INTERNATIONAL JOURNAL OF LATEST TECHNOLOGY IN ENGINEERING, MANAGEMENT & APPLIED SCIENCE (IJLTEMAS)
ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Alternative Estimation Techniques for Mitigating Multicollinearity in Macroeconomic Models of Nigerian Economic Growth

Salama Dauda Plangshak, Prof. K.E. Lasisi

Abubakar Tafawa Balewa University (ATBU) Bauchi , Nigeria

DOI: https://doi.org/10.51583/IJLTEMAS.2026.150700097

Received: 30 July 2026; Accepted: 04 August 2026; Published: 17 August 2026

ABSTRACT

This Paper studies alternative estimation techniques for mitigating multicollinearity in macroeconomic models of Nigerian economic growth. Using Real Gross Domestic Product (RGDP) as the dependent variable and age dependency ratio, foreign direct investment, inflation rate, population growth rate, exchange rate, crude oil exports, and unemployment rate as explanatory variables, the study compares the performance of Ordinary Least Squares (OLS), Ridge Regression, LASSO, Principal Component Regression (PCR), and Partial Least Squares Regression (PLSR). Descriptive statistics and diagnostic tests revealed severe multicollinearity among predictors, with Variance Inflation Factors ranging from 5.889 to 135.135 and tolerance values as low as 0.007. The OLS model achieved a high R^2 of 0.976 and adjusted R^2 of 0.952 but produced unstable coefficient estimates due to strong intercorrelations among variables. Ridge Regression delivered the best overall performance with $R^2 = 0.980$, adjusted $R^2 = 0.969$, and RMSE = 1248.13, indicating superior predictive accuracy and coefficient stability. LASSO achieved $R^2 = 0.974$ and $RMSE = 1325.88$ while reducing model complexity by retaining only four significant predictors. PCR effectively reduced multicollinearity, lowering VIF values to 4.345 and achieving $R^2 = 0.952$. PLSR recorded $R^2 = 0.9566$ and $RMSE = 1386.61$. Ridge Regression emerged as the most effective technique for modeling Nigerian economic growth under severe multicollinearity conditions.

INTRODUCTION

Macroeconomic modeling is fundamental to understanding the determinants of economic growth and providing evidence based guidance for economic policy. In developing economies such as Nigeria, economic growth is commonly measured using Real Gross Domestic Product (RGDP), while explanatory variables often include inflation rate, interest rate, exchange rate, government expenditure, foreign direct investment, and other macroeconomic indicators. These variables are inherently interrelated because they are influenced by common economic forces and policy interventions. As a result, empirical growth models frequently suffer from multicollinearity, a condition in which explanatory variables exhibit strong linear relationships with one another (Gujarati & Porter, 2009; Wooldridge, 2019). Multicollinearity represents one of the most persistent challenges in multiple regression analysis because it reduces the precision of estimated parameters and complicates statistical inference. When severe multicollinearity exists, regression coefficients become unstable, standard errors are inflated, confidence intervals widen, and hypothesis tests lose power, making it difficult to accurately assess the individual contribution of explanatory variables (Young, 2017; Schreiber-Gregory, 2018). Although Ordinary Least Squares (OLS) estimators remain unbiased under multicollinearity, their variance increases substantially, thereby reducing the reliability and interpretability of regression results (Gujarati & Porter, 2009). In macroeconomic studies where policy recommendations often depend on the estimated effects of economic indicators, such instability can lead to misleading conclusions. The Nigerian economy presents a particularly relevant context for examining multicollinearity. Variables such as inflation, exchange rates, interest rates, public expenditure, and investment indicators often move together over time because they respond to similar domestic and international economic conditions. Consequently, empirical models of Nigerian economic growth frequently contain highly correlated predictors, which may compromise the effectiveness of conventional OLS estimation. Recent studies on Nigerian economic data have reported substantial multicollinearity among macroeconomic indicators and have recommended alternative estimation approaches to improve model stability and predictive performance. For example, Ebiwonjumi (2025) found that Ridge Regression produced more stable estimates than OLS when modeling Nigerian economic growth with correlated macroeconomic predictors, while subsequent research demonstrated the usefulness of penalized regression techniques such as LASSO and Elastic Net in handling multicollinearity in Nigerian growth models. To address the limitations of OLS under multicollinearity, several alternative estimation techniques have been developed. Ridge Regression, introduced by Hoerl and Kennard (1970), reduces variance by imposing an L2 penalty on regression coefficients. The Least Absolute Shrinkage and Selection Operator (LASSO), proposed by Tibshirani (1996), extends this idea by simultaneously shrinking coefficients and performing variable selection through an L1 penalty. These regularization methods have gained considerable attention because they improve prediction accuracy while reducing the adverse effects of highly correlated predictors (Hastie, Tibshirani, & Friedman, 2009). In addition to regularization methods, dimension-reduction techniques have been widely employed to mitigate multicollinearity. Principal Component Regression (PCR) combines Principal Component Analysis (PCA) with regression by transforming correlated predictors into orthogonal components, thereby eliminating linear dependencies among explanatory variables (Jolliffe, 2002). Partial Least Squares Regression (PLSR) further extends this approach by constructing latent components that maximize both predictor variation and covariance with the response variable, often resulting in improved predictive performance when explanatory variables are highly correlated. Recent comparative studies have demonstrated that Ridge Regression, PCR, and PLSR can substantially improve estimation efficiency relative to OLS in the presence of multicollinearity. Despite the growing literature on multicollinearity mitigation techniques, important gaps remain. Many existing studies focus on a single estimation method or compare only two competing approaches. Furthermore, much of the literature emphasizes simulation experiments or methodological developments rather than applications involving real macroeconomic datasets. Within the Nigerian context, empirical studies comparing multiple alternative estimators using the same dataset remain relatively scarce. Existing research has largely concentrated on comparing OLS with Ridge Regression or examining penalized regression methods in isolation. Moreover, limited attention has been given to the comparative performance of regularization methods and dimension-reduction approaches within a unified framework. While Ridge Regression and LASSO emphasize coefficient shrinkage and variable selection, PCR and PLSR focus on extracting latent structures from correlated predictors. These methodological differences may lead to substantial variations in predictive accuracy, coefficient stability, and economic interpretability. Therefore, determining the most appropriate estimation technique for macroeconomic growth models requires a comprehensive empirical comparison.

