KASU JOURNAL OF MATHEMATICAL SCIENCES (KJMS) VOL. 2, ISSUE 2, December 2021
ISSN 2734-3839 (Print), ISSN 2735-962X (E-Copy)
http://www.journal.kasu.edu.ng/index.php/kjms

KASU JOURNAL OF MATHEMATICAL SCIENCES

Multicollinearity Regularization Using Lasso and Ridge Regression on Economic Data

[*]Amusan Ajitoni Simeon[1], Adeshina Ibrahim Olaiya[2]
[1,2]Department of Statistics, School of Applied Sciences, Federal Polytechnic, P.M.B. 231, Ede, Nigeria.
Amusan.ajitoni@federalpolyede.edu.ng[1], Adeshinaibrahim10@gmail.com[2]

Abstract
One of the most important assumptions to consider before the adoption of a multiple regression model is the presence or otherwise of multicollinearity in the predictor variables. A Multicollinearity-free model will yield a reliable result as well as strike balance between the biasness of the estimator and the extent of variation in the predictive power of such model. This study adopts real sector data that are highly correlated, to predict the Gross Domestic Product (GDP) of Nigeria over a period of thirty-five years. The Variance Inflation Factor (VIF) from the Ordinary Least Square (OLS) regression shows that five variables out of nine predictor variables are highly correlated, the condition which renders the regression coefficient of the OLS unreliable. Ridge regression (L2) was adopted using a shrinkage value (λ) of 3.4 to penalize each of the regression coefficients. The Least Absolute Selection and Shrinkage Operation (LASSO) regression (L1) was employed to select the most significant coefficients to be included in the model. The best model from the LASSO regression indicates that industrial activities, construction, food index, population, and inflation positively affect gross domestic product of Nigeria while Electricity rate has a Negative impact on the GDP.

Keywords: Mulitcollinearity, Ordinary Least Square, Ridge regression, LASSO regression, Shrinkage parameter.

1. Introduction

One of the major assumptions of multiple linear regression is the absence of a condition termed as multicollinearity. Multicollinearity as a condition, exists when independent variables are highly correlated in a regression model resulting in the possibility of one variable being linearly predicted from the others with a substantial degree of accuracy. Mathematically expressed as:

$$\boldsymbol{X_{2i} = \lambda_0 + \lambda_1 X_{1i}}$$

In this situation, the coefficient estimates of the multiple regression may change erratically in response to small changes in the model or the data(Pereira et al., 2016). The adverse effect of Multicollinearity cannot be overemphasized as the estimated regression coefficients (b_1,b_2…….b_n) tend to have large sampling variability thereby making the standard errors large. Consequently, the conditions will appear that there is no linear relationship between the affected independent variables and the dependent variables however the inference will be wrong. Multicollinearity can also result in overfitting model (a condition of having a lower bias but a

higher variance) or underfitting (a condition of having a lower variance but a higher bias) which would tend to inflate the forecast from such fitted model or deflate it leading to a less adequate conclusion made from such model(Ahrens et al., 2019). Bias which indicates the deviation of an estimator from its targeted parameter is the difference between the true population parameter and the expected estimator. It is given as:

$$\boldsymbol{Bias}\left(\widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}}\right) = \left(\widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}}\right) - \boldsymbol{\beta}$$

Variance on the other hand is an error from sensitivity to small fluctuations in the training set. It measures the spread, or uncertainty, in these estimates. This is given as

$$\boldsymbol{Var}\left(\widehat{\boldsymbol{\beta}}_{\boldsymbol{OLS}}\right) = \boldsymbol{\sigma}^{\boldsymbol{2}}(\boldsymbol{X}'\boldsymbol{X})^{-\boldsymbol{1}}$$

Where the unknown error variance σ^2 is estimated from the residuals as

$$\widehat{\boldsymbol{\sigma}}^{\boldsymbol{2}} = \frac{\boldsymbol{e}'\boldsymbol{e}}{\boldsymbol{n}-\boldsymbol{m}}$$

