2009). In other words, there exists an endogeneity problem in one or more equations that make up the system. Ignoring the endogeneity problem in this case can lead to inconsistent estimates. The Two-Stage Least Squares (2SLS) is however used to correct for the simultaneity or endogeneity problem. Combining the 2SLS with the Seemingly Unrelated Regression (SUR) method results in a simultaneous estimation of the system of equations by the Three-Stage Least Squares (3SLS) method (Zellner & Theil, 1962). The SUR method is also a system estimator that captures the problem of error correlation across the system of simultaneous equations, see Yahya *et al.* (2008). More so, as pointed by Henningsen and Hamann (2008), if all regressors are exogenous, then system of simultaneous equations can be consistently estimated by ordinary least squares (OLS) and seemingly unrelated regression (SUR).

In the literature, FDI and FA have been noticeably employed as dependent and independent variables. Besides, a common feature of some published research on FA–FDI relationship was either based on panel data or multivariate time series techniques. For instance, Aluko (2020) applied a dynamic panel threshold estimation methodology to study the effects of institutional quality and financial development on the FA-FDI relationship in Africa. Results from his study showed that the effectiveness of foreign aid in attracting foreign direct investment is greater in countries with better institutional quality and sound financial development**.** Ndambendia and Njoupouognigni (2010) employed the techniques of pooled mean group (PMG) estimator to show that there is strong evidence of the positive impact of foreign aid and foreign direct investment on economic growth in 36 Sub-Saharan Africa countries.

In his study, Arazmuradov (2011) used Seemingly Unrelated Regression Equation (SURE) techniques to investigate the relationship between foreign aid (Official Development Assistance (ODA)), foreign direct investment (FDI) and their effect on domestic investment in five landlocked and emerging economies of Central Asia. His findings revealed that the ODA-FDI nexus is present in countries with low per capita GDP and economic growth. To investigate whether Foreign Aid increases Foreign Direct Investment, Selaya and Sunesen (2012) used the Generalized Method of Moments (GMM) to show that aid invested in complementary inputs draws in FDI, while aid invested in physical capital crowds it out. Moreover, Ozekhome (2017) employed the Generalized Method of Moments (GMM) Estimator to investigate the impact of foreign aid and foreign direct investment (FDI) on economic growth in the Economic Community of West African States (ECOWAS) region. His empirical results revealed that trade openness; foreign direct investment, real gross domestic capital formation, human capital, and lagged real GDP (a measure of previous economic growth/market size) are the principal drivers of economic growth in ECOWAS countries.

Bhavan *et al.*, (2011) investigated whether or not foreign aid attracts foreign direct investment (FDI) in South Asian economies using various econometric techniques such as co-integration and Granger causality analyses. Results from their study suggested that both aid in the shape of physical capital and aid for human capital and infrastructure development serve as complementary factors to foreign direct investment rather than being substitutable in South Asian economies. In their study, Trevino and Upadhyaya (2003) examined the effectiveness of FA and FDI in five developing Asian countries (India, Indonesia, Nepal, Philippines and Thailand) using pooled annual time series data from 1990 to 1999. Their findings indicated that both FA and FDI have positively contributed to economic growth, but that the impact of FDI is greater than that of foreign aid for relatively open economies in the sample. To fathom the effectiveness of foreign aid, foreign direct investment, and economic freedom for selected 28 Asian countries, Tiwari (2011) employed static and dynamic panel data techniques to show that an increase in fiscal freedom, financial freedom and domestic capital stock were significant factors positively affecting economic growth while freedom from corruption, inflow of foreign direct investment and foreign aid were significant factors that negatively affect economic growth. Furthermore, Kimura and Todo (2010) used gravity equation-type estimation to show that foreign aid in general does not have any significant effect on FDI. Their further findings later revealed that for differences in the size of aid effects across donor countries, foreign aid from Japan in particular has a vanguard effect i.e., Japanese aid promotes FDI from Japan but does not attract FDI from other countries.

Little has been made to investigate the Foreign Aid–Foreign Direct Investment relationship in developing countries, especially in Nigeria. To fill the gap in the literature, this study therefore examines the Foreign Aid – Foreign Direct Investment relationship in Nigeria using six related level stationary time series variables namely; Foreign Aid (FA), Foreign Direct Investment (FDI), Inflation (INF), Population Growth (PG), Trade (TR) and Real Interest Rate (RI). We address this research gap by using Multiple Linear Regression (MLR) modeling techniques; which produce a non-spurious model whenever a level stationary time series is regressed on another level stationary time series. Furthermore, we formulate a two-equation model to further check if there is simultaneity between Foreign Aid (FA) and Foreign Direct Investment (FDI) variables. This work contributes significantly to the existing literature in two aspects. First, we are the first to fit a single-equation regression model to analyze the FA-FDI relationship in Nigeria. Second, we were able to show that, the endogeneity or simultaneity problem also exists between FA and FDI despite being level stationary series.

2. Methodology

2.1 Data Description

This work used annual time series data from 1970 to 2019 on Foreign Direct Investment (FDI), Foreign Aid (FA), Inflation (INF), Population Growth (PG), Trade (TR) and Real Interest Rate (RI) to investigate the two-way or simultaneous relationship between FA and FDI variables as shown in Table 1 (World Governance Index, 2019).

Table 1**.** Description of variables for the study.

Variable	**Acronym**	**Unit of Measurement**
Foreign Direct Investment	FDI	% of GDP
Foreign Aid	FA	% of GNI
Inflation	INF	Annual %
Population Growth	PG	Annual %
Trade	TR	% of GDP
Real Interest Rate	RI	% of GDP

2.2 Unit Root Analyses

The first part of the analysis focused on the stochastic properties of the series by testing FDI, FA, INF, PG, TR and RI variables for the presence of unit roots using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests, see Garba *et. al* (2020). The essence of doing this was to properly determine if truly these time series variables were stationary at levels or not. It was discovered based on the benchmark-criterion of weak stationary series (see Kwiatkowski *et al.*, 1992). The flowchart shows in Figure 1 below is a guide for model selection of time series variables.

