***Original Research Article***

# PRINCIPAL COMPONENT ANALYSIS OF NIGERIAN ECONOMY FROM 2006 – 2017


## ABSTRACT:

*The economy of a developing country like Nigeria depends mainly on crude oil exploration, currently other factors such as Agriculture, other available mineral resources, industrialization etc are not properly harnessed to improve the country's GDP. A detailed study of the effect of ten economic factors on Nigeria's economy was investigated and Principal Component Analysis was employed to explain the relationship, distribution and effect of the factors among the various sectors of Nigeria. It was discovered that a strong positive significant association existed between building and manufacturing, construction, wholesale and retail, transportation, communication, utilities, real estate and community social services, while a negative relationship existed between cruel oil and Agriculture. The component with the highest effect on the Nigerian economy was Agriculture, followed by crude oil and manufacturing/industrialization. The PCA has suggested retaining two components i.e. Agriculture and cruel oil. It was recommended that Nigeria Government should consider Agriculture first as the factor with the first effect on its economy before cruel oil and crude oil and manufacturing/industrialization.*



Keywords: Principal component analysis, Gross Domestic Product, Eigenvectors, Eigenvalues and Agriculture.




## 1.0 Introduction

The best means of measuring a country's economy is through Gross Domestic Product (GDP), It is a macroeconomic aggregate that is globally accepted as an economic indicator used in measuring the performance and strength of an economy. Movement of economic aggregates such as the GDP and their associated Price and Volume measures are used to evaluate the overall performance of the economy and hence judge the relative success or failure of economic policies pursued by governments.

Nigeria is a middle income mixed economy and emerging market with expanding manufacturing, financing service, communications and technology entertainment sectors. It is ranked as the 21$^{st}$ –largest economics in the world in terms of nominal GDP and the 20$^{th}$ largest

in terms of purchasing power panty. It is the largest economy in Africa, its re-emergent manufacturing sector became the largest on the continent in 2013, and it produces large proportion of goods & services for the west African sub-continent (world bank, 2015). Oil is a major source of energy in Nigeria and the world in general. Oil being the mainstay of the Nigerian economy plays a vital role in shaping the economic and political destiny of the country. Although Nigeria's oil industry was founded at the

beginning of the century, it was not until the end of the Nigeria civil war (1967 – 1970) that the oil industry began to play a prominent role in the economic life of the country. Nigeria can be categorized as a country that is primarily rural, which depends on primary product exports (especially oil products). Since the attainment of independence in 1960 it has experienced ethnic, regional and religious tensions, magnified by the significant disparities in economic, educational and environmental development in the south and the north. These could be partly attributed to the major discovery of oil in the country which affects and is affected by economic and social components. Crude oil discovery has had certain impacts on the Nigeria economy both positively and adversely. On the negative side, this can be considered with respect to the surrounding communities within which the oil wells are exploited. Some of these communities still suffer environmental degradation, which leads to deprivation of means of livelihood and other economic and social factors. Although large proceeds are obtained from the domestic sales and export of petroleum products, its effect on the growth of the Nigerian economy as regards returns and productivity is still questionable. Odularu, G. O. (2008)

Agricultural sector is seen as an engine that contributes to the growth of the overall economy of Nigeria, despite these efforts the sector is still characterized with low yields, low level of inputs and limited areas under cultivation due to government dependence on mono-cultural economy based on oil. Agriculture has been the backbone of the economy in Nigeria providing employment

And source of livelihood for the increasing population it accounts for over half of the GDP of the Nigerian economy as at independence in 1960. However, the role it plays in the regional and economic development of the country has diminished over the years due to the dominant role of the crude oil sector in the economy. With the increasing food demand in Nigeria, the country has Available natural resources and potential for increasing the volume of crop production towards meeting the food and nutritional requirement of the rapidly increasing population and guarantee

food security in the country. Therefore, the source of national wealth is essentially agriculture. Orji-Okoro (2010).

