



# Sectoral Contribution to Nigeria's Gross Domestic Product (GDP) Growth Rate: A Study of Multicollinearity in Aggregated Time Series Data

Ikenna V. Ejiba<sup>1\*</sup> and Olugbenga K. Omolade<sup>1</sup>

<sup>1</sup>Department of Agricultural Economics, University of Ibadan, Nigeria.

## Authors' contributions

*This work was carried out in collaboration between both authors. Author IVE designed the study, wrote the protocol and wrote the first draft of the manuscript. Authors IVE and OKO managed the literature searches, analyses of the study. Both authors made contributions to the work. Both authors also read and approved the final manuscript.*

## Article Information

DOI: 10.9734/JSRR/2016/26364

### Editor(s):

(1) Nidal Rashid Sabri, Ex. Dean of College of Economics (2005-2011), Birzeit University, Palestine.

### Reviewers:

(1) Bader Said Hamdan, Al-Azher University, Palestine.

(2) Imam Mukhlis, State University of Malang, Indonesia.

Complete Peer review History: <http://sciencedomain.org/review-history/14808>

Original Research Article

Received 13<sup>th</sup> April 2016  
Accepted 18<sup>th</sup> May 2016  
Published 28<sup>th</sup> May 2016

## ABSTRACT

The study examined aggregated GDP time series data for the period 1991 – 2010 to analyze the contributions of each sector to Nigeria's GDP growth; and observe deficiencies inherent in the use of Ordinary Least Square (OLS) technique in GDP studies. Secondary data was used for the study. The data was sourced from the Central bank of Nigeria (CBN), and the National Bureau of Statistics (NBS). Analysis was carried out using graphs, tables, OLS regression technique, pairwise correlation, and Generalized Ridge Regression Analysis (GRR). Findings show an increasing trend in aggregate GDP for the study period, with an average GDP growth rate of 4.98%. Severe multicollinearity was detected in the GDP time series data using OLS regression with poor relationships, and high  $R^2$  value but few significant variables. The pairwise correlation coefficients also showed the presence of moderate to severe multicollinearity in the aggregate GDP variables. As a remedial measure, Generalized Ridge Regression analysis was employed. The GRR result showed a more well behaved model compared to the OLS implying a more precise estimation of the regression coefficients for the GRR model, which show a better influence of each sector's contribution to the GDP growth rate within the twenty-year period. Therefore, it would be more

\*Corresponding author: E-mail: [victorejiba@yahoo.com](mailto:victorejiba@yahoo.com);

appropriate for researchers in GDP time series studies to account for the problem of multicollinearity by utilizing Ridge regression technique for estimation since all Ridge regression models give more robust estimates than the OLS models.

*Keywords: Time series data; GDP; OLS technique; multicollinearity; GRR model.*

## 1. INTRODUCTION

The various sectors of any economy have a contribution to the development of any economy. This is to say that no matter how small the contribution of any sector to the national income of that economy, it adds up to the aggregate income of the economy and thus contributing directly or indirectly to the gross domestic earnings of such economy [1]. Gross Domestic Product (GDP) is referred to as the value of final goods and services produced in the economy during a given period [2]. Since independence in 1960, the Nigerian economy remains weak, narrow and externally-oriented with primary production activities of agriculture and mining and quarrying (including crude oil and gas) accounting for about 65% of the GDP and over 80% of government revenues. In addition, the primary production activities account for over 90% of foreign exchange earnings and 75% of employment [3]. In contrast, secondary activities comprising manufacturing and building and construction, which traditionally have greater potential for employment generation, broadening the productive base of the economy and generating sustainable foreign exchange earnings and government revenues account for a mere 4.14 and 2.0% of gross output, respectively.

Over the last seven years, certain changes have taken place in the structure of output in the economy; prominent among these changes is the entry of the telecommunications sector which has witnessed explosive and sustained real GDP growth. In spite of this, the Nigerian economy continues to grapple with a number of challenges that has hampered efforts at economic transformation. The economy is yet to achieve the necessary structural changes required to jump-start rapid and sustainable growth and development, aside disarticulated and narrow productive base, while the sectoral linkages in the economy are also weak [4]. Although the economy experienced respectable GDP growth rates, averaging over 6.5% per annum between 2006 and 2012 [5] this growth neither brings commensurate employment nor reduce the poverty level experienced in the country.

For effective economic planning and economic growth however, government require GDP time series studies as they are important and give policy makers ample information on the performance of government. Time series data and observations including GDP, stock prices, money supply, consumer price index, annual homicide rate and automobile sales figures are given over time; according to [6]. Unlike the arrangement in cross sectional data, the chronological ordering of observations in a time series conveys potentially important information. A key feature of time series data that makes it more difficult to analyze than cross sectional data is the fact that economic observations can rarely, if ever, be assumed to be independent across time. Therefore, forecasting economic relationships may be plagued by estimation problems including the problem of multicollinearity. Therefore, an empirical investigation which attempts to measure the validity of the theory against the observable data [7] may be required.