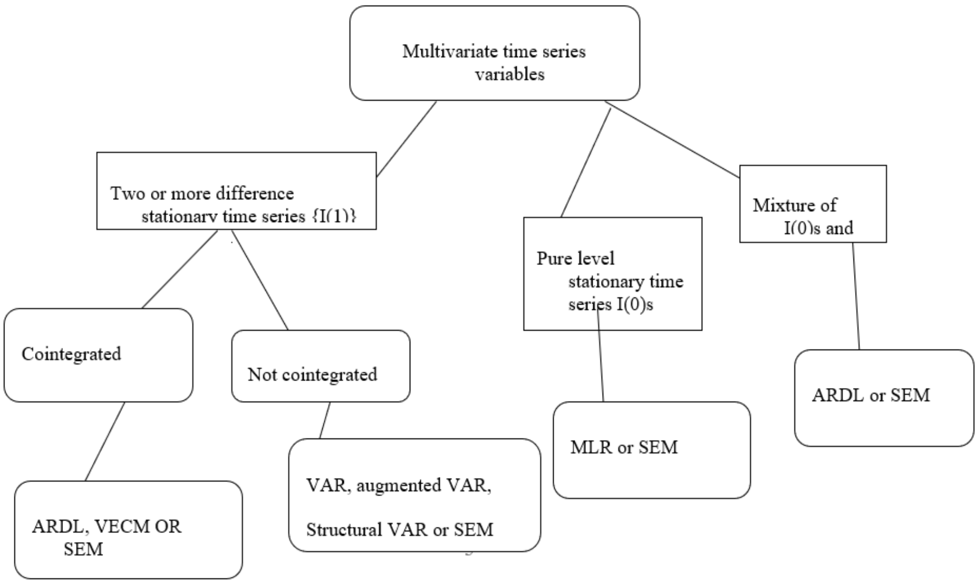

Figure 1. Flow chart for model selection.

From the Flowchart in Figure 1, it is seen that when the time series in a study are all I(1)s and are cointegrated, the ARDL, VECM or SEM model is desirable for analyzing the time series data depending on the objectives of the study. However, if multivariate time series are not cointegrated, then either VAR, augmented VAR, structural VAR or SEM is appropriate for analyzing the time series depending on the objectives of the study. Moreover, if multivariate time series are found to be all I(0)s, then either MLR or SEM is suitable for analyzing the series. Besides, if multivariate time series are mixtures of I(0) and I(1), then either ARDL or SEM is desirable for analyzing the time series depending on the objective of the study. More importantly, if there is at least one differenced stationary time series of order two {i.e. I(2)} in a multivariate time series setting, then SEM is apposite for analyzing such multivariate time series data since it is not perturbed by order of integration of time series.

2.3 Model Specification

To study the foreign aid–foreign direct investment relationship in Nigeria, we considered a dynamic regression model of the form:

$$\mathrm{FDI}_t = \beta_0 + \beta_1 \mathrm{FA}_t + \beta_2 \mathrm{INF}_t + \beta_3 \mathrm{PG}_t + \beta_4 \mathrm{TR}_t + \beta_5 \mathrm{RI}_t + u_t \quad (1)$$

Where: FDI$\mathrm{FDI}_{t_t}$is the dependent variable, FAINF_{tt}, INF_t,PG_t, TR_t, and RI_t are independent variables or input series, $\underline{\beta}$ =$\underline{\beta} = (\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$ is a vector of regression parameters to be estimated and u_t is the random error term of the model with the assumption that $u_t \sim (0, \sigma^2)$.

FDI_t = FDI at current time period t (2019), FA_t = FA at current time period t (2019), INF_t = INF at current time period t (2019), PG_t = PG at current time period t (2019), TR_t = TR at current time period t (2019) and RI_t = RI at current time period t (2019).

If the FDI, FA, INF, PG, TR and RI variables are all stationary at levels {i.e. they are all I(0)s} and there is no simultaneity or two-way relationship between FA and FDI variables, then the Time Series Regression in equation (1) is appropriate for modelling the relationship between the FDI and input variables. Otherwise, we consider other multivariate time series techniques or systems of simultaneous equations models.

However, to investigate if there is a two-way or simultaneous relationship between FA and FDI variables, we formulate a two-equation model from equation (1) as follows:

$$\mathrm{FDI}_{1t} = \beta_{10} + \gamma_{11}\mathrm{INF}_{1t} + \gamma_{12}\mathrm{PG}_{2t} + \gamma_{13}\mathrm{TR}_{3t} + \gamma_{14}\mathrm{RI}_{4t} + u_{1t} \quad (2)$$

$$\mathrm{FA}_{2t} = \beta_{20} + \beta_{21}\mathrm{FDI}_{1t} + u_{2t} \quad (3)$$

Where: $\mathrm{FDI}_{1t}$and $\mathrm{FA}_{2t}$are current endogenous variables, $\mathrm{INF}_{1t}$$\mathrm{INF}_{1t}$, PG_{2t}, $\mathrm{TR}_{3t}$and RI_{4t} are current exogenous variables, u_{1t} and u_{2t} are assumed to be well-behaved with a multivariate normal distribution u~NID(0, $\sum$), β_{10}, β_{20}, β_{21}, γ_{11}, γ_{12}, γ_{13} and $\gamma_{14}$are the structural parameters of the two-equation model.

2.4 Model identification

In a system of Simultaneous Equation Model (SEM), identification precedes estimation. An equation belonging to a system of SEM is identified if it has a unique statistical form. However, for identification of the entire SEM, the model must be complete and each equation in the model must be identified. The identification of SEM can be reflected by its rank and order conditions (Hausman, 1983).

