MODELING ECONOMIC GROWTH FOR NIGERIA USING ROBUST STATISTICAL MODELS

UNIVERSITY OF KWAZULU-NATAL

INYUVESI YAKWAZULU-NATALI

Ayooluwade Ebiwonjumi

August, 2024

Modeling Economic Growth for Nigeria using Robust Statistical Models

by

Ayooluwade Ebiwonjumi

A thesis submitted to the

University of KwaZulu-Natal

in fulfilment of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

STATISTICS

Thesis Supervisor: Prof. Retius Chifurira

Thesis Co-Supervisor: Prof. Knowledge Chinhamu

UNIVERSITY OF KWAZULU-NATAL

SCHOOL OF MATHEMATICS, STATISTICS AND COMPUTER SCIENCE

WESTVILLE CAMPUS, DURBAN, SOUTH AFRICA

Declaration- Plagiarism

I, Ayooluwade Ebiwonjumi hereby declare that:

i. The research reported in this thesis, except where otherwise indicated, is my original research.

ii. This thesis has not been submitted for any degree or examination at any other university.

iii. This thesis does not contain other persons' writing unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:

(a) Their words have been re-written, but the general information attributed to them has been referenced.

(b) Where their exact words have been used, their writing has been placed inside quotation marks and referenced.

iv. This thesis does not contain text, graphics or tables copied and pasted from the internet, unless specifically acknowledged, and the source is detailed in their dissertation/thesis and the References section.

[redacted]
Ayooluwade Ebiwonjumi
(Student)

18/08/2024
Date

…………………………….
Prof. Retius Chinfurira
(Supervisor)

,,,,,,,,,,,,,,,,,,,,,,,,,,
Date

……………………………….
Prof. Knowledge Chinhamu
(Co-Supervisor)

……………….
Date

Disclaimer

This document describes work undertaken as a PhD programme of study at the University of KwaZulu-Natal (UKZN). All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the institution.

ACKNOWLEDGEMENTS

Indeed, a dream can become reality with vision, focus, determination, dedication and God's grace. This thesis could not be completed without the assistance of many people. It is difficult to mention all those who have assisted in the course of the research, but I will mention a few. Firstly, I give all glory and adoration to God Almighty as the Alpha and the Omega, the beginning and the ending for all things, notably, for the gift of life, wisdom, knowledge and understanding, and for his infinite mercies, protection, grace and provision.

I am super grateful to my thesis supervisors, Prof. R. Chifurira and Prof. K. Chinhamu for their academic guidance, cooperation, constructive criticism and encouragement in the course of the research. On many occasions, they read all the chapters meticulously with care and painstaking efforts to draw my attention to matters that needed correction in the drafts. Sincerely, all their efforts are highly appreciated and I pray God blessings upon them shall be abundant. I also appreciate the encouragement and advice of Prof. Z. Temesgen during their visit to Nigeria. I prayed that God will help him and his household. I also send heartfelt appreciation to Mr. J. Hammujuddy from University of KwaZulu Natal for English editing my thesis.

My appreciation to Prof. I. A. Ajibefun for making this dream possible the day he approved and issued my appointment letter as an Assistant Lecturer in the Department of Mathematical Sciences, Adekunle Ajasin University Akungba Akoko, Ondo State, Nigeria. I also appreciate Dr. and Dr. (Mrs.) Obalade, Dr. Amusa, Associate Prof. O. M. Olarewaju, Dr. and Mrs. Adeduro for all your assistant and support during the admission and registration process for this program. My prayer is that, none of you will ever lack help.

To my parents and siblings, I appreciate you all for your prayers and support all through the period of this program. I also thank all members of staff in the Department of Mathematical Sciences, Adekunle Ajasin University Akungba Akoko, Ondo State, Nigeria.

ABSTRACT

Economic growth is one of the most important goals of macroeconomic policy-making, but measuring it is not easy. This study aimed at measuring economic growth for Nigeria using robust statistical models. In this study, Gross Domestic Product (RGDP) is used as a proxy for economic growth and is modelled using selected predictors, namely internal debt (INDT), external debt (EXDT), interest rate (RINR), an exchange rate (REXR), and trade openness (OPEN). Quarterly RGDP index collected from the Central Bank of Nigeria for the period 1986 to 2022 was used in this study. Exploratory data analysis (EDA) revealed the linear relationship between the RGDP and the predictors. EDA also revealed the presence of multicollinearity and outliers in the predictors. In the presence of outlier and multicollinearity, this study utilizes the ridge regression, robust principal component regression, partial least square regression, average centered penalized regression, gaussian process regression and the coupler FMKL-GLD quantile regression. The performance and the efficiency of the adopted methods were evaluated using forecasting accuracy metrics, namely the root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). In using robust PCR, it can be asserted that the robust principal component regression (M-estimator) technique was an efficient and optimal technique for predicting RGDP. Specifically, PC1 and PC2 account for 35.39% and 22.15% in RGDP. In PLS, non-cross-validated and cross-validated PLS selection methods were used. Thus, 91.5% of the variance in economic growth drivers were explained by the five components generated and selected from the non-cross-validated PLS method. While, 72.6% of the variance in economic growth drivers were explained by the two components generated and selected from the cross-validated PLS method. Hence, after the cross-validation and extraction, the first and second components were efficient and optimally predicted 63.1% and 18.4% economic growth. In the average centered penalized regression model, the performance of the LASSO, ridge and elastic net techniques were compared. Using the least value of the forecasting metric values, the LASSO-average centered penalized regression was robust. The result of the best-performing average-centered penalized regression model indicated that INDT, RINR, REXR and OPEN positively contributed to the RGDP by 4.27%, 0.40%, 0.49% and 0.52% respectively. while, EXDT decreases RGDP by 0.97%. In the fitted gaussian process regression, the main effect of INDT, EXDT, RINR, REXR and OPEN for predicting RGDP were 38.30%, 12.20%, 1.10%, 2.00%, and 1.20% respectively which were increased after the independent re-sampling to 56.30%, 6.90%, 3.10%, 2.80%, and 2.10% for predicting RGDP. The estimated performance of FMKL-GLD quantile model techniques revealed that FMKL-GLD 50Q model was efficient for examining and predicting economic growth in Nigeria. Thus, INDT, RINR, REXR and OPEN positive contribution to RGDP were 17.94%, 29.42%, 7.99% and 145.10% respectively. Meanwhile, RGDP, as a result of EXDT was reduced by 3.92%. Therefore, government and policymakers should properly harness the benefit of trade openness to engender international patronage for economic growth. Also, coupler FMKL-GLD 50Q quantile regression technique and gaussian process regression method are the most efficient predictive statistical methods to deal with multicollinearity and outliers. However, Nigeria's economy had gone through various seasons, thus, a further study can be done to investigate structural breaks and propose appropriate model(s)

Keywords: economic growth, multicollinearity, outliers, robust principal component regression, FMKL-GLD quantile regression and Gaussian process regression method

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACF	Autocorrelation function
ADF	Augmented Dickey-Fuller
ARDL	Autoregressive distributed lag
CBN	Central Bank of Nigeria
EXDT	External debt
FMKL-GLD	Freimer–Mudholkar–Kollia–Lin- generalized lambda distribution
GLD	Generalized lambda distribution
GMM	Generalized method of moment
GPR	Gaussian process regression
INDT	Internal debt
KPSS	Kwiatkowski-Phillips-Schmidt-Shin
LASSO	Least absolute shrinkage and selection operator
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ML	Maximum likelihood
MSE	Mean square error
OLS	Ordinary least square
Open	Trade openness/ degree of economy openness
PACF	Partial autocorrelation
PLS	Partial least square
PCR	Principal component regression
REXR	Exchange rate
RGDP	Economic growth rate
RINR	Interest rate
RMSE	Root mean square error
VAR	Vector autoregressive
VECM	Vector error correction model
VIF	Variance inflation factor

RESEARCH OUTPUTS

All journals considered for the publication were journals recommended by the University of KwaZulu-Natal. The following papers have been published or undergoing peer review for publication, stemming from this thesis

1. **Ebiwonjumi, A**., Chifurira, R. and Chinhamu, K. (2022). An Efficient Estimation Technique for investigating Economic Growth and its Determinants for Nigeria in the Presence of Multicollinearity, *International Journal of Finance & Banking Studies,* 11(1), 107-119

2. **Ebiwonjumi, A.**, Chifurira, R. and Chinhamu, K. (2023). A Robust Principal Component Analysis for Estimating the Parameters of Economic Growth in Nigeria in the Presence of Multicollinearity and Outlier, *Journal of Statistics Application and Probability,* 12(2), 611-627.

3. **Ebiwonjumi, A.**, Chifurira, R. and Chinhamu, K. (2023). New Perspectives on Predicting Economic Growth in the Presence of Multicollinearity and Outlier, *Journal of Applied Mathematics and Information Sciences,* 17(4), 599-613.

4. **Ebiwonjumi, A.**, Chifurira, R. and Chinhamu, K. (2024). A FMKL-GLD Quantile Method for Predicting Economic Growth in Nigeria in the Presence of Multicollinearity and Outlier, *Statistics, Optimization and Information Computing,* 12, 1526-1542.

5. **Ebiwonjumi, A.**, Chifurira, R. and Chinhamu, K. (2024). Short-term Forecasts of Economic Growth for Nigeria using Average Centered Penalized Regression Method, *Research Sqaure,* pp.1-34. DOI:https://doi.org/10.21203/rs.3.rs-4623084/v1.

6. **Ebiwonjumi, A.**, Chifurira, R. and Chinhamu, K. (2024). Gaussian Process Regression Method for Exploring Economic Growth for Nigeria under the violation of Least Square Assumption, *International Journal of Science, Mathematics, and Technology Learning,* 30(1), 690-702.

CHAPTER 1

INTRODUCTION

1.1 Background to the Study

Economic growth is the sustained increase in a nation's per capital output or net national product over an extended period (Imimole and Imoughele, 2012, Ullah and Rauf, 2013). This concept can also be understood as a quantitative rise in the monetary value of goods and services produced by an economy in a given year. The most common metric for measuring economic growth is the percentage change in the gross domestic product (GDP) or gross national product (GNP) (Dwivedi, 2004). It is a critical macroeconomic objective that governments worldwide economies, including Nigeria. The main aim of achieving economic growth heavily relies on its oil and gas sector. However, since gaining independence in 1960, the country's overall economic performance has been weak and disappointing (Ebiwonjumi, *et al.* 2022). Despite these challenges, economic growth remains central to Nigeria's policy and development strategies. It is seen as a means to transform and restructure critical sectors of the economy. Sustainable economic growth requires a nation's capacity to invest and effectively utilize its resources (Nyoni and Bonga, 2017). Thus, governments must focus on policies supporting long term growth by fostering investment and promoting efficient use of available resources.

The growth of Nigeria economy as at 1990 was 8.2% and decreased to 5.4%, 4.6% and 3.5% in 2000, 2001 and 2002 respectively. This further increased to 9.6% in 2003 and decrease to 5.8% in 2005 and increased marginally to 6.4% and 7.3% in 2008 and 2011 respectively. The economic growth rate averaged 7% from 2012 to 2014, fell to 2.7% in 2015 and to -1.6% in 2016. This was rebounded to 0.8% in 2017, 1.9% in 2018, and then plateaued to 2.0% in the first half of 2019. In 2021, the economic growth rates stood at 1.1% due to the impact of pandemic in 2020. The projected growth rates for 2023 and 2024 were 2.8% and 3.0% respectively (International monetary fund, 2023). The Nigerian central bank tends to responsible for calculating and estimating the economic growth rate of the country. Investors tend to invest in economics which show upward and sustainable trajectory. Hence, the need to have accurate and reliable economic growth rate.

1.2 Literature Review

Onyeiwu (2012) explored how monetary policy affects Nigerian economic growth using ordinary least square (OLS) estimation technique. The findings showed that monetary policy stimulates gross domestic product. Thus, it emphasized that monetary policy can be used to create an investment-friendly environment and the money market should strive to provide financial instruments that meet the needs of increasingly numerous players. Adejoke, *et al.* (2013) investigated the effects of trade openness and the financial investment on the economic growth in Nigeria (1960-2011). The result obtained using ordinary least square (OLS) method revealed that a positive and statistical significance relationship exists among trade openness, foreign direct investment (FDI) and economic growth in the long run. Thus, it was stressed that a structural trade-oriented policies should be considered in enhancing economic growth via high flows of export in accruing more foreign proceed to boost growth. Kasidi and Said (2013) investigated the impact of external debt and economic of growth in Tanzania (1990-2010). The study revealed the significant impact of the external debt and debt service on GDP growth. Whereas total external debt stock has a positive effect and debt service payment has a negative effect on economic growth. Iya and Aminu (2013) examined the impact of debt burden on the Nigerian economy (1970-2007). Ordinary least square (OLS) was used to examine the relationship between debt burden and growth of the Nigeria economy. The result showed a negative relationship between debt stock of internal and external; and gross domestic product, meaning that an increase in debt stock will lead to a reduction in growth rate of Nigerian economy.

Da'silva and Ehinomen (2014) examined the relationship between economic openness and growth productivity in Nigeria (1970-2010). An OLS regression technique used for the study revealed that trade openness and the economic growth were positively and significantly related. Thus, it can be emphasized that the need for an appropriate use of export led revenue generation and economic diversification will catalyse economic. Hussain and Asghar (2014) empirically examined the causal connection among financial development, openness, and economic growth in developing countries for the period 1978 to 2012. The study employed a panel unit root test, panel cointegration test, and panel causality test, and Augmented Dickey Fuller as techniques of analysis. The finding of the study indicated strong evidence of long-run relationship between financial development and economic growth and a bi-directional causality between financial development and foreign direct investment. Moreever, trade openness had a positive and statistically significant impact in all countries. Sulaiman and Migiro (2014) investigated the nexus between growth of Nigerian economy and monetary policy. The study found that monetary policy supports economic growth, and the study also found that economic growth is unrelated to monetary policy. Thus, emphasized the need for regulatory framework to strengthen financial sector contribution to the efficiency of the monetary policies. Adigwe, *et al*. (2015) studied how monetary policies in Nigeria affect the country's economic growth using the ordinary least square technique. It was observed from the study that monetary policy promotes economic growth. The study emphasized the need for monetary policy to foster enabling investment environment.

