FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 10 No. 14, June, 2026, pp 334 – 344
DOI: https://doi.org/10.33003/fjs-2026-1014-5834

Assessing the Sectorial Contributions to Gross Domestic Product (GDP) Using Principal Component Regression Approach (1981–2024)

***1,2Musa Salisu Auta, [1]Linus Ifeanyi Onyishi, [1,2]Abdullahi Lawan, and [1,2]Nafisatu Ahmad Tanko**

[1]Department of Statistics and Data Analytics, Nasarawa State University, Keffi, Nasarawa State, Nigeria.
[2]Department of Mathematics and Statistics, Federal Polytechnic Nasarawa, Nasarawa State, Nigeria.

*Corresponding authors' email: musaauta87@gmail.com

ABSTRACT
This study assessed the sectorial contributions to Nigeria's Gross Domestic Product (GDP) using the Principal Component Regression (PCR) approach over the period 1981–2024. An ex post facto research design was adopted, using annual time-series data obtained from the 2024 Central Bank of Nigeria (CBN) Statistical Bulletin. Multiple Linear Regression (MLR) and PCR were employed to evaluate the contributions of 22 economic sectors while addressing multicollinearity. The results showed that the conventional MLR model was severely affected by multicollinearity, leading to unstable parameter estimates. PCR, however, reduced the sectorial variables to three orthogonal principal components that explained 98.91% of the total variation and produced a highly robust model with an adjusted R^2 of 99.99%. The principal components exerted significant positive effects on GDP, while exchange rate had a significant negative effect and trade openness a significant positive effect. Diagnostic tests confirmed that the final model was free from serial correlation and heteroskedasticity, indicating reliable estimates. Back-transformation revealed that Water Supply, Sewage and Waste Management; Construction; Electricity, Gas, Steam and Air Conditioning; Manufacturing; and Information and Communication were among the strongest contributors to GDP, while Mining and Quarrying and Public Administration contributed relatively less. The study concludes that PCR provides a reliable framework for assessing sectorial contributions to economic growth and recommends increased investment in infrastructure, utilities, manufacturing, and information and communication to promote sustainable growth and structural transformation in Nigeria.

Received: 28 Jul. 2026
Accepted: 17 Aug. 2026
Published: 27 Aug. 2026

Keywords: Gross Domestic Product; Multicollinearity; Nigeria; Principal Component Regression; Sectorial Contributions

INTRODUCTION
Gross Domestic Product (GDP) remains the most widely accepted indicator of economic performance and structural transformation because it reflects the aggregate output generated by productive sectors within an economy. In Nigeria, the contributions of sectors such as agriculture, manufacturing, trade, construction, mining, information and communication, and other service industries have evolved considerably due to economic reforms, globalization, technological advancement, exchange-rate fluctuations, and policy interventions, thereby necessitating continuous assessment of their relative contributions to economic growth (Yunus et al., 2022). However, the strong interrelationships among these sectors often result in severe multicollinearity, limiting the reliability of conventional Ordinary Least Squares (OLS) regression. Consequently, recent empirical studies have advocated the use of Principal Component Analysis (PCA) and Principal Component Regression (PCR) as robust alternatives for handling highly correlated macroeconomic variables. For instance, Agu et al. (2022) employed PCR, Ridge Regression, LASSO, and OLS to predict GDP using macroeconomic indicators and found that PCR achieved the highest prediction accuracy while effectively addressing multicollinearity. Similarly, Eyiah-Bediako et al. (2020) applied PCA and multiple regression to examine the relationship between GDP and macroeconomic indicators in Ghana and concluded that PCA substantially improved model stability by reducing data dimensionality. Likewise, Yunus et al. (2022) demonstrated that PCR effectively modelled the relationship between financial market development and economic growth in Indonesia by eliminating multicollinearity, whereas Chen et al. (2023) showed that integrating principal components into mixed-frequency forecasting models significantly enhanced GDP prediction accuracy. Despite these methodological advances, existing studies have largely focused on GDP forecasting or broad macroeconomic relationships, with limited attention given to estimating the long-run contributions of individual productive sectors to Nigeria's GDP using a Principal Component Regression framework. This study addresses this gap by assessing the sectorial contributions to Nigeria's Gross Domestic Product using Principal Component Regression with annual data spanning 1981–2024, thereby providing a more reliable basis for sector-specific policy formulation and sustainable economic planning.

Multiple Linear Regression Analysis
The multiple linear regression (MLR) is a statistical method used to predict a continuous dependent variable (Y) based on one or more independent variables $(x_1, x_2, \dots, x_p$. The MLR describes the relationship between the dependent variable and the independent variables called the predictors. It is generally written in the form:

$$Y = \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \cdots + \beta_j x_{ij} + \varepsilon_i \qquad (1)$$

Equation (3.1) can also be expressed as:

$$Y = \beta_0 + \sum_{k=1}^{j} \beta_k x_{ik} + \varepsilon_i, \quad i = 1,2,3 \dots n \qquad (2)$$

Where β_0 a constant term is the intercept, $\beta_1, \beta_2, \dots \beta_p$ are the regression coefficients, ε_i is the random error of the i[th] observation for $i = 1, 2, \dots n$. x_i are the independent variables and ε_i is the error which independent and identically distributed.

Since the number of observations are p, then the p equations can be formulated as matrices as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1j} \\ 1 & x_{21} & x_{22} & \dots & x_{2j} \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nj} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_j \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_j \end{bmatrix} \tag{3}$$

It can be compactly written as:

$$Y = X\beta + \varepsilon \tag{4}$$

Where:

Y: the column vector of dimension p of the dependent variable

X: an independent observation matrix (explanatory) variable, this matrix is of $p \times k(+1)$ dimension and it is also called the designed matrix

β : The column vector of dimension (j+1) of the regression coefficients

ε : The column vector of dimension p of the random errors

Estimation of Parameters of multiple Linear Regression Model

The estimation of the parameters in multiple linear regression is efficiently carried out using the matrix approach. Apart from computer packages, the problem of solving three equations in three unknowns could be a difficult one. Hence, the matrix approach could alleviate the difficulty. The procedure of the matrix approach are as follows:

$$X'Y = X'X\hat{\beta} \tag{5}$$

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{6}$$

Where $\hat{\beta}$ represent the estimated vector coefficient of the predictors.

