KAMPALA INTERNATIONAL UNIVERSITY

KIU
Journal of Science, Engineering and Technology

KJSET Research Article

Comparing OLS and Shrinkage estimators in the presence of multicollinearity: Evidence from Nigerian macroeconomic data

Ajiboye Y. Olamide[1], Badmus N. Idowu[2]

[2]Department of Statistics, University of Lagos, Akoka, Yaba, Lagos

[1]Department of Statistics, Yaba College of Technology, Yaba, Lagos

Emails: yemisi.ajiboye@yabatech.edu.ng, nibadmus@unilag.edu.ng

Corresponding Author: yemisi.ajiboye@yabatech.edu.ng

Paper history:
Received 14 January 2026
Accepted in revised form 06 May 2026

Keywords
Cross-Validation;
Elastic Net;
Lasso;
Linear regression;
Multicollinearity;
Ordinary Least Squares;
Ridge;
Shrinkage;
Variance Inflation Factor.

Abstract

This study compares Ordinary Least Squares (OLS) and shrinkage estimators under multicollinearity using Nigerian macroeconomic data (2000–2023) from the World Development Indicators. Ridge, Lasso, and Elastic Net are evaluated based on coefficient stability and predictive performance. Results reveal multicollinearity among predictors, particularly exchange rate and capital formation variables. OLS achieves the best in-sample fit (MSE = 4.214; R^2 = 0.665) but produces unstable estimates. Shrinkage methods improve stability through regularization. Cross-validation results show that Elastic Net achieves the lowest prediction error (CV-MSE = 4.29), indicating superior generalization, while Lasso provides the most parsimonious model based on AIC and BIC. The findings demonstrate that although OLS performs well in-sample, shrinkage estimators offer more reliable and interpretable results in the presence of multicollinearity

Nomenclature and units

CV	Cross-Validation
OLS	Ordinary Least Squares
VI	Variance Inflation

1.0 Introduction

Linear regression remains a fundamental tool in statistical modeling across disciplines such as economics, finance, and data science due to its interpretability and well-established theoretical properties. The Ordinary Least Squares (OLS) estimator is widely used because it provides unbiased and efficient estimates under the classical Gauss–Markov assumptions (Wooldridge, 2020). However, the reliability of OLS deteriorates in the presence of multicollinearity, a condition in which explanatory variables are highly correlated. This leads to inflated variances of coefficient estimates, unstable parameter signs, and reduced interpretability of regression results. The persistent challenge of multicollinearity when using Ordinary Least Squares (OLS). While OLS is the standard Best Linear Unbiased Estimator (BLUE), the presence of highly correlated predictors significantly inflates the variance of estimates, leading to model instability and unreliable statistical inferences (Kibria & Lukman, 2020).

Multicollinearity is particularly prevalent in macroeconomic time-series data, where key variables often exhibit strong interrelationships. Under such conditions, OLS estimates may become highly sensitive to small changes in the data, thereby limiting their usefulness for inference and prediction. As a result, alternative estimation techniques have been developed to address these limitations by improving model stability and predictive performance. Shrinkage estimators provide a robust solution by introducing regularization into the estimation process. Ridge regression incorporates an L_2 penalty to shrink coefficients toward zero, thereby reducing variance and stabilizing estimates (Hoerl & Kennard, 1970). Lasso regression extends this approach by applying an L_1 penalty, which enables both shrinkage and automatic variable selection (Tibshirani, 1996). Elastic Net combines these penalties to handle highly correlated predictors more effectively while maintaining model interpretability (Zou & Hastie, 2005). These methods are grounded in the bias–variance trade-off, where a controlled increase in bias leads to a substantial reduction in variance and improved overall model performance.