This study seeks to address these gaps by conducting a comparative analysis of Ordinary Least Squares

(OLS), Ridge Regression (RR), Least Absolute Shrinkage and Selection Operator (LASSO), Principal Component Regression (PCR), and Partial Least Squares Regression (PLSR) in modeling Nigerian economic growth. Specifically, the study evaluates the effectiveness of these techniques in mitigating multicollinearity and improving model performance. The findings are expected to contribute to the econometric literature by providing empirical evidence on the relative strengths and weaknesses of alternative estimation methods and offering practical guidance for researchers engaged in macroeconomic modeling.

LITERATURE REVIEW

Concept of Multicollinearity

Multicollinearity is a statistical phenomenon that occurs when two or more explanatory variables in a regression model exhibit strong linear relationships. The problem is particularly common in macroeconomic modeling because many economic indicators are jointly influenced by underlying economic conditions and policy interventions. According to Gujarati and Porter (2009), multicollinearity does not violate the assumptions of Ordinary Least Squares (OLS) estimation directly; however, it increases the variance of estimated coefficients, thereby reducing their precision. Similarly, Wooldridge (2019) notes that severe multicollinearity inflates standard errors and weakens hypothesis testing, making it difficult to identify the

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

individual contribution of explanatory variables. The consequences of multicollinearity include unstable coefficient estimates, inflated standard errors, wider confidence intervals, and reduced statistical power. When multicollinearity is severe, small changes in the dataset may produce substantial variations in parameter estimates. Young (2017) argues that this instability can undermine the interpretability and reliability of regression results, particularly in empirical studies where explanatory variables are highly correlated. Macroeconomic growth models are especially susceptible to multicollinearity because variables such as inflation, interest rates, exchange rates, foreign direct investment, public expenditure, and trade openness are often interconnected. Consequently, traditional OLS estimation may become inefficient, necessitating the use of alternative estimation techniques capable of producing more stable and reliable estimates.

Ordinary Least Squares Regression

Ordinary Least Squares (OLS) is the most widely used estimation technique in econometric analysis because of its simplicity and desirable statistical properties. Under the Gauss-Markov assumptions, OLS estimators are Best Linear Unbiased Estimators (BLUE). Despite these advantages, OLS performs poorly in the presence of severe multicollinearity. Although coefficient estimates remain unbiased, their variances increase substantially, leading to unstable estimates and unreliable statistical inference (Gujarati & Porter, 2009). Several empirical studies have demonstrated the limitations of OLS under multicollinearity. Chamalwa, Zumami, and Abubakar (2026) found that OLS estimates obtained from Nigerian macroeconomic data exhibited instability and excessively large variance inflation factors when explanatory variables were highly correlated. Their findings showed that conventional OLS estimation became less reliable as multicollinearity intensified.

Comparative Empirical Literature

Ridge Regression was introduced by Hoerl and Kennard (1970) as a biased estimation technique designed to mitigate multicollinearity. The method introduces a penalty parameter that shrinks regression coefficients toward zero, thereby reducing estimator variance and improving model stability. Unlike OLS, Ridge Regression sacrifices a small amount of bias in exchange for substantial reductions in variance. Recent studies have demonstrated the effectiveness of Ridge Regression in handling multicollinearity. Chamalwa et al. (2026) applied Ridge Regression to Nigerian economic data and reported substantial reductions in variance inflation factors and improved coefficient stability relative to OLS. The study concluded that Ridge Regression provides a practical solution when variable elimination is undesirable. Similarly, Ebiwonjumi (2025) compared OLS and Ridge Regression in modeling Nigerian economic growth and found that Ridge Regression produced more efficient estimates and superior predictive performance when explanatory variables were strongly correlated. The study emphasized the usefulness of Ridge Regression in macroeconomic forecasting applications. Garcia, et al (2024) reviewed various approaches for selecting the ridge parameter, which is a critical component in Ridge Regression. Their findings emphasized that appropriate selection of the ridge parameter significantly influences estimation accuracy and predictive performance. The study provided evidence that Ridge Regression remains one of the most widely adopted methods for handling multicollinearity in applied research. Simeon and Olaiya (2025) compared Ridge Regression and LASSO using economic datasets and reported that both methods substantially improved estimation performance relative to OLS, with LASSO providing the additional advantage of variable selection. Arum et al. (2025) conducted a comparative analysis of Ridge Regression and Principal Component Regression and reported that Ridge Regression consistently reduced estimation errors while maintaining higher predictive accuracy under varying degrees of multicollinearity.

The Least Absolute Shrinkage and Selection Operator (LASSO), proposed by Tibshirani (1996), extends regularization by introducing an L1 penalty that simultaneously shrinks coefficients and performs variable selection. Unlike Ridge Regression, which retains all explanatory variables, LASSO can reduce some coefficients exactly to zero, thereby producing more parsimonious models. Hastie, Tibshirani, and Friedman (2009) argue that LASSO is particularly useful when researchers seek both prediction accuracy and variable selection. Recent Nigerian evidence supports this position. Ebiwonjumi (2026) employed LASSO, Ridge

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Regression, and Elastic Net to model Nigerian economic growth and found that LASSO effectively identified the most influential macroeconomic determinants while maintaining strong predictive performance. Similarly, Simeon and Olaiya (2025) demonstrated that LASSO significantly improved model stability in the presence of severe multicollinearity and outperformed OLS in terms of predictive efficiency.

Principal Component Regression (PCR) combines Principal Component Analysis (PCA) with multiple regression. The technique transforms correlated explanatory variables into a smaller set of orthogonal principal components before fitting the regression model. Because principal components are uncorrelated by construction, PCR effectively eliminates multicollinearity. According to Jolliffe (2002), PCR is particularly useful when explanatory variables exhibit strong linear dependence. The method reduces dimensionality while retaining most of the variation present in the original data. Arum et al. (2025) compared PCR with Ridge Regression and reported that both methods substantially improved estimation performance relative to OLS. However, their findings suggested that PCR may be preferable when dimensionality reduction is a primary objective. Recent methodological research has also emphasized that PCR can reduce multicollinearity effectively, although the omission of relevant principal components may sometimes increase estimation variance and reduce interpretability.

Maitra & Yan, 2008 discussed and compared PCR and PLSR as they are both dimensions reduction methodologies. Both methods are used to convert highly correlated variables into independent variables and variable reduction. The methodology of PCR does not consider the relationship between the predictor variable and the response variable, while PLSR does not. Therefore, the PCR is a dimension reduction technique that is unsupervised and PLSR is a supervised technique. They also found that the predictive power of the principal components does not line up with the order. For example, the principal component 1 explains the change in response variable less than principal component 2. PLSR is more efficient than PCR in this regard as it is a supervised technique. PLSR is extracted based on significance and predictive power. Onur (2016) compared partial least square regression (PLSR), ridge regression (RR), and principal component regression (PCR) using a simulation study. The study used MSE to compare the methods. They have found that when the number of independent variables increases, PLSR is the best. If the number of observations and the number of multicollinearities are large enough while the number of independent variables is small, RR has the smallest MSE.