$$\boldsymbol{e} = \boldsymbol{y} - \boldsymbol{X}\widehat{\boldsymbol{\beta}}$$

Common conditions that indicate the presence of multicollinearity include insignificant regression coefficients for the affected variables, large VIF value, significance of chi-square hypothesis from farrah-Glauber's test, higher correlation value in the correlation matrix, and data perturbation. This study adopted the use of VIF to detect the effect of multicollinearity. To correct the effect of multicollinearity, there are varieties of procedures ranging from the method of dropping one of the Multicollinear variables, obtain more data, using lagged variables, imposing regularization factor (shrinkage parameter) and so on. This study adopts the latter (regularization factor) procedure to adjust for the effect of multicollinearity in the analysis. One of the standard variable selection procedures in multiple linear regression is to use a penalization technique in least-squares (LS) analysis. In this setting, many different types of penalties have been introduced to achieve variable selection. It is well known that LS analysis is sensitive to outliers, and consequently outliers can present serious problems for the classical variable selection procedures. Regularization method has been identified as one of the effective ways of modelling modern data that are highly correlated with regard to the dependent variables, high-dimensional, sometimes noisy, and often contains a lot of unimportant predictors due to correlation extent (Lever et al., 2016). Regularization methods improve the predictive error of the model by reducing the variability in the estimates of regression coefficients by shrinking the estimates towards zero. A previously proposed a rank-based adaptive lasso-type penalised regression estimator and a corresponding variable selection procedure for linear regression models by (Turkmen & Ozturk, 2016), concluded that the proposed estimator and variable selection procedure are robust against outliers in both response and predictor space. A further of the novel 'deterministic Bayesian lasso' algorithm for computing the lasso solution by (Rajaratnam et al., 2016) was developed by considering a limiting version of the Bayesian lasso.

It was firstly observed that the Bayesian lasso improves as sparsity decreases and multicollinearity increases however in non-sparse and high multicollinearity settings the algorithm proposed can offer substantial increases in computational speed over co-ordinate wise algorithms. It was also concluded after a rigorous theoretical analysis that the deterministic Bayesian lasso algorithm converges to the lasso solution and it leads to a representation of the lasso estimator which shows how it achieves both L1 and L2 types of shrinkage simultaneously. In addition to the previously cited literatures, this paper adopted the ridge regression for correcting the effect of multicollinearity and lasso regression for variable selection to formulate an appropriate model which can be used for forecasting.

2. Material and Method

Using the real statistics toolpack (https://www.real-statistics.com/free-download/real-statistics-resource-pack/) in Excel 64-bit version, the analysis of this study was conducted on dataset that was extracted from the central bank of Nigeria (CBN) for a consecutive period of thirty-five years (1985-2019). The cross-sectional dataset comprises ten variables which include agriculture, industry, construction, trade, services, food index, population growth, inflation, electricity and total GDP. All the variables are continuous with the first nine variables considered as explanatory variables for the last variable (Total GDP). Ordinary least squares (OLS) regression produces regression coefficients that are unbiased estimators of the corresponding population coefficients with the least variance. However, there may be a model with less variance (that has smaller sum of square error, SSE), but at the cost of added bias. Considering the two scenario below:

Scenario 1: There are many independent variables, especially when there are more variables than observations.
Scenario 2: Data is close to multicollinearity, in which case small changes to X can result in large changes to the regression coefficients.
Here the OLS model over-fits the data (captures the data very well), but is not so good at forecasting based on new data. In these cases, Ridge and LASSO Regression can produce better models by reducing the variance at the expense of adding bias.

In this study, we adopted the second scenario where we have multicollinearity condition existing among the variables used that is the explanatory variables will be highly correlated. Ridge regression is the modifications of the least squares method with the used of biased estimators of the regression coefficients (Choi et al., 2020). Although, it has biased estimators, it only has a small biased substantially more precise than an unbiased estimator. Therefore the estimator is preferred since it often has a larger probability of being close to the true parameter value. Ridge regression estimator of the coefficient $\boldsymbol{\beta}$ is found by solving for $\mathbf{b}_{\boldsymbol{R}}$ in the equation

$$(\boldsymbol{X}'\boldsymbol{X} + \boldsymbol{\delta I})\boldsymbol{b}_{\boldsymbol{R}} = \boldsymbol{X}'\boldsymbol{y}$$

$\delta \geq 0$ is often referred to as a shrinkage parameter. Thus, the solution for ridge estimator is given by

$$\boldsymbol{b_R = (X'X + \delta I)^{-1} X'y}$$

Note that in previous explanation, we used $\boldsymbol{\lambda}$ to denote the shrinkage parameter. But in the documentation of Adnan (2006) he used δ therefore $\delta = \boldsymbol{\lambda = t}$ thus we can write the b_R as:

$$\boldsymbol{B = (X^T X + \lambda I)^{-1} X^T Y}$$

Recall that the covariance matrix of the OLS coefficients is expressed as

$$\boldsymbol{cov(B) = (X^T X + \lambda I)^{-1} \sigma^2}$$

Similarly the covariance matrix of ridge regression is given as:

$$\boldsymbol{cov(B) = \sigma^2 (X^T X + \lambda I)^{-1} X^T X (X^T X + \lambda I)^{-1}}$$

Here the λI term is considered to be the ridge (that is values added to the main diagonal of $X^T X$).