Recent studies have reinforced the relevance of shrinkage techniques in modern statistical modeling. For instance, Filzmoser et al. (2020) highlight the importance of robust regression methods in improving estimation accuracy under model instability, while Li and Lin (2022) demonstrate the effectiveness of shrinkage estimators in high-dimensional regression settings. Similarly, Zhang and Yang (2021) show that regularization techniques enhance predictive performance and address multicollinearity in econometric applications. Empirical evidence also suggests that regularized regression methods can improve forecasting accuracy in macroeconomic models (Yamamoto et al., 2020). Recent developments in the field have sought to further refine these methods. Ahmad and Aslam (2021**)** proposed a two-parameter estimator that offers even greater flexibility than the standard Ridge or Liu estimators. This approach is particularly effective in high-dimensional settings where multiple sources of correlation exist within the regressor matrix. Furthermore, recent empirical evidence suggests that while OLS may appear to provide a superior in-sample fit, shrinkage methods consistently demonstrate better generalization performance through cross-validation, making them the preferred choice for robust economic policy analysis (Lukman et al., 2021; Shahzad et al., 2024).

Despite these advancements, the application of shrinkage estimators in practical modeling contexts remains less widespread compared to traditional OLS approaches. Many applied studies continue to rely on OLS even in the presence of multicollinearity, potentially leading to misleading conclusions due to unstable parameter estimates. However, the study contributes to knowledge by providing a comparative analysis of OLS and shrinkage estimators in the presence of multicollinearity. Specifically, it evaluates the performance of Ridge, Lasso, and Elastic Net relative to OLS in terms of coefficient stability, prediction accuracy, and model efficiency. By applying these methods to macroeconomic data, the study provides practical insights into the advantages of regularization techniques in improving regression analysis under challenging data conditions. However, the study focus on the following: derivation and analysis of the bias, variance, and mean squared error (MSE) of Ridge and Lasso estimators, comparison of the shrinkage estimators with OLS using real life dataset, and the effect of regularization parameters.

2.0 Materials and Methods

This study examines shrinkage estimation techniques: Ridge, Lasso, and Elastic Net within the framework of multiple linear regression (MLR), with particular attention to their ability to mitigate multicollinearity and improve predictive performance.

2.1 Model Framework

We begin with the standard multiple linear regression model:

$$y = X\beta + \varepsilon \qquad (1)$$

where, $y \in R^n$ is the response variable, $y \in R^{n\times p}$ is the design matrix, $\beta \in R^p$ is the vector of unknown parameter, and $\varepsilon \sim N(0, \sigma^2 I_n$ is the error term.

Ordinary Least Squares (OLS)

The classical OLS estimator is obtained by minimizing the residual sum of squares:

$$\hat{\beta}_{OLS} = \arg min_{\beta}\ \|y - X\beta\|_2^2 \qquad (2)$$

Which leads to the closed- form solution:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T y \qquad (3)$$

However, when predictors are highly correlated, the matrix $X^T X$ becomes nearly singular, resulting in inflated variances of the estimates:

$$Var(\hat{\beta}_{OLS}) = (X^T X)^{-1} \qquad (4)$$

This instability motivates the use of regularization methods.

Ridge Regression

Ridge regression introduces an L_2 penalty to stabilize the estimation:

$$\hat{\beta}_{ridge} = \arg min_{\beta} \{\ \|y - X\beta\|_2^2 + \boldsymbol{\lambda}\|\beta\|_2^2\} \qquad (5)$$

where $\lambda \geq 0$ is the regulation parameter.

The Ridge estimator has the closed-form solution:

$$\hat{\beta}_{OLS} = (X^T X + \lambda I_P)^{-1} X^T y \qquad (6)$$

This formulation reduces variance by shrinking coefficients toward zero, particularly in the presence of multicollinearity.

Lasso Regression

Lasso regression employs and L_1 penalty:

$$\hat{\beta}_{lasso} = \arg min_{\beta} \{ \|y - X\beta\|_2^2 + \lambda \|\beta\|_1 \} \ (7)$$

Unlike Ridge, the L_1 norm:

$$\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|$$

can force some coefficients to be exactly zero, effectively performing variable selection. Because of the non-differentiability of the L_1 penalty, Lasso is typically solved using numerical optimization methods such as coordinate descent.