### 1.1 Conceptual Review

Shantha [8] defined multicollinearity as the existence of a linear relationship among of the independent variables. This occurs when too many variables have been included in the model and a number of different variables measure similar phenomena. It is defined in terms of a lack of independence, or of the presence of interdependence – signified by high inter-correlations within a set of variables, and under this view can exist quite apart from the nature, or even the existence of a dependency relationship between X and a dependent variable Y [9]. Its severity causes difficulty in making prediction inferences and estimations as well as selecting an appropriate set of variables for the model. In such cases, any inferences based on the parameter estimations of the model become invalid [10]. [11] noted that multicollinearity may be due to the data collection method employed, constraints on the model or in the population being sampled, model specification such as adding polynomial terms to the regression model, and an over determined model which is defined as a model with more explanatory variables than

the number of observations. It must thus suffice to emphasize that multicollinearity is both a symptom and a facet of poor experimental design [12]. It is a statistical issue which constitutes a threat and often a very serious threat both to the proper specification and to the effective estimation of the type of structural relationships commonly sought through the use of regression techniques because it inflates the variance of ordinary least squares estimator and a numerical issue in the sense that the small errors in input may cause large errors in the output. The presence of serious multicollinearity would reduce the accuracy of the parameter estimate in a linear regression models. Many authors even suggest that an examination for the existence of multicollinearity should be routinely performed as an initial step in regression analysis.

As rightly opined by [10], multicollinearity causes major interpretative problems in regression analysis, such as wrong sign problem, produces unstable and inconsistent estimates of parameter, insignificant regression coefficients where in fact it is significant and it is thus very essential to investigate and detect the presence to reduce the destructive effects of multicollinearity.

## **1.2 Literature Review**

A number of studies on GDP have been carried out in Nigeria. [13] studied the role of Stock Market development on economic growth in Nigeria using a 15-year time series data from (1994-2008) using Ordinary Least Square technique. The result revealed a negative, but non-significant relationship between GDP and Stock market capitalization; while there was a positive relationship between GDP and liquidity. To examine the severity of multicollinearity, collinearity diagnostic was performed for the independent variables which gave Tolerance values of 0.033, 0.34 and 0.495 respectively suggesting the presence of multicollinearity among the variables. Finally, the authors concluded that “we should view with caution the notion that stock market size is not significant for economic growth since multicollinearity exists in the data used for this analysis”.

Ahungwa et al. [14] examined Trend Analysis of the contribution of Agriculture to the GDP of Nigeria between the periods (1960-2012). The study utilized multiple regression analysis specified as Double- Log model to examine the

contribution of Agriculture to Gross Domestic Product (GDP). The study showed a statistical significant relationship between GDP and both agriculture and industrial sector; with  $R^2$  value of 99%.

Oziengbe [15] studied the relative impacts of Federal Capital and recurrent expenditures on Nigeria's Economy (1980-2011) using multiple regression, co-integration and Error Correction Model (ECM). The OLS results show that only the signs on the coefficients of government expenditure (GOVEXP) and remittance (REMIT) conform to a priori expectations. The signs on the other variables, including foreign direct investment (FDI), and official development assistance and foreign aid (ODAAID) do not conform to a priori expectations. The study concluded that within the period under review, government expenditure and personal remittances from Nigerians abroad contributed significantly to the size of the nation's economy, while the effects of FDI and ODAAID on the nation's economy, showed negative signs and were not significant. The study also drew attention to the fact that “we must be cautious not to insinuate that foreign aid, ODA and foreign direct investment pose dangers to the economy as they could be vital ingredients for economic growth and development if properly harnessed”. This assertion therefore brings to light the question of the appropriateness of the techniques utilized in these studies; as according to [8] using classical multiple regression analysis may result to various symptoms of multicollinearity including overstated regression coefficients, incorrect signs, and highly unstable predictive equations. It is in the light of these that this study aims to draw attention about the dangers apparent in the use of OLS technique in examining GDP time series data; bearing in mind the possibility of the existence of severe multicollinearity which could render inference less reliable - in spite of the contention by [16] that multicollinearity is God's will and not a problem with OLS or statistical technique.

## **2. METHODOLOGY**

### **2.1 Study Area**

The study was conducted in Nigeria, a country in West Africa with population of over 160 million people and sharing land borders with the Republic of Benin in the West; Chad and Cameroun in the East; and Niger in the North. Its coast lies on the gulf of Guinea in the south, and

it borders Lake Chad to the Northeast. Nigeria has an area of 923,768 square kilometers and lies between latitude 4° and 14°N and longitudes 3° and 14°E of the Greenwich meridian, which is entirely within the tropical zone.

## 2.2 Data

The study made use of secondary data obtained from the published documentation of National Bureau of Statistics (NBS) and Central Bank of Nigeria (CBN) bulletin on the Gross Domestic Product (GDP) of Nigeria over a 20 year interval (1991 – 2010). Sectoral contribution of the identified sectors and sub – sectors of the nation’s GDP (at 1984 Constant Factor Cost in ₦ Million) were considered for data collection. The fifteen (15) identified sectors includes: agriculture, crude petroleum, mining and quarrying, manufacturing, utilities, building and construction, transportation, communication, wholesale and retail trade, hotel and restaurants, finance and insurance, real estate and business services, housing, community social and personal services. From the total GDP and from sub – sectors, contribution of each sector to the aggregate national GDP was obtained for the periods under the study. From the aggregate GDP obtained from each year and according to [5] a computation of the GDP growth rate was done as given by the expression;

$$\frac{Y_t - Y_{t-1}}{Y_{t-1}} \times 100 \quad (1)$$

Where:  $Y_t$  = aggregate GDP in year t;  $Y_{t-1}$  = aggregate GDP in year t – 1 (previous year)

## 2.3 Analytical Technique

The analytical techniques employed include Ordinary Least Square (OLS) regression, and Generalized Ridge Regression Analysis (GRR).