In 2014, Nigeria changed its economic analysis to account for rapidly growing contributors to its GDP such as telecoms, banking and its film industry (economist.com, 2017). There are so many factors that affect Nigeria's GDP, they include; agriculture, crude petrol, natural gas, manufacturing, building and construction, wholesale and retail, transportation, communication, utilities, real estate, community social services, interest rate, foreign private investment, foreign exchange rate, inflation rate, Oil revenue, Federal government expenditure, Leisure preference, non-marketed activities, environmental quality and resource depletion, underground economy; quality of life etc. This study tends to employ the Principal Component Analysis (PCA) to determine the factors with the highest effect on Nigeria economy over the period of thirteen (13) years (i.e. 2006 – 2017).

## 2.0 Literature Review

There are vast Literatures on Principal Component Analysis of different country's economies, they include; India Sanjaya Kumar Lenka (2015), sub-Saharan Africa, Egbetunde (2010), Morocco, Tunisia and Turkey, Inoubli (2011).

## 3.0 Material and Methods

### 3.1 Principal Component Analysis

Principal component analysis (PCA) is one of the most frequently used multivariate data analysis and it can be consider as a projection method which project observations from a p-dimensional space with p-variable to a k-dimensional space (where K<P) so as to conserve the maximum amount of information (information is measured here through the total variance of the scatter plots) from the initial dimensions. This method transforms a set of highly correlated variable into the new sets which are uncorrelated but contain almost all information in the original data. The new set of variables are often fewer in number than original variables. The features of PCA are; it incorporate no error terms in its structure and hence it is a mathematical procedure. It transforms a set of variable into a new set of uncorrelated variables, It is appropriate when variables are on equal footing.

### 3.2 Derivation of PCA

Given a P dimensional random vector $X_i = (X_1, X_2, \ldots, X_p)$ with mean vector $\mu_i = (\mu_1, \mu_2, \ldots, \mu_p)$ and dispersion matrix $\sum$, PCA seeks a new set variable $Z_1, Z_2, \ldots, Z_p$ (P' often fewer than P) so that:

$$Z_j = a_{1j}X_1 + a_{2j}X_2 + \ldots + a_{pj}X_p = a_{ij}X \tag{3.1}$$

Where; $j = 1,2,\ldots, p$ *and* $a_{ij} = (a_{1j}, a_{2j}, \ldots, a_{pj})$ are coefficients; Thus; $Z_j$'s are linear combination with coefficient $a_{1j}, a_{2j}, \ldots, a_{pj}$ and

$$Z_{ij} = (a_{1j}, a_{2j}, \ldots, a_{pj}) \begin{bmatrix} X_1 \\ X_2 \\ . \\ .X_p \end{bmatrix} \tag{3.2}$$

And $Z_1 = a_{i1}X, Z_2 = a_{i2}X, \ldots, Z_p = a_{ip}X$ (3.3)

The procedures try to obtain $a_{ij}$ so that $Z_j$, the $j^{th}$ PC of X will have the following properties

(i)The Z's are orthogonal

(ii)Each Z capture the maximum variable remaining in X, hence we maximize the variation in Z subject to the constraint $a_i a = 1$

For instance if Z, is the first PC we seek for a', such that

$(i) Var(Z_i) = Var(Z_i X) = a_i Var(X) a_1 = a_{i1}, \sum a_1$ is a maximum

$(ii) a_{1i1} a_i = \sum a^2_1 = 1$

If $Z_2$ is the second PC after determining $Z_1$, and it's uncorrelated with $Z_1$, we seek for $a_2$ such that;

$(i) Var(Z_2) = a_{i2} Var(X) a_2 = a_{i2} \sum a_2$ is a maximum

$(ii) a_{i1} a_2 = \sum a^2 2 = 1$

$(iii) a_{i1} a_2 = 0$, that is, $Z_1$ and $Z_2$ are orthogonal

The procedure continue in this way to select the $j^{th}$ PC such that;

$(i) Var(Z_j) = Var(a^1{}_j X) = a_{ij} \sum a_j$ is a maximum

$(ii) a_{ij} a_j = \sum a^2{}_j = 1$

$(iii) a_{ij} a_m = 0, j \neq m$ i.e., $Z_j$'s are orthogonal

### 3.3 The procedure of finding the first pc

To find the first PC, we seek for $a_1$, such that

$$Z_i = a_{11} X_1 + a_{21} X_2 + \ldots + a_{p1} X_p = a_{i1} X \quad (3.4)$$

is the PC of X subjected to the constant.