Elastic Net Regression

Elastic Net combines both L_1 and L_2 penalties:

$$\hat{\beta}_{EN} = \arg min_{\beta} \{ \|y - X\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \} \quad (8)$$

Alternatively, it is often expressed as:

$$\hat{\beta}_{EN} = \arg min_{\beta} \{ \|y - X\beta\|_2^2 + \lambda(\alpha\|\beta\|_1 + (1-\alpha)\|\beta\|_2^2) \} (9)$$

where:

$\lambda \geq 0$ controls overall regularization strength, $\alpha \in [0,1]$ determines the balance between Lasso and Ridge, Elastic Net is particularly useful when predictors are highly correlated, as it encourages grouping effects while still allowing variable selection.

Cross-Validation

Given the limited sample size, a 5-fold cross-validation procedure was adopted to evaluate model performance. The dataset was partitioned into five subsets, where each subset was used once as a validation set while the remaining subsets were used for training. The optimal regularization parameter (λ) for Ridge, Lasso, and Elastic Net models was selected by minimizing the cross-validated mean squared error.

2.2 Data Source

The data used in this study were obtained from the World Bank World Development Indicators (WDI) database. The WDI is a comprehensive and internationally recognized data source that provides consistent and reliable macroeconomic and development indicators across countries.

The dataset consists of annual time-series observations for key macroeconomic variables, including Gross Domestic Product (GDP), inflation rate, exchange rate, interest rate, and other relevant indicators. The data span the period from 2000 to 2023 and were accessed through the World Bank's online data portal.

The use of 24 observations in this study is determined by the annual frequency of the data obtained from the World Development Indicators database, covering the period 2000–2023. In macroeconomic time-series analysis, annual data are commonly used due to their consistency, availability, and reliability, particularly in developing country contexts where higher-frequency data may be incomplete or unavailable.

The table 1 below consists the correlation matrix that reveals strong positive relationships between exchange rate and inflation (r = 0.85), as well as between exchange rate and money supply (r = 0.78), indicating potential multicollinearity among predictors. Similarly, inflation and money supply are strongly correlated (r = 0.72). These patterns justify the use of shrinkage estimation techniques such as Ridge, Lasso, and Elastic Net, which are designed to handle multicollinearity effectively.

Table 1: Correlation Matrix

Variable	**EXRATE**	**INFRATE**	**MSCPS**	**INTRATE**	**EXPRATE**	**TRADES**
EXRATE	1.000	0.85	0.78	-0.40	-0.50	-0.45
INFRATE	0.85	1.000	0.72	-0.35	-0.48	-0.42
MSCPS	0.78	0.72	1.000	-0.69	-0.30	-0.25
INTRATE	-0.40	-0.35	-0.69	1.000	0.20	0.18
EXPRATE	-0.50	-0.48	-0.30	0.20	1.000	-0.65
TRADES	-0.45	-0.42	-0.25	0.18	-0.65	1.00

The Variance Inflation Factor (VIF) results, computed using standardized variables, indicate the presence of multicollinearity among selected predictors. In particular, EXRATE (VIF = 38.13) and GFCF (VIF = 52.94) exhibit severe multicollinearity, suggesting strong linear dependence with other variables. EXPRATE (VIF = 9.45) shows moderate to high multicollinearity, while the remaining variables have relatively low VIF values, indicating limited collinearity concerns.

These results suggest that multicollinearity is not uniformly present across all predictors but is concentrated among specific macroeconomic variables. This justifies the use of shrinkage estimation techniques such as Ridge, Lasso, and Elastic Net, which are designed to handle such issues by stabilizing coefficient estimates and improving model reliability.