### 2.3.1 Contribution of sectors to aggregate GDP

The model employed was derived using Ordinary Least Square (OLS) to fit the implicit function  $Y_t = f(X_{1t}, X_{2t}, \dots, X_{15t}, u_t)$  into the following different functional forms.

- i. Linear functional form:  $Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_{15} X_{15t} + u_t$
- ii. Semi-log functional form:  $\ln Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_{15} X_{15t} + u_t$
- iii. Double-Log functional form:  $\ln Y_t = \beta_0 + \beta_1 \ln X_{1t} + \beta_2 \ln X_{2t} + \dots + \beta_{15} \ln X_{15t} + u_t$

Where

- $\beta_0$  = constant term
- $\beta_1 \dots \beta_{15}$  = Coefficients of sectoral contribution to GDP
- $Y_t$  = Aggregate National GDP Growth Rate (%)
- $X_1$  = Agriculture Sector’s Contribution (₦ Million)
- $X_2$  = Crude Petroleum Sector’s Contribution (₦ Million)
- $X_3$  = Mining and Quarrying Sector’s Contribution (₦ Million)
- $X_4$  = Manufacturing Sector’s Contribution (₦ Million)
- $X_5$  = Utilities Sector’s Contribution (₦ Million)
- $X_6$  = Building and Construction Sector’s Contribution (₦ Million)
- $X_7$  = Transport Sector’s Contribution (₦ Million)
- $X_8$  = Communication Sector’s Contribution (₦ Million)
- $X_9$  = Wholesale and Retail Sector’s Contribution (₦ Million)
- $X_{10}$  = Hotel and Restaurants Sector’s Contribution (₦ Million)
- $X_{11}$  = Finance and Insurance Sector’s Contribution (₦ Million)
- $X_{12}$  = Real Estate and Business Service Sector’s Contribution (₦ Million)
- $X_{13}$  = Housing Sector’s Contribution (₦ Million)
- $X_{14}$  = Community, Social and Personal Services (₦ Million)
- $X_{15}$  = Producer of Government Services Sector’s Contribution (₦ Million)

### 2.3.2 Detecting multicollinearity in the GDP time series data

In detecting multicollinearity, OLS regression analysis was conducted and the  $R^2$  value and the number of significant variables were observed. In addition, pairwise correlation of the selected sectors of the economy, Tolerance, Variance Inflating Factor (VIF), and Condition Index (CI) was also used in detection of multicollinearity.

### 2.3.3 Remedial measures for multicollinearity

The remedial measures adopted in the study include:

- (1) **Acquisition of additional data (analysed by OLS) to the initial sample:** Addition of time series data of 10 years (1981 – 1990) were added to the initial dataset of (1991 – 2010) to give a total of 30 years.

- (2) Transformation of variables.
- (3) The data is transformed by taking the second difference of the series (differences of successive values of the variables) to account for non-stationarity of the time series data. The original (linear) OLS regression equation is given as:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \dots + \beta_{15} X_{15t} + u_t \quad (2)$$

The first differenced form is therefore given as:

$$Y_t - Y_{t-1} = \beta_1 (X_{1t} - X_{1t-1}) + \beta_2 (X_{2t} - X_{2t-1}) + \dots + \beta_{15} (X_{15t} - X_{15t-1}) + v_t \quad (3)$$

Where  $v_t = u_t - u_{t-1}$

The second differenced form is:

$$Y_t - Y_{t-1} - Y_{t-1} - Y_{t-2} = \beta_1 (X_{1t} - X_{1t-1} - X_{1t-1} - X_{1t-2}) + \beta_2 (X_{2t} - X_{2t-1} - X_{2t-1} - X_{2t-2}) + \dots + \beta_{15} (X_{15t} - X_{15t-1} - X_{15t-1} - X_{15t-2}) + w_t \quad (4)$$

Where  $w_t = u_t - u_{t-1} - u_{t-1} - u_{t-2}$ .

- (4) Other remedial measures used in the study include dropping of one or more of the collinear variables from the model, and re-specification of the model; that is re-specifying the linear functional form of the OLS model stated above into a Log-Linear model given as:

$$\ln Y_t = \beta_0 + \beta_1 \ln X_{1t} + \beta_2 \ln X_{2t} + \dots + \beta_{15} \ln X_{15t} + u_t \quad (5)$$

- (5) **Undertaking General Ridge Regression Analysis (GRR):** A comparison is made between regression estimates obtained

from the GRR analysis and the OLS estimates of the initial sample, in terms of the predictive accuracy of model.

### 3. RESULTS AND DISCUSSION

Table 1 gives the average contribution of each sector to the aggregate national GDP growth rate, as well as trend in aggregate GDP, and its growth rate over the period (1991 – 2010) under review.

#### 3.1 Sectoral Contribution to Aggregate GDP

From Table, 1, the agricultural sector on the average over the years give a maximum contribution of ₦ 50,458.3740 million to aggregate GDP, followed by wholesale and retail sector (39,128.5460 ₦ Million), finance and insurance sector (₦14284.4360 million), and crude petroleum sector (₦ 13,321.8150 million). Mining and quarrying sector has the minimum contribution to aggregate GDP among all the economic sectors; while GDP over the years grew at an average rate of 4.977%.