Saaed (2015) investigated trade openness and financial development and their causal link to economic growth in Kuwait (1977 -2012). The econometrics techniques adopted on the variables such as gross domestic product, trade openness and financial development were cointegration and granger causality tests. In the result, it was discovered that both trade openness and financial development positively and significantly influenced the growth of the economy. Thus, the reformation of financial system leading to promotion of trade openness policy that can boost economic growth. Hamdan (2016) examined the impact of exports and imports on economic growth in 17 Arab Nations using annual data (1995-2013). The study adopted GDP as the dependent variable, while export and import were used as the independent variables. The results of the analysis showed that export and import had positive effect on economic growth. Nelson, *et al.* (2016) examined trade openness and exchange rate fluctuations in Nigeria with time series data spanning (1984-2013). In the study, an OLS method couple with the Augmented Dickey-Fuller (ADF) test for stationarity technique was adopted. Findings revealed that trade openness had positive impact on exchange rate fluctuations or volatility. The result also indicated a unidirectional causality between trade openness and exchange rate fluctuations. Nwannebuike, *et al.* (2016) investigation on external debt and economic growth in Nigeria using the econometric technique on the collected data revealed a positive relationship between external debt and economic growth in the short run but a negative in the long run. This was supported by Jilenga, *et al*. (2016) where Autoregressive distributed lag model (ARDL) and Bound tests were used for analyzing data sample from Tanzanians between 1971-2011. The results obtained revealed that a positive association exists between external debt and economic growth in the long run.

Okorontah and Odoemena (2016) investigated the effect of exchange rate fluctuation on economic growth in Nigeria using secondary data (1986-2012). Real gross domestic product as proxy for economic growth, while exchange rate, money supply and inflation were used as the independent variables. The time series data on the identified variables were sourced from CBN Statistical Bulletin. The OLS technique, Johansen cointegration test and error correction mechanism were used as the tools for data analysis and the results showed insignificant relationship between exchange rate and economic growth. Ukwueze and Okoronkwo (2016) examined the impact of trade openness on the economic growth in Nigeria (1961-2012). A vector autoregressive (VAR) method was employed for the study. Findings from the analysis revealed causality from exchange rate, trade openness, manufacturing output, and the size of public sector to output growth; trade openness, inflation and the manufacturing output had significant impact on exchange rate; the size of the public sector led to changes in inflation. The shocks as a result of inflation had little impact on growth output, the shocks caused by exchange rate had severe impact on growth output. Also, the shocks of trade openness significantly affected the growth output thus emphasizing that the significant impact of volume of international trade growth output. Afolabi, *et al.* (2017) empirically evaluated how international trade affects growth of Nigeria's economy using ordinary least square technique. From the study, it was revealed that government expenditure, interest rate, import and export were positively related with the economic growth while it was also observed that foreign direct investment and exchange rate negatively and significantly affected the growth of Nigeria's economy. Thus, it can be stressed that the country's trade should not be limited to primary and oil exports, but to the promotion of non-primary exports and non-oil exports.

Ajayi and Aluko (2017) evaluated how efficient monetary and fiscal policy are in Nigeria. The OLS estimation technique employed for the study revealed that export and money supply growth significantly stimulate economic growth. Also, the study found that monetary policy stimulates the growth more than fiscal policy. Ndubuisi (2017) focused on how externally acquired debt impacted the growth of the economy in Nigeria spanning 1985-2015. The data gathered were explored by least square regression technique, ADF stationary test, cointegration and error correction. Findings revealed the impact of servicing an externally incurred debt in growing the economy was negative and insignificant in Nigeria. Also, externally secured debt was found to be positive and significant in impacting the index of growing economy. The causality relationship between externally acquired debt and economic growth was unidirectional in the long-run. Al Kharusi and Ada (2018) in their study determined the relation of externally borrowed funds by the government have with economic growth between (1990-2015). The need for the study was spurred by an uninterrupted increase in externally borrowed funds basically to finance yearly budget. The statistical technique adopted for the analysis done was Autoregressive Distributed Lag cointegration method. Findings revealed a negative and significant effect of externally borrowed funds by the government in growing the economy. Furthermore, gross fixed capital impact in growing the economy was positive and significant as indicated during the period under investigation.

Akinmulegun and Falana (2018) assessed variability in exchange rate and output from the industries in Nigeria (1986-2015). A gross domestic product was used to capture growth of industrial output, while variability of exchange rate, inflation, interest rate, and net exports served as explanatory variables. Data extracted from the NBS and CBN Statistical Bulletin were analyzed using ADF and Philip Perron (PP) unit root, co-integration, granger causality and vector error correction model (VECM). Results showed that variability of exchange rate caused growth of industrial output. The authors stated further that the positive and significant effect of variability if exchange rate on the growth of the industrial output was more visible when compared to other variables. Thus, establishing the necessity for exchange rate capabilities to improve the growth of industrial output. Ufoeze *et al.* (2018) investigated exchange rate variability on growth of the economy in Nigeria using data spanning (1970-2012). In the study, dollar to naira rate, inflationary rate, money supply and oil revenue were used as the explanatory variables, while the response variable, gross domestic product was used to capture the growth of the economy. CBN Statistical Bulletin was the source of data and OLS regression technique was used for the analysis. Findings showed that to determine the growth of the economy uncontrolled exchange rate was better than fixed exchange rate. Moh'd AL-Tamimi and Jaradat (2019) analyzed the effect of debt incurred externally on the growth of the economy in Jordan from 2010 to 2017. The results obtained using descriptive and OLS analytic methods revealed that debt incurred externally was negatively and significantly affected the growth of the economy. Obayori *et al.* (2019) examined the impact of externally acquired debt on the growth of the economy in Nigeria spanning the period (1980-2016). The data were sourced from CBN bulletin and analyzed with the aid of generalized method of moment (GMM) technique. Findings showed the relationship between externally acquired debt and economic growth was positive and significant.

Oguntegbe and Alexander (2019) analyzed the contribution of lending rate and dollar to naira rate on economic attractiveness. Data collected was on lending rate, dollar to naira rate and gross domestic product the proxy for economic attractiveness (1981-2016). The OLS technique was adopted for the analysis and found that both lending rate and dollar to naira rate had impact on competitiveness of the economy and as such, the stability of dollar to naira rate need to be rigorously pursued by the government. Sunday and Ahmed (2019) investigated vigorous impact of degree of openness in relation to the growing economy in Nigeria considering the period 1980-2016. Secondary data sourced from the CBN Statistical Bulletin were explored. Analytic and diagnostic techniques adopted were: unit root, co-integration and error correction model. As a result, it was discovered trade openness had negative impact on growth of the economy. Thus, established the overriding effect of imports on exports in Nigeria and to address this, a concerted effort must be engaged by government to make diversification through export a priority in growing the economy. Afolabi, *et al.* (2020) researched on the impact of trade policy on the growth of Nigeria's economy using the autoregressive distributed lag (ARDL) technique. The study found that price-based variables and adjusted trade ratio positively influence gross domestic product in both long and short run. In the long run, dynamic responses showed that gross domestic product responded positively to trade policy. Therefore, there is need for policy makers to implement policies aimed at promoting international trade and innovations.

Ishola and Titiloye (2020) examined the effect of fiscal and monetary policies on growth of Nigeria economy using ARDL technique. The study found that supply of money vis-à-vis government spending cum revenue stimulates Nigerian economic growth. Hence, it can be stressed that government should allow expansionary monetary policy to stabilize economic growth. Lawali, *et al* (2020) sought to determine impact of the exchange rate on economic growth in Nigeria (1980-2019). Unit root, co-integration, and error correction model were statistical technique applied to explore the data extracted from CBN statistical bulletin. Findings, established a positive and significant impact of exchange rate on economic growth. Furthermore, the impact of degree of economy openness on economic growth was revealed to be negative. In view of this, exchange rate stability must be ensured through the monetary policies implementation. Export led diversification must be pursued and sustained to enhance economic growth. Ekor, *et al.* (2021) investigated whether external debt impair economic growth in Nigeria. The study was empirically carried out to assess the impact of foreign debt on the Nigerian economy. A dynamic ARDL model was adopted and the result indicated that in the long run, external debt accumulation and the associated debt servicing negatively affected the economy. In view of this, government must ensure aim for external debt acquisition is for infrastructure development which must be favourable and sustainable.

Ring, *et al.* (2021) investigated the impact of external debt on economic growth: The impact of institutional quality on economic growth was examined using panel generalized method of moment (panel GMM) least square analysis. A sample of twenty-three countries between 2011-2014 were used to examines the nexus between external debt and economic growth with institutional quality as a moderator. Findings revealed that institutional quality was a good moderator of the existing relationship between external debt and economic growth. It was further revealed that good governance practices have significant effect on economic growth. Thus, a good debt management and feasible policy must be prescribed as the keys to control external debt. Yusuf and Mohd (2021) evaluated debt incurred by the government and its effect on economic growth in Nigeria (1980-2018). Empirical analysis carried out were done using ARDL technique and findings showed that debt incurred externally negatively affected the growing of the economy thus, hampered the long-term growth despite the fact that in the short-term growth-enhancing was observed. Also, debt incurred internally inversely affected the growth of the economy in the short and long-term in comparison with externally acquired debt. Moreso, the servicing of acquired debt hampered the economy growing with a resultant consequence of debt overhang.

Abiodun and Uche (2023) examined the relationship between domestic public debt and economic growth in Nigeria based on type of regime between (1981-2019). An ARDL method was adopted as the econometric technique for the analysis of the data collected for the study. Findings showed that domestic public debt and economic growth were positively related under the military. It was further showed that in the long run, the two variables were negatively and insignificantly related during the civilian government. Also, the results revealed that budget deficit was positively and insignificantly related with economic growth. The results also showed that prime lending rate had positive and significant effect on the economic growth. Thus, the need for prudent and judicious used of the borrowed fund for sustainable public debt. Chika, *et al.* (2023) investigated the impact of rising external debt on the exchange rate in Nigeria spanning the period between (1980-2021). The secondary data used were collected on external debt, government spending, inflation rate and exchange rate volatility which were sourced from CBN statistical bulletin, DMO, and WDI. The ADF test and ARDL technique employed for the analysis showed that external debt had a negative and insignificant relationship with the exchange rate. Hence, the need for government and other policy makers to ensure that all incurred debt were channeled on feasible projects that will yield returns to serve and pay up the debt on or before maturity.

Luana, *et al.* (2023) examined the effect of governance quality on future economic growth of the emerging market and developed economies. Five emerging markets such as Brazil, Russia, India, China, and South Africa (BRICS) countries and three developed economies that include United States, Germany and Japan were considered. Other macroeconomic variables used in the study were government debt, external debt, current account balance, trade balance, budget balance, foreign exchange rate and short-term interest rate. A panel least square regression method was employed for the analysis of the data collected on the identified variables between (1996-2018). The results revealed the positive and significant impact of regulatory quality on economic growth. It was also revealed that rule of law had negative but insignificant impact on economic growth. Thus, emphasized the important of sound regulatory environment in stimulating economic growth. Boakye and Atuilik (2024) investigated the relationship between Ghana's external debt level, debt servicing and economic growth in developing economy. In the study, the time series data were gathered on the variables used between (1975-2021). The ADF test, PP test and Autoregressive Distributed Lag (ARDL) technique were employed for the analysis. Thus, from the results it was revealed that external debt stock had negative impact on economic growth. Also, the relationship between debt service and economic growth were positive and statistically insignificant. Hence, it can be emphasized the need for proper management of the external debt in enhancing the economy of the developing economy.

Kabemba and Kabwe (2024) evaluated the effect of public debt on economic growth in Zambia from (2011-2021). The variables used were prime lending rate, exchange rate, external debt stock, domestic debt stock and real gross domestic product. An autoregressive distributed lags model (ARDL) was used as the econometric technique for the analysis of the data collected for the study. Findings showed that public debt and other macroeconomic variables identified for the study had a significant impact on economic growth. Thus, it was stressed that policy makers need to ensure that the acquire debt should be for consumption rather it must be for productive purpose and as such enhance economic growth. Mokuolu, *et al*. (2024) examined the impact of public debt on economic growth in Nigeria between (1981-2021). In the study, an Autoregressive Distributed Lag (ARDL) model was used to investigate the impact of total domestic debt, total external debt, investment and government expenditure on economic development in Nigeria. The results revealed that all identified variables except government expenditure had an insignificant impact on economic growth in the long-run. It was further revealed that total domestic debt, government expenditure and inflation rate were positively related with economic growth while, total external debt, investment and gross domestic savings had negative effect on economic growth in the long-run. Thus, it can be stated that policy makers need to put in place suitable measures for the management of domestic debts and government should ensure that incurred national debts are properly spent to attract investment to the country.

Nyeche (2024) investigated the effect of exchange rate dynamics on economic growth in Nigeria spanning the period between (1985-2021). The secondary data gathered on real GDP, exchange rate, trade openness and external reserves and sources from the WDI and CBN Statistical Bulletin were used for the study. The unit root tests, cointegration and ARDL techniques were employed for the analysis. Findings revealed a long-run relationship between economic growth, exchange rate, trade openness and external reserves. Also, the ARDL revealed that the exchange rate and external reserves had positive and significant impact on real GDP while, trade openness had a positive and insignificant impact on real GDP. Thus, it was posited that government should develop policies that can stabilize exchange rate and the opening up of the economy to international trade with strategic agreements to protect domestic industries as well as promoting economic sectors with competitive advantage. Table 1.1 summarizes the related literature discussed in this section.

