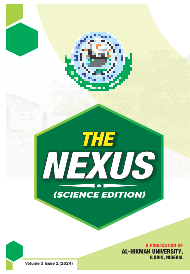
P-ISSN: 2811-2512
E-ISSN: 2811-2520

THE NEXUS

A Publication of Al-Hikmah University, Ilorin, Nigeria
(Science Edition)

Available online at https://www.alhikmahuniversity.edu.ng



TAX-ECONOMY RELATIONSHIP: TRADITIONAL TIME SERIES OR MULTIPLE LINEAR REGRESSION MODELS?

[1]Garba, M. K., [1]Akanni, S. B., [2]Kolawole, K. D., [3]Ojewale, T. T. and [1]Afolayan, R. B.

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

ARTICLE INFO

Article history:
Received: March 11, 2023
Revised: April 3, 2023
Accepted: April 6, 2023
Published online: May 30, 2024

Citation:
Garba, M. K., Akanni, S. B., Kolawole, K. D., Ojewale, T. T. and Afolayan, R. B. (2024). Tax-economy relationship: Traditional Time series or Multiple Linear Regression Models? The Nexus (Science Edition). 2(1): 51- 58.

***Corresponding Author:**
Garba, M. K. .
Department of Statistics,
University of Ilorin, Ilorin, Nigeria.
*e-mail: garba.mk@unilorin.edu.ng

ABSTRACT

In practice, virtually all form of Gross Domestic Products are usually difference stationary series of order one {I(1)} except Real GDP which is often difference stationary series of order two {I(2)}.Since no Traditional Time Series Models (TTSM) cannot adequately capture the dynamics of an I(2) variable in a multivariable time series settings, the system equation techniques(single or system estimators) such as the Indirect Least Squares (ILS), Two-Stage Least Squares (2SLS), Three-Stage Least Squares (3SLS)Seemingly Unrelated Regression (SUR) and Full Information Maximum Likelihood (FIML) estimators are desirable. However, comparison among these estimators using suitable selection criteria in order to determine the best estimator(s) for estimating the regression coefficients in the formulated model is crucial. To demonstrate this assertion empirically, this study therefore apply a Simultaneous Equation Model (SEM) to examine the Tax-Real GDP relationship under mixed order of integrations such as I(2) and I(1)s. Results from unit root tests established that Real GDP (Ly_{1t})is I(2)while Company Income Tax (Ly_{2t}), Petroleum Profit Tax (Ly_{3t}), Personal Income Tax (Lx_{1t}) and Value Added Tax (Lx_{2t}) are all I(1). Endogeneity tests carried out on the series further confirmed that there is no two-way causation among the three endogenous variables (i.e. Ly_{1t}, Ly_{2t} and Ly_{3t}) in the model. The findings show that despite the absence of simultaneity in the model, the Ordinary Least Squares (OLS) and Two-Stage Least Squares (2SLS) estimators though produced identical estimates which are spurious. The Three Stage Least Squares (3SLS) outperformed the Seemingly Unrelated Regression (SUR) reported the least values of the standard errors of the regression parameters. Projections using the 3SLS estimates further revealed that for every one percent increase in Lx_{1t} , Lx_{2t} and $Lx_{(t-1)}$), Ly_{1t} is expected to increase by 14.4%, 11.6% and 6.93% respectively.

Keywords: *Tax, Real GDP, Multiple Linear Regression, Traditional Time Series Models, Nigeria.*

INTRODUCTION

Theoretically, tax generates the state budget revenue which serves as a tool for regulating the economy (Nguyen, 2019). Therefore, the need to formulate better policies from tax-economy relationship requires that econometricians have a good knowledge of models selection processes for macroeconomic variables such as Gross Domestic Product (GDP) and tax variables. Moreover, the choice of model(s) for investigating the nexus between tax and economic growth relies mostly on the choice of GDP and order of integrations of the tax variables involved in a study. Different form of GDPs such as GDP Per Capita, GDP current, GDP, Real GDP (RGDP), etc. can be used to proxy for economic growth depending on the objectives of the study. In most cases, data sets on GDP are differenced stationary series of order one {I(1)} except that of RGDP which is often of order two {I(2)}.Gross Domestic Product measures the monetary value of final goods and services produced in a country in a given period of time which can either be quarterly or yearly (International Monetary Fund, 2022). Like other macroeconomic variables, tax variables which include Company Income Tax (CIT), Petroleum Profit Tax (PPT), Personal Income Tax (PIT) and Value Added Tax (VAT) are expected to have some order of integrations (ds) such that each d can assume any of the possible values 0, 1 and 2. Often times, the choice of unit root tests options such as whether to conduct the tests with drift only, with drift and trend, or none (i.e. no drift and no trend) is not a problem but consistency throughout the series at hand is required. To prevent these problems, researchers need to check the significance of drift, drift plus trend or none when subjecting time series to unit root analyses before proceeding to further analyses. Also, lack of understanding of the gimmicks for combining time series variables with order of integrations such as I(0) only, I(1) only or I(0) and I(1) can as well lead to model misspecification or wrong choice of model.

The application of Traditional Time Series Models (TTSM) to time series variables in a study requires that the series are all completely level stationary series {I(0)}, all completely I(1) or mixtures of I(0) and I(1) (see Sims, 1980; Johansen & Juselius, 1990; Toda & Yamamoto, 1995; Pesaran *et al.,* 2001; Gujarati & Porter, 2009).

Review of Literature

A number of researchers have investigated the tax-economic growth relationship under mixed order of integrations. For instance, Afolayan and Okoli (2015) used the Error Correction Model (ECM) and Granger causality tests to examine the impact of Value Added Tax (VAT) on Nigerian economic growth. The research findings from ECM revealed that there exists no long run relationship between real GDP and any of VAT, CIT, Custom and Exercise Duties (CED) and PPT. Further findings from the study using causality tests showed that though a positive and insignificant correlation exists between Value Added Tax (VAT) revenue and real GDP but there exists a unidirectional relationship which runs from real GDP to VAT revenue. The study concluded that :(i) VAT and CED have positive correlation with real GDP (ii) of all the variables, only PPT and CIT have significant impact on real GDP.