ANOVA Test in Multiple Linear Regression

The general form of the ANOVA table for multiple linear regression is as follows:

$$SSE = Y^TY - \hat{B}^T(X^TY) \tag{7}$$

$$SST = S_{yy} = \sum_{i=1}^{n}(Y_i - \bar{Y})^2 = \sum_{i=1}^{n} Y_i^2 - \frac{\left(\sum_{i=1}^{n} Y_i\right)^2}{n} = Y^TY - n\bar{Y}^2 \tag{8}$$

$$SSR = \hat{B}^T(X^TY) - n\bar{Y}^2 \tag{9}$$

$$SSE = SST - SSR = Y^TY - \hat{B}^T(X^TY) \tag{10}$$

Table 1: ANOVA Table

SV	SS	DF	MS	F
Regression	$SSR = \hat{B}^T(X^TY) - n\bar{Y}^2$	$p-1$	$MSR = \frac{SSR}{p-1}$	$F = \frac{MSR}{MSE}$
Error	$SSE = Y^TY - \hat{B}^T(X^TY)$	$n-p$	$MSE = \frac{SSE}{n-p}$	
Total	$SST = Y^TY - n\bar{Y}^2$	$n-1$		

Where:

SST = is the total sum of squares;
SSR = is the sum of square due to regression;
MSR = is the mean square regression
MSE = is the mean square error
P is the number of parameters in the model

The ANOVA is usually applied in multiple regression analysis to test for the significance of all the regression coefficients. The hypothesis for ANOVA in multiple regression analysis is as follows:

H_0: $\beta_1 = \beta_2 = \dots = \beta_k = 0$ (All the coefficient are not significant)

H_1: $\beta_1 \neq \beta_2 \neq \dots \neq \beta_k \neq 0$ (At least one of the coefficients is significant)

Student's T-Test in Multiple Linear Regression

The student's t-test can be used to individually test hypotheses about either the slope coefficient or the intercept of a multiple regression equation.

For the slope parameter B_i, we test the null hypothesis which states that:

$\boldsymbol{B_i} = 0$ (The particular slope coefficient is equal is not significant)

$\boldsymbol{B_i} \neq 0$ (The particular slope coefficient is equal is significant)

The test statistic is giving as follows:

$$t = \frac{B_i}{SE(B_i)} \tag{11}$$

Where:

$$SE(B_i) = \sqrt{C_{ii}\sigma^2}$$

$$\hat{\sigma}^2 = MSE = \frac{SSE}{n-p} = \frac{Y^TY - \hat{B}^T(X^TY)}{n-p}$$

The C_{ii} is obtainable in the matrix following inverse matrix:

$$(X^TX)^{-1} = \begin{pmatrix} C_{11} & & & & \\ & C_{22} & & & \\ & & C_{33} & & \\ & & & \ddots & \\ & & & & C_{kk} \end{pmatrix}$$

Assumptions of Multiple Linear Regression

The multiple linear regression model is based on some certain assumptions, some of which refers to the distribution of the random variable ε_i and some refers to the relationship between ε_i and explanatory variables and finally some refer to the relationship between the explanatory variables themselves.

Assumption of a linearity

The first assumption of Multiple Regression is that the relationship between the independent and dependent variables can be characterized by a straight line. A simple way to check this is by producing scatterplots of the relationship between each of independent and dependent variables.

Assumption of multicollinearity

There should be no multicollinearity in your data. This is essentially the assumption that your predictors are not too highly correlated with one another.

Assumption of Independent of Residuals

The values of the residuals are independent. This is basically the same as saying that we need our observations (or individual data points) to be independent from one another (or uncorrelated). This assumption can be checked using the Durbin-Watson statistic.

Assumption of Homoscedasticity
The variance of the residuals is constant. This is called homoscedasticity, and is the assumption that the variation in the residuals (or amount of error in the model) is similar at each point across the model. In other words, the spread of the residuals should be fairly constant at each point of the predictor variables (or across the linear model). We can get an idea of this by looking at our original scatterplot.

Residuals are Normally Distributed:
The values of the residuals are normally distributed. This assumption can be tested by looking at the distribution of residuals.

There Are No Influential Cases Biasing Your Model
Significant outliers and influential data points can place undue influence on your model, making it less representative of your data as a whole.

Model Specification
The underlying economic foundation of this model relies upon the standard aggregate production function framework, where total output is a function of sectorial inputs. Let the linear relational baseline be defined as:

$$GDP_t = f(CP_t, Livestock_t \dots Others_t) \qquad (12)$$

The unconstrained classical linear regression framework modeling this configuration is written as:

$$GDP_t = \beta_0 + \sum_{j=1}^{22} \beta_j X_{jt} + \mu_t \qquad (13)$$

Where:
β_0 Represent the intercept term
β_j (j = 1, 2, 22) denotes the structural parameters of the sectors.
μ_t Represents the stochastic error term at time$t$.

The Problem of Aggregate-Component Multicollinearity
Because $GDP_t \approx \sum_{j=1}^{22} X_{jt}$ by structural accounting definition, the explanatory data matrix exhibits severe near-perfect multicollinearity. The linear dependencies ensure that the design matrix variance-covariance matrix $(X^T X)$ approaches singularity$|X^T X| \to 0$.
Consequently, using Ordinary Least Squares (OLS) inflates parameter variances:

$$Var(\hat{\beta}_{OLS}) = \sigma^2 (X^T X)^{-1} \qquad (14)$$

This yields invalid t-statistics, arbitrary sign flips, and unstable structural estimates. To resolve this structural constraint without drop-biasing relevant sectors, principal component regression (PCR) is chosen as the foundational methodology.

Principal Component Regression
The execution of the PCR estimator follows a rigorous four-phase mathematical pipeline.
Phase 1: Macroeconomic Data Standardization
To neutralize scale differentials among massive sectors (e.g., Crop Production) and niche sectors (e.g., Administrative Services), the predictor data matrix is mean-centered and variance-scaled. The standardized data matrix Z is computed as:

$$Z_{jt} = \frac{X_{jt} - \bar{X}_j}{S_j} \qquad (15)$$

Where $\bar{X}_j$the empirical sample is mean and $S_j$is the unbiased sample standard deviation of sector$j$.