2.4.1 Order Condition of Identification

This is a necessary but not sufficient condition of identification. The essence of this identification process is to determine if an equation in a system is just (exactly) identified, over-identified, or under-identified. Consider a model of G simultaneous equations. In order for an equation to be identified, the number of predetermined variables excluded from the equation must not be less than the number of endogenous variables included in that equation less one. To determine the order condition of identification of equations (2) and (3), we use a mathematical relation given by equation (4) as follows:

$$K - M \geq G - 1 \tag{4}$$

Where:

G = Total number of equations (total number of endogenous variables),

K = Total number of variables in the model (endogenous plus exogenous),

M = Number of variables (endogenous plus exogenous) in a particular equation.

The order condition of identification states that:

1. If $K - M = G - 1$, then the equation is just or exactly identified
2. If $K - M > G - 1$, then the equation is over-identified
3. If $K - M < G - 1$, then the equation is under-identified

2.4.2 Rank Condition of Identification

The rank condition investigates whether two or more equations are linearly dependent on each other, which would be the case if the sum of two equations would equal a third equation in the model. If that is the case, it is impossible to identify all structural parameters.

2.5 Reduced-form equations

The reduced-form equations express an endogenous variable solely in terms of predetermined variables and the stochastic disturbances. This allows the Ordinary Least Squares (OLS) estimator to consistently estimate the reduced-form parameters.

Since equation (2) contains only exogenous variables, its reduced-form equation is given by:

$$FDI_{1t} = \pi_{10} + \pi_{11} INF_{1t} + \pi_{12} PG_{2t} + \pi_{13} TR_{3t} + \pi_{14} RI_{4t} + v_{1t} \tag{5}$$

Where: $\pi_{10}=\beta_{10}$, $\pi_{11}=\gamma_{11}$, $\pi_{12}=\gamma_{12}$, $\pi_{13}=\gamma_{13}$, $\pi_{14}=\gamma_{14}$ and $v_{1t}=u_{1t}$.

To obtain the reduced-form equation for equation (2), equation (2) is substituted into equation (3) as follows:

$$FA_{2t} = \beta_{20} + \beta_{21}(\beta_{10} + \gamma_{11}INF_{1t} + \gamma_{12}PG_{2t} + \gamma_{13}TR_{3t} + \gamma_{14}RI_{4t} + u_{1t}) + u_{2t}$$

$$FA_{2t} = (\beta_{20} + \beta_{21}\beta_{10}) + (\beta_{21}\gamma_{11})INF_{1t} + (\beta_{21}\gamma_{12})PG_{2t} + (\beta_{21}\gamma_{13})TR_{3t} + (\beta_{21}\gamma_{14})RI_{4t} + (\beta_{21}u_{1t} + u_{2t})FA_{2t} = \pi_{20} + \pi_{21}INF_{1t} + \pi_{22}PG_{2t} + \pi_{23}TR_{3t} + \pi_{24}RI_{4t} + v_{2t} \quad (6)$$

Where: $\pi_{20}=\beta_{20}+\beta_{21}\beta_{20}\pi_{20} = \beta_{20} + \beta_{21}\beta_{10}$, $\pi_{21}=\beta_{21}\gamma_{11}\pi_{21} = \beta_{21}\gamma_{11}$, $\pi_{22}=\beta_{21}\gamma_{12}$, $\pi_{23}=\beta_{21}\gamma_{13}$, $\pi_{24}=\beta_{21}\gamma_{14}$ and $v_{2t}=\beta_{21}u_{1t}+u_{2t}v_{2t} = \beta_{21}u_{1t} + u_{2t}$

2.6. Hausman Specification Test for Overidentified Equation

The Hausman specification test developed by Hausman (1978) detects endogenous regressors in a regression model. Observably, FDI appears in equation (2) as a regressand but appears in equation (3) as a regressor. Thus, it is crucial to confirm whether FDI_{1t} is correlated with u_{2t}or not. If there is correlation between them, then FDI is endogenous. Otherwise, FDI is a regresssand. Moreover, the Hausman tests in this study followed the numerical examples of simultaneity tests between GDP and Money Supply in Gujarati and Porter (2009) in which money supply was regressed on the fitted values of GDP and v respectively. The following hypotheses were formulated with respect to equations (3):

H_0: Corr(FDI_{1t},u_{2t})≠0 (There is correlation between FDI_{1t} and u_{2t})

H_1: Corr(FDI_{1t},u_{2t})=0 (There is no correlation between FDI_{1t} and u_{2t})

Essentially, the Hausman test here involves the following two steps:

Step 1: To get rid of the likely correlation between FDI and FA, FDI was regressed on all the exogenous (predetermined) variables in the system. That is, we regress FDI on INF, PG, TR and RI as follows:

$$FDI_{1t} = \widehat{(\pi}_{10} + \hat{\pi}_{11}INF_{1t} + \hat{\pi}_{12}PG_{2t} + \hat{\pi}_{13}TR_{3t} + \hat{\pi}_{14}RI_{4t}) + \hat{v}_{1t} \quad (7)$$

Where: $\hat{v}_{1t}$ are the usual Ordinary Least Squares (OLS) residuals from equation (7). The equation (7) is re-written a follow:

$$\widehat{FDI}_{1t} = \hat{\pi}_{10} + \hat{\pi}_{11}INF_{1t} + \hat{\pi}_{12}PG_{2t} + \hat{\pi}_{13}TR_{3t} + \hat{\pi}_{14}RI_{4t} \quad (8)$$

Where:$\hat{v}_{1t}$is an estimate of the mean value of FDI conditional upon the fixed exogenous variables. Substituting equation (8) into equation (7) further yields:

$$FDI_{1t} = \widehat{FDI}_{1t} + \hat{v}_{1t} \quad (9)$$

Step 2: Here, we substitute equation (9) into the over-identified equation (3) as follows:

$$\mathrm{FA}_{2t} = \beta_{20} + \beta_{21}(\widehat{\mathrm{FDI}}_{1t} + \hat{v}_{1t}) + u_{2t}$$

$$\mathrm{FA}_{2t} = \beta_{20} + \beta_{21}\widehat{\mathrm{FDI}}_{1t} + (\beta_{21}\hat{v}_{1t} + u_{2t})$$

$$\mathrm{FA}_{2t} = \beta_{20} + \beta_{21}\widehat{\mathrm{FDI}}_{1t} + \hat{v}_{t} \tag{10}$$

Such that $\hat{v}_t = \beta_{21}\hat{v}_{1t} + u_{2t}$

Decision rule: Reject the null hypothesis if the t-value of $\hat{v}_{1t}$ is statistically significant. Otherwise, accept the null hypothesis.