Bernard *et al.* (2018) applied standard Vector Autoregressive (VAR) model to examine the effects of British Columbia's carbon tax on GDP. The research concluded that carbon taxes do not significantly affect GDP change in the province of British Columbia. Moreover, theory of Arthur Laffer techniques adopted by Dracea *et al.* (2009) indicated the existence of a correlation between the real GDP and the real tax incomes in Turkey (progressive tax system) as compared to Romania (flat tax system).Akinkunmi(2016)used Vector Error Correction Model (VECM) to show that there is a significant relationship between real GDP and real company income tax revenues in the long run. The research findings also show that the one-year lag of tax revenue varieties poses a significant influence on the various sources of tax revenues.

Lawler and Al-Sayegh (2019) employed the General Least Square (GLS) method of estimation to show that the impact of changes in tax revenues on changes in the GDP of Kuwait is insignificant. Mannan *et al.* (2022) applied Panel Autoregressive Distributed-Lag (ARDL) model and Wald tests to examine the impacts of economic and financial factors on tax revenue for South Asia. The study established that financial variables impact tax revenue, GDP per capita has a long-run positive relationship with tax revenue and per capita income relates negatively in the short run. Cebula and Foley (2012) examined the effects of economic freedom, regulatory quality and taxation on the growth rate of per capita real GDP for some Organization for Economic and Development (OECD) nations using the Partial Least Squares (PLS) estimation techniques. The research findings revealed that higher levels of economic freedom in an economy promote a higher growth rate of economic activity and hence yield a higher growth rate of per capita real GDP in that economy. The study also discovered that higher quality government regulation leads to a more efficient economic system, in large part by interfering less with market functioning and in part by not adding unnecessarily to the cost of conducting business in the marketplace, and thereby leads to a higher per capita real GDP growth rate. Furthermore, the study showed that the higher the taxation level/burden relative to GDP in an economy, the lower the growth rate of private sector spending and hence the lower the growth rate of per capita real GDP in that economy.

Meanwhile, Branson and Lovell (2001) utilized a linear programming model to estimate a combination of the tax burden and the tax mix which could maximize the rate of growth of real GDP. The study found that moving to such a tax structure would generate nearly a 17% increase in real GDP while this increase would yield a 6% reduction in tax revenue to the Treasury, delivering a 27% increase in purchasing power to the remainder of the economy. De Mooij and Nicodème (2008) investigated corporate tax policy and incorporation in the European Union (EU) using a panel of European data on legal form of business to analyze income shifting via incorporation. The study discovered that the revenue effects of lower corporate tax rates possibly induced by tax competition will partly show up in lower personal tax revenues rather than lower corporate tax revenues. Through Romer's (1986) one-sector representative agent model of endogenous growth analysis conducted by Chen (2020), it was found that US income inequality significantly deteriorates since the mid-1980s. The research findings also revealed that the inequality-growth nexus displays a positive slope before and after the implementation of the Tax Reform Act of 1986 (TRA-86). The study concluded that the slope of the inequality-growth nexus sharply declines between the pre- and post-TRA-86-reform periods, indicating less deterioration in real GDP per capita growth when pursuing a more equal income distribution after 1986.Findings from Ordinary Least Squares (OLS) techniques in the study of Nguyen (2019)revealed that tax has a positive impact on Vietnam's economic growth. The study reported that the indirect tax has a positive influence and promote Vietnam's economic growth, while the impact of the direct tax is invisible.

Following a Vector Error Correction Model (VECM) that was applied to some quarterly time series data during the period of 2002 Q1 to 2017 Q3,Taghizadeh–Hesary *et al.* (2020) though found that increase in money stock (m1) through Quantitative Easing (QE) and Quantitative and Qualitative Easing (QQE) policies of the Bank of Japan (BOJ) significantly increases the income inequality but the Japanese tax policy was effective in reducing the income inequality. Olusegun (2021) applied Error Correction Model (ECM) and Pair wise Granger Causality(PGC) techniques to study the determinants and sustainability of manufacturing sector performance in Nigeria. The study reported that increase in consumption tax, real exchange rate and liberation of the economy were the determinants of manufacturing sector performance while appreciation of Nigeria's currency and increase in tax rate with proportional improvement in infrastructural facilities are needed to sustain it. Basirat *et al.,* (2014) employed the Autoregressive Distributed Lag (ARDL) model to show that exchange rate, import, and the value-added of industry sector had a positive significant relationship with total tax revenues.

Turyanskyy *et al.* (2020) through the use of systematic and institutional methods found that the total tax burden on business in Ukraine reaches 41.5% of corporate profits, which exceeds similar indicators in most European countries. In a study, Habeeb (2022) utilized simple linear regression model approach to show that tax revenues significantly contribute to the overall growth expansion in the Iraqi economy. Sethi *et al.* (2020) explored the Cointegration and Rolling-Window Causality techniques to establish that a stable long-run parameter stability relationship exists between the series on tax reform, but growth-led taxation effects and tax-led-growth do not exist in India.

In a study conducted by Bernard and Kichian (2021) examined the impact of a revenue-neutral carbon tax on GDP dynamics in British Columbia using a time series approach. The study concluded that implementing revenue-neutral carbon taxation contributes to lowering harmful greenhouse gases into the atmosphere without hurting the economy. Ihalanayake (2012) studied the economic effects of tourism tax changes in Australia using computable general equilibrium model. The research findings established that tax changes are contractionary, as reflected in real consumption and GDP. Blanchard and Perotti (2002) investigated the impact of shocks to total tax revenues using an Structural Vector Autoregressive (SVAR) estimator. The study observed that an impact multiplier of 0.69 and a peak multiplier of 0.78 in quarterly data on US for the sampled period from 1947 to 1997. Mertens and Ravn (2013) developed an estimator in the Structural Vector Autoregressive (SVAR) framework to show that short run output effects of tax shocks are large and that it is important to distinguish between different types of taxes when considering their impact on the labor market and on expenditure components.