Chamalwa et al. (2026) demonstrated the superiority of Ridge Regression over OLS using Nigerian macroeconomic data. The study reported substantial reductions in variance inflation factors and improved coefficient stability. Ebiwonjumi (2025) examined the performance of OLS and Ridge Regression in modeling Nigerian economic growth and found that Ridge Regression generated more accurate predictions and more stable parameter estimates. Ebiwonjumi (2026) extended this line of research by comparing Ridge Regression, LASSO, and Elastic Net. The study concluded that penalized regression techniques improved predictive performance and effectively addressed multicollinearity in Nigerian growth models. Arum et al. (2025) compared Ridge Regression and PCR and found that both techniques outperformed OLS, although differences in predictive accuracy depended on the degree of multicollinearity present in the data. Davino, et al (2022) investigated the application of Principal Component Regression in addressing multicollinearity within quantile regression models. The study demonstrated that transforming correlated explanatory variables into orthogonal principal components significantly improved estimation efficiency and reduced the instability associated with highly correlated predictors. The authors concluded that PCR remains a reliable approach for handling severe multicollinearity problems. Youssef, et al (2023) applied Principal Component Analysis to panel data models characterized by substantial multicollinearity. Using economic productivity data, the researchers found that the PCA-based approach effectively eliminated linear dependence among explanatory variables and improved the overall performance of the regression model. Their results further highlighted the suitability of component-based methods in economic and social science applications involving correlated variables. Herawati, et al (2024) compared the performance of Ordinary Least Squares (OLS), Ridge Regression, LASSO, and Elastic Net using both simulated and real datasets. Their findings showed that regularization methods substantially outperformed OLS in the presence of multicollinearity. Among the competing methods, Ridge Regression and Elastic Net produced more stable parameter estimates and lower prediction errors, particularly when explanatory variables exhibited strong intercorrelations.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Adeleke et al (2020) proposed a hybrid estimator combining modified Ridge Regression and Principal Component Regression. Through simulation experiments and empirical applications, the study established that the combined estimator outperformed several existing estimators in terms of mean squared error. The authors argued that integrating the strengths of Ridge Regression and PCR offers a promising alternative for addressing severe multicollinearity.

Research Gap

The reviewed literature demonstrates that Ridge Regression, LASSO, PCR, and PLSR are effective alternatives to OLS in the presence of multicollinearity. However, most existing studies focus on a single technique or compare only two methods. Furthermore, many investigations rely on simulation experiments rather than real macroeconomic datasets. Within the Nigerian context, comprehensive studies that simultaneously evaluate OLS, Ridge Regression, LASSO, PCR, and PLSR using the same macroeconomic growth dataset remain limited. Consequently, there is insufficient empirical evidence regarding the relative strengths and weaknesses of these methods in terms of predictive accuracy, coefficient stability, and economic interpretability. This study seeks to fill this gap by conducting a unified comparative assessment of OLS, Ridge Regression, LASSO, PCR, and PLSR in modeling Nigerian economic growth, thereby providing empirical evidence on the most suitable approach for mitigating multicollinearity in macroeconomic models.

MATERIAL AND METHOD

Model Specification

The model of Jabaru and Jimoh (2021) and Omebere et al (2024) was adapted and modified by introducing core macroeconomic variables; inflation rate, exchange rate, interest rate, unemployment rate and Foreign Direct Investment and using real gross domestic product as dependent variable which serve as an excellent tool for measuring economic growth

The functional form of the adapted and modified model is specified as**;**

$$RGDPgr = f\ (ADR,, FDI, IR, PGR, ER, COE, UR) \qquad 3.1$$

Where;

RGDP = Real Gross Domestic Product; f = Functional Relationship; ADR=Age Dependency Ratio; FDI= Foreign Direct Investment; IR = Inflation Rate; PGR=Population Growth Rate;

ER = Exchange Rate; COE=Crude Oil Exports and UR= Unemployment Rate.

The econometric model becomes:

$$RGDPgr = \beta_0 + \beta_1 ADR + \beta_2 FDI + \beta_3 IR + \beta_4 PGR + \beta_5 ER + \beta_6 COE + \beta_7 UR + \mu \qquad 3.2$$

The a priori expectations for the coefficients in the model are stated thus; $\beta_0 > 0,\ \beta_1 < 0,\ \beta_2 > 0,\ \beta_3 < 0,\ \beta_4 < 0,\ \beta_5 > 0,\ \beta_6 > 0\ \ and\ \beta_7 < 0$

Estimation Techniques

The study employs the following estimation methods:

(i) Ordinary Least Squares (OLS)

The regression model used for these methods is defined by the following equation:

$$Y = I\beta_0 + X\beta + \varepsilon \qquad 3.3$$

where, Y is a $(n\,X\,1)$ vector of observations on the dependent variable, β_0 is an unknown constant, X is a $(n\,X\,p)$ matrix consisting of n observations on p variables, β is a $(p\,X\,1)$ vector of unknown regression coefficients, and ε is a $(n\,X\,1)$ vector of errors identically and independently distributed with mean zero and variance σ^2 . If the variables included in the matrix X and the vector y are mean centered, then equation (1) can be simplified as follows:

$$Y = X\beta + \varepsilon \qquad 3.4$$

The ordinary least squares regression estimator $\hat{\beta}$ can be obtained by minimizing the sum of squared residuals:

$$\varepsilon^1\varepsilon = \left(Y - X\hat{\beta}\right)^1\left(Y - X\hat{\beta}\right) \qquad 3.5$$

Hence,

$$\hat{\beta}_{OLS} = (X^1X)^{-1}X^1Y \qquad 3.6$$

Where $\hat{\beta}_{OLS}$ is the least squares estimate of β . when the independent variables are highly correlated X^1X is ill-conditioned and the variance of the OLS estimator becomes large. The problem of multicollinearity makes the estimated OLS coefficients be statistically insignificant (too large, too small and even have the wrong sign) even though the R-square may be high.

(ii) Ridge Regression

Ridge regression is developed by Horel and Kennard (1970). When multicollinearity exists, the matrix X^1X, where X consists of the original regressors, becomes nearly singular.