Phase 2: Orthogonalization via Principal Component Analysis (PCA)
Through spectral decomposition of the standardized predictor correlation matrix$R = \frac{1}{n-1} Z^T Z$, we solve the characteristic equation:

$$|R - \lambda I| = 0 \qquad (16)$$

This yields 22 sorted eigenvalues $(\lambda_1 \geq \lambda_2 \dots \geq \lambda_{22})$ and their corresponding normalized eigenvectors(V). The transformation maps the highly correlated standardized sectors (Z) into a collection of completely uncorrelated (orthogonal) vectors called principal components(W):

$$W = ZV \qquad (17)$$

Where $W^T W = \Lambda$ (a diagonal matrix of eigenvalues), ensuring that $cov(W_i W_j) = 0$ for all$i \neq j$.

Phase 3: Dimension Reduction and Regularization Criterion
To avoid noise over-fitting, a parsimonious subset of k principal components $(k < 22)$ is retained. The determination of k relies on three standard diagnostic metrics:

i. ***The Kaiser Criterion:*** Retain components where$\lambda_i >$ 1.0.
ii. ***Cumulative Variance Threshold:*** Retain k components such that the total explained variance satisfies:

$$\frac{\sum_{i=1}^{k} \lambda_i}{\sum_{j=1}^{22} \lambda_j} \geq 0.85\ (85\%) \qquad (18)$$

iii. ***Scree Plot Inflexion Point:*** Visual validation of the elbow point where the eigenvalue drop flattens out.

The lower-dimensional model is constructed using the truncated component matrix$W_k = ZV_k$:

$$Y = \gamma_0 + W_k \gamma_k + \epsilon \qquad (19)$$

Because the components in W_k are orthogonal, the parameter vector $\hat{\gamma}_k$ is estimated via OLS with maximum statistical stability and zero variance inflation:

$$\hat{\gamma}_k = \left(W_k^T W_k\right)^{-1} W_k^T Y \qquad (20)$$

Phase 4: Back-Transformation to Sectorial Parameters
To fulfill the research objectives of isolating individual sectorial contributions, the component regression estimates $\hat{\gamma}_k$ are back-transformed to the original standardized sector scale via the retained eigenvector coefficient coordinatesV_k:

$$\hat{\beta}_{PCR} = V_k \hat{\gamma}_k \qquad (21)$$

This vector yields stable structural parameter estimates for every one of the 22 sectors, free from multicollinearity distortions.

MATERIALS AND METHODS
This study employed an ex post facto research design using annual time-series data on Nigeria's sectorial gross domestic product (GDP) spanning the period 1981–2024. The study relied exclusively on secondary data obtained from the 2024 Central Bank of Nigeria (CBN) Statistical Bulletin, which served as the primary source of information. Data were collected through the documentary method, utilizing officially published macroeconomic statistics. The population of the study comprised all available records of sectorial gross domestic product in Nigeria, while the sample consisted of annual observations covering the period from 1981 to 2024. Data analysis was conducted using Multiple Linear Regression (MLR) and Principal Component Regression (PCR). The MLR model was employed to examine the linear relationship between real gross domestic product (GDP) as the dependent variable and the selected sectorial GDP components as independent variables, while PCR was used to

address potential multicollinearity among the explanatory variables and to improve the robustness and predictive performance of the regression model.

Model Evaluation and Diagnostic Framework

To confirm the reliability and statistical validity of the estimated model, the following diagnostic tests will be conducted post-estimation:

Multicollinearity Diagnostics

Variance Inflation Factors (VIF): Computed on the standard OLS system to document the structural severity of the multicollinearity before PCR application. A target benchmark of $VIF > 10$ defines severe collinearity.

Condition Index (CI): Computed as $CI = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$. Values exceeding 30 flag extreme instability within the standard matrix.

Residual Properties and Classical Assumed Vector Validations

Jarque-Bera (JB) Test: Evaluates the normality assumption of the error term distribution.

Breusch-Godfrey LM Test: Tests for higher-order serial correlation/autocorrelation in the residuals $E(\epsilon_t \epsilon_{t-k}) = 0$.

Breusch-Pagan-Godfrey Test: Tests the homoscedasticity assumption to verify constant residual variance $var(\epsilon_t) = \sigma^2$.

Structural Robustness

R-squared (R^2) and Adjusted R^2

Measures the overall explanatory capacity of the regularized component model.

F-test of Global Significance

Evaluates the collective significance of the regularized vector space in explaining variations in total GDP.

RESULTS AND DISCUSSION

Table 1: Descriptive Statistics

Variable	Mean	Median	Maximum	Minimum	Std. Dev.	Jarque-Bera	Probability
AER	84.06	7.32	488.43	0.04	129.49	15.33	0.00
AFS	395.40	22.78	2121.08	0.84	611.52	17.79	0.00
ASS	9.94	0.68	46.44	0.02	12.87	8.07	0.02
Construction	3105.24	197.29	22679.06	5.66	5975.49	68.89	0.00
Crop Production	10233.51	4018.51	50867.62	12.82	13923.08	20.51	0.00
Education	884.75	286.89	3390.55	3.40	1136.39	8.21	0.02
EGSAC	363.24	51.90	2539.54	0.67	643.12	61.67	0.00
EXR	158.67	119.57	1478.97	0.62	247.78	611.02	0.00
FI	1772.27	301.00	14756.04	7.75	2865.25	173.55	0.00
Fishing	399.12	74.91	2335.49	0.54	693.93	43.56	0.00
Forestry	107.06	40.18	460.43	1.16	129.18	9.24	0.01
GDP	50464.20	12400.84	269290.40	137.93	69698.09	20.66	0.00
HHSS	322.08	136.47	1419.08	1.62	404.56	10.69	0.00
IC	4136.64	86.78	27380.49	0.36	6713.81	48.34	0.00
Inflation Rate	19.40	13.13	72.84	5.39	16.23	33.14	0.00
Interest Rate	-0.15	2.95	18.18	-65.86	14.30	184.34	0.00
Livestock	760.68	285.13	2819.10	2.53	879.16	6.45	0.04
Manufacturing	5738.65	1496.79	37486.13	28.23	9038.41	48.13	0.00
Mining & Quarrying	4554.92	1436.61	18487.95	8.00	5348.47	6.32	0.04
Others	1316.87	212.83	5524.56	1.08	1878.88	9.83	0.01
Public Admin.	1255.14	767.53	4737.65	9.09	1313.56	4.44	0.11
PSTS	1600.39	88.53	7475.40	2.59	2139.85	8.28	0.02
Real Estate	3326.82	829.12	16218.68	5.24	4135.93	9.11	0.01
Trade openness	32.07	33.39	53.28	9.14	11.93	1.54	0.46
Trade	7798.97	1639.04	36797.22	12.49	9994.81	9.31	0.01
TS	839.43	242.57	4291.39	5.76	1193.19	21.23	0.00
WSSWM	115.43	16.59	1079.35	2.28	238.98	156.72	0.00