From this empirical review of related studies, it is observed that some of studies erroneously analyzed the tax-real GDP relationship under mixed order of integrations such as I(1)s and I(2)s using either the Traditional Time Series Models (TTSM) or regression techniques.To capture the tax-real GDP dynamics under such mixed order of integrations, this study employed the techniques of Simultaneous Equation Model (SEM) to model multivariate time series data on Nigerian real GDP and some forms of taxes. This approach as reported by Gujarati and Porter (2009) is not perturbed by orders of integration.

3. Model comparative analysis

For easy representation, the Real Gross Domestic Product (RGDP), Company Income Tax (CIT), Petroleum Profit Tax (PPT), Personal Income Tax (PIT) and Value Added Tax (VAT) are denoted with y_1, y_2, y_3, x_1 and x_2 respectively. The time series variables were subjected to unit root analyses to determine the inherent order of integration of each series. Carrying out these tests is germane in order to formulate the necessary models for the data.

3.1 Model Specification

The formulated system of Simultaneous Equation Model (SEM) for the study is stated as equations (1), (2) and (3):

$$y_{1t} = \beta_{10} + \gamma_{11} x_{1t} + \gamma_{12} x_{2t} + \gamma_{13} x_{t-1} + u_{1t} \quad (1)$$

$$y_{2t} = \beta_{20} + \beta_{21} y_{1t} + u_{2t} \quad (2)$$

$$y_{3t} = \beta_{30} + \beta_{31} y_{1t} + u_{3t} \quad (3)$$

Where: y_{1t}, y_{2t} *and* y_{3t} represent current endogenous variables, x_{1t}, x_{2t} and $x_{(t-1)}$ represent the predetermined variables, u_{1t}, u_{2t} and u_{3t} represent disturbance terms which are well-behaved multivariate normal distribution u~NID(0, Σ), $\beta_{10}, \beta_{20}, \beta_{30}, \beta_{21}, \beta_{31}, \gamma_{11}, \gamma_{12}$ and γ_{13} represent the structural parameters of the simultaneous equation model which shall be estimated accordingly.

3.2 Model identification

Here, each equation embedded in the three-equation model was subjected to order and rank conditions of identification. The former, which is a necessary condition of identification identifies if an equation in the system is exactly identified, over identified or under-identified. If an equation is exactly identified, then Indirect Least Squares (ILS) estimator is appropriate for estimating the structural parameters of the model (see, López-Espín *et al.,* 2012).However, if an equation is over identified, then its parameters can be appropriately estimated using the Two-Stage Least Squares (2SLS) estimator (see, López-Espín *et al.,* 2012).The order condition of identification is determined for each equation in the three-equation model using the mathematical relation specified as equation (4):

$$\text{If } K\text{-}M \geqslant G\text{-}1 \quad (4)$$

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

K = Total number of variables in the model.
M = Number of variables in a particular equation.

According to Gujarati and Porter (2009), ILS and 2SLS estimators are usually the best for estimating the structural parameters of both exactly identified and over-identified equations respectively. However, none of the estimators are suitable for estimating the coefficients of an under-identified equation in a system of equation.

On the other hand, the sufficient condition of identification tells us if the system is estimable or not.

The system is estimable if all the embedded equations are estimable. Otherwise, the system is not estimable. To determine the rank condition of identification for each equation in the system, equations 1 to 3 are re-stated as equations 5 to 7 as follows:

$$y_{1t} - \beta_{10} - \gamma_{11} x_{2t} - \gamma_{12} x_{3t} - \gamma_{13} x_{(t-1)} = u_{1t} \quad (5)$$

$$y_{2t} - \beta_{20} - \beta_{21} y_{1t} = u_{2t} \quad (6)$$

$$y_{3t} - \beta_{30} - \beta_{31} y_{1t} = u_{3t} \quad (7)$$

3.2.1 Rank condition for equation (1)

Equation (1) excludes only two variables y_{1t} and y_{3t}. For this equation to be identified, at least one non-zero determinant of order 2-by-2 must be obtained from the coefficients of the variables excluded from equation (1) but included in equations (2) and (3). Otherwise, the equation is not identified by the rank condition.

3.2.2 Rank condition for equation (2)

In equation (2), four variables y_{3t}, x_{1t}, x_{2t} and $x_{(t-1)}$ are excluded from the equation. This equation is identified if and only if at least one non-zero determinant of order 2-by-2 is formed from the coefficients of the variables excluded from equation (2) but included in equations (1) and (3). Otherwise, the equation (2) is not identified by the rank condition.

3.2.3 Rank condition for equation (3)

In the same vein, equation (3) excludes y_{2t}, x_{1t}, x_{2t} and $x_{(t-1)}$. The equation is identified if and only if at least one non-zero determinant of order 2-by-2 is obtained from the coefficients of the variables excluded from equation (3) but included in equations (1) and (2). Otherwise, the equation (3) is not identified by the rank condition.

3.3 The reduced-form equations

Here, the assumed endogenous variables y_{1t}, y_{2t}, and y_{3t} are expressed in terms of the predetermined variables for each equation. For instance, equation (1) contains only predetermined variables on its right-hand side and its reduced-form equation is given by

$$y_{1t} = \pi_{10} + \pi_{11} x_{1t} + \pi_{12} x_{2t} + \pi_{13} x_{(t-1)} + v_{it} \quad (8)$$

Where $\pi_{10} = \beta_{10}$, $\pi_{11} = \gamma_{11}$, $\pi_{12} = \gamma_{12}$, $\pi_{13} = \gamma_{13}$ and $v_{it} = u_{it}$

Since equation (2) contains an endogenous variable y_{1t} on its right-hand side, its reduced-form equation is derived by substituting equation (1) into equation (2) as follows:

$$\left.\begin{aligned} y_{2t} &= \beta_{20} + \beta_{21}(\beta_{10} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + \gamma_{13}x_{t-1} + u_{1t}) + u_{2t} \\ y_{2t} &= (\beta_{20} + \beta_{21}\beta_{10}) + (\beta_{21}\gamma_{11})x_{1t} + (\beta_{21}\gamma_{12})x_{2t} + (\beta_{21}\gamma_{13})x_{t-1} + (\beta_{21}u_{1t} + u_{2t}) \\ y_{2t} &= \pi_{20} + \pi_{21}x_{1t} + \pi_{22}x_{2t} + \pi_{23}x_{t-1} + v_{2t} \end{aligned}\right\} \quad (9)$$