2.7 Durbin-Watson Statistic and Critical Values

In econometrics framework, the issues of autocorrelation or serial correlation often surface in the residuals of usual econometrics models such as the OLS, ILS, 2SLS, etc (see Gujarati & Porter, 2009). Consequently, it is important to check for the presence of serial correlation in the residuals of these estimators in order to prevent spurious results. However, this study checked the serial correlation in the residuals of equation (3) since it connects the variables of interests which are FA and FDI. Also, equation (3) is over-identified which means that 2SLS estimator is a candidate that can estimate the FA-FDI relationship in equation (3). Moreover, for proper determination of serial correlation status of the residuals of equation (3) under 2SLS estimator, this study therefore compares the Durbin-Watson statistic (DW) with the Farebrother (1980) critical values. The following hypothesis was formulated with respect to the residuals of equation (3):

H_0: There is zero serial correlation in the residuals of the model vs H_1: Not H_0

Decision rule: Reject H_0 if the DW statistic < Durdin-Watson Lower Bound (dL). Otherwise accept H_0. However, if the DW statistic is between dL and Durdin-Watson Upper Bound (dU), the test is inconclusive.

2.8 System estimators

Unlike the OLS, ILS and 2SLS estimators which estimate the parameters of equations embedded in a SEM equation-by-equation, the system estimators however estimate these parameters jointly and are expected to outperform these single-equation estimators if and only if the equations in the system are connected. Usually, equations is a system can be said to be connected either through correlation between an explanatory variable and the error term (endogeneity problem) or through contemporaneously correlation structures of the error terms across the systems of an equation (see Yaya et al., (2008); Zellner, 1962; Koutsoyiannis, 2003). For adequate estimations of the study's two-equation model (i.e., equations 2 and 3), the Three Stage Least Squares (3SLS) and Seemingly Unrelated Regression (SUR) estimators are also employed to estimate the regression coefficients or parameters of the SEM.

2.8.1 Seemingly Unrelated Regression (SUR) estimators

If the disturbances of equations embedded in a system of simultaneous equations are contemporaneously correlated across the system, a Generalized Least Squares (GLS) estimation leads to an efficient estimator for the coefficients (Henningsen & Hamann, 2008). This GLS estimator is termed the SUR estimator and it is developed by Zellner (1962). The only problem associated with this GLS is that the covariance matrix of the disturbance terms of the GLS is usually unknown. Instead, the Feasible Generalized Least Squares (FGLS) is used since it is based on an estimated covariance matrix of the disturbance terms. Henningsen & Hamann (2008) also noted if all the regressors in a system are exogenous, then such system can be conveniently estimated by OLS, Weighted Least Squares (WLS) and SUR. The mathematical relation for estimating the vector of a dependent variable y_i and exogenous variables X_I is stated as equation (11).

$$\hat{\beta} = (X'\hat{\Omega}^{-1}X)^{-1}X'\hat{\Omega}^{-1}y \tag{11}$$

The covariance matrix of the estimator in equation (11) is given by

$$\text{Cov}[\beta] = (X'\hat{\Omega}^{-1}X) \tag{12}$$

However, the estimator in equation (11) becomes the SUR estimator by replacing the $\hat{\Omega}$ in equations (11) and (12) with $\hat{\hat{\Sigma}} \otimes I_T$. Thus, the SUR estimator is of the form:

$$\hat{\beta} = (X'(\hat{\hat{\Sigma}} \otimes I_T)^{-1}X)^{-1}X'(\hat{\hat{\Sigma}} \otimes I_T)^{-1}y \tag{13}$$

Where $\hat{\hat{\Sigma}}$ is the estimated covariance matrix of the disturbance terms.

2.8.2 Three Stage Least Squares (3SLS) estimators

When the disturbances are contemporaneously correlated and there is endogeity in a system of simultaneous equations, the 3SLS developed by Zellner and Theil (1962) is desirable. The FGLS version of the 2SLS leads to consistent and asymptotically more efficient estimates (Henningsen & Hamann, 2008). However, the new fitted regressors are obtained using equation (14):

$$\hat{X}_i = Z_i(Z_i'Z_i)^{-1}Z_i'X_i \tag{14}$$

The 3SLS in this case is another version of equation (11) which is stated as equation (15)

$$\hat{\beta} = (\hat{X}'\hat{\Omega}^{-1}\hat{X})^{-1}\hat{X}'\hat{\Omega}^{-1}y \tag{15}$$

Also,

$$\hat{X} = \begin{bmatrix} \hat{X}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \hat{X}_G \end{bmatrix} \tag{16}$$

And

$$\widehat{Cov}[\hat{\beta}] = (\hat{X}'\hat{\Omega}^{-1}\hat{X})^{-1} \quad (17)$$

3. Results

In this section, the results of the analyses carried out on the Foreign Direct Investment (FDI), Foreign Aid (FA), Inflation Rate (INF), Population Growth (PG), Trade (TR) and Real Interest Rate (RI) series using GRETL are presented.