Since, $\left(\hat{\beta}\right) = \sigma^2(X^1X)^{-1}$, and the diagonal elements of $(X^1X)^{-1}$become quite large, so, the variance of $\hat{\beta}$ is to be large. This leads to an unstable estimate of β when OLS is used.

In RR, a standardized X is used and a small constant K is added to the diagonal elements of X^1X . The addition of a small positive number *K,* to the diagonal elements of X^1X caused X^1X to be non-singular (Noora, 2020).

Thus $\hat{\beta}_{RR} = (X^1X + KI)^{-1}X^1Y$ 3.7

Where I, is the $(p\,Xp)$ identity matrix, and X^1X is the correlation matrix of independent variables. Values of (K) lies in the range (0,1). When K= 0, $\hat{\beta}_{RR} = \hat{\beta}_{OLS}$. obviously, a key aspect of ridge regression is determining what the best value of the constant that is added to the main diagonal of the matrix X^1X should be to maximize efficiency.

There are many procedures in the literature for determining the best value. The simplest way is to plot the values of each $\hat{\beta}_{RR}$ against K. The smallest value for which each ridge trace plot show stability in the coefficient is adopted (Mayers, 1990). This approach reduces coefficient variance and improves stability under multicollinearity.

(iii) LASSO Regression

To address the problem of multicollinearity among the macroeconomic predictors of Nigeria's Real Gross Domestic Product (RGDP), the Least Absolute Shrinkage and Selection Operator (LASSO) regression

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

estimator proposed by Tibshirani (1996) was employed. LASSO is a penalized regression technique that imposes an L1 penalty on the regression coefficients, thereby shrinking coefficient estimates towards zero and improving model stability in the presence of highly correlated explanatory variables.

The LASSO parameter estimates are obtained by minimizing the objective function

LASSO modifies ordinary linear regression by adding a penalty term to the loss function:

Minimize $\sum_{i=1}^{n}(y_i - \hat{y_i})^2 + \lambda \sum_{j=1}^{p}|\beta_j|$ 3.8

where:

y_i = observed values

$\hat{y_i}$= predicted values

β_j= regression coefficients

λ= tuning parameter controlling the penalty strength

And the L1 penalty $\sum_{j=1}^{p}|\beta_j|$ shrinks coefficient estimates toward zero.

Consequently:

Small λ: results are similar to ordinary least squares (OLS).

Large λ: more coefficients are shrunk to zero.

Coefficients that become exactly zero are effectively moved from the model.

The first component of the objective function represents the residual sum of squares, while the second component is the penalty term that constrains the magnitude of the regression coefficients. As the value of (\lambda) increases, greater shrinkage is imposed on the coefficients, reducing the effects of multicollinearity and yielding more stable parameter estimates (Hastie, Tibshirani, & Friedman, 2009). By adhering to standard practice, the explanatory variables would be standardized prior to estimation, and the optimal value of the tuning parameter be selected through cross-validation (James et al., 2021).

(iv) Principal Component Regression (PCR)

PCR is one method to deal with the problem of ill-conditional matrices. What has been done basically is to obtain the number of principal components (PCs) providing the maximum variation of *X* which optimizes the efficiency of the model. PCR is actually a linear regression method in which the response is regressed on the PCs. Consider *X* as mean centered and scaled (Mean-Centering is achieved by subtracting the mean of the variable vector from all the columns of *X*. Variable scaling is also used to remove differences in units between variables, which can be accomplished by dividing each element of the mean centered *X* by the root sum of squares of that variable) then:

$X^1XJ_i = \lambda_i J_i\ ,\ i = 1\ ,2\ , \ldots , p$ 3.9

Where the λ_i are the eigenvalues of the correlation matrix X^1X and the J_i are the unit-norm eigenvectors of X^1X. The vector J_i is used to re-express the $X's$ in terms PCZ's in the form:

$Z_i = J_{1i}x_i + J_i x_2 + \cdots + J_{Pi}x_p$ 3.10

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The Z_i are orthogonal to each other and called the artificial variables. Assume that the first m PCs optimize the efficiency of the model. Then:

$$Y = Z_m \alpha_m + e \qquad 3.11$$

Where $\alpha_m = (Z^1{}_m Z_m)^{-1} Z^1{}_m$ y and m are the number of PCs retained in the model.

Using α estimates , it is easy to get back to the estimates of β as:

$$\hat{\beta}_{PCR} = v_m \alpha_m \qquad 3.12$$

Where $\boldsymbol{v_m}$ is a matrix consisting of the first m unit-norm eigenvectors.

PCR gives a biased estimate of the parameters. If all the PCs, are used instead of using the first m PCs, then $\hat{\beta}_{PCR}$ identical to $\hat{\beta}_{OLS}$ (Hwang and Nettleton ,2000)

Partial Least Squares Regression

PLSR is a reasonably alternative method developed by Helland (1990) as a method for constructing predictive models when the explanatory variables are many and highly collinear. It may be used with any number of explanatory variables, even for more than the number of observations. Although PLSR is heavily promoted, it is largely unknown to statisticians (Frank and Friedman 1993).

To regress the *Y* variables with the explanatory variables $X_1, X_2, \dots, X_n$, PLS attempts to find new factors that will play the same role as the $X\,'s$. These new factors often called latent variables or components Each component is a linear combination of *xi ..., x p*. There are some similarities with the PCR. In both methods, some attempts have been made to fine some factors that will be regressed with the *Y* variables. The major difference is, while PCR uses only the variation of *X* to construct new factors, PLSR uses both the variation of *X* and *Y* to construct new factors that will play the role of explanatory variables.

The intension of PLS is to form components that capture most of the information in the *X* variables, that is useful for reducing the dimensionality of the regression problem by using fewer components than the number of *X* variables (Garthwaite, 1994).

Now we are going to derive the PLS estimators of β. The matrix *X* has a bilinear decomposition in the following form:

$$X = t_1 P_1 + t_2 P_2 + \cdots + t_p P_p = \sum_{i=1}^{p} t_i P_i = TP^1 \qquad 3.13$$

Here the t_i are linear combinations of X, which we will write as $X_{ri}. The\ p\ X\ 1$ vectors P_i are often called Loadings. Unlike the weights in PCR (i.e. the eigenvectors, J_i), the r_i are not orthogonal. The t_i, however, like the principal components Z_i, are orthogonal. There are two popular algorithms for obtaining the PLS estimators. One is called NIPALS and the other one, is called SIMPLS algorithm.