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

Also, equation (3) contains the same endogenous variable y_{1t} on its right-hand side, its obtained is derived by substituting equation (1) into equation (3) as follows:

$$\left.\begin{aligned} y_{3t} &= \beta_{30} + \beta_{31}(\beta_{10} + \gamma_{11}x_{1t} + \gamma_{12}x_{2t} + \gamma_{13}x_{t-1} + u_{1t}) + u_{3t} \\ y_{3t} &= (\beta_{30} + \beta_{31}\beta_{10}) + (\beta_{31}\gamma_{11})x_{1t} + (\beta_{31}\gamma_{12})x_{2t} + (\beta_{31}\gamma_{13})x_{t-1} + (\beta_{31}u_{1t} + u_{3t}) \\ y_{3t} &= \pi_{30} + \pi_{31}x_{1t} + \pi_{32}x_{2t} + \pi_{33}x_{t-1} + v_{3t} \end{aligned}\right\} \quad (10)$$

Where: $\pi_{30} = \beta_{30} + \beta_{31}\beta_{10}$, $\pi_{31} = \beta_{31}\gamma_{11}$, $\pi_{32} = \beta_{31}\gamma_{12}$ and $v_{3t} = \beta_{31}u_{1t} + u_{3t}$

3.4 Hausman specification test for over identified equations (2) and (3)

In this section, the existence of endogeneity problem was investigated in the over identified equations (2) and (3) using the Hausman (1978) specification tests. The three-equation model was formulated in such a way that y_{1t} appeared on the right-hand side of equations (2) and (3) as regress ands after having appeared as a regressor in equation (1). As a result,y_1t is assumed to be endogenous. To check the existence of endogeneity in equations (2) and (3), the following hypotheses were formulated with respect to the two equations

H01:There is no correlation between y_{1t} and u_{2t}{i.e. Corr(y_{1t},u_{2t})≠0}

H02: There is no correlation between y_{1t} and u_{3t}{i.e. Corr(y_{1t}, u_{3t})≠0}

The test statistic is given by

$$t_{\alpha} = \frac{\dot{\alpha}}{SE(\alpha)}$$

Reject the null hypothesis if the calculated value of t of $\hat{v}_t$ and $\hat{v}_{4t}$ in equations (14) and (15)is statistically significant. Otherwise, we do not reject the null hypothesis.

If the null hypothesis of no endogeneity, we employed single-equation estimators. Otherwise, we employ system estimators will be employed. Otherwise, system estimators will be utilized.

There are basically two steps is carrying out Hausman's test. These steps are:

Step 1: The possible correlation between y_{1t}, y_{2t} and y_{3t}will be checked by regressing y_{1t} on x_{1t}, x_{2t} and $x_{(t-1)}$ using the reduced-form equation (8)

$$y_{1t} = (\hat{\pi}_{10} + \hat{\pi}_{11}x_{1t} + \hat{\pi}_{12}x_{2t} + \hat{\pi}_{13}x_{t-1}) + \hat{v}_{1t} \quad (11)$$

Where: $\hat{v}_{1t}$ are the residuals of the Ordinary Least Squares (OLS) estimator from equation (11). The estimated version of equation (11) is specified as equation (12)

$$\hat{y}_{1t} = \hat{\pi}_{10} + \hat{\pi}_{11}x_{1t} + \hat{\pi}_{12}x_{2t} + \hat{\pi}_{13}x_{t-1} \quad (12)$$

Where: $\hat{y}_{1t}$ is an estimate of the mean value of y_{1t} conditional upon the fixed x_{1t}, x_{2t} and $x_{(t-1)}$ predetermined variables. Now substituting equation (12) into equation (11) yields

$$y_{1t} = \hat{y}_{1t} + \hat{v}_{1t} \quad (13)$$

Step 2: Substitution of equation (13) into equation (2) will be done to yield equation (14)

$$\left.\begin{aligned} y_{2t} &= \beta_{20} + \beta_{21}(\hat{y}_{1t} + \hat{v}_{1t}) + u_{2t} \\ y_{2t} &= \beta_{20} + \beta_{21}\hat{y}_{1t} + (\beta_{21}\hat{v}_{1t} + u_{2t}) \\ \boldsymbol{y_{2t}} &= \boldsymbol{\beta_{20} + \beta_{21}\hat{y}_{1t} + \hat{v}_t} \end{aligned}\right\} \quad 14$$

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

Also, equation (13) will be substituted in equation (3) to produce equation (15)

$$\left.\begin{aligned} y_{3t} &= \beta_{30} + \beta_{31}(\hat{y}_{1t} + \hat{v}_{1t}) + u_{3t} \\ y_{3t} &= \beta_{30} + \beta_{31}\hat{y}_{1t} + (\beta_{31}\hat{v}_{1t} + u_{3t}) \\ \boldsymbol{y_{3t}} &= \boldsymbol{\beta_{30} + \beta_{31}\hat{y}_{1t} + \hat{v}_{4t}} \end{aligned}\right\} \quad 15$$

Where: $\hat{v}_{4t} = \beta_{31}\hat{v}_{1t} + u_{3t}$

4. Data Analysis and Results

This section presents the results and interpretations of the variables used in the study.