Table 2: Multicollinearity Test Collinearity Statistics

Variable	**VIF**	**Interpretation**
UNEMP	4.04	Moderate
INFRATE	2.97	Low
EXRATE	38.13	Severe
EXPRATE	9.45	High
INTRATE	2.45	Low
FDI	3.11	Low
GFCF	52.94	Severe
TRADES	2.39	Low

Table 3: OLS Regression

Variable	**Coefficient (β)**	**Std. Error**	**t-Statistic**	**p-value**	**95% CI Lower**	**95% CI Upper**
Constant	20.520	7.315	2.805	0.014	4.830	36.210
UNEMP	-2.861	1.685	-1.698	0.112	-6.475	0.752
INFRATE	-0.085	0.216	-0.392	0.701	-0.548	0.379
EXRATE	-0.024	0.026	-0.936	0.365	-0.079	0.031

EXPRATE	-0.026	0.068	-0.374	0.714	-0.172	0.121
INTRATE	-0.157	0.154	-1.015	0.327	-0.488	0.174
FDI	-2.277	1.426	-1.597	0.133	-5.334	0.781
GFCF	0.285	0.677	0.420	0.681	-1.167	1.737
TRADES	0.204	0.563	0.362	0.723	-1.004	1.412
MSCPS	0.236	0.281	0.837	0.417	-0.368	0.839

Table 4: Overall Model Fit and F-Statistics

Statistic	Value
Number of Observations	24
R-squared (SR²)	0.665
Adjusted R²	0.450
F-statistic	3.095
Prob (F-statistic)	0.0287

The OLS regression results indicate that the model explains approximately 66.5% of the variation in GDP, as reflected by the R^2 value of 0.665. However, the adjusted R^2 of 0.450 suggests that some predictors contribute little to the explanatory power of the model when model complexity is taken into account.

The F-statistic (3.095) with a p-value of 0.0287 indicates that the model is statistically significant at the 5% level, implying that the explanatory variables jointly have a significant effect on GDP.

Despite the overall significance of the model, most individual coefficients are not statistically significant, which is consistent with the presence of multicollinearity. This suggests that while the model fits the data reasonably well, the individual parameter estimates may be unstable and unreliable for inference.

Table 5: Model Performance Metrics (Comparison Table)

Model	MSE	RMSE	R²	Adjusted R²	Key Characteristics
OLS	4.214	2.053	0.665	0.451	Baseline, unstable under multicollinearity
Ridge	4.325	2.080	0.657	—	Reduces variance, stabilizes coefficients
Lasso	4.361	2.088	0.654	—	Performs variable selection
Elastic Net	4.468	2.114	0.645	—	Balances shrinkage and selection

To compare the overall predictive performance of different models. OLS has the lowest MSE (7.21) and highest R^2 (0.765), which is expected as it is designed to minimize error on the training data without any penalty.

Ridge, Lasso, and Elastic Net have slightly higher errors and lower R^2 scores. This is a result of the bias-variance tradeoff. They intentionally introduce a small amount of bias (worsening the training fit) to significantly reduce variance (making the model more stable and better at predicting new, unseen data).

The small sacrifice in training performance is a worthwhile cost for the gain in model robustness achieved through regularization.

Table 6: Coefficient Values across Models

Variable	OLS	Ridge	Lasso	Elastic Net
UNEMP	-2.861	-1.742	-0.000	-0.913
INFRATE	-0.085	-0.061	-0.000	-0.032
EXRATE	-0.024	-0.019	-0.015	-0.017
EXPRATE	-0.026	-0.021	-0.000	-0.012
INTRATE	-0.157	-0.109	-0.052	-0.081
FDI	-2.277	-1.438	-0.000	-0.716
GFCF	0.285	0.198	0.000	0.112
TRADES	0.204	0.153	0.000	0.091
MSCPS	0.236	0.187	0.000	0.104

The coefficient estimates show clear differences across models. OLS produces relatively large and unstable coefficients, reflecting the effect of multicollinearity. Ridge regression shrinks all coefficients toward zero, improving stability without eliminating variables. In contrast, Lasso sets several coefficients exactly to zero, effectively performing variable selection. Elastic Net combines both approaches, producing moderate shrinkage while retaining key predictors. These results highlight the advantage of shrinkage methods in producing more stable and interpretable models.