Table 1.1: Summary of Empirical Related Literature Reviews

Authors/year	Title	Method/Model	Findings	Criticism/Gap
Onyeiwu (2012)	Explored how monetary policy affects Nigerian economic growth.	OLS estimation technique	monetary policy (interest and exchange rate) stimulates GDP	Preliminary diagnostic tests were not carried out
Adejoke, *et al.* (2013)	Investigated the effects of trade openness and the financial investment on the GDP in Nigeria	Ordinary least square (OLS) method	Trade openness and FDI had positive and significant impact on GDP	The independency of the variables and influential values were not examined
Iya and Aminu (2013)	Examined the impact of debt burden on the Nigerian economy 1970-2007	Ordinary least square technique	Negative impact debt stock of internal and external on GDP	OLS assumptions not investigated
Kasidi and Said (2013)	investigated the impact of external debt on GDP in Tanzania (1990-2010.)	Least square method	external debt and debt service had significant impact on GDP growth	OLS assumptions not investigated
Da'silva and Ehinomen (2014)	Examined the relationship between economic openness and GDP growth in Nigeria (1970-2010)	Ordinary least square (OLS) regression technique	Trade openness and the GDP growth were positively and significantly related	Outliers and influential values were not examined
Hussain and Asghar (2014)	Empirically examined the causal connection among financial development, openness and GDP in developing countries	Panel unit root test, panel cointegration test, and panel causality test	A bi-directional causality between financial development and FDI. Also, trade openness had positive impact in all countries.	The independency of the variables and influential values were not examined
Sulaiman and Migiro (2014)	Investigated the nexus between growth of Nigerian economy and monetary policy	Granger Causality	Monetary policy granger caused GDP.	OLS assumptions not investigated
Adigwe, *et al.* (2015)	Studied how monetary policies in Nigeria affect the country's economic growth.	Ordinary least square technique	Interest and exchange rate had positive impact on economic growth	Basic OLS assumptions not investigated

Saaed (2015)	Investigated trade openness and financial development and their causal link to GDP in Kuwait.	Cointegration and granger causality tests	Trade openness and financial development significantly influenced the GDP	OLS and its assumptions were not investigated in the study
Hamdan (2016)	Examined the impact of exports and imports on GDP in 17 Arab Nations (1995-2013)	Ordinary least square estimation technique	Export and import had positive effect on GDP.	OLS and its assumptions not investigated
Nelson, *et al.* (2016)	Examined trade openness and exchange rate fluctuations in Nigeria (1984-2013).	Ordinary Least Square (OLS) method and Augmented Dickey-Fuller (ADF) test	Trade openness had positive impact on exchange rate. Unidirectional causal relation also exists between the two variables	Multicollinearity and outliers in the data set were not examined
Nwannebuike, *et al.* (2016)	investigated external debt and GDP in Nigeria	Ordinary least square technique	external debt and GDP were negatively related.	Failed to examine OLS assumptions
Okorontah and Odoemena (2016)	Investigated the effect of exchange rate fluctuation on economic growth in Nigeria (1986-2012)	OLS technique, Johansen cointegration test and ECM	Insignificant relationship between exchange rate and GDP growth	Basic OLS assumptions not investigated
Ukwueze and Okoronkwo (2016)	Examined the impact of trade openness on the GDP in Nigeria (1961-2012)	Vector autoregressive (VAR) method	Causality from exchange rate, trade openness, GDP	Multicollinearity and outliers were not examined
Afolabi, *et al.* (2017)	Empirically evaluated how international trade affects GDP growth in Nigeria	Ordinary least square technique	It was revealed that interest rate, import and export positively related with the GDP	OLS assumptions were not investigated in the study
Ajayi and Aluko (2017)	Evaluated how efficient monetary and fiscal policy are in Nigeria.	Ordinary least square (OLS) estimation technique	Export and money supply significantly stimulate GDP	Multicollinearity and outliers were not examined
Ndubuisi (2017)	Focused on how externally acquired debt impacted the GDP in Nigeria 1985-2015	OLS regression, ADF test, cointegration and ECM	External debt had positive impact on GDP	Outliers and influential values were not examined

Al Kharusi and Ada (2018)	Determined the relation between externally borrowed funds and RGDP (1990-2015)	(ARDL) and cointegration method	Externally borrowing was negative and significantly affected RGDP	OLS assumptions not considered
Akinmulegun and Falana (2018)	Assessed variability in exchange rate and output from the industries in Nigeria (1986-2015).	ADF and PP unit root, co-integration, granger causality and VECM.	Exchange rate variability had positive and significant effect on the growth of industrial output	Failed to consider OLS assumptions
Ufoeze, *et al.* (2018)	Investigated exchange rate variability on growth of the economy in Nigeria (1970-2012).	Linear regression technique	To determine GDP growth floating exchange rate was better than fixed exchange rate.	The independency of the other explanatory variables not considered
Moh'd AL-Tamimi and Jaradat (2019)	Analyzed the effect of debt incurred externally on GDP in Jordan (2010-2017).	Descriptive and least square regression method	debt incurred externally was negatively and significantly affected GDP	Failed to examine OLS assumptions
Obayori, *et al.* (2019)	Examined the impact of external debt on the GDP in Nigeria (1980-2016)	Generalized Method of Moments technique	External debt and GDP were positively and significantly related	OLS assumptions were not examined
Oguntegbe and Alexander (2019)	Analyzed the contribution of lending rate and dollar to naira rate on economic attractiveness (1981-2016).	OLS technique	Lending rate and exchange rate had impact on competitiveness of the economy	Not considering the assumptions of OLS
Sunday and Ahmed (2019)	Investigated degree of openness and the GDP in Nigeria (1980-2016).	Unit root, co-integration and error correction model	Trade openness had negative impact on GDP	OLS assumptions were not investigated
Afolabi, *et al.* (2020)	Investigated on the impact of trade policy on the GDP growth in Nigeria	Autoregressive distributive lagged (ARDL) technique	GDP responded positively to trade policy.	OLS and its assumptions were not investigated
Ishola and Titiloye (2020)	Examined the effect of fiscal and monetary policies on GDP.	Autoregressive distributive lagged (ARDL) technique.	Money supply and government spending stimulates GDP.	OLS and its assumptions were not investigated

Lawali, *et al.* (2020)	Determined the impact of exchange rate on economic growth in Nigeria (1980-2019).	Unit root, co-integration, error correction model (ECM) and OLS	Exchange rate hade positive and impact on GDP. Economy openness had negative impact on GDP.	OLS assumptions not investigated
Ekor, *et al.* (2021)	Investigated whether does external debt impair GDP in Nigeria.	A dynamic auto-regressive distributed lag model	External debt accumulation affected the GDP	OLS assumptions were not investigated
Ring, *et al.* (2021)	Investigated the impact of external debt on economic growth	A panel GMM least square method	External debt had significant effect on GDP	OLS assumptions not investigated
Yusuf and Mohd (2021)	Evaluated debt incurred by the government and its effect on GDP in Nigeria	Autoregressive Distributed Lag (ARDL) technique	Debt incurred externally negatively affected GDP growth	OLS assumptions not investigated
Abiodun and Uche (2023)	Examined the relationship between domestic public debt and economic growth in Nigeria based on type of regime	Autoregressive Distributed Lag (ARDL) technique	domestic public debt and economic growth were positively related under the military	Not considering the assumptions of OLS
Chika, *et al* (2023)	Investigated the impact of rising external debt on the exchange rate in Nigeria	The ADF test and ARDL technique	external debt had a negative and insignificant relationship with the exchange rate.	OLS assumptions not investigated
Luana, *et al* (2023)	examined the effect of governance quality on future economic growth of the emerging market and developed economies	Panel least square regression method was employed for the study	Regulatory quality had positive and significant impact on economic growth.	Multicollinearity and outliers in the data set were not examined
Boakye and Atuilik (2024)	investigated the relationship between Ghana's external debt, debt servicing and economic growth	The study employed ADF test, PP test and Autoregressive Distributed Lag (ARDL) technique	external debt stock had negative impact on GDP. Also, the relationship between debt service and GDP were positive and insignificant.	Failed to examine OLS assumptions

Kabemba and Kabwe (2024)	Evaluated the effect of public debt on economic growth in Zambia from (2011-2021)	An ARDL was used as the econometric technique for the study	public debt and other variables identified had significant impact on GDP	Basic OLS assumptions not investigated
Mokuolu, *et al*. (2024)	Examined the impact of public debt on GDP in Nigeria between (1981-2021)	Autoregressive Distributed Lag (ARDL) model was used for the study	Total domestic debt, government expenditure and inflation rate were positively related with GDP	Multicollinearity and outliers were not examined
Nyeche (2024)	Investigated the effect of exchange rate dynamics on GDP in Nigeria (1985-2021)	The ADF tests, cointegration and ARDL techniques were employed for the study	Exchange rate and external reserves had positive and significant impact on real GDP	OLS assumptions were not investigated in the study

In summary, most of the studies reviewed using multiple variables for the prediction of economic growth for Nigeria failed to account for correlations between the variables and the variability of economic growth data. This study aims to predict economic growth for Nigeria, accounting for multicollinearity and outliers among the economic growth predictors. To the best of our knowledge there are limited studies on Nigerian economic growth taking into account multicollinearity and outliers.

However, in this study, spurious model and estimation of parameters would be avoided in order to obtain optimal and efficient estimate for valid and reliable prediction. Thus, various robust statistical methods that include ridge regression, robust principal component regression, partial least square method, average centered penalized regression technique, gaussian process, a machine learning approach and Coupler Freimer–Mudholkar–Kollia–Lin- generalized lambda distribution (FMKL-GLD) quantile regression methods were explored to estimate and predict economic growth.

1.3 Statement of the Problem

The presence of multicollinearity among economic drivers and outliers posed a real challenge in estimating and predicting economic growth rates. Accurate and reliable growth rates are important to investors, scholars and economists since the show the performance of the economy. A reliable and good model(s) is important in producing accurate economic growth rates. Thus, this study aims to propose models that predict economic growth rates accurately in the presence of multicollinearity and outliers.

1.4 Aim and Objectives of the Study

The main aim of this study is to estimate and predict real gross domestic product (RGDP) which is used as a proxy for economic growth with internal debt, external debt, interest rate, exchange rate and degree of economy openness as economic growth drivers in Nigeria using robust statistical methods. This is achieved by:

- Modelling RGDP using a ridge regression model;
- modelling RGDP using the robust principal component covariate regression model;
- modelling RGDP using a partial least square regression model;
- modelling RGDP using average centred penalized regression model;
- proposing a gaussian process regression model for predicting RGDP;
- proposing a coupler FMKL-GLD quantile regression model for RGDP and compare its relative performance with quantile regression model.

1.5 Significance of the Study

The results of this study are important to policy maker in the area of policy formulation and implementation involving the performance of the economy and also for investors interested in investing into the economy. The accurate and reliable prediction of the economic growth rates assist in managing the investment risk. A suitable and sustainable upward trajectory of economic growth rates makes investments profitable to the investors and enhance the economic outlook and performance aims of the policy makers. Thus, findings of this study are beneficial to any person interested in investing in the economy, as well as economists and researcher scholar in business and academic.

1.6 Scientific Contributions of the Study

The key contributions of this study are the used of robust statistical techniques in modelling RGDP for Nigeria. The contributions are as follows:

- Application of ridge regression in modelling quarterly RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as economic growth drivers.
- Application of robust principal component regression in modelling quarterly RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as drivers.
- Application of partial least square regression in modelling quarterly RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as drivers.
- Application of average centred penalized regression in modelling quarterly RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as economic growth drivers.
- Proposing and applying gaussian process regression model for RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as covariates.
- Proposing coupler FMKL-GLD and quantile regression and their application on modelling RGDP using internal debt, external debt, interest rate, exchange rate and trade openness as economic drivers.

1.7 Thesis Structure

The structure of the remaining chapters in this thesis is outlined as follows.

- **Chapter 2** presents the methodology used in this research, it describes the steps taken to collect and analyze the data, providing the foundation for the analyses conducted in the subsequent chapters.
- **Chapter 3** is dedicated to exploratory data analysis. This chapter discusses various statistical techniques employed to examine the data, including descriptive statistics analysis, correlation analysis, stationarity tests, granger causality tests, structural break tests, multicollinearity tests, outlier tests, normality tests, and volatility tests.
- **Chapter 4** focusses on modelling and estimating real gross domestic product (RGDP) in the presence of multicollinearity and outliers. Ridge regression, a robust approach for handling multicollinearity, is used generate accurate parameter estimates.
- **Chapter 5** continue the exploration of modelling RGDP, this time using robust principal component regression. This technique is designed to reduce the impact of multicollinearity and outliers while maintaining predictive accuracy.

- **Chapter 6**, employs partial least squares regression to model RGDP, providing an additional perspective on handling multicollinearity and outliers. The method balances model complexity with predictive power.
- **Chapter 7** explores a combination of penalized regression techniques, including LASSO, ridge and elastic net, to model and predict RGDP in in the presence of multicollinearity and outliers. These approaches offer flexibility and control over model complexity.
- **Chapter 8** introduces machine learning approach, Gaussian process regression, to predict RGDP. This chapter demonstrates the application of advanced computational methods to overcome the challenges posed by multicollinearity and outliers.
- **Chapter 9** presents a proposed quantile regression method, incorporating the FMKL-GLD distribution, to model RGDP while accounting for multicollinearity and outliers. This approach provides a deeper understanding of the distribution and relationships within the data.
- **Chapter 10** concludes the thesis with summary of the findings, insight gained, and recommendations for future research. This chapter also outlines potential implications for policymakers and further areas of study.

Together, these chapters offer a comprehensive exploration of modelling economic growth rates under challenging conditions, providing a range of techniques to address multicollinearity and outliers.

CHAPTER 2
METHODOLOGY

2.1 Introduction

In this chapter the statistical tests used in this study are discussed. These include tests used in data exploratory analysis and model diagnostic evaluation techniques.

2.2 Testing for Normality

To check for normality of a response variable, the Q-Q plot, the Jarque-Bera and Shapiro-Wilk tests are used and a description of these tests follow.

2.2.1 The Quantile-Quantile (Q-Q) plot

The Q-Q plot is a graphical technique used to assess the validity of the theoretical distributional assumption of a response variable, for example, a normal distribution or an exponential distribution (Velez, *et al*., 2015). Generally, the main idea is to calculate the theoretically expected values for each data point based on the distribution in question. If the data points fall approximately on a straight line, the data will consequently follow the assumed distribution. This approach is merely a visual check and, as such, subjective, but it enables one to understand whether the assumed distribution is plausible.