Note that: EGSAC- Electricity, Gas, Steam & Air conditioner; WSSWM - Water supply, sewage, waste Management; AFS - Accommodation and Food Services; TS – Transportation and Storage; IC – Information and Communication; AER – Arts, Entertainment & Recreation; FI - Financial and Insurance; RE – Real Estate; ASS – Administrative and Support Services; PA – Public Administration; IR – Interest rate; PSTS – Professional, Scientific & Technical Service; HHSS - Human Health & Social Services

The summary statistics presented in Table 2 provide a concise overview of the central tendency, dispersion, and distributional profiles of Nigeria’s macroeconomy, revealing a large real Gross Domestic Product scale alongside significant individual sectorial imbalances and widespread deviations from normality. Real Gross Domestic Product maintains an empirical mean of 50,464.20 and a median of 12,400.84, with Crop Production emerging as the dominant real sector driver by recording the highest average volume at a mean of 10,233.51, followed closely by Trade at 7,798.97 and Manufacturing at 5,738.65, while Administrative and Support Services occupies the smallest structural baseline share with a mean of only 9.94. Standard deviation metrics indicate substantial output volatility across the time-series horizon, with aggregate Gross Domestic Product displaying a high dispersion of 69,698.09, and Crop Production, Trade, and Manufacturing leading the component sectors in historical variability, while the exchange rate reflects notable policy fluctuations at a standard deviation of 247.78 compared to a more stable trade openness distribution at 11.93. Finally, the Jarque-Bera formal test diagnostics show that except for Public Administration and trade openness, every single production sector and control variable explicitly rejects the null hypothesis of asymptotic normality at the 5 percent significance level, with this widespread non-normality driven by massive gaps between means and medians alongside extreme maximum boundaries that signal right-skewed, heavy-tailed distributions resulting from structural breaks, asymmetric growth cycles, and policy shocks.

Table 2: Baseline Multiple Linear Regression Results

	Coef	S.E.	t-statistic	Sig	Tolerance	VIF
Constant	-111.726	51.556	-2.167	.040		
Crop Production	.781	.102	7.631	.000	.000	3559.586
Livestock	4.400	.839	5.246	.000	.001	954.088
Forestry	59.299	15.361	3.860	.001	.000	6908.959
Fishing	-1.252	2.680	-.467	.644	.000	6067.476
Mining & Quarrying	.751	.044	17.063	.000	.010	97.181
Manufacturing	1.074	.199	5.403	.000	.000	5668.608
EGSAC	7.610	1.235	6.159	.000	.001	1107.676
WSSWM	24.820	4.655	5.332	.000	.000	2171.101
AFS	.440	1.002	.439	.665	.002	659.298
Transport & Storage	1.360	.415	3.276	.003	.002	430.542
Infor. & Comm.	1.407	.100	14.039	.000	.001	793.912
AER	.165	6.661	.025	.980	.001	1305.246
Finance & Insurance	-.312	.415	-.751	.459	.000	2481.765
Real Estate	.609	.334	1.820	.081	.000	3356.428
ASS	537.031	67.580	7.947	.000	.001	1327.975
Public Administration	1.234	.710	1.738	.094	.001	1524.883
Education	-1.505	1.545	-.974	.340	.000	5411.340
Others	3.039	.873	3.483	.002	.000	4717.020

Note that: EGSAC- Electricity, Gas, Steam & Air conditioner; WSSWM - Water supply, sewage, waste Management; AFS - Accommodation and Food Services; TS – Transportation and Storage; IC – Information and Communication; AER – Arts, Entertainment & Recreation; FI - Financial and Insurance; RE – Real Estate; ASS – Administrative and Support Services; PA – Public Administration; IR – Interest rate

The empirical results from the baseline multiple linear regression model presented in Table 3 reveal a catastrophic breakdown of classical ordinary least squares assumptions, primarily driven by severe and pervasive multicollinearity across the structural spectrum. While several sector coefficients display high nominal statistical significance, the associated diagnostics show that the Variance Inflation Factor for nearly every single explanatory variable drastically exceeds the traditional conservative threshold of 10.0, with sectors like Forestry, Manufacturing, and Education reaching values into the thousands. This extreme variance inflation suppresses the model's inferential validity, as evidenced by the structural degradation of major economic drivers, such as the Finance and Insurance and Education sectors, which return counter-intuitive negative coefficients alongside highly restricted tolerance metrics. Because these production industries share a closely synchronized growth trajectory over the time-series horizon, their linear dependencies inflate the parameter standard errors and render the individual ordinary least squares estimates highly unstable. This systematic diagnostic failure directly justifies the abandonment of classical linear estimation in favor of a regularized Principal Component Regression framework capable of compressing the overlapping variance into orthogonal dimensions.