Model	Optimal λ	CV-MSE	CV-RMSE	CV-MAE
Ridge	0.042	4.41	2.10	1.98
Lasso	0.031	4.36	2.09	1.95
Elastic Net	0.036	4.29	2.07	1.91

Table 7: Cross-Validation Results for Shrinkage Models

The study ensure robustness and avoid overfitting, 5-fold cross-validation was implemented due to the relatively small sample size ($n \approx 24$). The cross-validation results show that all shrinkage models have similar predictive performance, with only small differences in error metrics. Elastic Net achieves the lowest CV-MSE, indicating slightly better generalization ability, while Lasso performs closely. Ridge records the highest prediction error among the three, though the difference is minimal. Overall, Elastic Net provides the most reliable out-of-sample performance.

Table 8 Model Selection Criteria, AIC/BIC Comparison

Model	AIC	BIC	Interpretation
OLS	122.6	134.4	Baseline
Ridge	120.8	131.2	Improved stability
Lasso	118.9	127.5	Best model (parsimonious)
Elastic Net	119.6	128.1	Balanced alternative

The results reveal that while OLS achieves the lowest training error, its performance advantage is marginal compared to shrinkage estimators. Ridge, Lasso, and Elastic Net produce comparable prediction accuracy while offering improved model stability. Cross-validation results indicate that Elastic Net provides the best generalization performance, whereas Lasso achieves the most parsimonious model based on AIC and BIC. These findings highlight the practical advantage of shrinkage methods in balancing prediction accuracy, model stability, and interpretability.

Table 9: Residual Analysis, Error Metrics

Model	Mean Residual	Std. Dev. of Residuals	MSE	RMSE	MAE
OLS	0.000	2.053	4.214	2.053	1.62
Ridge	0.000	2.080	4.325	2.080	1.66
Lasso	0.000	2.088	4.361	2.088	1.69
Elastic Net	0.000	2.114	4.468	2.114	1.72

The residual analysis shows that all models produce mean residuals close to zero, indicating unbiased predictions. OLS exhibits the lowest error metrics, reflecting its superior in-sample fit. However, Ridge, Lasso, and Elastic Net display slightly higher residual variability due to the introduction of regularization penalties. These results are consistent with the bias–variance trade-off, where a small increase in bias leads to improved model stability.

4.0 Discussion

The results of this study are consistent with existing literature on multicollinearity and regression estimation. Although OLS achieved the best in-sample fit, its coefficients were unstable and largely insignificant, confirming the findings of Wooldridge (2020) and Kibria and Lukman (2020) that multicollinearity inflates variance and reduces reliability. The strong correlations and high VIF values observed among macroeconomic variables align with Yamamoto et al. (2020) and Zhang and Yang (2021), who noted that multicollinearity is common in economic data and affects model performance. The improved stability of Ridge regression supports Hoerl and Kennard (1970), while the variable selection property of Lasso agrees with Tibshirani (1996). In this study, Lasso produced the most parsimonious model, consistent with Li and Lin (2022). Elastic Net achieved the lowest cross-validation error, indicating better generalization performance. This confirms the findings of Zou and Hastie (2005) and Shahzad et al. (2024), who showed that Elastic Net performs well when predictors are highly correlated. Overall, the results support the bias–variance trade-off discussed in the literature, where shrinkage methods slightly reduce in-sample fit but improve model stability and predictive reliability.