$$(i) Var\ (Z_i) = Var\ (a^1{}_1 X) = a_1{}^1 \sum a, is\ \ a\ \max\ \ imum$$

$$(ii)\, a^1{}_1 a = 1$$

To maximize var ($Z_1$) subject to $a^1{}_1 a_1$, we define the langrangian function.

$$L(a_1) = a^1{}_1 \sum a_1 - \lambda (a_{i1} a_1 - 1) \quad (3.5)$$

$\lambda$ is the Langrangian multiplier.

To maximize $L(a_1)$ we differentiate $L(a_1)$ partially with respect to $a_1$ and equate the result to zero

Thus,

$$\frac{\partial L(a_1)}{\partial a_1} = 2 a_1 \sum - 2 \lambda a_1 = 0 \quad (3.6)$$

And $$\left(\sum - \lambda I\right) a_1 = 0 \quad (3.7)$$

Where: $\lambda$ = Eigen vector of $\sum a_1$ is the corresponding Eigen vector of $\lambda$

The solution $a_1 = 0$ is a trivial solution and since $a_1$ cannot be zero (i.e. $a_1 \neq 0$) to have non-trivial solution then $$\left(\sum - \lambda I\right) = 0 \quad (3.8)$$

and implies that $$\sum = \lambda I \quad (3.9)$$

If (3.7) have non-trivial solution, then because of (3.8), $\lambda$ must be the characteristics root of $\sum$.

Hence we will have P characteristic root and P $a_i$'s which are vector since $\sum$ is a P x P dimensional matrix lets $\lambda_1, \lambda_2 \ldots, \lambda_p$ be the characteristic roots of $\sum$, then var($Z_1$) = $a'_1 \lambda a_1 = \lambda$ hencemax var ($Z_1$), is equivalent to max ($\lambda$). That is, if we have P$\lambda$'s we choose the maximum and Var ($Z_1$) = $\lambda_i$.

### 3.4 The Procedure for Finding the Second PC

Let $Z_2$ be the second PC of X, then $Z_2 = a_{2i} X$ .

We seek $a_2$ such that;

$$(i) Var\ (Z_2) = a_{2i} \sum a_2\ is\ a\ \max\ \ imum$$

$$(ii)\, a_{2i} a_2 = 1$$

$(iii)\, a_{1i} a_2 = 0$

The langrangian function is

$$L(a_2) = a_{2i} \sum a_2 - \lambda (a_{2i} a_2 - 1) - \theta a_2^{1} a_1 \quad (3.10)$$

Thus;

$$\frac{\partial L(a_2)}{\partial a_2} = 2 a_2 \sum - 2\lambda a_2 - \theta_{a1} = 0$$

$$= 2\left(\sum - \lambda I\right) a_2 - \theta_{a1} = 0 \quad (3.11)$$

Multiply by $a_i$ to get;

$$2 a_{1i} \sum a_2 - 2 a_{1i} \lambda a_2 - a_{1i} \theta_{a1} = 0$$

$$2 a_{1i} \sum a_2 - \theta = 0$$

$2a_1'' \lambda a_2 - \Theta = 0$

$$2 a_{1i} \lambda a_2 - \theta = 0$$

$$2(\sum - \lambda I) a_2 = 0$$

$$(\sum - \lambda I) a_2 = 0 \quad (3.12)$$

$$\sum a_2 = \lambda a_2$$

### 3.5 Interpretation of the Principal Components

The *loading* or the eigenvector $\alpha_j = \alpha_1, \alpha_2, \ldots, \alpha_p$, is the measure of the importance of a measured variable for a given PC. When all elements are positive, the first component is a weighted average of the variables and is sometimes referred to as measure of *overall crime rate*. Likewise, the positive and negative coefficients in subsequent components may be regarded as factor of the Nigeria economy components. The plot of the first two or three loadings against each other enhances visual interpretation.