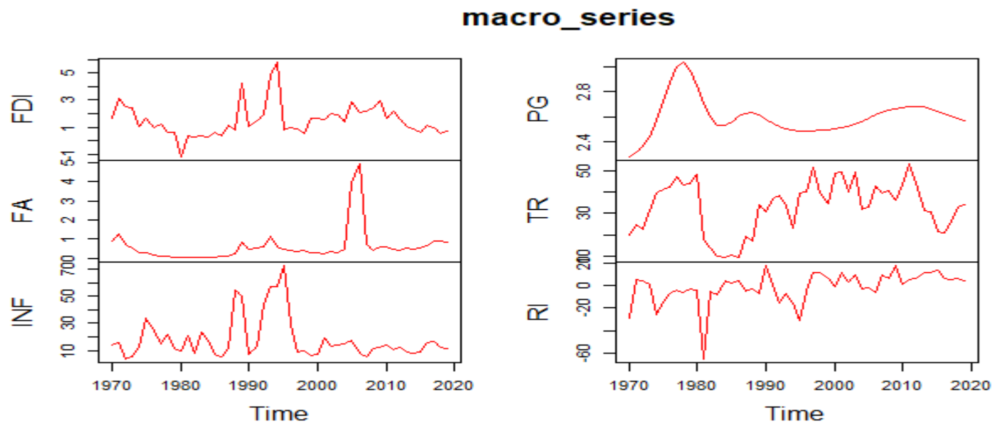

Figure 1. Time plots of Foreign Direct Investment (FDI), Foreign Aid (FA), Inflation (INF), Population Growth (PG), Trade (TR) and Real Interest Rate (RI).

Figure 1 reveals that FDI, FA, INF, PG, TR and RI series exhibit various trends (upward and downward) with fluctuations; which are indications of non-stationarity in each of the series. To correctly determine the order of integration of these series, each of the series was subjected to unit root analyses using the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests. The results reported by this unit root test are presented in Table 2 below.

Table 2. Unit root tests results for the series

Variable	KP-statistic	Critical value	d_{max}	Remark
FDI	0.10981	0.463	I(0)	Stationary
FA	0.265995	0.463	I(0)	Stationary
INF	0.188577	0.463	I(0)	Stationary
PG	0.07886	0.463	I(0)	Stationary
TR	0.167426	0.463	I(0)	Stationary
RI	0.751096	0.463	I(0)	Stationary

***Note**: Chosen level of significance is 5%, d_{max}=Max. order of integration*

From Table 2, the results reported by the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests showed that the null hypotheses of stationarity were accepted for FDI, FA, INF, PG, TR and RI series. This implies that these series are level stationary series {I(0)s}. Since all these series have been confirmed to be I(0)s, the time series regression model can be fitted to the datasets if and only if there is no simultaneity or two-way causation between FA and FDI variables.

Table 3. OLS estimates of time series regression (eqn 1).

Variable	Coeff	Std. Error	t-Statistic	p-value
Const	6.379023	2.659203	**2.398847**	**0.0208**
FA	0.386925	0.172841	**2.23862**	**0.0303**
INF	0.035623	0.010315	**3.453432**	**0.0012**
PG	-2.395208	1.046689	**-2.288366**	**0.027**
TR	0.014957	0.012858	1.163188	0.251
RI	0.016074	0.011484	1.39966	0.1686
R^2= 0.384065	DW=1.850857	P-value (F) = 0.000521		

From Table 3 the estimated reduced-form equation is stated as equation (18):

$$\mathrm{FDI}_{1t} = 6.379023 + 0.386925\mathrm{FA}_{1t} + 0.035623\mathrm{INF}_{2t} - 2.395208\mathrm{PG}_{3t} + 0.014957\mathrm{TR}_{4t} + 0.016074\mathrm{RI}_{5t} \quad (18)$$

Results reported in Table 3 showed that FA and INF contributed positively significantly to the FA (i.e. p-values < 0.05) whereas TR and RI did not (i.e. p-values > 0.05). On the other hand, only PG contributed negatively significantly to the FA (p-value < 0.05). The value of R^2 (= 0.384065) which is less than the Durbin-Watson (DW=1.850857) statistic indicates that the model is non-spurious. Besides, the Durbin-Watson statistic (DW=1.850857) is very close to 2. This implies that there is no evidence of serial correlation in the fitted model. The next thing is to check if there is a two-way causation between FA and FDI or not.

Table 4. Summary results of order condition of identification (equations 2 and 3).

Equation	Excluded variable (K-M)	Included endog. Vars. (G-1)	Remark
2	1	1	Exactly identified
3	4	1	Over-identified

Based on the results of order condition of identification presented in Table 4, equations (2) and (3) are exactly and over-identified equations respectively. However, these equations are also identified by the rank condition since their additions will give rise to a third equation as suggested by Kutsoyiannis (2003).

Table 5. Reduced-form estimates (Eqn 3).

Variable	Coeff	Std. error	t-Statistic	p-value
Const	7.75429	2.70012	2.8718	**0.00620*** **
INF	0.0350536	0.0107621	3.2571	**0.00214*** **
PG	-2.88835	1.06788	-2.7048	**0.00962*** **
TR	0.0194072	0.013258	1.4638	0.1502
RI	0.0156221	0.0119836	1.3036	0.19899
R^2=0.313912	*DW=1.706350*	*P-value (F) = 0.001679*		

From Table 5, the estimated reduced-form equation is stated as equation (19):

$$\text{FDI}_{1t} = 7.75429 + 0.0350536\text{INF}_{1t} - 2.88835\text{PG}_{2t} + 0.0194072\text{TR}_{3t} + 0.0156221\text{RI}_{4t} \quad (19)$$

The essence of reduced-form estimates presented in Table 5 is to further obtain the predicted value $\widehat{\text{FDI}}_{1t}$ and residuals $\hat{v}_t$; which was later used for the simultaneity test as described by Abrevaya and Hausman (2004). Table 6 presents the results of simultaneity tests conducted on the over-identified equation (3).