2.2.2 The Jarque-Bera Test

The Jarque-Bera (JB) test is a test for normality and is also used to test whether a series of observations is random and independent. The null hypothesis is that the response variable follows a normal distribution, and the test statistic is given by

$$JB = \frac{n-p+1}{6}\left[S^2 + \frac{1}{4}\,(K-3)^2\right], \qquad (2.1)$$

where n is the number of observations, S is the skewness, K is the kurtosis and p are the number of explanatory or independent variables. Under the null hypothesis of normality, the JB test statistic is asymptotically distributed as $\chi^2_{(2)}$. The null hypothesis of normality is rejected if the calculated test statistic exceeds a critical value from the $\chi^2_{(2)}$ distribution or p-value is less than the given significance level.

2.2.3 Shapiro-Wilk Test

The Shapiro-Wilk test (1972) has been found to be the most powerful normality test (Ghasemi and Zahediasl, 2012). The test statistic is given as:

$$SW = \frac{\left(\sum_{i=1}^{n} a_i y_{(i)}\right)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}, \qquad (2.2)$$

where $y_{(i)}$ are the ordered random sample values and a_i are constants generated from the covariances, variances and means of the sample of size n. For p-values of SW less 0.05, it indicates that the sample is normally distributed and thus the null hypothesis is accepted.

2.3 Test for Stationarity

Generally, a time series data set from various macroeconomic variables is non-stationary, as demonstrated by many studies, including Nelson and Plosser (1982), Stock and Watson (1988) and Campbell and Peron (1991). Therefore, the mean and variance of these time series economic variables are not independent of time. Granger and Newbold (1974) posited that to avoid the problem of spurious regression results associated with non-stationary time series model and to generate a possibility of long-run equilibrium relationship or stability as well as an optimal and efficient predictive model, a unit root test or test for stationarity on the economic variables under investigation was carried out using Augmented Dickey–Fuller (ADF), Philips-Peron (PP) and Kwiatkowski-Philips-Schmidt-Shin (KPSS) tests under the conditions of the model without trend and intercept, model without trend but with intercept and model with both trend and intercept. Thus, a description of the unit root model or stationary tests is as follows:

2.3.1 The Augmented Dickey-Fuller (ADF) Test

The Augmented Dickey-Fuller (ADF) test is an augmented version of the original Dickey-Fuller Test that was introduced by Dickey and Fuller (1979). The ADF test is used to test for a unit root in a time-series sample. The ADF test assumes the error term to be a white noise process. The null hypothesis is that returns have a unit root (exhibit non-stationary). The ADF has three cases, depending on the nature of the time series of the data being tested.

Case 1: no constant

The test equation is given as

$$\Delta Y_t = \varphi Y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_{k-1} \Delta Y_{t-k+1} + \varepsilon_t , \qquad (2.3)$$

where k is the lag order of the autoregressive process, $\delta_1, \dots, \delta_{k-1}$ are coefficients of lagged difference terms ΔY_{t-i} which are used to approximate the autoregressive moving average (ARMA) structure of the errors. The equation has no intercept and no trend.

Case 2: no trend

The test equation is expressed as

$$\Delta Y_t = \Phi + \varphi Y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_{k-1} \Delta Y_{t-k+1} + \varepsilon_t , \qquad (2.4)$$

where Φ is a constant and k is the lag order of the autoregressive process. The equation has an intercept (Φ) but no time trend.

Case 3: with trend

The test equation is expressed in form given as

$$\Delta Y_t = \Phi + \psi t + \varphi Y_{t-1} + \delta_1 \Delta Y_{t-1} + \cdots + \delta_{k-1} \Delta Y_{t-k+1} + \varepsilon_t , \qquad (2.5)$$

where Φ is a constant, ψ is the coefficient on a time and k is the lag order of the autoregressive process. The test equation has an intercept (Φ) and time trend ψt. In all the cases ε_t is the white noise error term. The test statistic is given as

$$ADF_T = \frac{\hat{\varphi}}{SE(\hat{\varphi})}, \qquad (2.6)$$

where $\varphi = \Theta - 1$ and Θ is the coefficient of the lagged term of an autoregressive model of order one model. If $\Theta = 1$, then there is unit root.

In the ADF test, $\varphi = 0$ is tested. Once a value of the test statistic is computed, it can be compared to the critical value of the Dickey-Fuller test. The corresponding *p*-value can be calculated. The lag order, in addition to the sample size, can affect the finite sample behaviour of the ADF test. Proper correction for the lag effect when implementing the ADF test is desirable. The number of augmenting lags is determined by minimizing the Schwarz Bayesian Criterion (SBI) or minimizing the Alkaike Information Criterion (AIC). This study uses SBI because the software automatically selects the appropriate lag length.

2.3.2 The Philip Perron (PP) Test

The Phillips-Perron (PP) test was introduced by Phillips and Perron (1988). The PP test is used to test for a unit root in a time-series sample. They proposed the nonparametric test statistics for the unit root null by using consistent estimates of variances. In its simplest form, there are three cases of PP test equation depending on the nature of the time series being tested.

Case 1: no constant

The test equation is expressed as

$$Y_t = \varphi Y_{t-1} + \varepsilon_t . \qquad (2.7)$$

The equation has no intercept and no time trend.

Case 2: no trend

The test equation can be expressed as

$$Y_t = \Phi + \varphi Y_{t-1} + \varepsilon_t . \qquad (2.8)$$

The equation has an intercept (Φ) but no time trend.

Case 3: with trend

The test equation is expressed in the form given by

$$Y_t = \Phi + \psi t + \varphi Y_{t-1} + \varepsilon_t . \qquad (2.9)$$

The test equation has an intercept (Φ) and time trend (ψt). In all cases ε_t is a white noise error term.

The two test statistics for each case are taken from Maddala and Kim (2004) given as follows:

Case 1: no constant

The test statistic can be expressed as:

$$W_\varphi = T(\hat{\varphi} - 1) - \frac{1}{2} \frac{(\mathrm{M}^2 - \mathrm{M}_e^2)}{T^{-2}\sum_1^T Y_{t-1}^2} , \qquad (2.10)$$

$$W_t = \frac{\mathrm{M}_e}{\mathrm{M}} t_{\hat{\varphi}} - \frac{1}{2} \frac{(\mathrm{M}^2 - \Sigma_e^2)}{\mathrm{M}(T^{-2}\sum_1^T Y_{t-1}^2)^{\frac{1}{2}}} . \qquad (2.11)$$

Case 2: no trend

The test statistic can be expressed in the form given as:

$$W_\varphi = T(\hat{\varphi} - 1) - \frac{1}{2} \frac{(\mathrm{M}^2 - \mathrm{M}_e^2)}{T^{-2}\sum_1^T (Y_{t-1} - \bar{Y}_{t-1})^2} , \qquad (2.12)$$

$$W_t = \frac{\mathrm{M}_e}{\mathrm{M}} t_{\hat{\varphi}} - \frac{1}{2} \frac{(\mathrm{M}^2 - \mathrm{M}_e^2)}{\mathrm{M}[T^{-2}\sum_1^T (Y_{t-1} - \bar{Y}_{t-1})^2]^{\frac{1}{2}}} , \qquad (2.13)$$

where $\bar{Y}_{t-1} = \frac{\sum_1^T Y_t}{T-1}$.

Case 3: with trend

The test statistic is given as:

$$W_{\varphi} = T(\hat{\varphi} - 1) - \frac{T^6}{24\, D_Z}\,(\mathrm{M}^2 - \mathrm{M}_e^2)\,, \tag{2.14}$$

$$W_t = \frac{\mathrm{M}_e}{\mathrm{M}}\, t_{\hat{\varphi}} - \frac{T^3\,(\mathrm{M}^2 - \mathrm{M}_e^2)}{4\sqrt{3}\; D_Z^{\frac{1}{2}}\mathrm{M}}\,, \tag{2.15}$$

where, $D_Z = det\,(Z'Z\}$ and the regressors are $Z = (1.\ t,\ Y_{t-1})$, M^2 is the Newey-West consistent estimator of σ^2 (Newey and West, 1987), M_e^2 is the consistent estimator of σ_e^2,

$$\mathrm{M}_e^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2 \text{ and } \mathrm{M}^2 = \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2 + \frac{2}{T}\sum_{r=1}^{\varrho} G_{r\varrho} \sum_{t=r+1}^{\varrho} \varepsilon_t\varepsilon_{t-r}\,,$$

where $G_{r\varrho} = 1 - \frac{r}{\varrho+1}$ and $\varepsilon_t\varepsilon_{t-r}$ is the estimator of the covariance between error terms. The limiting distributions of W_{φ} and W_t are identical to those of $P = T(\hat{\varphi} - 1)$ and the t-statistics, respectively with $\mathrm{M}^2 = \mathrm{M}_e^2$. Thus, the asymptotic critical values of the tests are the same as the asymptotic critical values tabulated by Fuller (1976).

2.3.3 The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Test

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was introduced by Kwiatkowski *et al*. (1992). The KPSS test is used to test for a stationary root in a time-series sample. The KPSS test assumes the error term to be white noise process.

Using the models given as:

$$Y_t = \psi'\Pi_t\varepsilon_t + e_t\,, \tag{2.16}$$

$$\epsilon_t = \varepsilon_{t-1} + \upsilon_t\,, \tag{2.17}$$

$$\upsilon_t \sim iid(0, \quad \sigma_\upsilon^2)\,,$$

where Π_t contains the deterministic components (constant or constant plus time trend). The test statistic is given by:

$$KPSS = \frac{1}{T^2}\,\frac{\sum_{t=1}^{T}\mathrm{H}_t^2}{\sigma_\infty^2}\,, \tag{2.18}$$

where $\mathrm{H}_t = \sum_{h=1}^{t}\hat{e}_h$ representing a partial sum of e_h, e_t is the residual of Y_t on a constant (case 1) and the time trend (case 2) $\hat{\sigma}_\infty^2$ represents a heteroscedasticity and autocorrelation consistent estimator of the variance of $\hat{e}_t$ (This is a Lagrange Multiplier test for constant parameters against a random-walk parameter). For testing the null hypothesis of level stationary instead of trend stationary, the test is constructed the same way except that e_t is obtained as a residual from a regression of Y_t on an intercept only.

2.4 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF)

Autocorrelation Function (ACF) is a function that shows the correlation between the observation at time t and the observation at previous times (Teusch, 2006). Dubrova and Arhipova (2004) and Toutenburg and Heumann (2008) emphasized that in considering the sample data, the autocorrelation function is required and, as such, can be expressed in the form given by:

$$\Gamma_k = \frac{\sum_{t=1}^{n-k}(y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}\,, \tag{2.19}$$

where $\bar{y} = \frac{\sum_{t=1}^{n} y_t}{n}$ is the sample mean.

The partial autocorrelation function is the correlation between Z_t and Z_{t-k} after the influence of the confounding variable Z_t and Z_{t+k} is removed. Partial autocorrelation coefficients are usually denoted by Φ_{kk} and given as

$$\Phi_{kk} = Cov(Z_t, Z_{t-k} | Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1})\ , \tag{2.20}$$

where Φ_{kk} is the correlation coefficient between two random variables Z_t and Z_{t-k} given that $Z_{t-1}, Z_{t-2}, \dots, Z_{t-k+1}$ exit (Bucur and Harja, 2012). Thus, Nikitin and Sosunova (2003) stated that the common method used in calculating the partial autocorrelation coefficients is the Yule-Walker equation defined by

$$\begin{array}{ccccc} P_1 = & \Phi_{k1} & \Phi_{k2}P_1 & \cdots & \Phi_{kk}P_{k-1}, \\ P_2 = & \Phi_{k1}P_1 & \Phi_{k2} & \cdots & \Phi_{kk}P_{k-2}, \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ P_k = & \Phi_{k1}P_{k-1} & \Phi_{k2}P_1 & \cdots & \Phi_{kk}, \end{array} \tag{2.21}$$

Partial autocorrelation coefficients can be estimated by using partial autocorrelation coefficients of the sample by changing the value P on the Yule-Walker equation with Γ, and counting for $K = 2, 3, \dots$ to get the value Φ_{kk} using Cramer's rules. Thus, Partial Autocorrelation Function (PACF) used to measure the level of closeness between y_t and y_{t+k} if the effect of the time lag $1, 2, \dots, k - 1$ is removed. The partial autocorrelation (Φ_{kk}) can be given as

$$\Phi_{kk} = \frac{\Gamma_k - \sum_{j=1}^{k-1} \Phi_{k-1}, j^{r_{k-j}}}{1 - \sum_{j=1}^{k-1} \Phi_{k-1}, j^{r_{k-j}}}\ , \tag{2.22}$$

where $j = 1, 2, \dots, k-1$ and $k > 1$.