The *score* is a measure of the importance of a PC for an observation. The new PC observations $Y_{ij}$ are obtained simply by substituting the original variables $X_{ij}$ into the set of the first PCs. This gives

$$Y_{ij} = \alpha'_{j1} X_{i1} + \alpha'_{j2} X_{i2} + \ldots + \alpha'_{jp} X_{ip} \quad (3.13)$$

$i = 1,2,\ldots, n, \quad j = 1,2,\ldots,p$

### 3.6 The Proportion of Variance

The proportion of variance tells us the PC that best explained the original variables. A measure of how well the first $q$ PCs of Z explain the variation is given by:

$$\psi_p = \frac{\sum_{j=1}^{q} \lambda_j}{P} = \frac{\sum_{j=1}^{q} Var\left(z_j\right)}{P}$$

A cumulative proportion of explained variance is a useful criterion for determining the number of components to be retained in the analysis. A Scree plot provides a good graphical representation of the ability of the PCs to explain the variation in the data.

## 4.0 Analysis and Results

In this section, the results of the analysis on the data of the factor that affects Nigeria economy from 2006 through 2017.

Table 1 summarizes the results of correlation between all the components, most of the correlations are well above 0.3 (a good indication that we will obtain a result). About 62.2% of the correlation coefficients among the variables are over 0.8 while about 33.3% of the correlation coefficient are negative.

Table 1. Correlation matrix of the factor affecting Nigeria's GDP

| Compone | Agric | Crude | Manu. | Build | Whole | Trans. | Comm. | Utility | Real | SCS |
|---|---|---|---|---|---|---|---|---|---|---|
| Agric | 1 | | | | | | | | | |
| crude | -0.038 | 1 | | | | | | | | |
| manu | 0.149 | -0.874 | 1 | | | | | | | |
| Build | 0.167 | -0.848 | 0.991 | 1 | | | | | | |
| Wholes | 0.271 | -0.817 | 0.95 | 0.979 | 1 | | | | | |
| Trans | 0.285 | -0.733 | 0.898 | 0.945 | 0.979 | 1 | | | | |
| Comm | 0.312 | -0.723 | 0.905 | 0.948 | 0.986 | 0.99 | 1 | | | |
| Utility | 0.256 | -0.781 | 0.963 | 0.955 | 0.919 | 0.867 | 0.885 | 1 | | |
| Real | 0.276 | -0.83 | 0.959 | 0.982 | 0.997 | 0.973 | 0.983 | 0.926 | 1 | |
| scs | -0.003 | -0.892 | 0.975 | 0.97 | 0.929 | 0.861 | 0.869 | 0.914 | 0.928 | 1 |

Table 2: KMO and Bartlett's Test

| | | |
|---|---|---|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | | .618 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 204.443 |
| | df | 64 |
| | Sig. | .000 |

The null hypothesis that the correlation matrix is an identity matrix was rejected at 5% level of significance (Bartlett's test of Sphericity; $\chi^2$ = 204.443, p-value =.000), this implies that the

correlation in the dataset are appropriate for factor analysis. Also, "Kaiser-Meyer-Olkin statistic = 0.618" revealed that adequate sampling is being used for this analysis.

Table 3: Extraction method: principal component analysis

| Economic Factors | Eigen values | Proportion (%) | Cumulative (%) |
|---|---|---|---|
| Agriculture | 8.377 | 83.768 | 83.768 |
| Crude Oil | 1.086 | 10.858 | 94.626 |
| Manufacture | .314 | 3.138 | 97.764 |
| Building | .171 | 1.708 | 99.472 |
| Whole/Retail | .028 | .284 | 99.756 |
| Transportation | .016 | .158 | 99.914 |
| Communication | .008 | .084 | 99.997 |
| Utility | .000 | .002 | 99.999 |
| Real Estate | 9.663E-5 | .001 | 100.000 |
| Community/Social | -3.912E-17 | -3.912E-16 | 100.000 |

Table 3 summarizes the ten principal components for the PCA. Each Eigen-value measures the variance accounted for by the corresponding principal component, and calculation of all possible Eigen-values permits all the variance of the original variables to be quantified. Principal components can be ranked according to their ability to explain variance in the original data set. A common approach is to select only those with Eigen-values equal to or greater than one or with at least 80% cumulative variance. The result shows that the Eigenvalues for Agriculture and Crude/Petrol i.e. 8.929 and 0.768 respectively have cumulative variance 84% and 89% respectively, which implies that; the most effective component of the Nigeria's economy is Agriculture then crude oil exploration.