Table 6. Results of Hausman specification tests (eq 3).

Variable	**Coeff**	**Std. error**	**t-ratio**	**p-value**
Const	0.331956	0.283543	1.1707	0.2476
$\widehat{\text{FDI}}_{1t}$	0.175292	0.172125	1.0184	0.3137
$\hat{v}_t$	0.264263	0.116428	2.2698	**0.02785****
R^2=0.116358	DW=1.066032	P-value (F) = 0.054638		

From Table 6, the results of simultaneity tests revealed that the fitted FDI (i.e., $\widehat{\text{FDI}}_{1t}$) do not statistically significantly affects the Foreign Aid (FA) at all levels of significance (i.e., p-value =0.31370>0.01, 0.05 and 0.1). On the other hand, the t-value of $\hat{v}_t$ is statistically significant (i.e., p-value = 0.02785) < 0.05), we cannot reject the hypothesis of simultaneity between FA and FDI. Besides, the Durbin Watson statistic (DW=1.066032) is substantially less than two; which means there is evidence of positive serial correlation. Furthermore, the null hypothesis of zero positive serial correlation in the residuals of equation (3) has been rejected since the DW statistic (=1.066032) is less than the dL (= 1.07). The value of 1.07 was obtained from Farebrother (1980) critical values table.

From the same Table 6, the estimated model of the Hausman specification tests equation is stated as equation (20):

$$\text{FA}_{2t} = 0.331956 + 0.175292\widehat{\text{FDI}}_{1t} \quad (20)$$

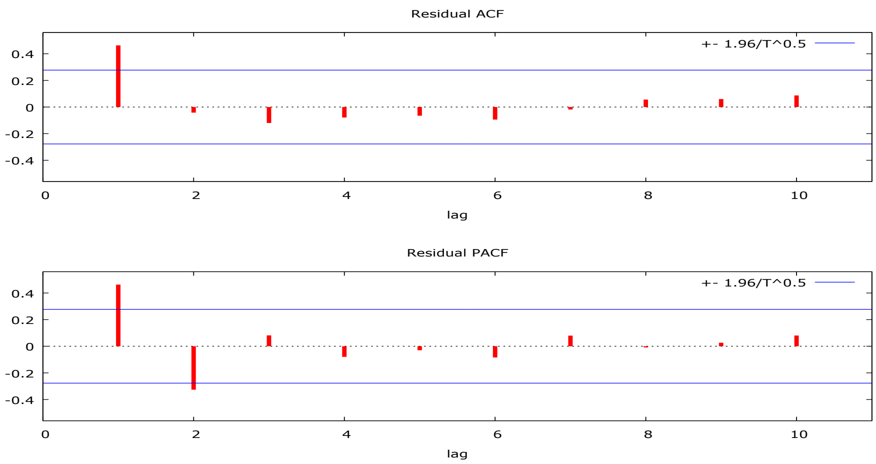

Figure 2. Residual Correlogram of the fitted equation (13)

The plots in Figure 2 show that there is one significant spike at lag 1 for the Autocorrelation Function (ACF) and there are two significant spikes at lags 1 and 2 for the Partial Autocorrelation Function (PAFC); which further confirms the presence of positive serial correlation in the residual of equation (20).

Table 7a. OLS and 2SLS estimates of the exactly identified equation (2)

Single-Equation Estimators								
Ordinary Least Squares (OLS)					**Two-Stage Least Squares (2SLS)**			
Var	**Coeff.**	**Std. Error**	**t-ratio**	**p-value**	**Coeff.**	**Std. Error**	**Z**	**p-value**
C	7.7543	**2.7001***	2.8718	0.0062***	7.7543	**2.7001***	2.8720	0.0041***
INF	0.0351	**0.0108***	3.2571	0.0021***	0.0351	**0.0108***	3.2570	0.0011***
PG	-2.8884	**1.0679***	-2.7048	0.0096***	-2.8884	**1.0679***	-2.7050	0.0068***
TR	0.0194	**0.0133***	1.4638	0.1502	0.0194	**0.0133***	1.4640	0.1432
RI	0.0156	**0.0120***	1.3036	0.1990	0.0156	**0.0120***	1.3040	0.1924
R^2=0.313912, DW= 1.706350and P-value (F) = 0.001679					R^2= 0.313912, DW= 1.706350 and P-value = 0.001679			

***Note**: The best results as reported by the standard errors of the regression parameters of the estimators are asterisked (*).*

Summary of results reported in Table 7a shows that the OLS and 2SLS estimates are identical; which means that any of the two estimators is appropriate for estimating the exactly identified equation (2) since it contains only exogenous variables. The OLS and 2SLS reported the same R^2 (= 0.313912); which means that 31.4% variation in FDI was explained INF, PG, TR and RI. Also, the standard errors of the structural parameters for these estimators are the same.

Table 7b. 3SLS and SUR estimates of the exactly identified equation (2)

System-Equation Estimators								
Three-Stage Least Squares (3SLS)					**Seemingly Unrelated Regression (SUR)**			
Var	**Coeff.**	**Std. Error**	**Z**	**p-value**	**Coeff.**	**Std. Error**	**t-ratio**	**p-value**
C	7.4973	**2.5500***	2.9400	0.0033***	7.6844	2.5601	3.0020	0.0044 ***
INF	0.0361	**0.0102***	3.5470	0.0004 ***	0.0355	**0.0102***	3.740	0.0011 ***
PG	-2.7824	**1.0091***	-2.7570	0.0058***	-2.8601	1.0125	-2.8250	0.0070 ***
TR	0.0183	**0.0125***	1.4620	0.1437	0.0191	0.0126	1.5190	0.1359
RI	0.0162	**0.0113***	1.4300	0.1528	0.0158	0.0114	1.3920	0.1706
R^2= 0.313600					R^2= 0.313881			

***Note**: The best results as reported by the standard errors of the regression parameters of the estimators are asterisked (*).*

Table 7b revealed that the 3SLS estimator outperformed the SUR, OLS and 2SLS estimators in estimating the structural coefficients of the exactly identified equation (2) since it reported the least values of the standard errors for the exogenous variables. The value of R^2(= 0.313600) reported by the 3SLS showed that 31.4% variation in FDI was explained by INF, PG, TR and RI.