2.1: Variance inflation factors

The variance inflation factor (VIF) is a good indicator of extent of multicollinearity in the dataset. The higher the VIF from a regression, the severe the multicollinearity. The VIF for the Ridge regression coefficients is given by (Charles 2019):

$$\boldsymbol{VIF = (n - 1)(X^T X + \lambda I)^{-1} X^T X (X^T X + \lambda I)^{-1}}$$

This is equivalent to

$$\boldsymbol{VIF = (R + \lambda^* I)^{-1} R (R + \lambda^* I)^{-1}}$$

2.2: Estimating Ridge Regression Shrinkage Factor (λ)

A key aspect of Ridge regression is to find a good value for lambda. There are a number of approaches for estimating the parameter. One approach of estimating it is the use of ridge trace; were we plot the values of the coefficients for various values of lambda, with one plot for each coefficient. As we have noted previously, the higher the value of lambda the smaller the coefficient values. We are looking for the smallest value of lambda where the various coefficient plots stabilize. We want the smallest such value since this will introduce the least amount of bias. Another approach is to find a value of λ that ensures that all the VIF values are less than some designated value. As described in Collinearity, this value should be no larger than 10 (Christensen, 2018), although a value of one or less is desirable. There is also a procedure called k-fold cross-validation whereby the data is partitioned into k approximately equal-sized groups. Typically k = 5 or k = 10 is used. For any value of lambda and each value of j between 1 and k, we can calculate the Ridge regression coefficients based on the data in all the partitions except for the jth partition.

We then use these coefficients to forecast the y values of the data in the jth partition and calculate the residuals for each data element:

$$\boldsymbol{res(i,\lambda,j) = y_i - \sum_{j=1}^{k} b_j x_{ij}}$$

We now compute the error for the jth partition group as follows:

$$\boldsymbol{CVErr(\lambda,j) = \frac{1}{n}\sum_{n=1}^{k}[res(i,\lambda,j)]^2}$$

Finally, we calculate the cross validation (CV) error for the entire partition as:

$$\boldsymbol{CVErr(\lambda) = \frac{1}{k}\sum_{j=1}^{k} CVErr(\lambda,j)}$$

Our goal is to select the value of λ with the smallest CVErr (λ) value.

2.3: LASSO Regression

While Ridge regression addresses multicollinearity issues, it is not so easy to determine which variables should be retained in the model. These variables will converge to zero more slowly as lambda is increased, but they never get to zero (Darne & Charles, 2020).

LASSO, which stands for least absolute selection and shrinkage operator, addresses this issue since with this type of regression, some of the regression coefficients will be zero, indicating that the corresponding variables are not contributing to the model. This is the selection aspect of LASSO. Thus, LASSO performs both shrinkage (as for Ridge regression) but also variable selection. In fact, the larger the value of lambda, the more coefficients will be set to zero (Darne & Charles, 2020). For LASSO regression, we add a different factor to the ordinary least squares (OLS) SSE value as follows (Zhang et al., 2019):

$$\boldsymbol{SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 + \lambda\sum_{j=0}^{k}|b_j|}$$

There is no simple formula for the regression coefficients, similar to that of Ridge Regression Basic Concepts, for LASSO. Instead, we use the following iterative approach, known as cyclical coordinate descent. First, we note that this iterative approach can be used for OLS regression. This is based on the following property:

Property 1:

$$b_j = \frac{c_j}{z_j}$$

Given that:

$$z_j = \sum_{i=1}^{n} x_{ij}^2$$

And

$$c_j = \sum_{i=1}^{n} x_{ij}\left(y_i - \sum_{h \neq j} b_h x_{ih}\right)$$

In the cyclical coordinate descent algorithm, initially set all the b_j to some guess (e.g. zero) and then calculate b_0, b_1, …, b_k as described in Property 1. Continue to do this until convergence (i.e. the values don't change more than a predefined amount).
We use the same approach for LASSO, except that this time we use the following property.