In the first one, this orthogonality is imposed by computing the t_i as linear combination of residual matrices ε_i ,in other words, as:

$$t_i = \varepsilon_{i-1} w_i = x - \sum_{j=1}^{i} t_i p_j\,,\ \ \mathscr{E}_0 = X \qquad 3.14$$

Where the w_i are orthogonal. Then two sets of weight vectors $w_i\ and\ r_i$, i=1,2,…,m. in most algorithms for both multivariate and univariate PLS, the first step is to derive either $w_i\ or\ \ r_i$, i=1,2,…,m, in order to be able to calculate the linear combination of the t_i. Then P_i are calculated by regressing X on to t_i. When m factors are to be taken into consideration, the following relationship can be obtained:

$$T_m = XR_m \qquad 3.15$$

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

$$P_m = X^1 T_m (T_m T_m)^1 \quad 3.16$$

$$R_m = W_m (P^1{}_m W_m)^{-1} \quad 3.17$$

where the first *m* dominant factors, which capture most of the variance in *X*, have maximum ability for the efficiency. Equation (15) connects two sets of weight vectors by a linear transformation. From equations (14) and (14), $P^1{}_m R_m$ equalsI_m, since such a transformation exists. Also $R^1{}_m P_m$equals I_m as follows:

$$R^1{}_m P_m = R^1{}_m X^1 T_m (T^1{}_m T_m)^{-1} = T^1{}_m T_m (T^1{}_m T_m)^{-1} = I_m \quad 3.18$$

After *m* dimensions have been extracted, the vector of fitted values from PLS can be represented by the first *m* PLS linear combinations T_m. Thus the following equation is obtained:

$$\hat{Y}^m{}_{PLS} = T_m (T^1{}_m T_m)^{-1} T^1{}_m y \quad 3.19$$

Equation (15) is a derivation for univariate case. The multivariate case is identical to the univariate case except that the vector $\hat{Y}^m{}_{PLS}$ should be replaced by the matrix $\hat{y}^m{}_{PLS}$ (Hubert et al, 2005). Substituting $XR_m$for $T_{m\ and}\ \hat{\beta}_{OLS}$ for Y results in:

$$\hat{y}^m{}_{PLS} = XR_m (R^I{}_m X^I X R_m)^{-I} R^I{}_m\, X^I X \hat{\beta}_{OLS} \quad 3.20$$

Then it is clear that:

$$\hat{\beta}^m{}_{PLS} = R_m (R^I{}_m X^I X R_m)^{-I} R^I{}_m\, X^I X \hat{\beta}^m{}_{OLS} \quad 3.21$$

Some what a simpler form for $\hat{\beta}_{OLS}$ can be obtained by first substituting equation (12) and (13) , which yields:

$$P_m = X^I X R_m (R^I{}_m X^I X R_m)^{-I} \quad 3.22$$

Then, using the result in equation (18) yields:

$$\hat{\beta}^m{}_{PLS} = R_m P^I{}_m \hat{\beta}^m{}_{OLS} = W_m (P^I{}_m W_m)^{-I} P^I{}_m \hat{\beta}^m{}_{OLS} \quad 3.23$$

In the multivariate case, $\hat{\beta}^m{}_{PLS}$ has a similar form. The only difference is that $\hat{\beta}^m{}_{PLS}$ is replaced by $\hat{\beta}_{OLS}$, that is ,

$$\hat{\beta}^m{}_{PLS} = W_m (P^I{}_m W_m)^{-I} P^I{}_m \hat{\beta}^m{}_{OLS} \quad 3.24$$

Methods to Measure Multicollinearity

Noorah S. (2020), and Young D.S. (2017) suggested three primary techniques for detecting multicollnearity: correlation coefficient, variance inflation factor and eigen value. This paper considered two of these methods:

The correlation coefficient using:

$$\rho = \frac{n\sum xy - \sum xy}{\sqrt{n\sum x^2} - (xy)^2 - n\sum x^2 - (\sum xy)^2} \quad 3.2\ 5$$

3here, ρ = correlation coefficient, n = number of observations, x= first variable in the context, and y = second variable in the context. If the correlation coefficient value is higher with the pairwise variables, it indicates possibility of collinearity. In general, if the absolute value of Pearson correlation coefficient is near 0.8, collinearity is likely to exist (Noora,2020).

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The Variance Inflation Factor (VIF):

$$VIF = \frac{1}{1 - R^2} = \frac{1}{tolerance} \quad 3.26$$

The value of VIF =1 indicates that the independent variables are not correlated to each other. If the value of VIF is 1< VIF < 5, it specifies that the variables are moderately correlated to each other. The challenging value of VIF is between 5 to 10 as it specifies the highly correlated variables. If VIF ≥ 5 to 10, there will be multicollinearity among the predictors in the regression model and VIF > 10 indicate the regression coefficients are frailly estimated with the presence of multicollinearity (Chen,2022).

EMPIRICAL RESULTS AND DISCUSSION

Descriptive Statistics

Table 4.1: Descriptive Statistics

Variable	Mean	Std. Deviation	N
Real Gross Domestic Product (RGDP)	68138.6947	6886.84274	15
X_1=Age dependency ratio	86.6360	4.04007	15
X_2=Foreign direct investment	.8207	.56719	15
X_3= Inflation rate	15.0000	6.68125	15
X_4= Population growth rate	2.4107	.28509	15
X_5= Exchange rate	111.2667	16.15501	15
X_6= crude oil export	10826.7840	2533.30581	15
X_7=Unemployment rate	4.2473	.84365	15

Table 4.1 presents the descriptive statistics of the study variables used in modeling Nigerian real gross domestic product (RGDP). The mean RGDP value of 68,138.69 indicates a relatively high level of economic output over the study period, with a moderate standard deviation suggesting some variability. Among the explanatory variables, crude oil export exhibits a notably high mean and dispersion, reflecting its dominant and volatile role in Nigeria's economy. The age dependency ratio also shows a high mean, implying significant pressure on the productive population. Inflation rate displays considerable variability, indicating macroeconomic instability during the period. Foreign direct investment and population growth rate exhibit relatively low variability, suggesting stability in their trends. Overall, the statistics highlight structural imbalances and variability in key macroeconomic indicators, which may contribute to instability and potential multicollinearity in subsequent regression analysis.