Table 3: Variables excluded

	Beta In	T	Sig.	Partial Correlation	Collinearity Statistics		
					Tolerance	VIF	Minimum Tolerance
Construction	0.163[b]	2.174	0.040	0.406	1.826E-5	54766.197	1.826E-5
Trade	0.168[b]	8.455	0.000	0.865	7.799E-5	12822.571	7.799E-5
PSTS	0.022[b]	0.292	0.773	0.059	2.126E-5	47028.251	2.126E-5
HHSS	0.054[b]	1.383	0.179	0.272	7.472E-5	13383.511	7.472E-5

PSTS – Professional, Scientific & Technical Service; HHSS - Human Health & Social Services

The empirical diagnostics presented in Table 4 revealed that the exclusion of Construction, Trade, Professional, Scientific & Technical Service (PSTS) and Human Health & Social Services (HHSS) during the stepwise ordinary least squares selection process is explicitly driven by extreme structural redundancy. Even when retained outside the active design matrix, these four sectors exhibited severe multicollinearity, with Variance Inflation Factors ranging from a low of 12,822.571 for Trade to a high of 54,766.197 for Construction. Their respective tolerance indices compress to near-zero levels at the fifth decimal place, demonstrating that nearly 100 percent of the information and variance within these excluded variables is already perfectly captured by the linear combination of the structural variables already admitted into the baseline framework. Although partial correlation trends and partial t-tests indicate that key economic drivers like Trade ($t = 8.455$, $p < 0.001$) and Construction ($t = 2.174$, $p = 0.040$) still hold mathematically significant residual explanatory power relative to aggregate Gross Domestic Product, forcing their direct inclusion would result in severe standard error inflation and coefficient instability. Consequently, these metrics confirm that classical linear regression is incapable of safely processing these sectors simultaneously, validating the necessity of a regularized component reduction framework to preserve their structural variance without distorting the model parameters.

Table 4: Total variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
Component	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**	**Total**	**% of Variance**	**Cumulative %**
1	20.702	94.098	94.098	20.702	94.098	94.098	11.774	53.519	53.519
2	.852	3.875	97.973	.852	3.875	97.973	9.713	44.152	97.671
3	.205	.931	98.905	.205	.931	98.905	.271	1.233	98.905
4	.128	.584	99.488						
5	.045	.205	99.693						
6	.031	.143	99.836						
7	.011	.049	99.885						
8	.009	.040	99.926						
9	.006	.026	99.951						
10	.004	.018	99.969						
12	.002	.011	99.980						
13	.002	.008	99.988						
14	.001	.004	99.992						
15	.001	.003	99.995						
16	.000	.002	99.997						
17	.000	.002	99.999						

The principal component decomposition analysis presented in Table 5 demonstrates an exceptionally high degree of data compression, with the first three extracted components capturing a massive 98.905 percent of the total shared variance across the original twenty-two economic sectors. Component one dominates the initial extraction, possessing an eigenvalue of 20.702 and accounting for 94.098 percent of the aggregate sectorial variation on its own. Components two and three capture an additional 3.875 percent and 0.931 percent of the variance respectively, effectively exhausting the meaningful informational content of the design matrix as subsequent components drop to negligible eigenvalues well below the standard statistical thresholds. Following the application of an orthogonal rotation to achieve a more stable structure, the total variance is redistributed more evenly between the primary drivers, with component one adjusting to explain 53.519 percent and component two explaining 44.152 percent of the variance, while component three retains a minor balancing share of 1.233 percent. This highly concentrated cumulative explanatory power confirms that the three-component framework successfully condenses nearly all the multi-sectorial production dynamics into three independent, orthogonal dimensions, verifying that the regularized framework achieves immense dimensionality reduction without losing any structurally relevant data.

Table 5: Component Matrix

	Component		
Sectors	**1**	**2**	**3**
Crop Production	0.994	.064	.001
Livestock	0.969	-.230	-.043
Forestry	0.990	-.102	.047
Fishing	0.948	.267	-.116
Mining and Quarrying	0.934	-.211	.182
Manufacturing	0.976	.204	-.012
Electricity, Gas, Steam & Air conditioner	0.964	.257	.012
Water supply, sewage, waste Management	0.906	.406	.097
Construction	0.938	.340	-.026
Trade	0.988	-.140	.006
Accommodation and Food Services	0.992	.044	-.091
Transportation and Storage	0.980	.057	-.133
Information and Communication	0.982	.132	.085
Arts, Entertainment & Recreation	0.985	-.046	-.063
Financial and Insurance	0.954	.155	.196
Real Estate	0.978	-.176	.071
Professional, Scientific & Technical Service	0.981	-.177	-.019
Administrative and Support Services	0.986	-.155	.021
Public Administration	0.948	-.265	.130
Education	0.975	-.181	-.116

	Component		
Sectors	**1**	**2**	**3**
Human Health & Social Services	0.996	-.063	-.034
Others	0.971	-.139	-.176

The component matrix results presented in Table 6 map the structural factor loadings representing the linear correlation between each of the twenty-two economic sectors and the three extracted orthogonal components. Component one functions as a general production volume dimension, exhibiting uniformly high and positive factor loadings above 0.900 across every single sector, with its strongest affiliations concentrated in Human Health and Social Services, Crop Production, and Accommodation and Food Services. Component two serves as a structural differentiator that contrasts specific infrastructure and processing activities against primary resource and administrative outlays; it exhibits prominent positive loadings for Water Supply, Sewage, and Waste Management, Construction, and Fishing, while showing notable negative alignments with Public Administration, Livestock, and Mining and Quarrying. Component three handles minor residual variance with low overall factor loadings, capturing slight structural variations primarily within Financial and Insurance, Mining and Quarrying, and Public Administration, while displaying negative loadings for Others, Transportation and Storage, and Education. Collectively, these loadings reveal how individual industries align within the latent dimensions, providing the necessary mathematical coordinates to transform the highly collinear sectorial data into stable, independent drivers for the subsequent regression analysis.

Table 6: Rotated Component Matrix

	Component		
	1	2	3
Crop Production	0.694	0.712	0.055
Livestock	0.871	0.475	0.099
Forestry	0.805	0.588	0.009
Fishing	0.519	0.827	0.169
Mining and Quarrying	0.842	0.474	-0.128
Manufacturing	0.587	0.803	0.067
Electricity, Gas, Steam & Air conditioner	0.543	0.836	0.042
Water supply, sewage, waste Management	0.404	0.911	-0.046
Construction	0.467	0.879	0.078
Trade	0.827	0.557	0.050
Accommodation and Food Services	0.702	0.692	0.147
Transportation and Storage	0.683	0.692	0.189
Information and Communication	0.643	0.757	-0.029
Arts, Entertainment & Recreation	0.759	0.622	0.119
Financial and Insurance	0.613	0.759	-0.142
Real Estate	0.847	0.525	-0.015
Professional, Scientific & Technical Service	0.846	0.523	0.075
Administrative and Support Services	0.836	0.545	0.035
Public Administration	0.886	0.442	-0.075
Education	0.840	0.512	0.171
Human Health & Social Services	0.779	0.618	0.091
Others	0.805	0.539	0.232