2.5 Granger Causality Test

This technique was used to determine whether a time series helped predict or forecast another. Granger (1969) and Sims (1972) defined causality as the value of the independent variable(s) (x_t) such that, x_t have explanatory power in a regression of x_t on y_t. In general, let consider the information set F_t which has the form $(x_t, z_t, x_{t-1}, z_{t-1}, \dots, x_1, z_1)$, where x_t and z_t are vectors that includes scalars and z_t usually include y_t and further z_t may or may not include other variables than y_t. Thus, according to Sørensen (2005), it can be emphasized that x_t is Granger causes y_t with respect to F_t if the variance of the optimal linear predictor of y_{t+k} based on F_t has smaller variance than the optimal linear predictor of y_{t+k} based on $z_t, z_{t-1}, \dots, z_{t-k}$for any k. In other word, x_t is Granger causes y_t if x_t can be used to predict y_t at a particular stage in the future. It is often said that x_t Granger causes y_t and y_t Granger causes x_t. It does not mention anything about possible instantaneous correlation between x_t and y_t but if the innovation to y_t and the innovation to x_t are correlated then, there is instantaneous causality. One will usually find an instantaneous correlation between two-time series. However, since the causality (in the "real" sense) can go either way, one usually does not test for instantaneous correlation (Sims, 1972). However, one can find Granger causality in only one direction, and one may feel that the case for "real" causality is more robust if there is no instantaneous causality (Sims, 1972). Granger causality is straightforward to deal with in vector autoregressive model (VAR) models given as follows:

$$\begin{bmatrix} y_t \\ z_t \\ x_t \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \Pi_{11}^1 & \Pi_{12}^1 & \Pi_{13}^1 \\ \Pi_{21}^1 & \Pi_{22}^1 & \Pi_{23}^1 \\ \Pi_{31}^1 & \Pi_{32}^1 & \Pi_{33}^1 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \\ x_{t-1} \end{bmatrix} + \cdots + \begin{bmatrix} \Pi_{11}^k & \Pi_{12}^k & \Pi_{13}^k \\ \Pi_{21}^k & \Pi_{22}^k & \Pi_{23}^k \\ \Pi_{31}^k & \Pi_{32}^k & \Pi_{33}^k \end{bmatrix} \begin{bmatrix} y_{t-k} \\ z_{t-k} \\ x_{t-k} \end{bmatrix} + \begin{bmatrix} \mu_{t-k} \\ \mu_{t-k} \\ \mu_{t-k} \end{bmatrix}, (2.23)$$

Also assume that

$$\Sigma_u = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{13} \\ \Sigma_{12}' & \Sigma_{22} & \Sigma_{23} \\ \Sigma_{13}' & \Sigma_{23}' & \Sigma_{33} \end{bmatrix},$$

(2.23) is a general vector autoregressive model that has been partitioned into 3 sub-vectors such as: the y_t and the x_t vectors between which will test for causality and the z_t vector (which may be empty) based on condition that x_t does not Granger cause y_t with respects to the information set generated by z_t if either $\Pi_{13}^{j} = 0$ and $\Pi_{23}^{j} = 0$; $j = 1, 2, \ldots, k$ or $\Pi_{12}^{j} = 0$ and $\Pi_{13}^{j} = 0$; $j = 1, 2, \ldots, k$. It must be noted that this is the way one can test for Granger causality. Hence, one will use the vector autoregressive approach if an econometric hypothesis of interest states that x_t Granger causes y_t but y_t does not Granger cause x_t (Sims, 1972).

2.6 Multicollinearity

The linear regression model is the most widely used statistical technique for fitting functional relationship problems among variables. It helps explain observations of a response (dependent) variable with observed values of one or more explanatory (independent) variables. In achieving this, an OLS estimator is the most appropriate and well-known technique for estimating the model's parameters. According to Gunst and Mason (1980) as cited in Arowolo*et al*. (2016), it was posited that the OLS estimator contains some strong, attractive statistical properties under certain assumptions, which made it an appropriate, robust and popular estimator for linear models. The validity of the results, test statistics and confidence interval depend mainly on the degree to which the model's assumptions are met. However, in dealing with economic growth and its drivers, including internal debt, external debt, interest rate exchange rate and degree of economic openness, there is a tendency to violate assumptions because of the correlation among the aforementioned economic variables. Also, Oyewole and Agunbiade (2020) stressed that the dynamics in the observed value of the variables mentioned during the period of economic boom or recession, pandemics, insecurities and political unrest will lead to multicollinearity. Thus, ordinary least squares (OLS) will not produce the best results, particularly estimated variance, resulting in inefficient estimation of the model's parameters. The prediction and estimation of the model become biased, insufficient, inconsistent and inefficient (Agunbiade, 2012). Hence, there is a need for a robust statistical methods to model, estimate and predict economic growth for Nigeria using the identified covariates. Moreover, it can be emphasized that under the assumption of multicollinearity, correlation coefficients of independent variables are computed even though a strong correlation coefficient does not necessarily imply the presence of multicollinearity; it can be a suspect, and as such, it can be ascertained by checking the variance inflation factor (VIF) and conditional index number (CIN).

2.6.1 Variance Inflation Factor (VIF) Test for Multicollinearity

Multicollinearity occurs when there are several predictors or explanatory variables that are highly correlated with each other. For instance, $x_1, x_2, \dots, x_k$ that are highly correlated with other predictors or explanatory variables such as $x_2, \dots, x_k$ (Marquardt, 1970, Belsley *et al*., 1980 and Murray *et al*. 2012). This can lead to significant problems when adding or removing a predictor, as it may substantially change the estimated regression coefficients and consequently alter the conclusions drawn from the model. When multicollinearity is present, the sampling distribution of individual β_i coefficients may have inflated variances, resulting in large confidence intervals for some β_i, $i = 1, 2, \dots, k$ (Belsley *et al*., 1980). This invalidates the standard interpretation of β_i as the mean change in the response when x_I increased by one unit. For instance, if x_2 is highly correlated with x_3, holding x_3 constant while increasing x_2 becomes meaningless. Murray et al. (2012) noted that a formal method for detecting multicollinearity is the variance inflation factor (VIF). VIFs measure how much the variances of estimated regression coefficients are inflated compared to having uncorrelated predictors (Murray et al., 2012; Ebiwonjumi et al., 2022; Ebiwonjumi et al., 2023). To compute the VIF, one uses the standardized regression model.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + U_i \ , \tag{2.24}$$

was used. This can be written in vector and matrix form as:

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} X_1'X_1 & X_1'X_2 & \dots & X_1'X_n \\ X_2'X_1 & X_2'X_2 & \dots & X_2'X_n \\ \vdots & \vdots & \ddots & \vdots \\ X_n'X_1 & X_n'X_2 & \dots & X_n'X_n \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_n \end{bmatrix} \ , \tag{2.25}$$

2.6.2 Variance Inflation Factors Formular

In dealing with VIFs, it is noted that collinearity is illustrated using numeric predictors in regression models, showing the observed values of two or three predictors being literally co-linear or co-planar (Chatterjee and Price, 1977 and Belsley *et al*. 1980). Diagnostic techniques include pairwise correlation coefficients for collinearity involving only two predictors and condition indices and VIFs for relationships involving multi-variables or multiple predictors (Murray *et al*., 2012). This study focuses on VIFs because VIF formulas are given for each predictor, which is supposed to identify a particular predictor that contributed to a collinearity problem. For the multiple regression model with predictors, $X_i, i = 1, \dots, k$. VIFs are the diagonal elements (r^{ii}) of the inverse of the correlation matrix R_{kxk} of the k predictors (Chatterjee and Price 1977; Belsley *et al*. 1980). According to Murray *et al*. (2012), the VIF for the r^{th} predictor variable can be expressed as

$$\text{VIF}_i = r^{ii} = \frac{1}{1 - R_i^2}, i = 1, 2, \dots, k \ , \qquad (2.26)$$

where R_i^2 is the multiple correlation coefficient of the regression between X_i and the remaining $k-1$ predictors. According to Belsley *et al*. (1980), there is no clear cutoff point to distinguish between high and low VIFs. Several researchers, such as Hocking and Pendelton (1983) and Craney and Surles (2002), suggested that the typical cutoff values (or rules of thumb) for large VIFs are 5 or 10. According to Allison (1999) and Freund and Littell (2000), as cited in Khalaf and Iguernane (2016) opined that VIF increased the variability of the estimated coefficients. Thus, it overstated the estimated parameters' variance compared to what can be obtained when there was no correlation with any of the remaining variables present in fitting the model. Thus, VIF greater than 10 indicates a statistically significant multicollinearity. O'Brien (2007) recommended that well-known VIF rules of thumb, for instance VIFs greater than 5 or 10 or 30 should be treated with caution when making decisions to reduce collinearity by eliminating one or more predictors. Thus, it is suggested that researchers should also consider other factors, such as sample size, which influence the variability of regression coefficients.

2.7 Test for Outliers

Identifying and diagnosing the presence of extreme values and their influence is a key issue in modeling, interpreting, and generalizing the volatility of RGDP and the identified economic growth drivers in this study. Consequently, data exploration, description, and various diagnostic tests are essential. The tests that can be conducted for this purpose include Dixon's test, Grubbs' test, Cochran's C test, and Bartlett's test.

2.7.1 Grubbs' Test

The Grubbs' test for outliers was recommended by International Statistical Organization (ISO) because it allows the use of all the observed values for the variables under consideration in computing the statistic; that is, the observed values are used without deleting the extreme value(s) and as such it was used in this study. This test is carried out by computing the deviation between the value assumed to be an outlier and the mean value of the given dataset. The computed value then compares with the standard deviation obtained from the dataset. The value assumed to be an outlier is extremely higher or lower than the mean. Grubbs' test null hypothesis (H0) is that no outliers exist in the dataset under investigation. International Statistical Organization (1994) stated that the test statistic can be calculated:

$$\widehat{\mathrm{G}}_{\mathrm{m}} = \frac{|\mathrm{X}_{\mathrm{s}}^{*} - \overline{\mathrm{X}}|}{\mathrm{S}}, \tag{2.27}$$

$$\widehat{\mathrm{G}}_{\mathrm{m}} = \frac{|\text{Suspect value} - \overline{\mathrm{X}}|}{\mathrm{S}}. \tag{2.28}$$

where, $\overline{\mathrm{X}}$ and S are the mean and standard deviation repectively.

2.8 Anderson-Darling Test (Goodness-of fit-test)

The Anderson-Darling (AD) statistic is used to check for the goodness-of-fit of the data. The Anderson-Darling statistic is a modification of the Kolmogorov Simirnov (KS) test that gives more weight to the tails of the distribution than the KS test (Chalabi, 2012). The AD test was introduced by Anderson and Darling (1952). The AD test is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function. The null hypothesis is that the data follows the specified distribution. The AD test procedure is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function. The AD test statistic (A^2) is defined as

$$A^2 = m \ \frac{1}{m}\sum_{i=1}^{m}(2i-1)[lnF(Y_i) + \ln(1 - F(Y_{m-i+1}))], \tag{2.29}$$

where $y_{(1)} < y_{(2)} < \cdots < y_{(n)}$ is the ordered sample size *n*, and F(x) is the underlying theoretical cumulative distribution to which the sample is compared. The null hypothesis is that $y_{(1)} < y_{(2)} < \cdots < y_{(n)}$ comes from the underlying distribution $F(Y)$ that is, the data follows a specified distribution which is rejected at a chosen level of significance(α), if the A^2 test statistic is greater than the critical value. Critical values of AD test statistics depend on the specific distribution being tested. However, tables of critical values for many distributions are difficult to find (Thas, 2010). However, some statistical software, such as Minitab generates the *p*-values for the AD test. The AD test may be used to compare the goodness-of-fit of several distributions (Engmann and Cousineau, 2011).

2.9 Forecasting Evaluation Metrics

Forecasting metrics in this study, the mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are used to evaluate the forecasting performance of the prediction models employed.

The MSE is defined as:

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(Y_i - \hat{Y}_i\right)^2, \tag{2.30}$$

where Y_i= observed values, $\hat{Y}_i$ = set of estimated or predicted values and N = number of sample data or sample size. The RMSE statistic is defined as

$$R\text{MSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{Y}_i - Y_i\right)^2} \ . \tag{2.31}$$

The MAE statistic is defined as

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i - \hat{Y}_i\right| . \tag{2.32}$$

while, the MAPE statistic is defined by

$$\text{MAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{Y_i - \hat{Y}_i}{Y_i}\right| . \tag{2.33}$$

CHAPTER 3

DATA EXPLORATORY ANALYSIS

3.1 Introduction

In this chapter, the data source and exploration results are presented and discussed.

3.2 Data Source and Description

The data used in this study consist of quarterly macroeconomic data collected on economic growth rate (RGDP) and its drivers, including internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR), and trade openness (OPEN). The data spanning between 1986-2022 were sourced from the Central Bank of Nigeria (CBN) statistical bulletin.

3.3 Exploratory Data Analysis

Table 3.1 shows the descriptive summary of statistics of the data used in this study.

Table 3.1: Summary of Statistics

	RGDP(%)	INDT(%)	EXDT(%)	RINR(%)	REXR(%)	OPEN(%)
Panel A: Descriptive Statistics						
Mean	10.3046	6.6288	6.6573	3.1051	4.3175	0.1665
Maximum	11.1422	9.0867	8.4950	3.5860	6.3197	0.4600
Minimum	9.6316	3.3478	3.7246	2.4849	2.7763	0.0100
Std. Dev.	0.4503	1.4988	1.0655	0.1929	0.6951	0.1302
Skewness	0.3710	-0.4095	-0.1634	-0.5734	-0.1748	0.6710
Kurtosis	1.8572	2.5576	2.7076	4.1714	3.6271	2.1246
Panel B: Normality						
Jarque-Bera (*p*-value)	10.9083 (0.0043)	5.0906 (0.0785)	1.1299 (0.5684)	15.7879 (0.0004	3.0287 (0.2120)	15.0811 (0.0005)
Shapiro-Wilk (*p*-value)	0.9290 (0.0000)	0.9564 (0.0002)	0.9652 (0.0012)	0.9673 (0.0018)	0.9431 (0.0000)	0.8899 (0.0000)
Anderson-Darling (*p*-value)	3.2890 (<0.0050)	1.5270 (<0.0050)	1.6460 (<0.0050)	1.1320 (0.0060)	3.2930 (<0.0050)	5.6660 (<0.0050)
Panel C: Unit root or Stationary tests						
	RGDP	INDT	EXDT	RINR	REXR	OPEN
None (*p*-value)	0.1289 (0.7216)	0.2788 (0.7653)	-0.2548 (0.5927)	-0.2731 (0.5859)	-0.6103 (0.4513)	-0.8534 (0.3444)
ADF Intercept (*p*-value)	-1.7388 (0.4095)	-2.4017 (0.1433)	-3.0666 (0.0315)	-3.2733 (0.0181)	-2.4131 (0.1401)	-1.7166 (0.4205)