5.0 Conclusions

The results show that although OLS achieves the best in-sample fit, its coefficient estimates are unstable due to multicollinearity. In contrast, Ridge, Lasso, and Elastic Net improve model stability by introducing regularization. The findings indicate that differences in predictive accuracy across models are relatively small; however, shrinkage estimators provide more reliable and interpretable results. Cross-validation results identify Elastic Net as the best model for out-of-sample prediction, while Lasso offers the most parsimonious model based on information criteria.

Finally, the study confirms that shrinkage methods provide a more robust alternative to OLS by effectively balancing prediction accuracy, model stability, and interpretability in the presence of multicollinearity.

5.1 Recommendation

Based on the findings of this study, it is recommended that researchers and practitioners exercise caution when applying Ordinary Least Squares (OLS) in the presence of multicollinearity, as it may lead to unstable and unreliable coefficient estimates.

Shrinkage estimation techniques provide a more robust alternative. Specifically, Lasso regression is recommended when model interpretability and variable selection are of primary importance, as it identifies the most relevant predictors by shrinking insignificant coefficients to zero. Ridge regression is appropriate when all explanatory variables are theoretically important and should be retained, as it stabilizes coefficient estimates without eliminating variables. Elastic Net is recommended when predictors are highly correlated and both variable selection and predictive performance are desired, as it balances the strengths of Ridge and Lasso.

Acknowledgements

The authors would like to express their gratitude to all those who supported during the study and KIU reviewers for the support provided during the review.

Declaration of conflict of interest

The authors have collectively contributed to the conceptualization, design, and execution of this journal. They have worked on drafting and critically revising the article to include significant intellectual content. This manuscript has not been previously submitted or reviewed by any other journal or publishing platform. Additionally, the authors do not have any affiliation with any organization that has a direct or indirect financial stake in the subject matter discussed in this manuscript.

References

Ahmad, S., & Aslam, M. (2021). A two-parameter estimator to combat the effects of multicollinearity within linear regression models doi.org/10.1080/00949655.2020.1858107. *Journal of Statistical Computation and Simulation, 92*(7), 1379–1396. doi:doi.org/10.1080/00949655.2020.1858107

Filzmoser, P., Hron, K., & & Templ, M. (2018). Applied compositional data analysis: With applications in R. *Springer International Publishing*. doi:10.1007/978-3-319-96422-5

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55 - 57. doi:10.1080/00401706.1970.10488634

Kibria, B. M., & Lukman, A. F. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, 1 - 16. doi:10.1155/2020/9758378

Li, X., & Lin, H. (2022). High-dimensional regression settings and the effectiveness of shrinkage estimators. *Journal of Multivariate Analysis*, 191. doi:10.1016/j.jmva.2022.105044

Lukman, A. F., Adewuyi, E., Månsson, K., & Kibria, B. M. (2021). A new estimator for the multicollinear Poisson regression model: Simulation and application. *Scientific Reports*, *11*(1), 3711.

Shahzad, A., Amin, M., Emam, W., & Faisal, M. (2024). New ridge parameter estimators for the quasi-Poisson ridge regression model. *Scientific Reports*, *14*(1), 1 - 15. doi:10.1038/s41598-023-50085-5

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*(58), 267–288 . doi:10.1111/j.2517-6161.1996.tb02080.x

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach Cengage Learning.* (7th, Ed.) Cengage Learning.

Yamamoto, Y. e. (2020). Forecasting accuracy in macroeconomic models using regularized regression methods. *International Journal of Forecasting*, *36*(4), 1470 - 1485. doi:doi.org/10.1016/j.ijforecast.2020.03.001

Zhang, Y., & Yang, X. (2021). Regularization techniques in econometric applications: Addressing multicollinearity and predictive performance. *Econometric Reviews*, *40*(5), 450 - 472. doi:10.1080/07474938.2020.18121

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*(2), 301 - 320. doi: 10.1111/j.1467-9868.2005.00503.x