Figs. 1 and 2 show the trend in GDP growth rate and aggregate GDP over the period under review (1991 – 2010). The least GDP growth rate (2.899%) is observed in 1993 and the highest in year 2008 (9.547%). Fig. 2 shows an increasing trend in aggregate GDP yearly from the beginning of the 2000s. Year 1991 (N 113,508.70 million) has the least aggregate GDP while the highest aggregate GDP (N282, 049.71 million) was observed in year 2010.

**Table 1. Average growth rate and sectoral contribution in million Naira (1991 – 2010)**

Variable	Mean	Std. deviation
GDP growth rate	4.9768	5.23120
Agriculture	50458.3740	12558.66390
Crude petroleum	13321.8150	791.78954
Mining and quarrying	365.2500	71.09684
Manufacturing	11266.1650	16209.94585
Utilities	653.2515	94.22508
Building and construction	3285.9395	1675.88862
Transport	3520.0071	471.76561
Communication	428.5635	155.02794
Wholesale and retail	39128.5460	43177.44097
Hotel and restaurant	654.3225	212.94817
Finance and insurance	14284.4360	6676.99815
Real estate and business services	387.0830	119.02912
Housing	3219.6495	920.71447
Producers of government services	11585.6100	2712.76280
Community, social and personal services	2401.1095	1635.13185

*Sectoral contribution to GDP (1991-2010)*

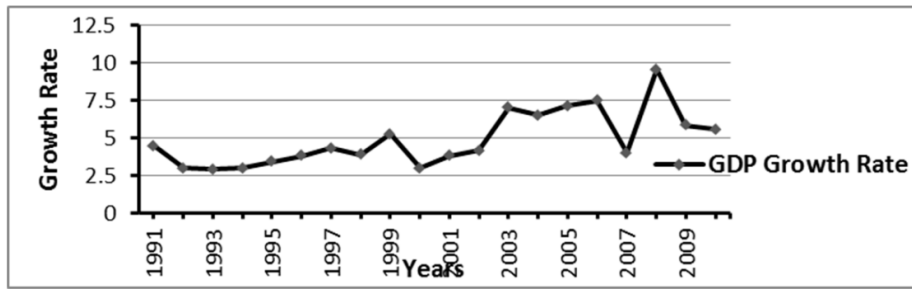


Fig. 1. Trend in GDP growth rate (1991-2010)

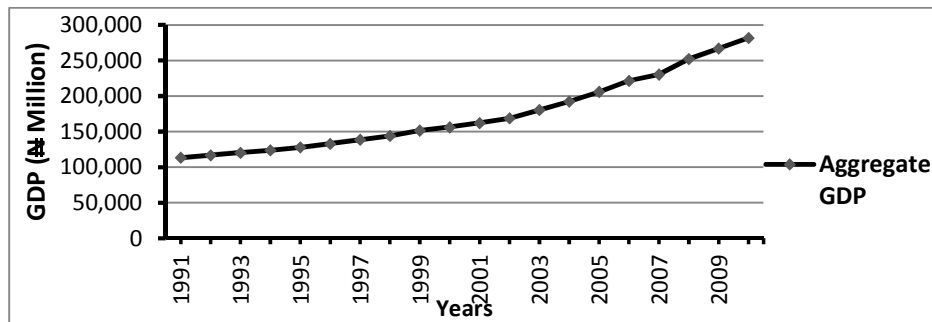


Fig. 2. Aggregate GDP trend (1991-2010)

### 3.2 Detecting Multicollinearity

#### 3.2.1 OLS regression analysis

Subjecting the variables to an OLS regression however gives the following outcome. From the result, only few t-ratios are statistically significant with high  $R^2$  ("classic" symptom of multicollinearity) and other values indicate the presence of strong multicollinearity. In accessing the individual sectoral contribution to GDP growth rate, the detection of multicollinearity, type and severity, Ordinary Least Square (OLS) model in linear functional form is employed. From the result, agriculture, utilities, finance and insurance sector show negative relationship with GDP growth rate (-1.532E-5, -.001, and -3.525E-5 respectively), while other sectors show positive relationship with growth rate of GDP. Manufacturing (.007) is significant at 1%, wholesale and retail sector is significant (.027), utilities (.078), and producers of government services (.089) are significant at 5%.  $R^2$  (.995) and its adjusted value (.974) show that the model is well behaved but few of the variables are significant indicating the presence of multicollinearity. This result obtained here shows a similar trait to those obtained in the previous GDP time series studies given above in the review of literature. From Table 2, the different

ways of detecting multicollinearity and its severity as reviewed in the literature indicated that building and construction, hotel and restaurants, housing and dwelling and community social and personal services have zero tolerance (TOL) values and hence exhibit greatest collinearity among variables of model. Only crude petroleum sector has VIF value (6.694) with below 10 indicating moderate collinearity, while other 14 sectors are highly collinear variables with building and construction (9225.752) being the highest.

#### 3.2.2 Pairwise correlation

The pairwise correlation estimate or correlation matrix shows the correlation among various sectors of the economy. High pairwise correlation with respect to the sectors contribution to GDP growth rate is one of the ways to determining the presence of multicollinearity. Table 3 shows few high pairwise correlation coefficients among explanatory variables while most coefficients are moderate or low correlation suggesting weak or no serious collinearity. This method is not usually reliable in identifying severe collinearity as it gives many low pairwise correlation coefficients. However, high pair-wise correlation is sufficient but not a necessary condition for the existence of the multicollinearity [17].