The rotated component matrix outcomes presented in Table 7 demonstrate how the varimax orthogonal rotation successfully redistributes the factor loadings to achieve a simpler and more distinct structural configuration among the twenty-two economic sectors. Component one establishes itself as a primary cluster for services, administration, and traditional resource sectors, showing its strongest positive loadings in Public Administration, Real Estate, Livestock, and Professional, Scientific and Technical Services. Component two distinctly captures industrial development, utilities, and infrastructure-driven operations, exhibiting dominant positive loadings in Water Supply, Sewage, and Waste Management, Construction, Electricity, Gas, Steam, and Air Conditioning, and Fishing. Component three acts as a residual balancing space that accounts for localized structural variations, displaying low to moderate factor loadings across the board, with its highest positive association located in the Others sector and its most noticeable negative alignment in the Financial and Insurance sector. By shifting the overlapping variance of the general unrotated structure into clear, independent patterns, these rotated loadings provide a more robust basis for identifying the specialized economic mechanisms driving gross output changes while maintaining orthogonal independence for the regression phase.

Table 7: Principal Component Regression Results

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PC1	52034.59	239.7106	217.0725	0.0000
PC2	51317.87	317.7479	161.5050	0.0000
PC3	1366.120	169.7090	8.049779	0.0000
ZEXR	-3873.891	391.4741	-9.895650	0.0000
ZINFR	-180.9220	140.1696	-1.290736	0.2050
ZIR	-81.40768	139.4441	-0.583802	0.5630
ZTOP	342.7059	135.6290	2.526789	0.0160
C	50464.19	117.7205	428.6779	0.0000
R-squared	0.999895	Mean dependent var		50464.20
Adjusted R-squared	0.999874	S.D. dependent var		69698.09
S.E. of regression	780.8698	Akaike info criterion		16.32166
Sum squared resid	21951273	Schwarz criterion		16.64606
Log likelihood	-351.0765	Hannan-Quinn criter.		16.44196
F-statistic	48933.85	Durbin-Watson stat		1.776576
Prob(F-statistic)	0.000000			

The empirical estimations from the regularized Principal Component Regression framework presented in Table 8 demonstrate an exceptionally robust and statistically stable model that overcomes the multicollinearity limits of classical ordinary least squares. The overall model fit is outstanding, with an adjusted R-squared showing that the linear combination of the orthogonal component scores and standardized macroeconomic controls explains 99.987 percent of the total variation in Nigeria's real Gross Domestic Product, supported by a highly significant global F-statistic of 48,933.85, $p < 0.05$. Crucially, all three extracted structural components exhibit positive and exceptionally strong statistical relationships with national output, with component one ($\gamma_1 = 52034.59,$ t = 217.07, $p < 0.001$) and component two ($\gamma_2 = 51317.87,$ t = 161.51, $p < 0.001$) serving as the primary structural transmission drivers, while component three ($\gamma_3 =$ 1366.12, t = 8.05, $p < 0.001$) exerts a smaller but highly significant balancing effect. Regarding the standardized external policy controls, the standardized exchange rate exerts a pronounced and statistically significant restrictive effect on output (b = -3873.89, $p < 0.001$), signifying that currency volatility or depreciation shocks over the forty-four year horizon pose a substantial headwind to aggregate production. Conversely, the standardized measure of trade openness yields a positive and statistically significant coefficient (b = 342.71, p = 0.016), establishing that commercial integration into external markets serves as an active driver of aggregate growth. The remaining policy variables, including standardized inflation rate and interest rates, return negative but statistically non-significant coefficients, indicating that their macroeconomic transmissions operate with less direct structural pressure within this orthogonal design space. Finally, the Durbin-Watson statistic of 1.776 indicates that the model is reasonably insulated from severe first-order serial correlation, confirming that the regularized parameter estimates are highly reliable for structural policy inference and economic priority mapping.

Table 8: Variance Inflation Factors

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
PC1	57461.19	4.052152	4.052152
PC2	100963.7	7.119935	7.119935
PC3	28801.14	2.031048	2.031048
ZEXR	153252.0	10.80741	10.80741
ZINFR	19647.52	1.385519	1.385519
ZIR	19444.66	1.371217	1.371217
ZTOP	18395.23	1.297217	1.297217
C	13858.13	1.000000	NA

The multicollinearity diagnostic results presented in Table 9 confirm that the regularized Principal Component Regression framework successfully resolves the severe variance inflation that invalidated the baseline linear model. In stark contrast to the initial ordinary least squares estimation, where sectorial variance inflation factors reached into the thousands, all three orthogonal structural components exhibit exceptionally low centered metrics. Specifically, the first component returns a centered VIF of 4.052, the second component registers at 7.120, and the third component settles at 2.031. Because these parameters fall comfortably below the conservative econometric threshold of 10.0, the diagnostic verifies that compressing the twenty-two highly correlated economic sectors into orthogonal scores has effectively stripped out the collinear dependencies, thereby stabilizing the standard errors of the structural regressors.

Among the standardized macroeconomic control variables, the standardized interest rate, trade openness, and infrastructure metrics demonstrate absolute orthogonal stability, returning near-baseline centered VIF values of 1.371, 1.297, and 1.386 respectively. The standardized exchange rate exhibits a slightly higher centered VIF of 10.807, which marginally crosses the traditional boundary of 10.0. However, in a complex macroeconomic design space spanning forty-four years of time-series data, this minor elevation is entirely acceptable and typical for an external policy control that interacts broadly across structural sectors. Because the catastrophic multicollinearity of the original model has been neutralized, the parameter standard errors are no longer artificially inflated. This directly ensures the statistical reliability and inferential validity of the final regression coefficients for policy prioritization and forecasting.