Property 2:

$$b_j = \begin{cases} \frac{c_j + .5\lambda}{Z_j} & c_j > +.5\lambda \\ \frac{c_j - .5\lambda}{Z_j} & c_j > +.5\lambda \\ 0 & otherwise \end{cases}$$

Where c_j and z_j are defined as in Property 1.

3. Result and discussion

3.1: Descriptive analysis

Table 1: Statistics of explanatory variables

	Agriculture	**Industry**	**Construction**	**Trade**	**Services**
Mean	7623116.165	6822586.292	1154757.098	5965171.024	12180501.960
Standard Error	1488857.484	1386083.388	290227.741	1285435.717	2677441.913
Range	28155731.797	27823781.651	6024962.670	23384099.534	48672082.134
Maximum	28189968.887	27874862.561	6031060.770	23401870.634	48755108.664
Minimum	34237.090	51080.910	6098.100	17771.100	83026.530
	food index	**pop.grow**	**Inflation**	**Electricity**	**Total GDP**

Mean	85.255	2.577	19.348	111.748	30297561.112
Standard Error	4.879	0.012	3.028	4.788	6456949.222
Range	92.380	0.236	67.447	83.863	127544554.530
Maximum	127.670	2.681	72.836	158.354	127736827.800
Minimum	35.290	2.445	5.388	74.491	192273.270

Table 1 above shows the result of descriptive analysis on the model variables, it can be deduced that the average annual revenue generated by Nigeria government from the agricultural sector is 7623116.165 million naira with an associated standard error of 1488857.484.

3.2: Correlation analysis

Table 2: Correlation matrix of explanatory variables

	Agric	***Indus***	***Construct***	***Trade***	***Services***	***food index***	***pop.grow***	***Inflation***	***Electricity***
Agric	1.000								
Indus	0.981	1.000							
Construct	0.970	0.965	1.000						
Trade	0.991	0.965	0.976	1.000					
Services	0.991	0.969	0.980	0.999	1.000				
food index	0.883	0.852	0.789	0.856	0.850	1.000			
pop.grow	0.406	0.388	0.272	0.385	0.386	0.309	1.000		
Inflation	-0.351	-0.340	-0.280	-0.321	-0.319	-0.376	-0.292	1.000	
Electricity	0.905	0.897	0.819	0.874	0.873	0.833	0.581	-0.271	1.000

Table 2 above reveals the degree of association between explanatory variables. I can be observed from the table that variable like "service" has a very high positive correlation of about (0.991) with "Agric", (0.969) with "Industry", (0.980) with "construction", and (0.999) with "trade"; however it has a low positive degree of association of about 0.386 with "population growth" and a low negative correlation of about (-0.319) with "inflation".

3.3: Least square regression

Table 3: Regression coefficients

	Coeff	***std err***	***t stat***	***p-value***	***lower***	***upper***	***vif***
Intercept	-8.7E+08	1.81E+08	-4.7879	6.46E-05	-1.2E+09	-4.9E+08	
Agriculture (Total)	-5.7233	4.2174	-1.3570	0.1868	-14.4094	2.9627	262.6869
Industry (Total)	-1.6511	1.9014	-0.8683	0.3934	-5.5672	2.2649	46.2779
Construction	29.0174	15.2879	1.8980	0.0692	-2.4685	60.5034	131.1589
Trade	3.3083	9.6320	0.3434	0.7341	-16.5292	23.1459	1021.318

Services (Total)	-0.2673	4.881177	-0.0547	0.9567	-10.3204	9.7855	1137.925
food index	1026722	293254.6	3.5011	0.0017	422752.7	1630691	13.6392
pop.grow	3.27E+08	68440507	4.7741	6.69E-05	1.86E+08	4.68E+08	4.3545
Inflation	106125.7	158296.3	0.6704	0.5087	-219892	432143.1	1.5305
Electricity	-261034	291080.8	-0.8967	0.3783	-860526	338458.2	12.9417

Table 3 above reveals the outcome of least square regression analysis on the raw variables. The coefficients of regression (column 2) obtained from the least square regression is observed to be very high with a value of about 1026722 for *"food index",* and 3.27×10^8 for *"population growth"* and so on. These higher values obtained are noticed to be as a result of the higher VIF values (column 8) associated with each coefficient.