**Table 2. OLS results for linear regression model (1991 – 2010)**

Variable	Coeff.	Std. error	t-stat	Prob.	TOL	VIF	Eigen
(Constant)	.589	1.026	.574	.597			14.561
Agriculture	-1.532E-5	.000	-.835	.451	.001	1257.876	.845
Crude petroleum	3.910E-5	.000	1.841	.139	.149	6.694	.503
Mining and quarrying	.002	.002	1.016	.367	.003	300.585	.051
Manufacturing	1.284E-5	.000	5.111	.007**	.025	39.289	.023
Utilities	-.001	.001	-2.357	.078*	.012	81.933	.009
Building and construction	.000	.000	-1.130	.322	.000	9225.752	.004
Transport	3.687E-5	.000	.197	.853	.005	183.902	.002
Communication	.001	.001	1.837	.140	.005	193.159	.001
Wholesale and retail	1.730E-5	.000	3.419	.027*	.001	1130.069	.001
Hotel and restaurants	.000	.003	-.044	.967	.000	8128.386	.000
Finance and insurance	-3.525E-5	.000	-1.293	.266	.001	784.351	.000
Real estate and business service	.000	.001	-.195	.855	.002	466.836	.000
Housing dwelling	.000	.000	-.591	.587	.000	2680.181	3.879E-5
Producers of government services	7.713E-5	.000	2.239	.089*	.005	206.728	2.570E-5
Community social and personal services	.000	.000	1.101	.333	.000	2886.967	5.183E-6

*R-square .995, Adjusted R-square .974, S.E of regression .028, Sum of Squared residuals .003, Durbin-Watson Statistics 2.821, \*P < 0.05; \*\*P < 0.001*

### 3.3 Remedial Measures to Multicollinearity

#### 3.3.1 Acquisition of additional data

As one of the remedial measures to reducing multicollinearity, additional data from 1981 – 1990 was added to the initial time-series data to make a total of 30 years. OLS regression result for this measure is shown in Table 4 to assess the extent to which multicollinearity severity has been reduced by comparing it with OLS regression results in Table 2. With the acquisition of additional data, mining and quarrying, hotel and restaurants, finance and insurance, real estate and business service sectors and housing dwelling (-.001, -.001, -5.138E-5 and -.010 respectively) show negative relationship with GDP growth rate while the other sectors show positive relationship with GDP growth rate. Real estates and business service sector (-.010) exhibit greatest influence on GDP growth rate similar to that of Table 2. Crude petroleum (.001), manufacturing (.000), wholesale and retail (.006), real estate and business service (.003), producers of government services (.001), community social and personal services (.001) are positive and significant at 1%. Transport

(.017) sector, finance and insurance (.041), and mining and quarrying sector (.063) are significant at 5%. Also,  $R^2$  (.979) and its adjusted value (.953) are high in Table 6 relative to that of Table 4.

Different indicators pointed out to the fact that collinearity among the sectors was reduced as a result of the added data. In Table 4 after acquisition of additional data, more coefficients (9) are significant than in Table 2 before acquisition of additional data. More coefficients in Table 4 have tolerance (TOL) values (greater than zero) than in Table 2. Thus explanatory variables before acquisition of additional data exhibit greater correlation than after the acquisition of additional data. Also, more VIF values in Table 4 are closer to 10 than in Table 2, though none of the VIF values are below 10. Condition Index (CI) value (211.821) above 30 indicated a severe case of multicollinearity but less severe compared to CI value (275.559) before the acquisition of additional variables. Therefore, acquisition of additional data to the initial dataset made collinearity among the sectors less severe.

Table 3. Pairwise correlation among sectors (1991 – 2010)

Variable	Community social	Manufacturing	Crude petroleum	Transport	Mining quarrying	Utilities	Communication	Prod. of Gov.	Hotel & rest.	Fin. & insurance	Real estate	Agriculture	Wholesale & retail	Housing	Building & const
Community social	1.00														
Manufacturing	-0.35	1.00													
Crude petroleum	0.27	-0.46	1.00												
Transport	0.19	-0.30	0.69	1.00											
Mining quarrying	-0.32	0.67	-0.50	-0.63	1.00										
Utilities	-0.05	0.21	-0.07	0.21	0.01	1.00									
Communication	0.08	-0.17	-0.42	-0.84	0.26	-0.34	1.00								
Prod.	0.61	-0.27	0.13	-0.26	0.02	-0.67	0.49	1.00							
Hotel & rest.	-0.57	0.78	-0.70	-0.61	0.76	0.31	0.15	-0.42	1.00						
Fin. & insurance	-0.45	0.87	-0.50	-0.42	0.49	0.24	-0.02	-0.34	0.81	1.00					
Real estate	-0.21	-0.05	-0.23	-0.46	0.21	-0.27	0.55	0.14	-0.01	-0.07	1.00				
Agriculture	0.18	-0.38	0.69	0.96	-0.68	0.14	-0.76	-0.20	-0.65	-0.47	-0.44	1.00			
Wholesale & retail	0.13	0.12	0.29	0.34	0.36	0.29	-0.37	-0.05	-0.07	-0.28	0.01	0.28	1.00		
Housing	-0.88	0.32	-0.56	-0.46	0.34	0.05	0.18	-0.54	0.70	0.49	0.19	-0.47	-0.38	1.00	
Building & const.	0.42	-0.82	0.67	0.56	-0.84	-0.39	-0.12	0.35	-0.95	-0.74	-0.07	0.62	-0.19	-0.52	1.00