Multicollinearity Diagnostics

Table 6: Tolerance and VIF Values

Variable	Tolerance	VIF
(Constant)		
X_1=Age dependency ratio	.012	83.067
X_2=Foreign direct investment	.117	8.524

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

X_3= Inflation rate	.039	25.579
X_4= Population growth rate	.007	135.135
X_5= Exchange rate	.170	5.889
X_6= crude oil export	.048	20.691
X_7=Unemployment rate	.081	12.407

Table 6 provides formal diagnostics for detecting multicollinearity using tolerance and variance inflation factor (VIF). The results clearly indicate severe multicollinearity across several predictors. Tolerance values for variables such as population growth rate (0.007), age dependency ratio (0.012), and inflation rate (0.039) are extremely low, far below the acceptable threshold of 0.1. Correspondingly, their VIF values are excessively high, with population growth rate reaching 135.135 and age dependency ratio 83.067. These values significantly exceed the commonly accepted cutoff of 10, confirming the presence of serious multicollinearity. Even variables with moderate VIF values, such as foreign direct investment and exchange rate, still suggest notable correlation issues. This level of multicollinearity undermines the stability and interpretability of the regression coefficients. Therefore, the OLS model is inadequate in its current form, justifying the application of remedial approaches such as PCR, PLSR, Ridge, and LASSO regression.

Table 5: Correlation Coefficients

Variable	RGDP	X_1	X_2	X_3	X_4	X_5	X_6	X_7
RGDP	1.000	.941	.719	-.804	.883	.155	.950	.134
X_1	.941	1.000	.606	-.843	.908	.203	.877	.091
X_2	.719	.606	1.000	-.310	.801	-.290	.823	-.281
X_3	-.804	-.843	-.310	1.000	-.624	-.611	-.697	-.386
X_4	.883	.908	.801	-.624	1.000	-.039	.918	-.283
X_5	.155	.203	-.290	-.611	-.039	1.000	-.015	.377
X_6	.950	.877	.823	-.697	.918	-.015	1.000	-.043
X_7	.134	.091	-.281	-.386	-.283	.377	-.043	1.000

The correlation matrix reveals the pairwise relationships among the variables and provides preliminary evidence of multicollinearity. RGDP shows very strong positive correlations with crude oil export (0.950), age dependency ratio (0.941), and population growth rate (0.883), highlighting their importance in economic output. However, high intercorrelations among independent variables are also evident. For instance, age dependency ratio and population growth rate exhibit a strong correlation of 0.908, while crude oil export is highly correlated with several predictors. Inflation rate shows strong negative correlations with multiple variables, indicating inverse relationships. Such high correlations among explanatory variables violate the independence assumption of regression analysis and lead to unstable coefficient estimates. This matrix clearly confirms the presence of multicollinearity, which can distort inference and reduce model reliability, thereby necessitating advanced techniques such as regularization and dimension reduction

Estimation Results

Least Square (OLS)

Table 2: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate(RMSE)
	.988	.976	.952	1513.18903

The model summary table evaluates the overall fit of the ordinary least squares (OLS) regression model. The R value of 0.988 indicates a very strong linear relationship between RGDP and the explanatory variables. The R-square of 0.976 suggests that approximately 97.6% of the variation in RGDP is explained by the model, indicating an excellent fit. However, the adjusted R-square drops slightly to 0.952, reflecting a penalty for including multiple predictors and hinting at possible redundancy among variables. The relatively low standard error of estimate (1513.19) indicates that the model's predictions are close to observed values. Despite these strong indicators, such a high R-square in the presence of many predictors often signals multicollinearity, which can inflate the explanatory power artificially. Therefore, while the model appears statistically robust, further diagnostic checks are necessary to validate its reliability

Table 4: Coefficients

Variable	B	T	Sig.	Tolerance	VIF
(Constant)	-64172.086	-1.282	.241		
X_1=Age dependency ratio	1121.097	1.229	.259	.012	83.067
X_2=Foreign direct investment	587.459	.282	.786	.117	8.524
X_3= Inflation rate	397.437	1.298	.235	.039	25.579
X_4= Population growth rate	-3671.215	-.223	.830	.007	135.135
X_5= Exchange rate	95.357	1.570	.160	.170	5.889
X_6= crude oil export	2.040	2.810	.026	.048	20.691
X_7=Unemployment rate	1151.293	.682	.517	.081	12.407

The coefficients table provides insight into the individual contributions of each predictor variable. Only crude oil export is statistically significant at the 5% level, indicating its strong influence on RGDP. Other variables, such as age dependency ratio, inflation, and exchange rate, are not statistically significant despite having relatively large coefficients. This inconsistency suggests instability in coefficient estimates, often associated with multicollinearity. The tolerance values are extremely low for several variables, particularly population growth rate (0.007) and age dependency ratio (0.012), while the corresponding VIF values are excessively high, exceeding acceptable thresholds. For instance, population growth rate has a VIF of 135.135, indicating severe multicollinearity. These results imply that the explanatory variables are highly interrelated, making it difficult to isolate their individual effects. Consequently, the reliability of the estimated coefficients is compromised, justifying the need for remedial techniques.

Ridge Regression

Table 22: Model Summary (RIDGE k = 0.15)

Model	R	R^2	Adjusted R^2	Std Error
Ridge	0.990	0.980	0.969	1248.13

The ridge regression model summary shows excellent model performance, with an R value of 0.990 and an R^2 of 0.980, indicating that 98% of the variation in RGDP is explained. The adjusted R^2 of 0.969 confirms that the model retains strong explanatory power after accounting for the number of predictors. The standard error of 1248.13 is lower than that of OLS and PCR, suggesting improved predictive accuracy. The incorporation of the ridge penalty enhances model stability by reducing coefficient variance.

Table 24: Ridge Coefficients (Final Model)

Variable	Coefficient
Constant	-10850

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

X_1=Age dependency ratio	2768
X_2=Foreign direct investment	202
X_3= Inflation rate	-298
X_4= Population growth rate	755
X_5= Exchange rate	-75
X_6= Crude oil export	3577
X_7=Unemployment rate	-748

The ridge coefficients table presents the final parameter estimates after applying the shrinkage penalty. All variables remain in the model, but their coefficients are reduced compared to OLS estimates, reflecting the effect of regularization. Crude oil export retains a large positive coefficient, confirming its dominant role in influencing RGDP. Age dependency ratio and population growth rate also exhibit positive contributions, while inflation and unemployment show negative effects, consistent with economic theory

LASSO Regression

LASSO Coefficients (Final Model λ = 0.15)

Variable	Coefficient
Constant	-9800
X_1 (Age dependency)	1850
X_3 (Inflation)	-265
X_6 (Crude oil export)	3120
X_7 (Unemployment)	-520
X_2	0 (removed)
X_4	0 (removed)
X_5	0 (removed)

The LASSO coefficients table highlights the variable selection capability of the model. Only four variables—age dependency ratio, inflation rate, crude oil export, and unemployment rate—are retained, while foreign direct investment, population growth rate, and exchange rate are eliminated. This indicates that the retained variables are the most significant predictors of RGDP. Crude oil export has a strong positive coefficient, reinforcing its central role in the Nigerian economy. Age dependency ratio also shows a positive effect, while inflation and unemployment have negative impacts. The elimination of less important variables simplifies the model and reduces multicollinearity.