3.4 Ridge Regression analysis

Table 4: Ridge trace

Lambda	**0**	**0.0034**	**0.034**	**0.34**	**3.4**	**34**	**340**
Agriculture (Total)	-1.320	-1.261	-0.936	-0.198	0.090	0.106	0.052
Industry (Total)	-0.354	-0.359	-0.358	-0.234	-0.019	0.084	0.049
Construction	1.304	1.265	1.051	0.471	0.141	0.095	0.048
Trade	0.659	0.558	0.287	0.225	0.196	0.120	0.052
Services (Total)	-0.111	-0.010	0.243	0.278	0.197	0.119	0.052
food index	0.776	0.765	0.698	0.494	0.268	0.121	0.049
pop.grow	0.598	0.592	0.564	0.478	0.335	0.159	0.042
Inflation	0.050	0.051	0.058	0.071	0.037	-0.019	-0.018
Electricity	-0.194	-0.198	-0.228	-0.282	-0.100	0.083	0.048

Table 4 above indicates the of ridge regression associated with different value of λ (ridge penalty). For a ridge penalty of 0, the regression model is given by:

$$\boldsymbol{Gdp = -1.32(Agric) - 0.35(Industry) + 1.30(Construction) + 0.66(trade) - 0.11(service) + 0.78\,(food\ index) + 0.598(pop.growth) + 0.05(Inflation) - 0.194(Electricity)}$$

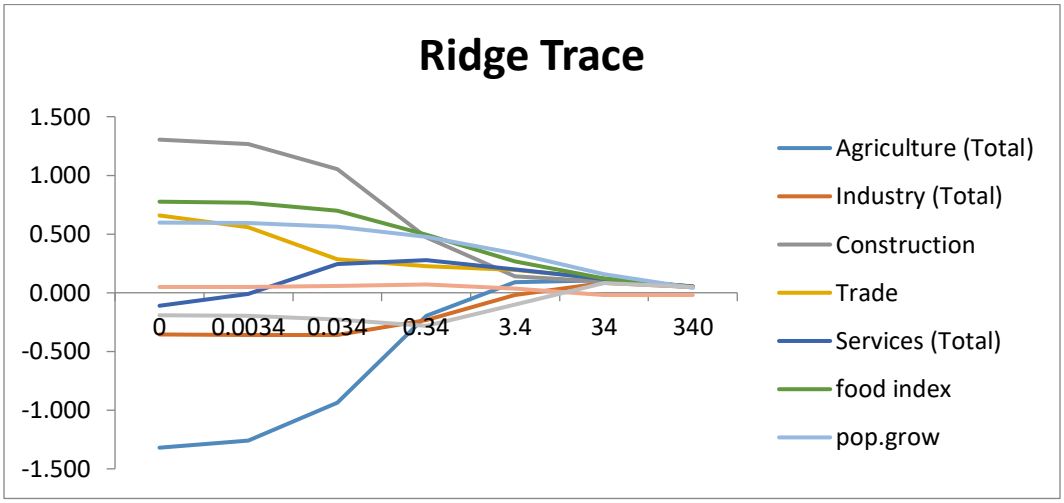

Figure 4.1: Ridge trace

Figure 4.1 above shows the plot of ridge trace presented in table 4, it can be observed that at a ridge penalty of 0.34, the regression coefficients starts to converge and more streamlined at a value of 3.4, 34 and 340.

Table 6: Ridge regression coefficient table

	coeff	***std err***	*t stat*	*p-value*	*lower*	*upper*	*vif*
Agriculture (Total)	0.090	0.044	2.046	0.051	0.000	0.180	0.332
Industry (Total)	-0.019	0.093	-0.204	0.840	-0.209	0.171	1.475
Construction	0.141	0.066	2.138	0.042	0.005	0.276	0.745
Trade	0.196	0.060	3.254	0.003	0.072	0.320	0.627
Services (Total)	0.197	0.053	3.691	0.001	0.087	0.307	0.493
food index	0.268	0.102	2.638	0.014	0.059	0.477	1.779
pop.grow	0.335	0.077	4.339	0.000	0.177	0.494	1.030
Inflation	0.037	0.074	0.501	0.621	-0.115	0.189	0.939
Electricity	-0.100	0.106	-0.941	0.355	-0.319	0.119	1.954

Table 6 above shows the result of ridge regression analysis at a ridge penalty (λ) of 3.4. The ridge penalty of 3.4 suffice us in correcting the multicollinearity deficiency among the explanatory variable as it yields a VIF values lesser than 10. However some of the coefficients (Agric, Industry, Inflation, and Electricity) are not statistically significant.