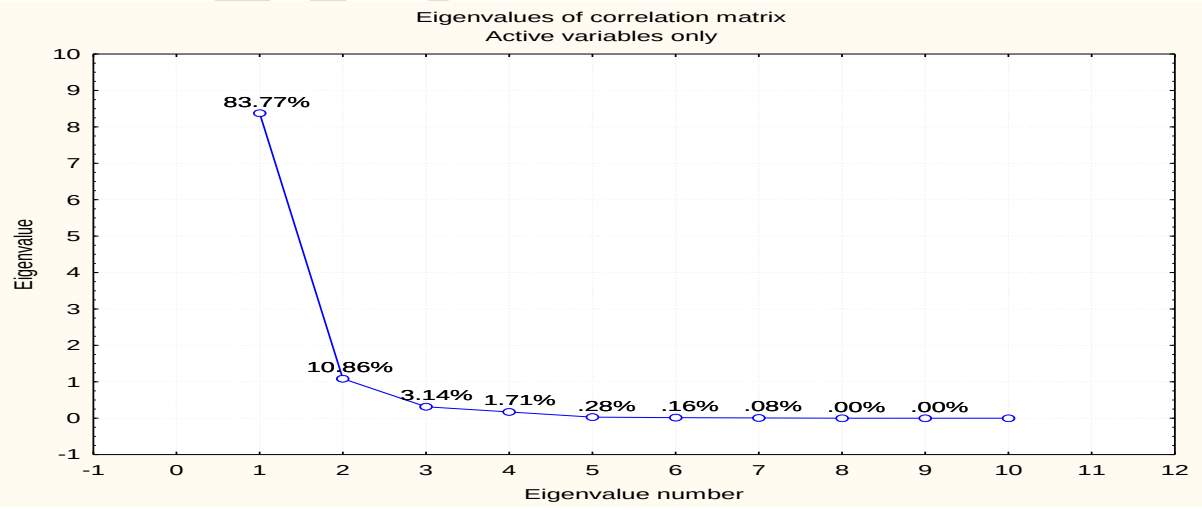


Fig 1: Scree Plot of the economic factors

The scree plot above present the graphical eigenvalue of the Principal components; the first component has the highest eigenvalue of up to 8.37 and a variability of 83.77% is Agriculture followed by the Crude oil exploration with an eigenvalue of 1.09 and a 10.89% variability.

Table 4: Component Matrix of the economic factors

| | Component | |
|---|---|---|
| Economic Factors | 1 | 2 |
| Agriculture | .231 | .954 |
| Crude Oil | .859 | .245 |
| Manufacture | .982 | -.098 |
| Building | .995 | -.062 |
| Whole/Retail | .991 | .060 |
| Transportation | .957 | .112 |
| Communication | .962 | .138 |
| Utility | .951 | .028 |
| Real Estate | .994 | .061 |
| Community/Social | .957 | -.250 |