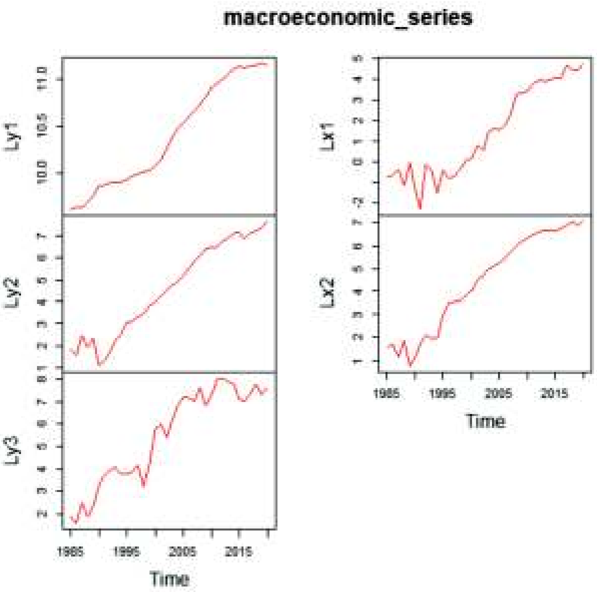


Figure 1: Time plots of Logarithmic Transformations of Real GDP (Ly1), Personal Income Tax (Ly2), Petroleum Profit Tax (Ly3), Company Income Tax (Lx2) Value Added Tax (Lx3) from 1985 to 2020.

Visual inspection of the time series plots in Figure 1 revealed that all the logged series exhibited upward trends with fluctuations which is an indication of non-stationarity in all the series. As a result, these series are expected to be difference stationary series of certain order of integration respectively.

Table 1: Unit root analyses for Ly1, Ly2, Ly3, Lx1 and Lx2

Augmented-Dickey-Fuller (ADF) Test				
Variable	ADF-Statistic	Critical Value	P-value	Order of Integration
Ly1 with drift+intercept	-4.954922	-3.562882	0.0020	I(2)
Ly2 with none	-2.43669	-1.951332	0.0164	I(1)
Ly3 with none	-5.81268	-1.951	<0.0001	I(1)
Lx1 with none	-6.42992	-1.951332	<0.0001	I(1)
Lx2 with drift	-7.341383	-2.951125	<0.0001	I(1)
Phillips-Perron (PP) Test				
Variable	PP-Statistic	Critical Value	P-value	Order of Integration
Ly1 with drift+intercept	-14.99295	-3.552973	<0.0001	I(2)
Ly2 with none	-6.457131	-1.951	<0.0001	I(1)
Ly3 with none	-5.81268	-1.951	<0.0001	I(1)
Lx1 with none	-7.719893	-1.951	<0.0001	I(1)
Lx2 with none	-7.389665	-2.951125	<0.0001	I(1)

Table 1 has the reports of the unit root analyses conducted on the series. The unit root test results confirmed that only the Real GDP (Ly1) is difference stationary series of order two {I(2)} whereas Personal Income Tax (Ly2), Petroleum Profit Tax (Ly3), Company Income Tax (Lx2) and Value Added Tax (Lx3) were confirmed to be difference stationary series of order one {I(1)}.Both ADF and PP tests carried out at 5% level of significance agree that all the series are of order one except real GDP which is of order 2.

Table 2: Summary results of order condition of identification for equations (1), (2) and (3)

Equation	Exog, variables excluded (K-M)	Endog. Variable included less one (G-1)	Remarks
1	2	2	Exactly identified
2	4	2	Over-identified
3	4	2	Over-identified

Summary of results of order condition of identification presented in Table 2 shows that equation (1) is exactly identified while equations (2) and (3) are over identified respectively.

Table 3: Matrix table for rank conditions of identification for equations (1), (2) and (3)

	Coefficients of variables in the equations						
Equation	1	y_{1t}	y_{2t}	y_{3t}	y_{t-1}	x_{2t}	x_{3t}
(1)	$-\beta_{10}$	1	0	0	$-\gamma_{11}$	$-\gamma_{12}$	$-\gamma_{13}$
(2)	$-\beta_{20}$	$-\beta_{21}$	1	0	0	0	0
(3)	$-\beta_{30}$	0	$-\beta_{32}$	1	0	0	0

Rank condition for equation (1)

$$|A| = \begin{vmatrix} 1 & 0 \\ -\beta_{32} & 1 \end{vmatrix} = 1 \neq 0$$

Since one non-zero determinant of order 2-by-2 is formed from the coefficients of the variables excluded from equation (1) but included in equations (2) and (3), equation (1) is identified by the rank condition.

Rank condition for equation (2)

$$|B_1| = \begin{vmatrix} 0 & -\gamma_{11} \\ 1 & 0 \end{vmatrix} = +\gamma_{11} \neq 0, |B_2| = \begin{vmatrix} 0 & -\gamma_{12} \\ 1 & 0 \end{vmatrix} = +\gamma_{12} \neq 0, |B_3| = \begin{vmatrix} 0 & -\gamma_{13} \\ 1 & 0 \end{vmatrix} = +\gamma_{13} \neq 0$$

$$|B_4| = \begin{vmatrix} -\gamma_{11} & -\gamma_{12} \\ 0 & 0 \end{vmatrix} = 0, |B_5|, = \begin{vmatrix} -\gamma_{11} & -\gamma_{13} \\ 0 & 0 \end{vmatrix} = 0, |B_6| = \begin{vmatrix} -\gamma_{12} & -\gamma_{13} \\ 0 & 0 \end{vmatrix} = 0$$

Equation (2) is also identified by the rank condition of identification since at least one non-zero determinant of order 2-by-2 is formed from the coefficients of the variables excluded from equation (2) but included in equations (1) and (3).

Rank condition for equation (3)

$$|C_1| = \begin{vmatrix} 1 & -\gamma_{11} \\ -\beta_{21} & 0 \end{vmatrix} = +\beta_{21}\gamma_{11} \neq 0, |C_2| = \begin{vmatrix} 1 & -\gamma_{12} \\ -\beta_{21} & 0 \end{vmatrix} = +\beta_{21}\gamma_{12} \neq 0,$$

$$|C_3| = \begin{vmatrix} 1 & -\gamma_{13} \\ -\beta_{21} & 0 \end{vmatrix} = +\beta_{21}\gamma_{13}, |C_4| = \begin{vmatrix} -\gamma_{11} & -\gamma_{12} \\ 0 & 0 \end{vmatrix} = 0, |C_5| = \begin{vmatrix} -\gamma_{11} & -\gamma_{13} \\ 0 & 0 \end{vmatrix} = 0$$

$$|C_6| = \begin{vmatrix} -\gamma_{12} & -\gamma_{13} \\ 0 & 0 \end{vmatrix} = 0$$

Lastly, equation (3) is identified by the rank condition of identification since at least one non-zero determinant of order 2-by-2 is obtained from the coefficients of the variables excluded from equation (3) but included in equations (1) and (2).