	Trend and Intercept (*p*-value)	-0.9638 (0.9445)	-2.4636 (0.3456)	-2.8327 (0.1883)	-3.3995 (0.0558)	-2.3384 (0.4101)	-1.4117 (0.8533)
	None (*p*-value)	0.1536 (0.7292)	0.6836 (0.8622)	-0.0608 (0.6608)	-0.1193 (0.6409)	-1.0967 (0.2464)	-0.9186 (0.3168)
PP	Intercept (*p*-value)	-1.3870 (0.5871)	-2.1825 (0.2136)	-2.9184 (0.0458)	-2.6605 (0.0836)	-3.0596 (0.0320)	-1.3626 (0.5989)
	Trend and Intercept (*p*-value)	0.7660 (0.9997)	-1.4859 (0.8300)	-2.4979 (0.3288)	-3.5294 (0.0401)	-2.9405 (0.1532)	-0.9936 (0.9407)
KPSS	Intercept (*p*-value)	1.0204 (>0.0500)	1.0134 (< 0.0100)	0.4143 (<0.1000)	0.0818 (>0.0500)	0.1434 (>0.0500)	0.8028 (<0.0100)
	Trend and Intercept (*p*-value)	0.2134 (>0.0500)	0.2683 (<0.0100)	0.2344 (<0.0100)	0.0773 (>0.0500)	0.147765 (<0.0500)	0.2103 (<0.0500)

Table 3.1 shows the descriptive results for the macroeconomic variables (RGDP, INDT, EXDT, RINR, REXR, and OPEN) considered for this investigation. ‘The average value of RGDP during the period under investigation is 10.3046, ranging from 9.6315 to 11.1422. The mean values of INDT and EXDT are 6.6288 and 6.6572, which range between 3.3478 to 9.0867 and 3.7245 to 8.4950, respectively. Meanwhile, the average RINR, REXR, and OPEN values are 3.1051, 4.3175, and 0.16645, respectively. ‘It is observed that RINR ranged from 2.4849 to 3.5860, REXR ranged from 2.7763 to 6.3197, and OPEN ranged from 0.01 to 0.46 during the period under study. ‘The values 0.4503, 1.4988, 1.0655, 0.1929, 0.6951 and 0.1302 reveal the rate at which RGDP, INDT, EXDT, RINR, REXR and OPEN deviate from their respective mean values. The skewness and kurtosis results presented in Table 3.1 explain the nature of the distribution and shape of the RGDP and its identified drivers. The skewness results show that the RGDP (0.3710) and OPEN (0.6709) are positively skewed. That is, the variables are skewed to the right of the mean. It is also found from the results that INDT (-0.4095), EXDT (-0.1634), RINR (-0.5734) and REXR (-0.1748) are negatively skewed. This implies a skew to the left of the mean. Also, the kurtosis results reveal that RGDP and its identified drivers under consideration are platykurtic with a kurtosis coefficient index less than 3, which implies a lower peak and lighter tail curve, except for RINR and REXR, which are leptokurtic which emphasizes higher peak and heavier tail curve for the economic growth drivers beyond the level of normal distribution.

We test for stationarity of the RGDP and its drivers. In this study, the stationarity test is carried out using ADF, PP and KPSS tests under the model conditions without trend and intercept (None), intercept but no trend and with both trend and intercept. The results are presented in Table 3.1, panel C. In Table 3.1, the test results are reported for the ADF, PP and KPSS test statistics for the model without both trend and intercept (None), a model with intercept but no trend and the model with both trend and intercept. In determining the exact order for the stationary tests, the SIC automatically select the appropriate lag length based on the size of the data sets used for this study. The results of the stationarity for RGDP and its drivers are determined at critical values of 10%, 5% and 1%, respectively. Thus, ADF and PP tests reveal that at zero level of difference denoted by (0) for the RGDP(0) INDT(0), RINR(0), REXR(0) and OPEN(0) are not stationary under the identified model conditions except REXR(0) which is stationary at 5% and 10% as revealed using PP technique under the condition that the model with intercept but no trend and KPSS method at the same level of significance but under the model condition that contains both trend and intercept reveals that REXR(0) is stationary.

Also, a model with an intercept but no trend condition using the KPSS technique reveals that RGDP(0) is stationary at 10%, 5% and 1% levels of significance, respectively. The KPSS method shows that INDT(0) is stationary at 10%, 5% and 1% for a model with intercept and trend conditions. In the results presented in Table 3.1, EXDT(0) is also stationary at 10% and 5% significance levels when the model assumes intercept but no trend situation as revealed using ADF and PP methods. Meanwhile, the KPSS method shows that EXDT(0) is stationary at 10%, 5% and 1% significance levels when the model is fitted with intercept and trend. The KPSS method also reveals that OPEN(0) is stationary at 10%, 5% and 1% levels of significance, respectively, when the model contains the intercept but no trend and when the model contains both intercept and trend, OPEN(0) is stationary at 10% and 5% level of significance respectively. From the results, it can be established that KPSS is effective and optimal in determining the stationarity of the variable(s) at level [(0)] with a model that satisfied intercept but no trend condition and with a model with both intercept and trend condition. Following the stationarity test, the normality test is carried out, and in Table 3.1 panel B, the results that show the distribution of the economic growth (RGDP) and its identified drivers in this study are presented and discussed.

From Table 3.1, as presented in panel B, the Shapiro-Wilk test statistic, which is the most efficient test for normality, shows that RGDP, INDT, EXDT, RINR, REXR and OPEN values are 0.9290, 0.9564, 0.9652, .09673, 0.9431 and 0.8899 respectively. Thus, a Shapiro-Wilk values that are closer to one and are statistically significant at a 5% level reveals that all the RGDP and its identified drivers under investigation are from the normally distributed population. The probability of Jarque-Bera values, which are 0.0042 < 0.05, 0.0003 < 0.05 and 0.0005 < 0.05, respectively, reveal that RGDP, RINR and OPEN are normally distributed, which may contain outliers as indicated in Table 3.1. In addition, the Q-Q plots for the RGDP and its identified drivers are shown in Figure 3.1

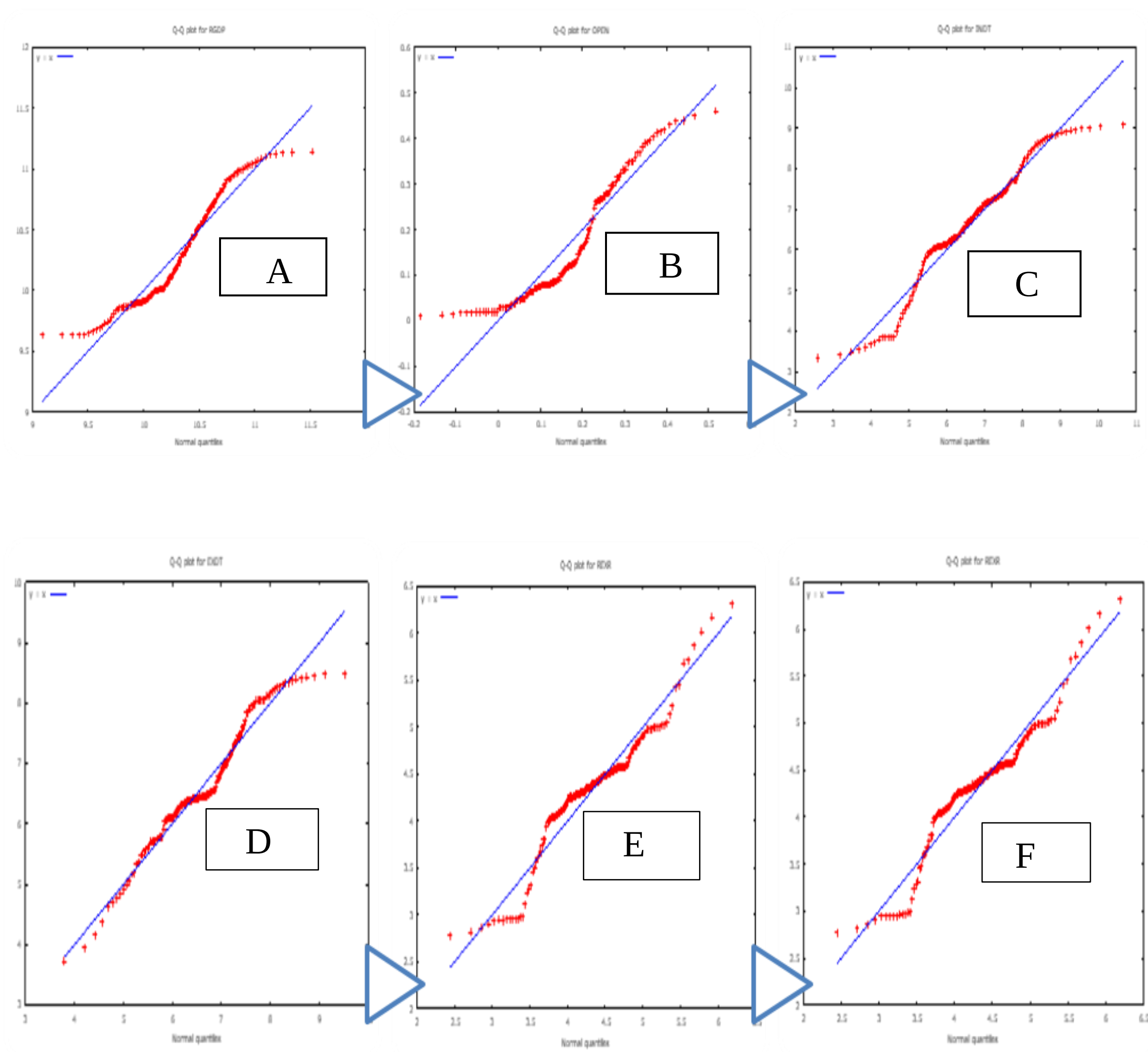
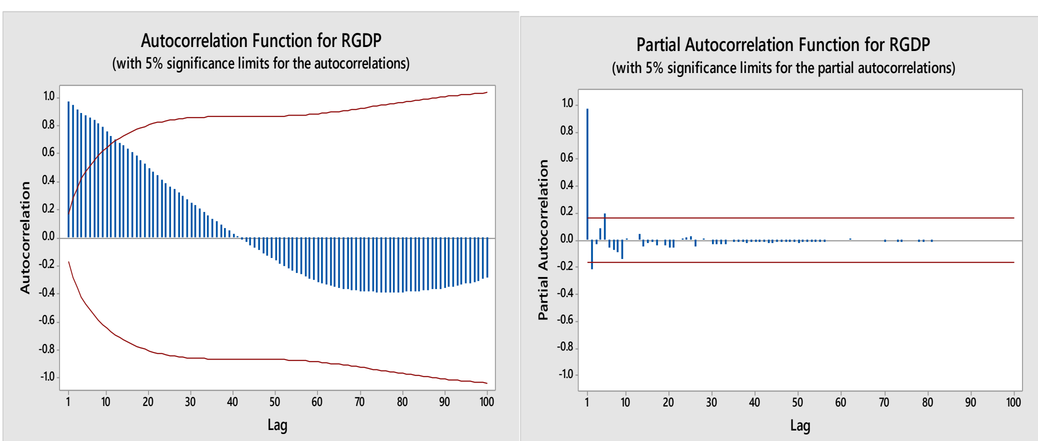


Figure 3.1: The Q-Q Plot showing the Normality of RGDP and its Drivers

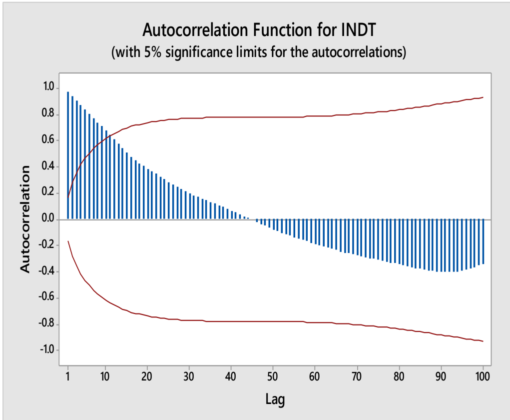
In Figure 3.1, we present the Q-Q plots for the RGDP indicated by [A] and its drivers: ‘internal debt (INDT) labelled as [B], external debt (EXDT) labelled as [C], interest rate (RINR) labelled as [D], exchange rate (REXR) labelled as [E] and trade openness (OPEN) labelled as [F]. Thus, it is revealed that a greater proportion of the observed values denoted by the dotted red lines for the RGDP and its identified drivers are concentrated on the blue straight line and, as such, indicate that RGDP and its identified drivers are from a normal distributed population that may contain outliers. Thus, further diagnostic tests are needed to ascertain the presence of outliers. Another exploratory and diagnostic evaluation to be considered in this study is investigating serial correlation in the data using autocorrelation function and partial autocorrelation function plots of the RGPD and its drivers or predictors.

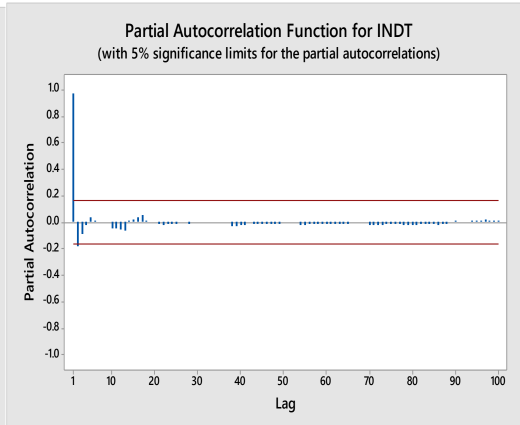
3.4 Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) Plots

In order to check for serial correlation of the data used in the study, we use the ACF and PACF plots. The ACF and PACF plots for the data used in the study are shown in Figure 3.2. The left panels in Figure 3.2 show the ACF plots and the right panels show the PACF plots.