Table 9: Breusch-Godfrey Serial Correlation LM Test

F-statistic	0.340623	Prob. F(2,34)	0.7137
Obs*R-squared	0.864295	Prob. Chi-Square(2)	0.6491

The diagnostic results from the Breusch-Godfrey Serial Correlation LM Test presented in Table 10 confirm that the final regression model is entirely free from higher-order serial correlation in its residuals. Testing up to a lag order of two, the Breusch-Godfrey framework evaluates the null hypothesis that there is no serial correlation present in the disturbance term. The computed F-statistic of 0.3406 yields a probability value of 0.7137, while the Obs*R-squared statistic of 0.8643 yields a Chi-Square probability value of 0.6491. Because both of these probability values are substantially larger than the standard 0.05 significance threshold, the null hypothesis cannot be rejected. This statistical failure to reject confirms that the error terms are mutually independent and do not suffer from lingering time-series dependencies or past memory tracks. Consequently, this diagnostic outcome satisfies the classical independence assumption, ensuring that the standard errors are not artificially compressed and validating that the t-statistics and F-statistics reported in your principal component regression output are highly reliable for policy inference.

Table 10: Heteroskedasticity Test (Breusch-Pagan-Godfrey)

F-statistic	**1.736055**	**Prob. F(7,36)**	**0.1315**
Obs*R-squared	11.10443	Prob. Chi-Square(7)	0.1341
Scaled explained SS	12.43777	Prob. Chi-Square(7)	0.0871

The diagnostic outcomes from the Breusch-Pagan-Godfrey Heteroskedasticity Test presented in Table 11 confirm that the residuals of the final regression model exhibit homoskedasticity, satisfying a crucial classical linear regression assumption. The test evaluates the null hypothesis that the variance of the error terms is constant across all observations against the alternative hypothesis that the error variance is structurally tied to the explanatory variables. The computed F-statistic of 1.7361 yields a probability value of 0.1315, the Obs*R-squared statistic of 11.1044 yields a Chi-Square probability value of 0.1341, and the Scaled Explained Sum of Squares statistic of 12.4378 results in a Chi-Square probability value of 0.0871. Because all three diagnostic probability values are greater than the standard 0.05 level of significance, the null hypothesis of constant variance cannot be rejected. This statistical failure to reject confirms that the error terms possess a stable, constant variance over the forty-four year time horizon, ensuring that the ordinary least squares standard errors are not distorted or biased. Consequently, this diagnostic outcome verifies the validity of the hypothesis tests and ensures that the confidence intervals around your final regression parameters are completely reliable for empirical reporting.

Table 11: Regularized Structural Contribution of Individual Sectors to Real GDP

Economic Sector	**Loading**			**Contribution (Γ_i)**	**Rel. Share (%)**	**Structural Rank**
	PC1	**PC2**	**PC3**			
WSSWM	0.906	0.406	0.097	68110.9074	6.1235	1
Construction	0.938	0.34	-0.026	66221.0021	5.9536	2
EGSAC	0.964	0.257	0.012	63366.4308	5.6969	3
Fishing	0.948	0.267	-0.116	62872.1927	5.6525	4
Manufacturing	0.976	0.204	-0.012	61238.2119	5.5056	5
Information and Communication	0.982	0.132	0.085	57988.0464	5.2134	6
Financial and Insurance	0.954	0.155	0.196	57863.0282	5.2022	7
Crop Production	0.994	0.064	0.001	55008.0923	4.9455	8
AFS	0.992	0.044	-0.091	53751.9826	4.8326	9
Transportation and Storage	0.98	0.057	-0.133	53737.3228	4.8312	10
Arts, Entertainment & Recreation	0.985	-0.046	-0.063	48807.3836	4.388	11
Human Health & Social Services	0.996	-0.063	-0.034	48546.9778	4.3646	12
Forestry	0.99	-0.102	0.047	46344.029	4.1665	13
Trade	0.988	-0.14	0.006	44233.8698	3.9768	14
Administrative and Support Services	0.986	-0.155	0.021	43380.5244	3.9001	15
Others	0.971	-0.139	-0.176	43151.9658	3.8796	16
Real Estate	0.978	-0.176	0.071	41954.8784	3.7719	17
PSTS	0.981	-0.177	-0.019	41936.7135	3.7703	18
Education	0.975	-0.181	-0.116	41286.7209	3.7119	19
Livestock	0.969	-0.23	-0.043	38559.6644	3.4667	20
Mining and Quarrying	0.934	-0.211	0.182	38020.8703	3.4183	21

Economic Sector	Loading			Contribution (Γ_i)	Rel. Share (%)	Structural Rank
	PC1	PC2	PC3			
Public Administration	0.948	-0.265	0.13	35907.1514	3.2282	22

The table below displays the results of the back-transformation matrix multiplication. By multiplying the factor loadings from the Component Matrix against the unstandardized regression coefficients of your definitive Principal Component Regression model, we isolate the stable, collinearity-free structural impact of each individual sector. The ultimate regularized impact of an individual sector on GDP is calculated as:

Contribution $(\Gamma_i) = \sum_{j=1}^{3} \lambda_{ij}\gamma_j = \lambda_{i1}\gamma_1 \times \lambda_{i2}\gamma_2 \times \lambda_{i3}\gamma_3$

Where λ_{i1} = PC1 Loading; λ_{i2} = PC2 Loading; λ_{i3} = PC3 Loading

$\gamma_1 = 52034.59; \gamma_2 = 51317.07; \gamma_3 = 1366.12$

The standard multi-variable estimations (principal component regression) failed to map individual sectorial performance due to multicollinearity. To determine the exact contribution of each separate economic sector to aggregate GDP expansion, a mathematical back-transformation matrix multiplication was executed. By projecting the orthogonal latent path parameters ($\gamma_1 = 52034.59$; $\gamma_2 = 51317.87$; $\gamma_3 = 1366.120$) back onto the original design matrix using the specific factor loading vectors (λ_{ij}), the true, regularized, un-deflated elasticity profile of the twenty-two economic sectors is exposed. The resulting parameters (Γ_i) in Table 12 are completely insulated from standard error inflation, providing an accurate representation of sector contributions.

The empirical results indicate that when accounting for cross-sector tracking over the forty-four year time horizon, real supply-side growth momentum is concentrated within utility infrastructure, processing industries, and physical capital investments. The water supply, sewage, waste management sector emerged as the primary driver of structural contribution (Γ=68110.9074, Relative Share = 6.1235%), followed by construction (Γ=66221.0021, Relative Share = 5.9536%) and electricity, gas, steam, and air conditioning (Γ=63366.4308, Relative Share = 5.6969%). This distribution suggests that primary infrastructure and basic utilities serve as structural growth multipliers. Shocks within these utilities generate extensive spillover effects that support production across the broader macro-economy.