\*Pairwise correlation among GDP variables



**Table 4. Linear OLS regression results with added data (1981 – 2010)**

Variable	Coeff.	Std. error	t-stat	Prob.	TOL	VIF	Eigen
(Constant)	2.143	.636	3.370	.006**			14.286
Agriculture	1.271E-6	.000	.162	.874	.003	320.956	.910
Crude petroleum	5.673E-5	.000	4.277	.00**	.060	16.643	.552
Mining and quarrying	-.001	.001	-2.048	.063*	.007	149.859	.146
Manufacturing	9.392E-6	.000	5.122	.000**	.065	15.423	.052
Utilities	4.267E-5	.000	.089	.930	.012	86.106	.023
Building and construction	.000	.000	1.023	.326	.000	2492.726	.018
Transport	.000	.000	2.754	.017*	.022	45.057	.006
Communication	.000	.000	1.047	.316	.015	68.415	.003
Wholesale and retail	8.705E-6	.000	3.288	.006**	.004	252.245	.002
Hotel and restaurants	-.001	.001	-1.494	.161	.002	656.478	.001
Finance and insurance	-5.138E-5	.000	-2.295	.041*	.001	717.713	.001
Real estate and business service	-.010	.003	-3.694	.003**	.000	2594.510	.000
Housing dwelling	.000	.000	-3.005	.011*	.001	1849.874	9.560E-5
Producers of government services	.000	.000	4.557	.001**	.005	209.706	3.912E-5
Community social and personal services	.001	.000	4.160	.001**	.000	2855.092	2.027E-5

*R-square .979, Adjusted R-square .953, S.E .034, Sum of Squared residuals .013, F-Statistic 37.419 Durbin-Watson Statistics 2.734*

*\* Significant at 1%, \*\* Significant at 5%,*

*\*P < 0.05; \*\*P < 0.001*

### **3.3.2 Re-specification of the model**

One remedial measure that produced similar results to this multicollinearity remedial measure in Table 4 is re-specifying the model into a log-linear regression model as shown in Table 5. OLS results from the log-linear model such as  $R^2$  (.993), adjusted  $R^2$  (.965), number of significant variables (9) and CI value (291.465) indicated severe multicollinearity, less severe than that of the linear model (403.722) but more severe (211.821) than after additional data acquisition.

### **3.3.3 Dropping variables from the model**

As this is a severe case of multicollinearity, another remedial measure is dropping highly collinear explanatory variables from the model to reduce collinearity among them. Variable were dropped based on tolerance and CI values of each of the variables leaving 5 sectors as shown in Table 6. From the table, utilities (-.002), and

transport (-9.835E-5) show negative relationship, while other sectors show a positive relationship with GDP growth rate. Manufacturing (.000), utilities (.006), and communication (.001) are significant at 1% with relatively high  $R^2$  (.834) and adjusted  $R^2$  (.775). Tolerance values for all the variables are well above zero (*i.e.* variables are less collinear) compared to results in Tables 2, 4 and 5. VIF values are well above 10 indicating far less collinearity by each of these variables. CI value is (19.619) which is lesser than 30 and all the CI values of any of the remedial measures. All these indicate that the multicollinearity reduced to is a moderate one unlike in the previous cases. Thus, dropping of some highly collinear variables, better than any other remedial measures has greatly reduced severe multicollinearity. However as often seen, exclusion of these variables also led to a model specification error as some of the excluded variables are important contributors to aggregate GDP and GDP growth rate.

**Table 5. OLS regression results for log-linear model (1991 – 2010)**

Variable	Coeff.	Std. error	t-Stat	Prob.	TOL	VIF	Eigen
(Constant)	58.757	49.501	1.187	.301			14.561
Agriculture	-5.675	5.284	-1.074	.343	.001	828.790	.845
Crude petroleum	5.941	1.807	3.289	.030*	.171	5.848	.503
Mining and quarrying	-3.527	3.781	-.933	.404	.004	263.801	.051
Manufacturing	2.850	.557	5.114	.007**	.021	47.026	.023
Utilities	-11.509	3.261	-3.529	.024*	.009	112.617	.009
Building and construction	-6.104	6.554	-.931	.404	.000	4577.921	.004
Transport	-24.851	7.371	-3.371	.028*	.002	476.324	.002
Communication	7.700	1.756	4.385	.012*	.005	208.992	.001
Wholesale and retail	-5.864	2.932	-2.000	.116	.000	4022.258	.001
Hotel and restaurants	16.131	7.371	2.189	.094*	.000	2135.355	.000
Finance and insurance	7.854	2.292	3.427	.027*	.001	766.417	.000
Real estate and business service	-4.708	1.495	-3.150	.035*	.008	126.904	.000
Housing dwelling	.050	10.308	.005	.996	.000	4260.068	3.879E-5
Producers of government services	12.119	2.953	4.105	.015*	.005	214.252	2.570E-5
Community social and personal services	1.638	1.668	.982	.382	.001	753.878	5.183E-6

*R-square .993, Adjusted R-square .965, S.E .191, Sum of Squared residuals .147, F-Statistic 36.438 Durbin-Watson Statistics 3.281*  
*\*P < 0.05; \*\*P < 0.001*