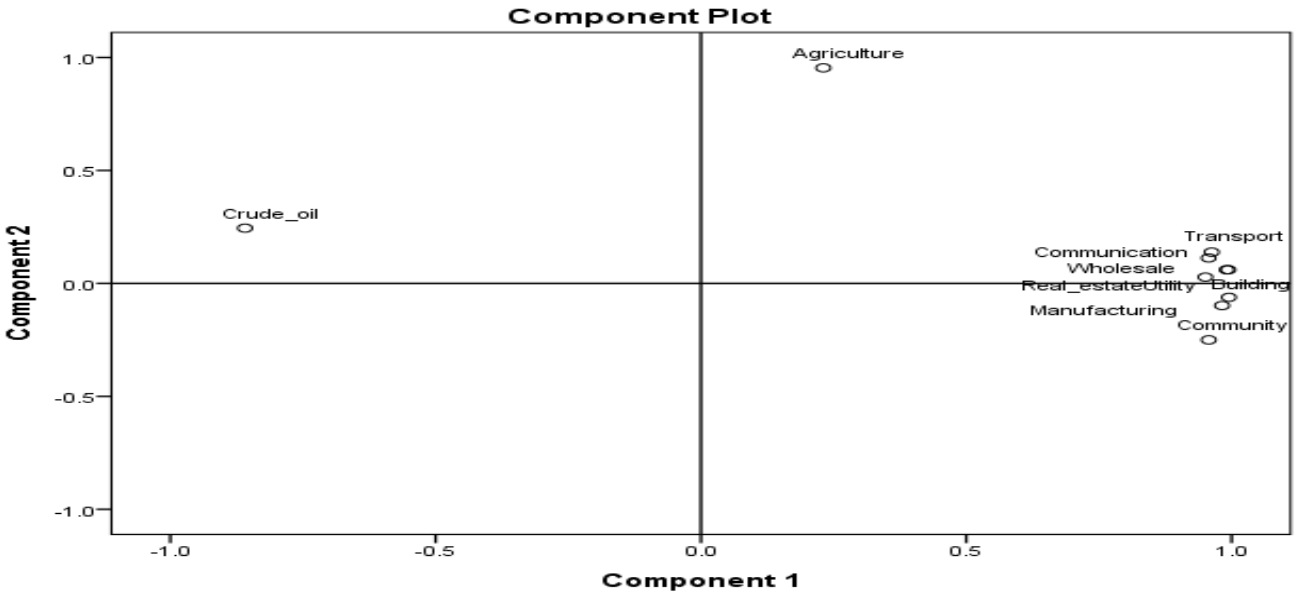


Fig 2: Component Plot of the economic factors

The result and chart above show the expected pattern with high positive and high negative loadings on the first factor. Agriculture and Crude oil have a low and negative loading in the first factor but high and positive in the second factor. This implies the Agriculture and crude oil will both lead to an improvement in the Nigeria's economy (both had positive coefficients in components 1 and 2), all the other economic factors also had positive coefficients at the both components except manufacturing, social and community that had negative coefficients at the second components.

Table 5: Communalities of the economic factors

| Economic Factors | Initial | Extraction |
|---|---|---|
| Agriculture | 1.000 | .963 |
| Crude Oil | 1.000 | .798 |
| Manufacture | 1.000 | .974 |
| Building | 1.000 | .994 |
| Whole/Retail | 1.000 | .986 |
| Transportation | 1.000 | .928 |
| Communication | 1.000 | .945 |
| Utility | 1.000 | .905 |
| Real Estate | 1.000 | .991 |
| Community/Social | 1.000 | .978 |

The communalities result shows the proportion of each economic factor's variance that can be explained by the principal components, most of the economic factor communalities are closer to one indicating high percentage of variability is attributed to the model.

## 5.0 Discussion and Conclusion

Principal component analysis of economic factors affecting the Nigeria economy was conducted. The first and second principal components had Eigen-values greater than one and a cumulative variance of 83.77% and 94.63%, respectively. According to the coefficient of each variable in the combination, Agriculture, crude oil, building and construction, wholesale and retails, transportation, communication, utility and real estate would cause Nigeria's economic to

improve while increase in manufacturing and community/social life would lead to decline in Nigeria's economy.

**Reference**

(1) Akinlo, A.E., Egbetunde, T. (2010). "Financial Development and Economic Growth: The Experience of 10 Sub-Saharan African Countries Revisited", *the Review of Finance and Banking*, 2(1), pp. 17 – 28.

(2) Inoubli, C. (2011). "Does financial development impact on growth? Empirical evidence with threshold effect in the MENA Region", *ERF 17th Annual Conference Paper*. Economic research forum.

(3) Lam, J.C., Wan, K.W., Wong, S.L. and Lam, T.T (2010). "Principal component analysis and long-turn building energy simulation correlation", Energy Conversion and Management, vol. 51, Jan. 2010, pp. 135-139.

(4) Odularu, G. O. (2008). Crude Oil and Nigerian Economic Performance, Oil and Gas Business, pp. 1 – 30.

(5) Orji-Okoro, I. (2010). Analysis of the Contribution of Agricultural Sector on the Nigerian Economic Development.

(6) World Bank Annual report (2015)