Table 4: Estimates of reduced-form equations

Variable	Coefficient	SE	t-ratio	P-value	Other Statistic
			Equation (8)		
π_{10}	9.75251	0.0675304	144.4	<0.0001	R^2=0.976742
Lx_{1t}	0.161235	0.0193927	8.314	<0.0001	DW=0.644084
Lx_{2t}	0.0976158	0.0202694	4.816	<0.0001	F-stat =433.954
Lx_{t-1}	−0.101277	0.0238539	−4.246	0.0002	P-valuefor F-stat=2.16e-25
			Equation (9)		
Variable	Coefficient	SE	t-ratio	P-value	Other Statistic
π_{10}	1.42960	0.227561	6.282	<0.0001	R^2= 0.981814
Lx_{1t}	0.413918	0.0653487	6.334	<0.0001	DW= 2.12064
Lx_{2t}	0.576681	0.0683029	8.443	<0.0001	F-stat = 557.85
Lx_{t-1}	−0.0861405	0.0803816	−1.072	0.2922	P-valuefor F-stat 4.79e-27
			Equation (10)		Other Statistic
Variable	Coefficient	SE	t-ratio	P-value	
π_{10}	233.826	390.336	0.5990	0.5535	R^2= 0.769990
Lx_{1t}	376.249	112.093	3.357	0.0021	DW= 1.18700
Lx_{2t}	36.4323	117.160	0.3110	0.7579	F-stat = 34.592
Lx_{t-1}	−192.824	137.879	−1.398	0.1719	P-valuefor F-sta 5.15e-10

Table 4 reports the estimates of all the reduced-form equations (8), (9) and (10). The reduced-form equation (8) were further used to obtain the predicted $\hat{y}_{1t}$; which was later used to conduct the endogeneity tests for equations (2) and (3).

Table 5: Results of endogeneity tests for equations (2) and (3)

Variable	Coefficient	SE	t-Statistic	P-value	Other Statistics
			Eq. (14)		
β_{20}	-563.601	74.7513	-7.5397	<0.0001	R^2 = 0.87545
$\hat{y}_{1t}$	0.025797	0.00172063	14.9928	<0.0001	DW=0.49576
$\hat{v}_t$	0.00473175	0.0121869	0.3883	0.70039	F-stat=112.4668
					P-valuefor F-stat=3.35e-15
			Eq. (15)		Other Statistics
Variable	Coefficient	SE	t-Statistic	P-value	
β_{30}	-763.039	169.537	-4.5007	0.00008	R^2 = 0.8012
$\hat{y}_{1t}$	0.0438972	0.00390242	11.2487	<0.00001	DW=1.03998
$\hat{v}_{4t}$	-0.0430725	0.0276401	-1.5583	0.12899	F-stat=64.481
					P-value for F-stat=5.95e-12

Results reported in Table 5 showed that $\hat{v}_t$ is not statistically significant since its associated p-value = 0.70039is greater than α = 0.05 which means that there is no simultaneity between y_{2t} and y_{1t}. Also, from the same Table 5, $\hat{v}_{4t}$ is also not statistically significant because its p-value = 0.12899 is greater than α = 0.05 which also implies that there is no simultaneity between y_{3t} and y_{1t}.

Table 6: Estimates of Equation (1)

Single-equation Estimators									
OLS					2SLS				
Parameter	Coeff.	SE	t-ratio	P-value	Parameter	Coeff.	SE	Z	P-value
β_{10}	9.75251	0.0675304*	144.417	<0.00001	β_{10}	9.75251	0.0675304*	144.417	<0.00001
γ_{11}	0.161235	0.0193927*	8.3142	<0.00001	γ_{11}	0.16124	0.0193927*	8.3142	<0.00001
γ_{12}	0.0976158	0.0202694*	4.8159	0.00004	γ_{12}	0.09762	0.0202694*	4.8159	<0.00001
γ_{13}	-0.101277	0.0238539*	-4.2457	0.00018	γ_{13}	-0.1013	0.0238539*	-4.2457	0.00002
System-equation Estimators									
3SLS					SUR				
Parameter	Coeff.	SE	Z	P-value	Parameter	Coeff.	SE	t-ratio	P-value
β_{10}	9.69203	0.0479092*	202.3	<0.00001	β_{10}	9.69243	0.048477	199.9	<0.00001
γ_{11}	0.143634	0.0133401*	10.77	<0.00001	γ_{11}	0.14392	0.0135283	10.64	<0.00001
γ_{12}	0.115738	0.0137149*	8.439	<0.00001	γ_{12}	0.11557	0.0139322	8.295	<0.00001
γ_{13}	-0.06929	0.0161097*	-4.301	<0.00001	γ_{13}	-0.0697	0.0163682	-4.26	0.0002

Table 6 shows that OLS and 2SLS produced identical estimates. Though, Real GDP (Ly_{1t}), Personal Income Tax (Ly_{2t}) and Petroleum Profit Tax (Ly_{3t}) are not endogenous in nature as confirmed by the simultaneity test but OLS and 2SLS are not appropriate for estimating the parameters of the equation the Durbin-Watson Statistic (DW = 0.644084) is considerably less than two which indicates a positive serial correlation.

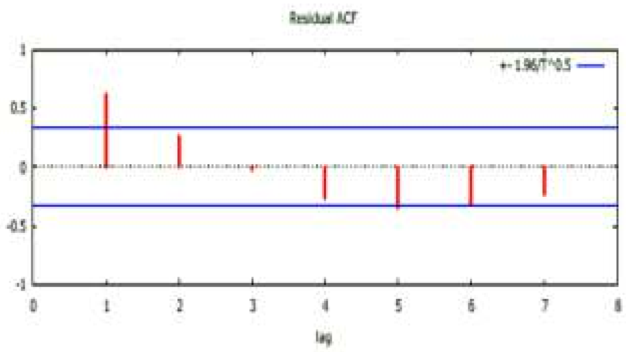
However, only the 3SLS estimator reported the least value of the standard error which means that the 3SLS is the best estimator at estimating the parameters of the exactly identified equation. Besides, the p-values of all the estimated regression coefficients in equation (1) are statistically significant. Consequently, Lx_{1t} and Lx_{2t} have positive impact on Ly_{1t} whereas $Lx_{(t-1)}$ has negative impact on Ly_{1t}.