**Table 6. Linear OLS regression results of included variables (1991 – 2010)**

Variables	Coeff.	Std. error	t-stats	Prob.	TOL	VIF	Eigen.
(Constant)	.387	.369	1.048	.313			5.332
Crude petroleum	4.099E-5	.000	1.609	.130	.914	1.094	.585
Manufacturing	1.013E-5	.000	7.244	.000**	.724	1.381	.071
Utilities	-.002	.000	-4.043	.001**	.230	4.355	.007
Transport	-9.835E-5	.000	-1.645	.122	.468	2.139	.004
Communication	.001	.000	5.124	.000**	.286	3.502	.002

*R-square .834, Adjusted R-square .775, S.E .084, Sum of Squared residuals .099, F-Statistic 14.070 Durbin-Watson Statistics 2.145*  
*\*\*P < 0.001*

### **3.3.4 Transformation of variables**

As a result of specification bias inherent from the previous measure (dropping of variables), transforming the variables is another preferable way by which multicollinearity can be reduced. In this study, the variables were transformed by taking the second difference of the time series data. The time series data was transformed by differencing because these variables tend to move in the same direction over time and are non-stationary. Thus, transforming the variables

by differencing helps to minimize variables linear dependence as well as non-stationarity or unit root problem. Using the differenced series, agriculture, crude petroleum, mining and quarrying sector, manufacturing, utilities, hotel and restaurants sector, finance and insurance sector and housing and dwelling sector (-7.786E-6, -3.533E-5, -.002, -2.269E-6, -.001, -.004, -1.120E-5, and -.002 respectively) show negative relationship with GDP growth rate while the other sectors show positive relationship with GDP growth rate. Manufacturing (.071), wholesale and

retail (.068), finance and insurance sector (.089) are the few sectors significant at 5%, fewer than the number of significant sectors in the other remedial measures discussed above.  $R^2$  (.961) and its adjusted value (.688) are high and with few significant variables as shown in Table 7. Other indicators showed that multicollinearity was reduced as a result of transforming the variables than in other remedial measures, even though it has the fewest significant variables. There are more coefficients in Table 7 with tolerance (TOL) values (greater than zero) than in Table 2 and 5 except for Table 4 and 6. Thus explanatory variables before transforming the variables exhibit greater interrelationship than after the variable transformation. Table 7 VIF values also have similar observations like that of the tolerance values. Condition Index (CI) value (74.131) above 30 indicated a severe case of multicollinearity but far less severe compared to CI values of the other remedial measures except for dropping of variables.

### 3.3.5 Ridge regression analysis results

The estimate obtained from ridge regression analysis show that all the sectors are positively related to GDP growth rate in line with a priori expectations as indicated in Table 8. This estimate is called a generalized ridge estimate [17]. From the results, crude oil sector's contribution to the GDP growth rate, manufacturing, finance and insurance was significant at 1%; while mining and quarrying, and communication was significant at 5%. The  $R^2$  value (.837) shows a good fit as the sectors explain 83.7% of the observed variation in the GDP growth rate. Hence, the concern of multicollinearity as observed and illustrated in the GDP time series studies above could be reduced using the ridge regression technique. According [8], all ridge regression models are better than Ordinary Least Square when the multicollinearity problem exists, and the best model is the Generalized Ridge Regression because it has

**Table 7. Linear OLS regression results for differenced series (1991 – 2010)**

Variable	Coeff.	Std. error	t. Stat	Prob.	TOL	VIF	Eigen
(Constant)	.062	.009	-1.204	.352			6.364
Agriculture	-7.786E-6	.000	.378	.742	.002	473.043	3.496
Crude petroleum	-3.533E-5	.000	2.186	.160	.051	19.520	2.182
Mining and quarrying	-.002	.002	-1.715	.229	.005	203.026	1.440
Manufacturing	-2.269E-6	.000	3.560	.071*	.005	212.804	.788
Utilities	-.001	.001	.849	.485	.012	84.055	.759
Building and construction	.001	.000	.405	.725	.002	430.776	.380
Transport	.000	.000	.977	.432	.002	405.919	.232
Communication	.000	.001	.513	.659	.006	171.658	.158
Wholesale and retail	1.642E-5	.000	3.623	.068**	.015	68.158	.088
Hotel and restaurants	-.004	.003	-1.506	.271	.004	278.452	.058
Finance and insurance	-1.120E-5	.000	-3.132	.089*	.019	52.077	.025
Real estate and business Service	.002	.001	-.051	.964	.018	54.122	.018
Housing dwelling	-.002	.001	.338	.767	.005	184.552	.009
Producers of government services	7.469E-5	.000	-1.196	.863	.074	13.443	.001
Community social and personal services	.000	.000	2.405	.138	.029	34.792	.001

*R-square .961, Adjusted R-square .688, S.E .107, Sum of Squared residuals .023,*