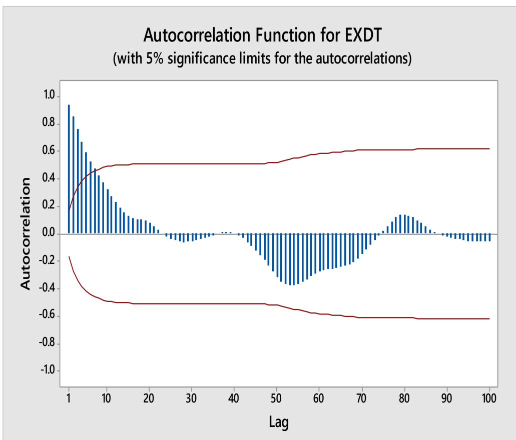
(A) (B)

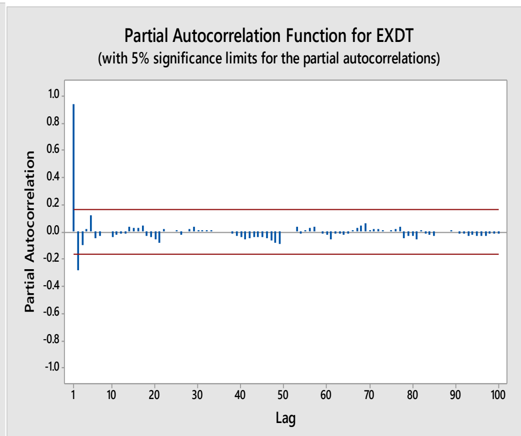

Autocorrelation Function for INDT
(with 5% significance limits for the autocorrelations)
Autocorrelation
Lag


Partial Autocorrelation Function for INDT
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
Lag

(C)

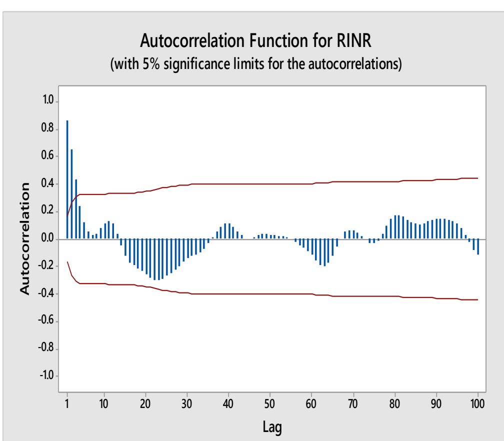
(D)

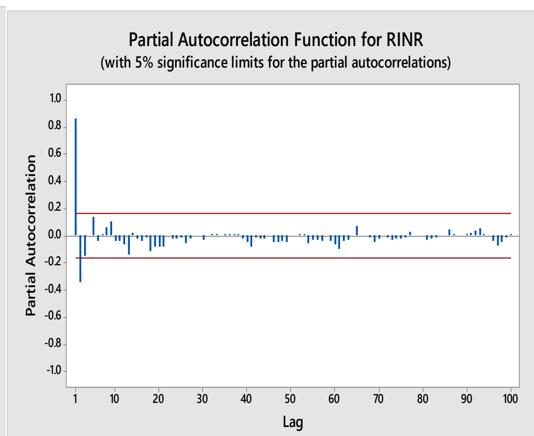

Autocorrelation Function for EXDT
(with 5% significance limits for the autocorrelations)
Autocorrelation
Lag


Partial Autocorrelation Function for EXDT
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
Lag

(E)

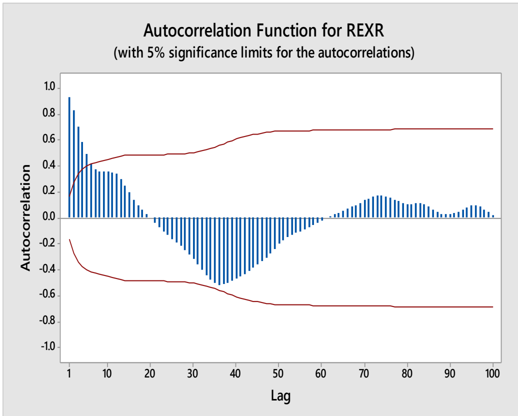
(F)

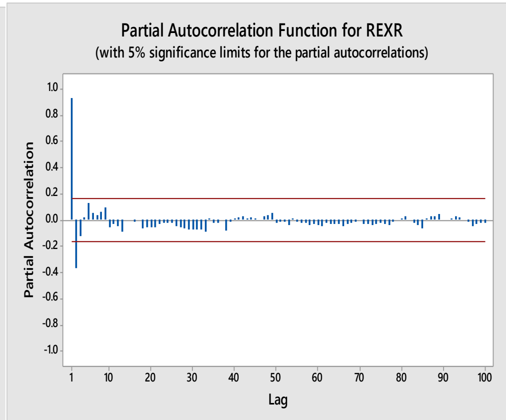

Autocorrelation Function for RINR
(with 5% significance limits for the autocorrelations)
Autocorrelation
Lag


Partial Autocorrelation Function for RINR
(with 5% significance limits for the partial autocorrelations)
Partial Autocorrelation
Lag

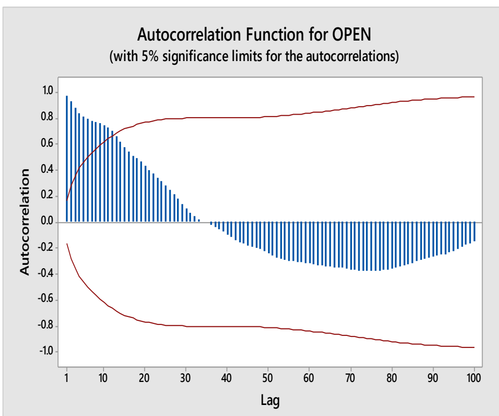
(G)

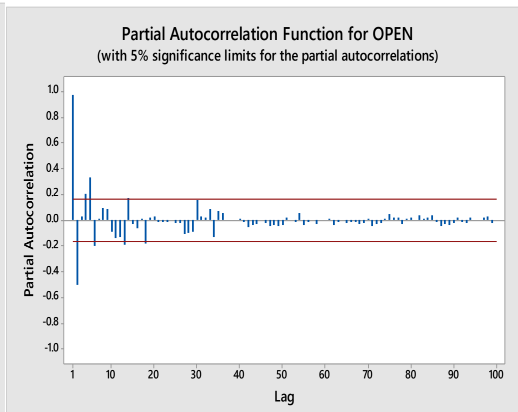
(H)





(I) (J)





(K) (L)

Figure 3.2: Panel A: ACF of RGDP, Panel B: PACF of RGDP Plots and follows by its Identified Drivers ACF and PACF respectively.

Figure 3.2 contains plots of the autocorrelations and partial autocorrelations labelled as A and B for the RGDP, C and D for INDT, E and F for EXDT, G and H for RINR, I and J for REXR and K and L for OPEN respectively. This is done to examine and test the serial correlation in RGDP and its identified drivers in this study. The plots of the aforementioned macroeconomic variables spread are highly persistent, assuming an AR(k) because the autocorrelations consist of damped exponentials and sinusoidal expressions. The partial autocorrelation is also varied negatively and positively across various lags order process used in this study. Thus, the RGDP and its drivers under consideration follow an autoregressive model of order k [AR(k)] as revealed by the plots. Hence, time series plots for RGDP and its identified drivers in Figure 3.2 are not constant or zerobut varies across the various lags order. Thus, it reveals that RGDP and its drivers are not stationary because the information decays toward zero exponentially and non-zero through lags order process, as indicated by the ACF and the PACF plots. After establishing the stationarity of the data used in this study, it is imperative to examine the strength of the relationship between the RGDP and its drivers or predictors. Thus, the Kendall tau test for the strength of the relationship is carried out, and the results are shown in Table 3.2.

Table 3.2: Kendall's tau test statistic and *p*-value in parenthesis Correlation

		RGDP	INDT	EXDT	RINR	REXR	OPEN
RGDP	Kendall statistic	1.0000	0.8640	0.3840	0.0880	0.0410	0.7190
	p-value	.	(0.0000)	(0.0000)	(0.1220)	(0.4740)	(0.0000)
INDT	Kendall statistic		1.0000	0.4390	0.0260	0.0460	0.7310
	p-value		.	(0.0000)	(0.6490)	(0.4140)	(0.0000)
EXDT	Kendall statistic			1.0000	0.2120	-0.2090	0.2480
	p-value			.	(0.0000)	(0.0000)	(0.0000)
RINR	Kendall statistic				1.0000	-0.3500	-0.0440
	p-value				.	(0.0000)	(0.4430)
REXR	Kendall statistic					1.0000	0.1740
	p-value					.	(0.0020)
OPEN	Kendall statistic						1.000
	p-value						.

The Kendall tau test presented in Table 3.2 shows evidence of a relationship or association among the RGDP and its identified drivers, which include INDT, EXDT, RINR, REXR, and OPEN, which are under consideration in this study. Table 3.2 shows that INDT, EXDT, RINR, REXR and OPEN are positively related to the RGDP in Nigeria. The results in Table 3.2 reveal that the strength of the relationship of the INDT, EXDT, RINR, REXR and OPEN with RGDP are 86.4%, 38.4%, 8.8%, 4.1% and 71.9%, respectively. Thus, it can be emphasized that INDT and OPEN show greater strength in driving the RGDP. Also, the linear relationship between RGDP and its identified drivers is examined using the plot in Figure 3.3.

3.5 Relationship that exists among the RGDP Drivers

In order to investigate the relationship between the RGDP and its drivers, the plot in Figure 3.3 is used. Thus, from Figure 3.3, it is observed that the trend lines show the existing relationship among the economic growth drivers considered in this study. This is carried out to clearly view the relation pathway between economic growth (RGDP) and its drivers.

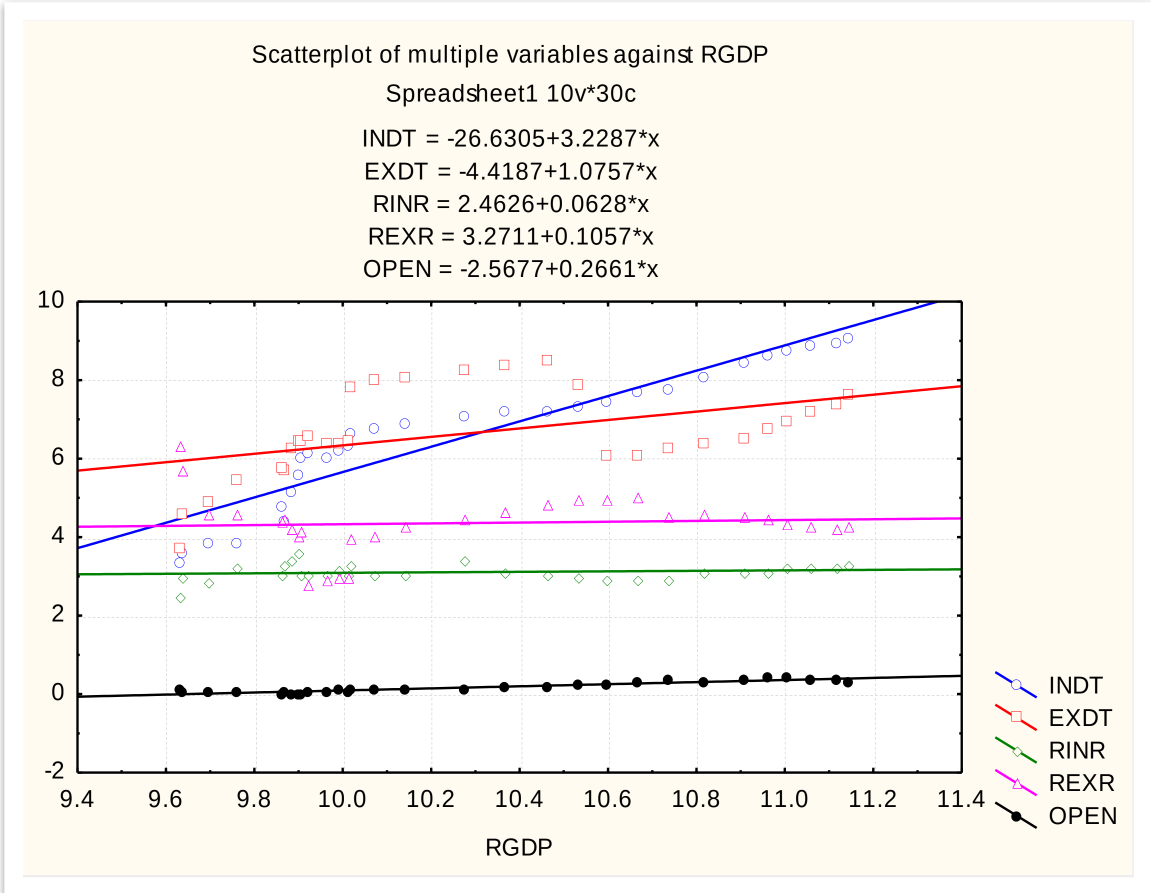


Figure 3.3: The Plot showing the Relationship that exist among the Economic Growth Drivers

The plot presents in Figure 3.3 shows the trend line for mean and variance of the economic growth drivers (explanatory variables) that include: INDT, EXDT, RINR, REXR and OPEN and how related they are in the period under study in determining RGDP. Thus, it reveals that a relationship exists among the identified economic growth drivers to be used in examining RGDP in this study. After establishing the stationarity and the relationship among the identified economic growth drivers to be used in predicting RGDP in this study, it is important to examine the direction of causal relationship among the variables. The plot presented in Figure 3.3 shows the trend line for mean and variance of the economic growth drivers (explanatory variables) that include INDT, EXDT, RINR, REXR and OPEN and how related they are in the period under study in determining RGDP. Thus, it reveals that a relationship exists among the identified economic growth drivers to be used in examining RGDP in this study. After establishing the stationarity and the relationship among the identified economic growth drivers to be used in predicting RGDP in this study, it is essential to examine the direction of causal relationship among the variables. This is done using the Granger causality test, and the results are presented in Table 3.3.