Estimated structural equation (1) using single-equation estimators:

OLS: $Ly_{1t} = 9.75251+0.161235\ Lx_{1t}+0.0976158Lx_{2t}-0.101277\ Lx_{(t-1)}$ (16)

2SLS: $Ly_{1t}=9.75251+0.161235\ Lx_{1t}+0.0976158Lx_{2t}-0.101277Lx_{(t-1)}$ (17)

$R^2 = 0.976742$, DW= 0.644084 and P-value (F) = 2.16e-25



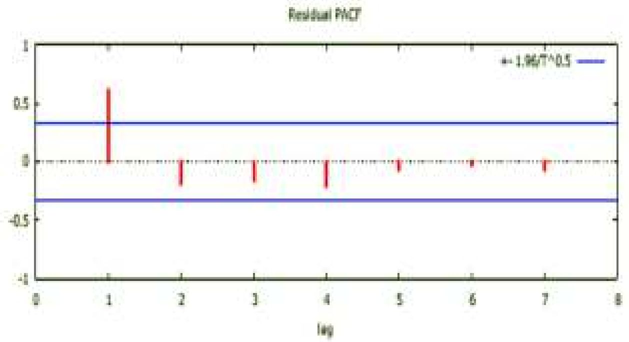


Figure 2: Residual correlogram for the OLS and 2SLS estimators in equation (1)

Figure 2 shows that there is one significant spike each at both the residuals ACF and PACF; which further confirmed the existence of positive serial correlation in the OLS and 2SLS estimators.

Estimated structural equation (1) using system estimators:

3SLS: Ly_{1t}= 9.69203+0.143634 Lx_{1t}+ 0.115738 Lx_{2t}-0.06929$Lx_{(t-1)}$ (18)

SUR: Ly_{1t} = 9.69243+0.14392Lx_{1t} + 0.11557 Lx_{2t} - 0.0697 $Lx_{(t-1)}$ (19)

R^2 = 0.975086

Model interpretation

The regression coefficient R^2 (0.975086) indicates that 97.6% variation in Ly_{1t} was explained by Lx_{1t}, Lx_{2t} and $Lx_{(t-1)}$ respectively.

The estimates in Table 6 tell us that for every one percent increase in Personal Income Tax (Lx_{1t}) there is an associated 14.4% increase in Real GDP (Ly_{1t}), for every one percent increase in Value Added Tax (Lx_{2t}) there is an associated 11.6% increase in Ly_{1t} and that for one percent increase in lagged one value of VAT ($Lx_{(t-1)}$) there is an associated 6.93% decrease in Ly_{1t}.

5. Summary of Findings

This work evaluated the choice of models for investigating the Tax-Real Gross Domestic Product relationship in Nigeria under mixed order of integrations.

Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) test results presented in Table 1 revealed that Personal Income Tax (Ly2), Petroleum Profit Tax (Ly3), Company Income Tax (Lx2) and Value Added Tax (Lx3) became stationary after first differences {I(1)s} whereas Real Gross Domestic Product (Ly1) became stationary after second difference {I(2)}. Consequently, the usual multivariate time series models cannot be employed to investigate the Tax-Real GDP relationship since there are different orders of integrations.

The Hausman specification test carried out produced the results presented in Table 5showed that the null hypotheses of simultaneity were not only rejected for Ly2 and Ly1 but rejected for Ly_{3t} and Ly_{1t} respectively. This means that Ly1, Ly2 and Ly3 are not endogenous in nature.

Results of the exactly identified equation (1) reported in Table 5 showed that Three Stage Least Squares (3SLS) estimator outperformed the Seemingly Unrelated Regression (SUR), OLS and Two Stage Least Squares (2SLS) estimators since it yielded the least value of the standard errors.

For the 3SLS estimates reported in Table 6, the regression coefficient R^2 (= 0.975086) indicated that 97.6% variation in Ly_{1t} was explained by Lx_{1t}, Lx_{2t} and $Lx_{(t-1)}$ respectively. Lastly, the same Table 6 estimates showed that for every one percent increase in Personal Income Tax (Lx_{1t}) there is an associated 14.4% increase in Real GDP (Ly_{1t}). Moreover, for every one percent increase in Value Added Tax (Lx_{2t}) there is an associated 11.6% increase in Ly_1t and that for one percent increase in $Lx_{(t-1)}$ there is an associated 6.93% decrease in Ly_{1t}.

6. Conclusion and Recommendations

This paper provides evidences that Multiple Linear Regression (MLR) and traditional multivariate time series models cannot be used to investigate the Tax-Real Gross Domestic Product (Ly_{1t}) relationship when Ly_{1t}, Company Income Tax (Ly_{2t}), Petroleum Profit Tax (Ly_{3t}), Personal Income Tax (Lx_{1t}) and Value Added Tax (Lx_{2t}) under mixed order of integrations I(2) and I(1)s respectively. Also, comparison with 3SLS was made on the ground that it combines the SUR and 2SLS estimators for its estimation. Thus, the 3SLS estimates the regression coefficients jointly efficiently compare to other estimators. Based on the results of analysis and summary of findings, though the Ly_{1t}, Ly_{2t} and Ly_{3t} are not endogenous in nature but the single-equation estimators such as the Ordinary Least Squares (OLS) and Two Stage Least Squares (2SLS) were found to be inappropriate for investigating the Tax-Real Gross Domestic Product (Ly_{1t}) relationship in the sense that there estimates revealed a positive serial correlation. However, the Three Stage Least Squares (3SLS) is adjudged the efficient estimator for estimating the Nigeria's tax-real GDP dynamics.