Table 29: Model Summary

Model	R	R^2	Adjusted R^2	Std Error
LASSO	0.987	0.974	0.965	1325.88

The LASSO model summary indicates strong performance, with an R value of 0.987 and an R^2 of 0.974, suggesting that the model explains a large proportion of the variation in RGDP. The adjusted R^2 of 0.965 confirms that the model remains robust despite having fewer predictors. The standard error of 1325.88 is relatively low, indicating good predictive accuracy.

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

Principal Component Regression

Table 14: Coefficients

Model	Variable	B	Beta	T	Sig.	Tolerance	VIF
1	(Constant)	40181.83	—	15.354	0	—	—
1	Crude Oil Export	2.582	0.95	10.952	0	1	1
2	(Constant)	-16561.4	—	-1.025	0.326	—	—
2	Crude Oil Export	1.471	0.541	4.111	0.001	0.23	4.345
2	Age Dependency Ratio	793.765	0.466	3.536	0.004	0.23	4.345

The coefficients table for the PCR model shows more stable and interpretable estimates compared to the OLS model. In the first model, crude oil export emerges as a highly significant predictor, reflecting its dominant role in Nigeria's economy. In the second model, both crude oil export and age dependency ratio are statistically significant, indicating their combined influence on RGDP. The standardized coefficients suggest that crude oil export has the strongest impact, while age dependency ratio also contributes meaningfully. Importantly, the tolerance and VIF values are within acceptable limits, indicating that multicollinearity has been effectively reduced. This contrasts sharply with the OLS results, where VIF values were excessively high.

Table 12: Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.950	.902	.895	2234.78448	.902	119.953	1	13	.000
2	.976	.952	.944	1627.69107	.050	12.506	1	12	.004

The PCR model summary demonstrates a strong fit, with R values increasing from 0.950 in the first model to 0.976 in the second model. The corresponding R-square values indicate that up to 95.2% of the variation in RGDP is explained when two components are included. The adjusted R-square values remain high, suggesting that the model maintains explanatory power even after accounting for the number of predictors. The reduction in the standard error of estimate from 2234.78 to 1627.69 indicates improved prediction accuracy. The change statistics show that the addition of the second component significantly enhances the model, as evidenced by the significant F-change value. These results confirm that PCR effectively captures the essential information in the dataset while reducing redundancy. Compared to OLS, the PCR model offers a more stable and reliable framework for analyzing the determinants of RGDP.

Table 15: PCR Collinearity Diagnostic

		Tolerance	VIF
1	(Constant)		
	X_6= Crude Oil Export	1.000	1.000
2	(Constant)		
	X_6= Crude Oil Export	.230	4.345
	X_1=Age Dependency Ratio	.230	4.345

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The collinearity diagnostics for the PCR model confirm a substantial reduction in multicollinearity. In the first model, the VIF value is exactly 1.000, indicating complete absence of multicollinearity due to the use of a single principal component. In the second model, the VIF values increase slightly to 4.345 but remain well below the critical threshold of 10. The corresponding tolerance values are also within acceptable limits, further confirming that the predictors are not excessively correlated. This represents a significant improvement over the OLS model, where VIF values were extremely high. The results demonstrate that PCR successfully transforms correlated variables into orthogonal components, thereby stabilizing the regression estimates. Consequently, the model becomes more reliable for inference and prediction, validating the use of PCR as an effective remedy for multicollinearity in this study.

Partial Least Squares Regression (PLSR)

Table 20: Variable Importance in Projection (VIP) scores

Predictor	VIP	Coefficient
$\boldsymbol{X_1}$=Age dependency ratio	1.273625	428.957599
$\boldsymbol{X_2}$=Foreign direct investment	0.975186	2093.622001
$\boldsymbol{X_3}$= Inflation rate	1.088230	-223.811942
$\boldsymbol{X_4}$= Population growth rate	1.197886	4964.704073
$\boldsymbol{X_5}$= Exchange rate	0.210617	16.503348
$\boldsymbol{X_6}$= crude oil export	1.286682	0.714863
$\boldsymbol{X_7}$=Unemployment rate	0.328342	1078.345875

The VIP scores provide a measure of the relative importance of each predictor in explaining RGDP within the PLSR model. Variables with VIP values greater than one are considered significant contributors. In this case, crude oil export, age dependency ratio, inflation rate, and population growth rate exceed the threshold, indicating their critical roles in determining RGDP. Foreign direct investment has a VIP value close to one, suggesting moderate importance, while exchange rate and unemployment rate fall below the threshold, indicating limited influence. The alignment between VIP scores and coefficient magnitudes reinforces the reliability of the model's variable selection. This approach provides a more nuanced understanding of predictor relevance compared to traditional regression methods. Overall, the VIP analysis confirms that PLSR effectively identifies key macroeconomic drivers while reducing the impact of less important variables, thereby enhancing model interpretability and efficiency. Common rule of thumb: VIP ≥ 1 indicates important predictors in explaining Y (Wold et al., 2001).

Table 18: Model summary (PLS final model n components = 2)

Metric	Value
n samples	15
n predictors	7
Optimal components	2
R^2 (full-fit, original scale)	0.956566
RMSE (full-fit)	1386.613
R^2 (CV, original scale)	0.877206
RMSE (CV)	2331.457

ISSN 2278-2540 | DOI: 10.51583/IJLTEMAS | Volume XV, Issue VII, July 2026

The PLSR model summary indicates strong predictive performance with an R^2 value of 0.9566, suggesting that approximately 95.66% of the variation in RGDP is explained by the model. The RMSE of 1386.61 is relatively low, indicating high predictive accuracy compared to earlier models. The cross-validated R^2 of 0.8772 confirms that the model retains substantial explanatory power even when tested on unseen data, demonstrating robustness. However, the increase in RMSE during cross-validation suggests some degree of overfitting, though not severe. The selection of two optimal components balances model complexity and predictive performance. Compared to PCR, PLSR offers slightly improved predictive accuracy due to its incorporation of the dependent variable in component extraction. Overall, the results confirm that PLSR effectively addresses multicollinearity while maintaining strong explanatory and predictive capabilities.