Table 8a. 2SLS estimates of Over-identified Equation (3).

Single-Equation Estimators				
Two-Stage Least Squares (2SLS)				
Var	**Coeff.**	**Std. Error**	**Z**	**p-value**
C	0.3320	0.2823	1.1760	0.2397
FDI	0.1753	0.1714	1.0230	0.3064
R^2= 0.112911, DW= 1.000988 and P-value = 0.306383				

Based on the results reported in Table 8a, the FDI do not have significant effect on FA. (P-value = 0.3064>0.05). This means that the 2SLS estimator is not appropriate for estimating the structural parameters of this model.

Table 8b. 3SLS and SUR estimates of Over-identified Equation (3)

System-Equation Estimators								
Three-Stage Least Squares (3SLS)					**Seemingly Unrelated Regression (SUR)**			
Var	**Coeff.**	**Std. Error**	**Z**	**p-value**	**Coeff.**	**Std. Error**	**t-ratio**	**p-value**
C	0.3320	0.2766	1.2000	0.2301	0.268922	**0.1800***	1.4930	0.1419
FDI	0.1753	0.1680	1.0440	0.2965	0.217188	**0.0936***	2.3190	0.0247**
R^2=0.112911					R^2=0.112911			

***Note**: The best results as reported by the standard errors of the regression parameters are asterisked (*).*

$$\mathrm{FA}_{2t} = 0.268922 + 0.217188\widehat{\mathrm{FDI}}_{1t} \tag{21}$$

As shown in Table 8b, the SUR estimator outperformed the 2SLS and 3SLS estimators, since the p-value (=0.0247) of the FDI variable is less than the 0.05 chosen level of significance. That is, the FDI contributes positively significantly to FA in Nigeria. In other words, FA is projected to increase by 0.217188 units on average for each one unit increase in the FDI. Moreover, the value of R^2 (= 0.112911) showed that 11.3% variation in FDI was explained by FA.

4. Discussion

This work has used system of simultaneous equation techniques to investigate the relationship between Foreign Aid and Foreign Direct Investment relationship in Nigeria as described by Hausman (2004). Our summary of results for the unit root analyses presented in Table 2 revealed that the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test accepts the null hypotheses of stationarity for all the FDI, FA, INF, PG, TR and RI series. This means that the FDI, FA, INF, PG and TR are truly level stationary series. Results presented in Table 3 revealed that the fitted Multiple Linear Regression (MLR) model is non-spurious since the value of (R^2=0.384065) is less than the value of Durbin-Watson (DW=1.850857) statistic. With respect to the results of model identification reported in Table 4, the two-equation model (i.e., equations

(2) and (3)) has been identified by both the order and rank conditions respectively. The order condition identified equations (2) and (3) as exactly and over-identified equations respectively.

The Hausman Specification Test reported in Table 6 revealed that the fitted FDI (i.e., $\widehat{FDI}_{1t}\widehat{FDI}_{1t}\widehat{FDI}_{1t}\widehat{FDI}_{1t}$) does not significantly influence the FA at all levels of significance (i.e., p-value = 0.31370>0.01, 0.05 and 0.1). The Hausman specification test accepts the null hypothesis of simultaneity between FA and FDI in the over-identified equation (3); which implies that the MLR is not appropriate for analyzing the FA-FDI relationship in Nigeria. However, the fitted model stated as equation (20) also suffers from positive serial correlation aside endogeneity/simultaneity problem; which is observable from the value of the Durbin Watson statistic (DW=1.066032) which is substantially less than two (see Table 6 and Figure 2). Estimated results of the exactly identified equation (2) presented in Table 7(a) showed that the single-equation estimators (OLS and 2SLS) do not only produce identical estimates which are consistent but also produce identical standard errors for each of the regression coefficients in the model. On the other hand, estimates from the system estimators reported in Table 7(b) showed that the 3SLS estimator is preferred to OLS, 2SLS and SUR at estimating the structural parameters of the equation because it reported the least values of the standard error for each of the regression coefficients. Finally, estimates of the 2SLS detailed in Table 8(a) disclosed that the single-equation estimator 2SLS is not appropriate for estimating the structural parameters of the overidentified equation (3) since FDI does not have a significant effect on the FA (p-value=0.3064>0.05). However, further reports of the estimates of the system estimators reported in Table 8(b) revealed that the SUR estimator is preferred to 2SLS and 3SLS since the FDI impacts the FA statistically significantly.

5. Conclusion

This study examines the foreign aid-foreign direct investment relationship in Nigeria using the system of simultaneous equation modelling techniques in which case the endogenous variables FA and FDI are level stationary variables using level stationary variables such as INF, PG, TR, and RI as instruments. Based on the results of the analysis and summary of findings, the system estimators such as the 3SLS and SUR outperformed the single-equation estimators such as OLS and 2SLS at estimating the structural parameters of the exactly and overidentified equations embedded in the two-equation model. The 3SLS is the best at estimating the regression coefficients of the exactly identified equation while the SUR is the best at estimating the parameters of the overidentified equation. Additionally, findings from the SUR estimates of the overidentified equation established that Nigeria's FA is projected to increase by 0.217188 units on average for each one unit increase in the FDI. The findings of this work therefore recommends that related level stationary time series variables should always be checked for endogeneity problems, serial correlation, and others in order to prevent misleading results which might result from non-sense regression.

Acknowledgements

The authors are thankful to all the researchers whose works were beneficial to this study. We are grateful for the reviewers' valuable comments which improved the manuscript.

Funding

No funding was involved in this research study.