The ridge regression model at λ of 3.4 is given by:

$$\boldsymbol{Gdp = 0.14\ (constr) + 0.196\ (trade) + 0.197\ (serv) + 0.268\ (food) + 0.335\ (pop)}$$

3.5: LASSO Regression

Table 7: Lasso trace

Lambda	**0**	**0.0034**	**0.034**	**0.34**	**3.4**	**34**	**340**
Agriculture (Total)	-1.320	-1.286	-1.113	0.000	0.000	0.000	0.000
Industry (Total)	-0.354	-0.360	-0.352	-0.214	0.000	0.000	0.000
Construction	1.304	1.278	1.191	0.455	0.000	0.000	0.000
Trade	0.659	0.554	0.507	0.000	0.000	0.380	0.000
Services (Total)	-0.111	0.000	0.000	0.327	0.536	0.000	0.000
food index	0.776	0.770	0.736	0.474	0.220	0.000	0.000
pop.grow	0.598	0.594	0.582	0.472	0.277	0.000	0.000
Inflation	0.050	0.050	0.053	0.070	0.000	0.000	0.000
Electricity	-0.194	-0.195	-0.210	-0.290	0.000	0.000	0.000

Table 7 above shows the LASSO regression result for each shrinkage parameter which is presented graphically in figure 4.2 below. It can be deduced that at the penalty value (λ) of 3.4 all non-significant coefficients (Agric, Industry, Inflation, and Electricity) are completely eradicated in inclusion with one significant coefficient (Trade); however a penalty of 34 will render all the coefficients zero except trade; indicating that a penalty value of 3.4 suffice in correcting multicollinearity deficiency among the explanatory variables.

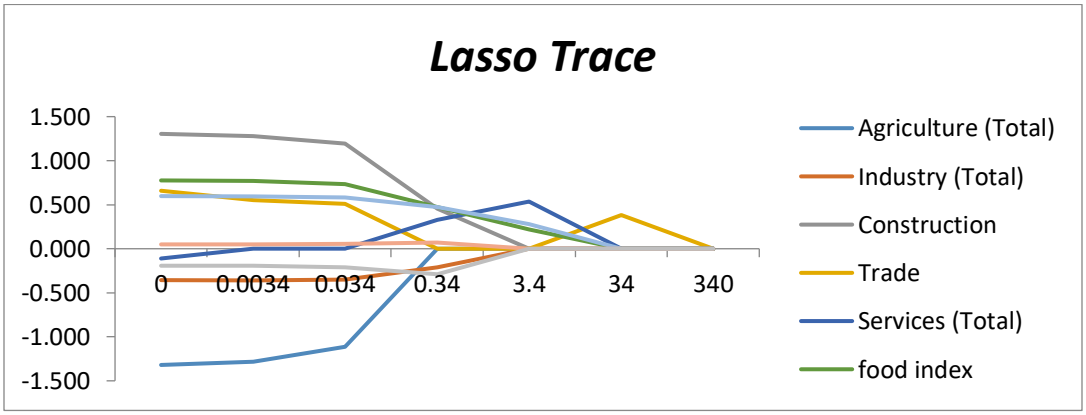

Figure 4.2

4. Discussion

Being a developed, developing or under-developed country is a factor of how a country can handle their diversified sector of social wellbeing which is a factor of economic activities, security, infrastructure, education, and standard of living. The variables included in the analysis of this study are substantial economic variables that are expected to contribute to GDP growth of a country. Agriculture as a sole factor has framed a central question in development economics for several decades (Chaudhary & Mishra, 2021). It is expected of any developed country to be able to cater for food index and spend largely in agriculture which will in turn improve the GDP by 14 to 19 percent within a short run of 5 years (McArthur & McCord, 2017); however opposite is the result of our findings in this study where we established that the agriculture of Nigeria is not significantly improving the GDP which might be due to mismanagement of funds allocated to the agricultural sector and misconduct in the sector.

As expected of a developed country like Malaysia, their industrial sector was documented to be contributing 25% of GDP growth and 60% of total export which indicates the significant impact of their industrial sector in GDP growth (Aziz & Azmi, 2017); however the case is again opposite in Nigeria wherein our model indicated that there is no significant impact of industrialization on Nigeria's GDP.