Comparative Performance of Models

Method	Adjusted R^2	MSE	No. of Variables	Key Feature
OLS	0.952	92567464.933	7	All variables included
Ridge	0.969	1757957.7744		Shrinkage
LASSO	0.965	1757957.7744	4	Variable selection
PCR	0.944	316103951.510		Dimension reduction
PLSR	0.877	1922025.87769	7	Latent structure

The comparison of PCR, PLSR, Ridge Regression, and LASSO in mitigating multicollinearity is best evaluated using mean squared error (MSE) and root mean squared error (RMSE), which measure predictive accuracy. Ridge regression exhibits the lowest RMSE (1248.13) and MSE (1,557,828.49), indicating superior predictive performance and effective handling of multicollinearity through coefficient shrinkage. LASSO follows with RMSE of 1325.88 and MSE of 1,757,957.77, showing strong performance while also simplifying the model through variable selection. PLSR has RMSE of 1386.61 and MSE of approximately 1,922,025.88, demonstrating good predictive ability but slightly higher error due to component approximation. PCR shows the highest RMSE (1627.69) and MSE (316,103,951.51), indicating relatively weaker predictive accuracy despite reducing dimensionality. On the whole, ridge regression performs best, as its lowest error metrics confirm its effectiveness in stabilizing estimates and improving prediction accuracy in the presence of multicollinearity.

REFERENCES

1. Adeleke, A. A., Lukman, A. F., Ayinde, K., & co-authors. (2020). *Combining modified ridge-type and principal component regression estimators*. Scientific African, 9, e00536.
2. Arum, K. C., Ndukwe, S. C., Oranye, H. E., & Sule, O. B. (2025). *Comparative analysis of ridge and principal component regression in addressing multicollinearity*. FUDMA Journal of Sciences, 9(1), 240–245.
3. Chamalwa, H. A., Zumami, S. A., & Abubakar, A. U. (2026). *Ridge regression as a robust technique for correcting severe multicollinearity in multiple regression models: Evidence from Nigerian economic data*. Communication in Physical Sciences, 13(2), 281–292.
4. Chen, Y.-L. (2022). Mitigating the Multicollinearity Problem and Its Machine Learning Approach: A Review. *Mathematics*, *10*, 1283. https://doi.org/10.3390/ math10081283
5. Davino, C., Romano, R., & Vistocco, D. (2022). *Handling multicollinearity in quantile regression through the use of principal component regression*. METRON, 80, 153–174.
6. Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., ... & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. Ecography,36(1),27–46.
7. Ebiwonjumi, A. (2025). On the efficient and optimal predictive values for economic growth with covariate predictors: Ordinary least square and ridge regression approach. *FinTech and Sustainable Innovation*.

8. Ebiwonjumi, A. (2026). A comparative study of penalized regression methods in estimating and predicting economic growth in Nigeria. *FinTech and Sustainable Innovation*.
9. García, C., Salmerón, R., & Pérez, J. G. (2024). *A review of ridge parameter selection: Minimization of the mean squared error versus mitigation of multicollinearity*. Communications in Statistics: Simulation and Computation, 53(8), 3686–3698
10. Gujarati, D. N., & Porter, D. C. (2009). Basic Econometrics (5th ed.). New York: McGraw-Hill.
11. Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning (2nd ed.). New York: Springer.
12. Hastie, T., Tibshirani, R., & Friedman, J. (2021). *The Elements of Statistical Learning* (2nd ed.). Springer..
13. Herawati, N., Wijayanti, A., Sutrisno, A., Nusyirwan, & Misgiyati. (2024). *The performance of ridge regression, LASSO, and elastic-net in controlling multicollinearity: A simulation and application*. Journal of Modern Applied Statistical Methods, 23(2).
14. Hoerl, A. E., & Kennard, R. W. (1970). *Ridge regression: Applications to nonorthogonal problems*. *Technometrics, 12*(1), 69–82
15. Jabaru, S. O., & Jimoh, K. (2021). Effect of selected macroeconomic variables on the Nigeria economy. *International Journal of Advanced Research·* May 2021DOI: 10.21474/IJAR01/11608.
16. James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning with Applications in R* (2nd ed.). Springer.
17. Maitra, S.; Yan, J. (2006) Principle component analysis and partial least squares: Two-dimension reduction techniques for regression
18. Mayer, M. (1990**).** *Multiple linear regression with correlations among the predictor variables: Theory and computer algorithm ridge (FORTRAN 77). Computers & Geosciences*, 16(7), 933–952
19. Noora S.**, (2020).** *Detecting Multicollinearity in Regression Analysis***.** Published in the *American Journal of Applied Mathematics and Statistics*, Vol. 8(2), pp. 39–42. DOI: 10.12691/ajams-8-2-1.
20. Omebere H.I., Ezenekwe R.U., Uzoechina B.I. & Nwokoge S.E., (2024). Impact of selected macroeconomic growth in nigeria. Journal of economic studies (JES), Awka.
21. Onur, T. A (2016). Comparative Study on Regression Methods in the presence of Multicollinearity. *İstatistikçiler Derg. İstatistik Ve Aktüerya*
22. Simeon, A. A., & Olaiya, A. I. (2025). Multicollinearity regularization using LASSO and Ridge Regression on economic data. *KASU Journal of Mathematical Science, 2*(2), 43–54.
23. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B*, 58(1), 267–288.
24. Wooldridge, J. M. (2019). Introductory Econometrics: A Modern Approach (7th ed.). Boston: Cengage Learni
25. Young, T.; Hazarika, D.; Poria, S.; Cambria, E. Recent trends in deep learning based natural language processing. *IEEE Comput. Intell. Mag.* **2018**, *13*, 55–75.
26. Young, D.S., *Handbook of regression methods*, CRC Press, Boca Raton, FL, 2017, 109-136.
27. National Bureau of Statistics (2023). *Nigerian domestic and foreign debt (Q4 2022)*
28. Youssef, A. H., Mohamed, E. S., & Abdel Latif, S. H. (2023). *Handling multicollinearity using principal component analysis with the panel data model*. EUREKA: Physics and Engineering, 1, 177–188.