Finally, the 3SLS estimates confirmed that for every one percent increase in Personal Income Tax (Lx_{1t}) there is an associated 14.4% increase in Real GDP (Ly_{1t}), for every one percent increase in Value Added Tax (Lx_{2t}) there is an associated 11.6% increase in Ly_{1t} and that for one percent increase in $Lx_{(t-1)}$ there is an associated 6.93% decrease in Ly_{1t}.

REFERENCES

Afolayan, S. M. and Okoli, M. N. (2015). The Impact of Value Added Tax on Nigerian Economic Growth (1994-2012): An Investigation. European Journal of Business and Management, 7(9), 226-235.

Akinkunmi, M. A. (2016). Dynamic Analysis of Structural Shifts of Fiscal Revenue in Nigeria, 1999-2016. International Journal of Economics and Finance, 8(11), 96-110.

Basirat, M., Aboodi, F. and Ahangari, A. (2014). Analyzing the effect of economic variables on total tax revenues in Iran. Asian Economic and Financial Review, 4(6), 755-767.

Bernard, J. T., Kichian, M. and Islam, M. (2018). Effects of

BC's Carbon Tax on GDP. USAEE Research Paper Series, 1-31.
Blanchard, O. and Perotti, R. (2002).An empirical characterization of the dynamic effects of changes in government spending and taxes on output. The Quarterly Journal of Economics, 117(4), 1329-1368.
Branson, J. and Lovell, C. A. (2001).A Growth Maximizing Tax Structure for New Zealand. International Tax and Public Finance, 8(2), 129-146.
Cebula, R. J. and Foley, M. (2012). A Panel Data Study of the Effects of Economic Freedom, Regulatory Quality, and Taxation on the Growth rate of Per Capita Real GDP. Journal of Public Finance and Public Choice, 30(1-3), 103-122.
Chen, S. H. (2020). Inequality-Growth nexus under Progressive Income Taxation. Journal of Macroeconomics, 65, 103234.
De Mooij, R. A. and Nicodème, G. (2008).Corporate Tax Policy and Incorporation in the EU. International Tax and Public Finance, 15(4), 478-498.
Definition and Statistics of GDP. Retrieved from https://www.imf.org/external/pubs/ft/fandd/basics/gdp.htm, 23 October, 2022.
Dracea, R., Cristea, M., Ionascu, C. andIrtes, M. (2009). Are there any Correlations between Fiscality Rate, GDP and Tax Incomes Flux? Case study Romania and Turkey. European Research Studies, Volume XII, Issue (2), 77-98.
Gujarati, D. N. and Porter, D. C. (2009).Basic Econometrics, 5th edition.McGraw-Hill/Irwin, New York.
Habeeb, H. B. (2022). Assessing The Contribution of Tax Revenues to Overall Economic Growth Expansion in The Iraqi Economy During the Course Of The Years 1990-2020. World Economics and Finance Bulletin, 15, 95-103.
Haberly, D. and Wójcik, D. (2015). Tax Havens and the Production of Offshore FDI: an Empirical Analysis. Journal of Economic Geography, 15(1), 75-101.
Ihalanayake, R. (2012). Economic Effects of Tourism Tax Changes in Australia: Empirical Evidence from a Computable General Equilibrium Model. Tourism Economics, 18(1), 181-202.
Johansen, S. and Juselius, K. (1990).Maximum Likelihood Estimation and Inference on Co integration—with applications to the Demand for Money. Oxford Bulletin of Economics and Statistics, 52(2), 169-210.
Lawler, K. A. and Al-Sayegh, F. A. (2019).Potential Tax Reforms and Kuwait Economic Growth.ВісникКиївськогонаціональногоуніверситетуім. ТарасаШевченка. Серія: Економіка, (1 (202)), 42-48.
López-Espín, J. J., Vidal, A. M. and Giménez, D. (2012). Two-stage Least Squares and Indirect Least Squares Algorithms for Simultaneous Equations Models. Journal of Computational and Applied Mathematics, 236(15), 3676-3684.
Mannan, A., Aziz, B. and Nawaz, Z. (2022). Impact of Economic and Financial Factors on Tax Revenue: Evidence from South Asia. Asian Finance Research Journal 3(2).1-18.
Mertens, K. and Ravn, M. O. (2013). The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States. American Economic Review, 103(4), 1212-47.
Nguyen, H. H. (2019). Impact of Direct Tax and Indirect Tax on Economic Growth in Vietnam. The Journal of Asian Finance, Economics and Business, 6(4), 129-137.
Olusegun, F. A. O. (2021). Determinants and Sustainability of Manufacturing Sector Performance in Nigeria: The Roles of Selected Macroeconomic Variables. Applied Finance and Accounting, 7(2), 31-40.
Pesaran, M. H., Shin, Y. and Smith, R. J. (2001).Bounds Testing Approaches to the Analysis of Level Relationships. Journal of Applied Econometrics, 16(3), 289-326.
Sethi, N., Mohanty, S., Sucharita, S. and Loganathan, N. (2020). Tax Reform and Economic Growth Nexus in India: Evidence from the Cointegration and Rolling-Window Causality. The Singapore Economic Review, 65(06), 1699-1725.
Sims, C. A. (1980). Macroeconomics and Reality. Econometrica: Journal of the Econometric Society, 48(1), 1-48
Statistics on taxation in Nigeria. Retrieved from https://www.lexology.com, 19 May, 2022.
Taghizadeh-Hesary, F., Yoshino, N. and Shimizu, S. (2020). The Impact of Monetary and Tax Policy on Income Inequality in Japan. The World Economy, 43(10), 2600-2621.
Toda, H. Y. and Yamamoto, T. (1995).Statistical Inference in Vector Autoregressions with Possibly Integrated Processes. Journal of Econometrics, 66(1-2), 225-250.
Turyanskyy, Y., Svydruk, I., Sydorchuk, O., Mitsenko, N. and Klepanchuk, O. (2020). Management Priorities of Tax Reform in Ukraine: Implementation of International Experience. Independent Journal of Management & Production (Special Edition ISE, S&P), 11, 640-660.