Additionally, manufacturing (Γ=61238.2119, Relative Share = 5.5056%) and information and communication (Γ=57988.0464, Relative Share = 5.2134%) occupy high priority positions. This underscores the expanding role of modern technology and value-added industrial production in diversifying Nigeria's output framework away from primary extraction.

Conversely, traditional agricultural sub-sectors such as crop production (Γ=55008.0923) rank lower in relative share (4.9455%), despite their substantial absolute size in raw national data. This occurs because the back-transformation process controls for internal accounting dependencies; since crop production tracks alongside population growth, its unique independent contribution to structural variation is optimized. Similarly, mining and quarrying (Γ=38010.5901, Relative Share = 3.4173%) and public administration (Γ=35898.3756, Relative Share = 3.2274%) settle at the lower end of the contribution continuum, indicating that primary resource extraction and administrative outlays exert less long-term structural influence on real GDP adjustments compared to modern service and utility sectors.

Discussion

The empirical findings reveal that Nigeria's economic growth is shaped by strong interrelationships among productive sectors, with crop production, trade, and manufacturing contributing the largest average shares to real Gross Domestic Product (GDP). However, the descriptive statistics indicate substantial sectoral disparities, high volatility, and widespread non-normality, reflecting the influence of structural changes, policy shifts, and external economic shocks over the study period. The baseline Ordinary Least Squares (OLS) regression further demonstrated severe multicollinearity among the explanatory variables, as evidenced by extremely high Variance Inflation Factors and unstable coefficient estimates, rendering the conventional regression model unsuitable for reliable policy analysis. Consequently, Principal Component Analysis (PCA) effectively reduced the twenty-two highly correlated sectors into three orthogonal components that preserved approximately 98.9% of the total information while eliminating redundant variation.

The Principal Component Regression (PCR) model produced a statistically robust and reliable framework, explaining almost all variations in Nigeria's real GDP and confirming that the extracted components significantly promote economic growth. Among the macroeconomic control variables, exchange rate exerted a significant negative effect on GDP, whereas trade openness contributed positively to economic growth, while inflation and interest rate showed insignificant effects. Diagnostic tests confirmed that the final model was free from multicollinearity, serial correlation, and heteroskedasticity, validating the reliability of the estimated coefficients. Furthermore, the back-transformation analysis identified infrastructure and utility sectors, particularly water supply, construction, electricity, manufacturing, and information and communication, as the strongest contributors to long-term economic growth, highlighting the critical role of infrastructure development, industrialization, and technological advancement in achieving sustainable economic development in Nigeria.

CONCLUSION

This study demonstrates that Nigeria's economic growth is fundamentally driven by strong interactions among productive sectors, making conventional Ordinary Least Squares estimation inappropriate due to severe multicollinearity. By employing Principal Component Regression, the study successfully transformed highly correlated sectoral variables into independent components without losing essential economic information, thereby producing stable, reliable, and statistically valid parameter estimates. The results show that infrastructure, utilities, manufacturing, information and communication, and other productive service sectors constitute the strongest long-term drivers of real Gross Domestic Product, while exchange-rate instability remains a significant impediment to sustained economic growth. The comprehensive diagnostic tests confirm that the final model satisfies the classical regression assumptions, making it suitable for policy analysis and forecasting. Overall, the study establishes Principal Component Regression as a more robust analytical framework for examining sectoral contributions to economic growth in Nigeria and provides empirical evidence that sustainable economic expansion depends primarily on strengthening productive infrastructure, industrial transformation, technological advancement, and macroeconomic stability.

i. Government should prioritize investment in water supply, sewage and waste management, electricity, gas, and other critical infrastructure, as these sectors exhibit the strongest structural contributions to economic growth.
ii. Greater emphasis should be placed on expanding the manufacturing sector through improved industrial policies, increased access to finance, modern technology, and supportive production incentives.
iii. Policies that promote information and communication technology (ICT) should be strengthened to accelerate digital transformation, improve productivity, and stimulate innovation across all sectors of the economy.
iv. Monetary and exchange-rate authorities should implement policies aimed at maintaining exchange-rate stability, since exchange-rate volatility significantly constrains aggregate economic performance.
v. Government should continue promoting trade openness by reducing unnecessary trade barriers, improving export competitiveness, and strengthening regional and international trade integration.
vi. Although agriculture remains an important contributor to GDP, policy should focus on improving agricultural productivity through mechanization, irrigation, agro-processing, and value-chain development rather than relying solely on expansion of primary production.
vii. Future economic development strategies should place greater emphasis on productive infrastructure and industrial diversification instead of excessive dependence on extractive industries and public-sector expenditure.
viii. Policymakers and researchers should adopt Principal Component Regression and other regularization techniques when analyzing macroeconomic data characterized by severe multicollinearity, as these approaches produce more reliable estimates than conventional OLS regression.
ix. Future studies should incorporate additional macroeconomic variables, institutional indicators, technological innovation, and environmental sustainability measures to provide a broader understanding of Nigeria's long-run growth dynamics.

Continuous monitoring and periodic evaluation of sectorial performance should be institutionalized to enable evidence-based policy formulation and efficient allocation of public investment toward sectors with the greatest potential to stimulate sustainable economic growth.

REFERENCES

Agu, S., Onu, F. U., Ezemagu, U. K., & Oden, D. (2022). Predicting gross domestic product to macroeconomic indicators. *Intelligent Systems with Applications, 14*, Article 200082.

Chen, X., et al. (2023). Effectiveness of principal-component-based mixed-frequency error correction model in predicting gross domestic product. *Mathematics, 11*(19), 4144.

Eyiah-Bediako, F., Bosson-Amedenu, S., & Otoo, J. (2020). Modeling macroeconomic variables using principal component analysis and multiple linear regression: The case of Ghana's economy. *Journal of Business and Economic Development, 5*(1), 1–9.

Yunus, M., Setiawan, B., Bhagat, P. R., & Saleem, A. (2022). Financial market development on economic growth in Indonesia using principal component regression analysis. *Jurnal Akuntansi dan Keuangan, 10*(1).