Another amazing sector of economy which improves the GDP significantly is the service sector which contributes 54 percent of Malaysia GDP (Aziz & Azmi, 2017). This sector's contribution was also established in our model to be significantly improving the GDP growth of Nigeria. It has a leading impact on Nigeria economy at the expense of industrialization.

The established model also suggest that population growth of Nigeria is a significant determinant of GDP growth contributing a whopping 33 per cent improvement to GDP for every 100 per cent increase in population. This similar scenario was equally noted in India where there is a correlation of 0.54 between the population growth and GDP which is subsequently associated with the 29% changes in the country's GDP (Goliuk, 2020).

In addition, the impact of inflation in improving the GDP is overwhelming especially in European countries where both inflation and GDP are sometimes considered as bidirectional. This significant impact was also established in our study where 100 percent increase in inflation rate will yield a small but significant 3.7% increase in GDP.

Lastly, it is expected that sufficiently large supply of electricity can ensure a higher level of economic growth; however the impact of Nigeria's electricity consumption on GDP as shown that enough has not being done by the government in this aspect to improve the economy as there is no significant impact of electricity on GDP of Nigeria.

5. Conclusion

Having considered the results from our study and compared them with other countries, we conclude that the government of Nigeria has more to adjust as regard some sectors that are lagging behind in significantly driving the economy of the country forward; sectors like Agriculture,

Industrialization, and Electricity are demanding attention so as to bring about a reform in the economy of the country as these sectors are the economic power of other countries. Although some sectors (Construction, Trade, Services, Food index, Population Growth) seems to be significantly driving the economy, however it is still recommended that Nigeria government should look into these sectors as report from our initial studies (least square model) stated that both services and trade are not significant leaving us with 5% probability (level of significance) that these two variables can actually not be significantly driving the economy forward in reality.

Reference

Ahrens, A., Hansen, C., & Schaffer, M. E. (2019). PDSLASSO & LASSOPACK: Stata module for post-selection and post-regularization OLS or IV estimation and inference.

Aziz, R., & Azmi, A. (2017). Factors affecting gross domestic product (GDP) growth in Malaysia. *International Journal of Real Estate Studies*, *11*(4), 61-67.

Chaudhary, K. K., & Mishra, A. K. (2021). Impact of Agriculture on Economic Development of Nepal using Statistical Model. *J Adv Res Alt Energ Env Eco*, *8*(2), 1-3.

Choi, N.-H., Shedden, K., Xu, G., Zhang, X., & Zhu, J. (2020). Comment: Ridge Regression, Ranking Variables and Improved Principal Component Regression. *Technometrics*, *62*(4), 451-455.

Christensen, R. (2018). Comment on "A note on collinearity diagnostics and centering" by Velilla (2018). *The American Statistician*, *72*(1), 114-117.

Darne, O., & Charles, A. (2020). Nowcasting GDP growth using data reduction methods: Evidence for the French economy. *Economics Bulletin*, *40*(3), 2431-2439.

Goliuk, V. (2020). THE EFFECT OF POPULATION DYNAMICS ON GDP GROWTH IN INDIA. *Економіка та держава*(4), 109-112.

Lever, J., Krzywinski, M., & Altman, N. (2016). Points of significance: Regularization. *Nature methods*, *13*(10), 803-805.

McArthur, J. W., & McCord, G. C. (2017). Fertilizing growth: Agricultural inputs and their effects in economic development. *Journal of development economics*, *127*, 133-152.

Pereira, J. M., Basto, M., & da Silva, A. F. (2016). The logistic lasso and ridge regression in predicting corporate failure. *Procedia Economics and Finance*, *39*, 634-641.

Rajaratnam, B., Roberts, S., Sparks, D., & Dalal, O. (2016). Lasso regression: estimation and shrinkage via the limit of Gibbs sampling. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, 153-174.

Turkmen, A., & Ozturk, O. (2016). Generalised rank regression estimator with standard error adjusted lasso. *Australian & New Zealand Journal of Statistics*, *58*(1), 121-135.

Zhang, R., Zhang, F., Chen, W., Xiong, Q., Chen, Z., Yao, H., Ge, J., Hu, Y., & Du, Y. (2019). A variable informative criterion based on weighted voting strategy combined with LASSO for variable selection in multivariate calibration. *Chemometrics and Intelligent Laboratory Systems*, *184*, 132-141.