*F-Statistic 3.284 Durbin-Watson Statistics 1.810*

*\*P < 0.05; \*\*P < 0.001*

**Table 8. Generalized ridge regression results for GDP data (1991–2010)**

Variables	Coeff.	Std. error	t-stats	Prob.
(Constant)	-2.566E-02			
Agriculture	1.584E-07	1.131E-02	0.766	.443470
Crude petroleum	1.865E-06	2.119E-02	0.304	.001331**
Mining and quarrying	7.791E-06	1.741E-02	0.139	.039679*
Manufacturing	9.824E-07	2.092E-02	3.318	.000906**
Utilities	2.038E-05	1.368E-02	0.612	.540802
Building and construction	1.660E-06	1.156E-02	1.049	.294174
Transport	8.844E-07	1.822E-02	0.100	.920477
Communication	4.292E-05	1.331E-02	2.179	.029347*
Wholesale and retail	7.697E-08	1.378E-02	1.052	.293021
Hotel and restaurants	7.028E-06	1.244E-02	0.525	.599872
Finance and insurance	1.3428E-05	1.5678E-02	2.731	.00689**
Real estate and business service	2.4578E-06	1.112E-02	0.867	.4561812
Housing dwelling	3.531E-06	1.159E-02	1.222	.221620
Producers of government services		1.296E-02	0.055	.956522
Community social and personal services	1.573E-08	1.133E-02	0.990	.322340

*R-square .837, S.E .172, Sum of Squared residuals 1.814,*

*\*P < 0.05; \*\*P < 0.001*

smaller MSE of estimators, smaller standard deviation for most estimators and has larger coefficient of determination.

#### 4. CONCLUSIONS

Every identified sector of an economy is believed to contribute to GDP from which the growth rate is estimated. Studies have been conducted to examine the contribution of each sector within the Nigerian economy and the relevance of each sector to the nation's economy. However, obtaining the estimate for such contribution often involve analysis of data involving multiple regression in which characteristic time series data are fitted into the regression model. Associated with multiple regressions is the problem of multicollinearity which may affect adversely result obtained leading to misleading inference. From the analysis conducted, ridge regression analysis would be more appropriate in such studies as it is a more reliable method to draw inference about the contributions of the various sectors to the GDP. The GRR model indicate a more well behaved model compared to the OLS implying a more precise estimation of the regression coefficients for the GRR model, which show a better influence of each sector's contribution to the GDP growth rate within the year 1991-2010 – a twenty-year period. Therefore, it would be more appropriate for researchers in GDP time series studies to account for the problem of multicollinearity by utilizing Ridge regression technique for estimation since all Ridge regression models

give more robust estimates than the OLS models.

#### ACKNOWLEDGEMENTS

The authors wish to acknowledge Adeola Olajide (PhD) and Ekeh Catherine O. for their comments and contribution towards the preparation of this manuscript.

#### COMPETING INTERESTS

Authors have declared that no competing interests exist.

#### REFERENCES

1. Abdulraheem A. Sectoral contribution to Gross Domestic Product (GDP) in Nigeria (1977-2005). Nigeria Institute of International Affairs, 13/15 Kofo Abayomi Road, Victoria Island, Lagos Nigeria; 2011.
2. Blanchard O. Macroeconomics, updated edition. Pearson Education International, Fifth Edition; 2013.
3. NBS. The Nigerian economy: Annual performance report. The presidency, national planning commission report. Abuja; 2011.
4. Oluwakiyesi T. Construction industry report: A haven of opportunities vitiva research (online); 2011. Available:[t.oluwakiyesi@vetiva.com](mailto:t.oluwakiyesi@vetiva.com) (Accessed 3rd January, 2012)

5. CBN. Real sector developments; Central banks of Nigeria annual reports; 2012. Available:<http://www.cenbank.org> (Accessed 20th January, 2016)
6. Jefferey W. Introductory econometrics- A modern approach. Second edition. 2004; 8-9.
7. Greene H. Econometric analysis. Fifth edition. Mc. Graw Hill Publishers. Pearson education incorporated, Upper Saddle River, New Jersey. 2002;25.
8. Shantha Kumari SS. Multicollinearity estimation and elimination. Journal of Contemporary Research in Management. 2008;3(1).
9. Gujarati DN. Basic econometrics, Fifth Edition, McGraw-Hill, New York; 2004.
10. Midi A, Bagheri H, Rahmatullamon AH. Monte carlo simulation study on high leverage collinearity-enhancing observation and its effect on multicollinearity pattern. Sains Malaysiana. 2011;40(12):1437–1447.
11. Montgomery DC, Peck EA, Viving GG. Introduction to linear regression analysis. 3<sup>rd</sup> ed. New York; John Wiley and Sons; 2001.
12. El-Dereny M, Rashwan NI. Solving multicollinearity problem using ridge regression models. International Journal of Mathematical Sciences. 2011;6(12):585-600.
13. Alajekwu UB, Achugbu AA. The role of stock market development on economic growth in Nigeria. A time series analysis. African Research Review. 2012; 6(1):24:(51-70).
14. Ahungwa GT, Haruna U, Rakiya YA. Trend analysis of the contribution of Agriculture to the Gross Domestic Product of Nigeria (1960-2012). Journal of Agriculture and Veterinary Science. 2014;7(1):Ver. iv.50-55.
15. Oziengbe SA. The relative impacts of federal capital and recurrent expenditures on Nigeria's economy (1980-2011). American Journal of Economics. 2013; 3(5):210-221. p-ISSN: 2166-4951 e-ISSN: 2166-496X
16. Blanchard OJ. Comments. Journal of Business and Economic Statistics. 1967;5:449-451.
17. Sang HL. Comparison of regression techniques for certain patterned matrices in the errors in variables model, Ph.D. Dissertation, American University, Washington. Second Edition. 2001;128.

© 2016 Ejiba and Omolade; This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Peer-review history:*  
*The peer review history for this paper can be accessed here:*  
<http://sciencedomain.org/review-history/14808>