Authors' Contribution

All authors have accepted responsibility for the content of this manuscript and gave the approval for its submission to be published in MJoC. Mohammed Kabir Garba conceived the study, provided data in part, and offered directions for statistical tools and techniques. He also validated and interpreted the results, edited the manuscript, and handled all correspondence with the editorial board. Saheed Busayo Akanni collected and cleaned the data, carried out all the analyses, and prepared the draft of the manuscript with contributions from the co-authors. Alabi Waheed Banjoko provided the codes and supervised the analysis of data, contributed to manuscript writing, and provided critical revision of the article. Toluwa Celestine Oladele contributed to the manuscript writing, proofread the manuscript draft, and provided inputs as a financial expert.

References

Abrevaya, J., & Hausman, J. A. (2004). Response error in a transformation model with an application to earnings-equation estimation. *The Econometrics Journal*, 7(2), 366-388.

Aluko, O. A., Chen, G. S., & Opoku, E. E. O. (2021). Is foreign direct investment globalization-induced or a myth? A tale of Africa. *International Journal of Finance & Economics*.

Aluko, O. A. (2020). The foreign aid–foreign direct investment relationship in Africa: The mediating role of institutional quality and financial development. *Economic Affairs*, 40(1), 77-84.

Arazmuradov, A. (2011). Foreign aid, foreign direct investment and domestic investment nexus in landlocked economies of Central Asia.

Bayo, A. K., Rafiu, A. B., Funmilayo, A. T. & Oluyemi, O. I. (2021). Investigating the Impact of Multicollinearity on Linear Regression Estimates. *Malaysian Journal of Computing (MJoC)*, *6*(1), 698-714.

Bhavan, T., Xu, C., & Zhong, C. (2011). The relationship between foreign aid and FDI in South Asian economies. *International Journal of Economics and Finance*, 3(2), 143-149.

Easterly, W. (2003). Can foreign aid buy growth? *Journal of economic Perspectives*, 17(3), 23-48.

Farebrother, R. W. (1980). The Durbin-Watson test for serial correlation when there is no intercept in the regression. *Econometrica: Journal of the Econometric Society*, 1553-1563.

Garba, M. K., Akanni, S. B., Yahya, W. B., Kareem, K. Y., & Afolayan, R. B. (2020). Modelling Effects of some Factors that Contribute to Cereals Yields in Nigeria using Toda-Yamamoto Techniques. *SLU Journal of Science and Technology*, 1(1), 50-56.

Garba, M. K. & Sikiru, A. O. (2022). Assessment of the Impacts of Oil Prices on Nigeria Economy using Cobb-Douglas Production Function. *Malaysian Journal of Computing (MJoC)*, *7*(1), 938-951.

Gujarati, D. N., & Porter, D. C. (2009). *Basic Econometrics* (International Edition). *New York: McGraw-Hills Inc*.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the econometric society*, 1251-1271.

Hausman, J. A. (1983). Specification and estimation of simultaneous equation models. *Handbook of econometrics*, 1, 391-448.

Henningsen, A., & Hamann, J. D. (2008). Systemfit: A package for estimating systems of simultaneous equations in R. *Journal of statistical software*, 23, 1-40.

Kimura, H., & Todo, Y. (2010). Is foreign aid a vanguard of foreign direct investment? A gravity-equation approach. *World Development*, 38(4), 482-497.

Koutsoyiannis, A. (2003): Theory of Econometrics; An Introductory Exposition of Econometric Methods. 2nd Edition; Palgrave New York.Madansky, A. (1976): Foundation of Econometrics. North

Kwiatkowski, D., Phillips, P. C., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1-3), 159-178.

Ndambendia, H., & Njoupouognigni, M. (2010). Foreign aid, foreign direct investment and economic growth in Sub-Saharan Africa: Evidence from pooled mean group estimator (PMG). *International journal of economics and finance*, 2(3), 39-45.

Okafor, I. G., Ugochukwu, U. S., & Chijindu, E. H. (2016). Foreign capital inflows and Nigerian economic growth nexus: A Toda Yamamoto approach. *European Journal of Accounting, Auditing and Finance Research*, 4(3), 16-26.

Ozekhome, H. O. (2017). Foreign aid, Foreign direct investment and economic growth in ECOWAS countries: are there diminishing returns in the aid-growth nexus. *West African Journal of Monetary and Economic Integration*, 17, 61-84.

Selaya, P., & Sunesen, E. R. (2012). Does foreign aid increase foreign direct investment? *World Development*, 40(11), 2155-2176.

Statistics on GDP rankings in Africa. *Retrieved from Visafrican.com, 17 May 2021*

Tiwari, A. K. (2011). Foreign aid, FDI, economic freedom and economic growth in Asian countries. *Global Economy Journal*, 11(3), 1850231.

Trevino, L. J., & Upadhyaya, K. P. (2003). Research Note on Foreign aid, FDI and Economic Growth: evidence from Asian countries. *Board of Advisers CHAIRPERSON*, 12(2), 119.

World Governance Index (2019). *Data on macroeconomic variables.* Retrieved from http://data.worldbank.org

Yaffee, R. A., & McGee, M. (2000). *An introduction to time series analysis and forecasting: with applications of SAS and SPSS*. Elsevier.

Yahya, W. B., Adebayo, S. B., Jolayemi, E. T., Oyejola, B. A., & Sanni, O. O. M. (2008). Effects of non-orthogonality on the efficiency of seemingly unrelated regression (SUR) models. *InterStat Journal*, (1), 1-29.

Zellner, A., & Theil, H. (1962). Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations." *Econometrica*, 30**(**1), 54{78.

Zellner, A., & Theil, H. (1992). Three-stage least squares: simultaneous estimation of simultaneous equations. In *Henri Theil's Contributions to Economics and Econometrics: Econometric Theory and Methodology* (pp. 147-178). Dordrecht: Springer Netherlands.