Table 3.3: Granger Causality Test

Null Hypothesis:	Test statistic	*p*-value	Remarks
INDT does not Granger Cause RGDP	2.5570	0.0813	INDT granger cause RGDP
RGDP does not Granger Cause INDT	0.1726	0.8416	H_0 not Rejected
EXDT does not Granger Cause RGDP	4.0794	0.0191	EXDT granger caused RGDP
RGDP does not Granger Cause EXDT	1.4128	0.2470	H_0 not Rejected
RINR does not Granger Cause RGDP	3.4239	0.0355	RINR granger caused RGDP
RGDP does not Granger Cause RINR	1.7148	0.1839	H_0 not Rejected
REXR does not Granger Cause RGDP	0.1200	0.8870	H_0 not Rejected
RGDP does not Granger Cause REXR	1.0038	0.3692	H_0 not Rejected
OPEN does not Granger Cause RGDP	14.8862	0.0000	OPEN granger caused RGDP
RGDP does not Granger Cause OPEN	8.1669	0.0005	RGDP granger caused OPEN
EXDT does not Granger Cause INDT	0.3523	0.7037	H_0 not Rejected
INDT does not Granger Cause EXDT	0.8429	0.4327	H_0 not Rejected
RINR does not Granger Cause INDT	0.9110	0.4045	H_0 not Rejected
INDT does not Granger Cause RINR	0.2414	0.7859	H_0 not Rejected
REXR does not Granger Cause INDT	1.3051	0.2746	H_0 not Rejected
INDT does not Granger Cause REXR	0.2271	0.7971	H_0 not Rejected
OPEN does not Granger Cause INDT	0.6511	0.5231	H_0 not Rejected
INDT does not Granger Cause OPEN	2.4288	0.0920	INDT granger caused OPEN
RINR does not Granger Cause EXDT	0.4637	0.6299	H_0 not Rejected
EXDT does not Granger Cause RINR	0.2613	0.7704	H_0 not Rejected
REXR does not Granger Cause EXDT	0.3998	0.6712	H_0 not Rejected
EXDT does not Granger Cause REXR	1.8060	0.1683	H_0 not Rejected
OPEN does not Granger Cause EXDT	0.0127	0.9873	H_0 not Rejected
EXDT does not Granger Cause OPEN	0.3336	0.7169	H_0 not Rejected
REXR does not Granger Cause RINR	2.8291	0.0620	REXR granger caused RINR
RINR does not Granger Cause REXR	0.7404	0.4788	H_0 not Rejected
OPEN does not Granger Cause RINR	0.6473	0.4788	H_0 not Rejected
RINR does not Granger Cause OPEN	0.6473	0.5251	H_0 not Rejected
OPEN does not Granger Cause REXR	1.2137	0.3003	H_0 not Rejected

REXR does not Granger Cause OPEN	1.5741	0.2110	H_0 not Rejected

From Table 3.3, the results show that INDT Granger causes RGDP at a 10% level of significance, and EXDT, RINR, and OPEN Granger cause RGDP at a 5% level of significance. However, REXR is found not to Granger cause RGDP at a 10% significance level. The Granger causality reveals a unidirectional causal relation between the INDT and RGDP and, as such, shows that the direction of the causal relation is from INDT to RGDP. Also, a unidirectional causal relationship is observed between the EXDT, RINR and RGDP. This implies a causal relation from EXDT and RINR to RGDP at p-value < 0.05. Thus, it can be stressed that EXDT and RINR Granger caused RGDP at a 5% significance level. Also, in Table 3.3, it is shown that a bidirectional causal relation exists between OPEN and RGDP. Thus, it indicates that both OPEN and RGDP Granger caused each other, and the causal relationship is significant as the associated p-value < 0.01.

In Table 3.3, we observe that a unidirectional causal relation between INDT and OPEN and REXR and RINR. This implies that a causal movement from INDT to OPEN and REXR to RINR with associated p-value < 0.10. Hence, it can be asserted that INDT Granger cause OPEN and REXR Granger cause RINR at 10% level of significance. This also emphasizes the dependency of the identified economic growth drivers (explanatory variables) on each other. As such, there is evidence of correlation between the variables, as earlier revealed by the computation of the correlation coefficients among the explanatory variables under investigation in this research work. Therefore, it can be emphasized based on this result that the economic growth drivers under consideration that Granger causes RGDP can be used to predict economic growth, and the explanatory variable that Granger causes any of the other explanatory variable(s) are not independent and as such can be a suspect to multicollinearity problem in this study. Also, it must be stressed that the explanatory variables that do not Granger cause any other explanatory variable(s) are independent and, as such, cannot cause any multicollinearity problem.

In order to check for the independence of the economic growth drivers such as INDT, EXDT, RINR, REXR, and OPEN used as covariates for determining economic growth (RGDP), the correlations among the economic growth drivers are examined by using Pearson’s correlation coefficients present in Table 3.4.

Table 3.4: Pearson's Correlation test statistic and corresponding *p*-values in parenthesis

	INDT	EXDT	RINR	REXR	OPEN
INDT	1.0000	0.6074	0.1276	-0.0924	0.8201
p-value		(0.0000)	(0.1320)	(0.2760)	(0.0000)
EXDT		1.0000	0.3876	-0.2898	0.2534
p-value			(0.0000)	(0.0000)	(0.0020)
RINR			1.0000	-0.4594	-0.0090
p-value				(0.0000)	(0.9160)
REXR				1.0000	0.2416
p-value					(0.0040)
OPEN					1.0000
p-value					

The correlation coefficients in Table 3.4 show the extent of the relationship between the economic growth drivers under consideration, which include INDT, EXDT, RINR, REXR, and OPEN in Nigeria. From Table 3.4, we observe that INDT is found to be positively correlated with EXDT, RINR and OPEN with correlation coefficients of 0.6074, 0.1276 and 0.8201, respectively. The results also reveal a positive correlation between the EXDT and RINR, EXDT and OPEN, and REXR and OPEN with correlation coefficients of 0.3876, 0.2534 and 0.2416, respectively. Thus, a high or strong correlation between the INDT and OPEN reveals the need to test for multicollinearity to avoid fitting a spurious linear model based on estimation and prediction from the data set. Also, according to Granger (2003), it can be asserted that a high and significant correlation coefficient between any two independent variables or determinants, such as the one under consideration in this study, is evidence to suspect the presence of multicollinearity. Thus, examining and testing for multicollinearity in the data under investigation is necessary. This test is carried out using variance inflation factor (VIF), and the results are presented in Table 3.5.

Table 3.5: Variance Inflation Factor Results for the Independency Variables

Variables	VIF for OLS Regression
INDT	14.2657
EXDT	3.5290
RINR	1.4369
REXR	1.7571

OPEN	9.5644

In Table 3.5, we present the results of the variance inflation factor for each of the identified drivers of economic growth under study to show the economic growth driver(s) that is or are responsible for the multicollinearity problem in the data being used. Specifically, in Table 3.5, the results show that VIF of 14.2657 > 10.0000 and according to Allison (1999) and Freund and Littell (2000), as cited in Khalaf and Iguernane (2016) and Ebiwonjumi *et al.* (2022) indicate the presence and statistical significance of multicollinearity problem. The INDT causes this as an economic growth driver used in this study. Based on this, a further investigation using appropriate estimation techniques and predictive models as an alternative to the ordinary least square estimation method is considered and discussed in the subsequent chapters of this study.

3.6 Outliers Test

Since the skewness and kurtosis values for the data discussed in Table 3.1 under the summary of descriptive statistics indicated a departure from the normal distribution, it is imperative to check for outliers in the data. In this study, we utilized the Grubb's test. The results of the Grubb's test are shown in Table 3.6

Table 3.6: Test for Outliers using Grubb's Test

Variable	G-statistics	G-critical value
RGDP	1.8600	0.8210
INDT	2.1890	0.8210
EXDT	2.7520	0.8210
RINR	3.2150	0.8210
REXR	2.8800	0.8210
OPEN	2.2540	0.8210

In Table 3,6, we present Grubb's test results for outliers. The null hypothesis for this test is that there are no outliers in the dataset under consideration. This hypothesis is rejected if G-statistics > G-critical values of the RGDP and the various identified economic growth drivers considered for this study. Thus, from the results presented in Table 3.6, it is revealed that G-statistic values of 1.8600, 2.1890, 2.7520, 3.2150, 2.8800 and 2.2540 are greater than the G-critical values of 0.8210 for RGDP, INDT, EXDT, RINR, REXR, OPEN respectively. Based on this result, it can be asserted that there are outliers in the dataset under consideration. Also, in Figure 3.4, we present the plots that reveal the presence of outliers in the data used for this study.

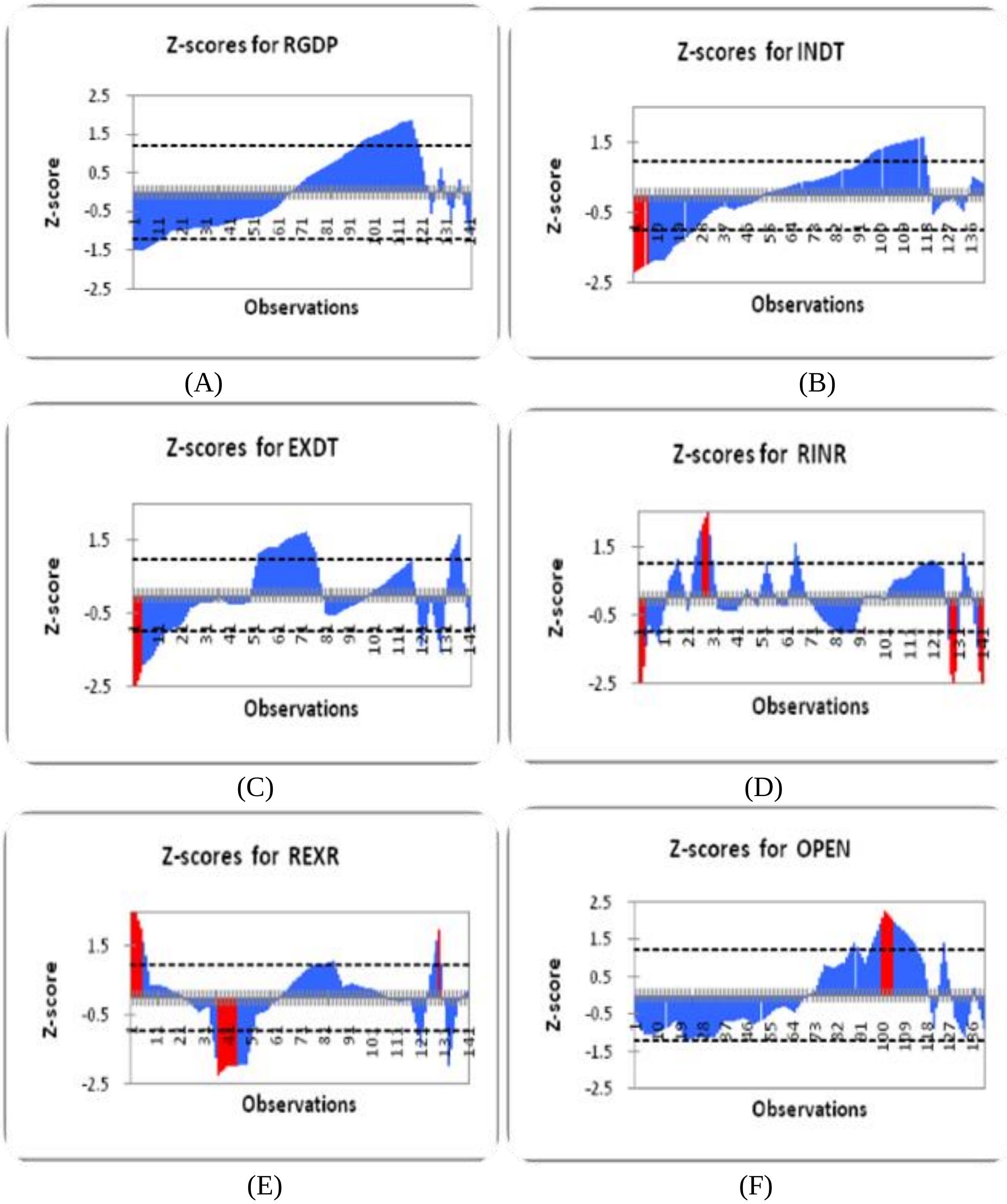


Figure 3.4: The Plots showing Outliers in RGDP and Identified Drivers

In Figure 3.4, A shows the plot for RGDP, B represents the plot for INDT, C represents the plot for EXDT, D shows the plot for the RINR, E is the plot for REXR, and F represents the plot for OPEN. In all the plots shown in Figure 3.4, outliers are indicated by the various points above and below the upper and the lower limits of the Z-score for plots RGDP and its drivers, INDT, EXDT, RINR, REXR and OPEN under investigation. Thus, in the following chapters, the multicollinearity problem and outliers shall be addressed using various robust statistical methods due to the inefficiency and instability of the predictive power of the ordinary least square method in the presence of multicollinearity and outliers in the data used for this study.

3.7 Summary of Exploratory Data Analysis

Exploratory data analysis revealed the following data properties such as: RGDP and OPEN are skewed to the right of the mean. INDT, EXDT, RINR and REXR are skewed to the left of the mean. Also, RGDP and its identified drivers are platykurtic thus implies the lower peak and lighter tails curve whle, RINR and REXR which are leptokurtic thus, emphasize higher peak and heavier tails curve of the economic growth drivers beyond the level of normal distribution. Thus, it indicates the presence of outliers in the data sets. Shapiro-Wilk and Jarque-Bera test statistics reveal that RGDP, INDT, RINR, and OPEN are normally distributed and may contain outliers. The Q-Q plots show that RGDP, INDT, EXDT, RINR, REXR and OPEN are from normal populations that may contain outliers. ADF and PP tests reveal that RGDP INDT, RINR, REXR and OPEN are not stationary except REXR. All these are the outlier's influence in the data set. The significant correlation coefficient among some of the identified economic growth drivers and the variance inflation factor result reveals the presence of multicollinearity. Thus, in order to build an efficient and robust statistical models for the prediction of RGDP using the INDT, EXDT, RINR, REXR and OPEN as economic drivers, we suggest the following models, which are capable of capturing most of the characteristics of the data.

1. Ridge regression method.
2. Robust principal component regression method.
3. Partial least square regression technique.
4. Average centered penalized regression least square method.
5. Gaussian process regression method, a machine learning approach.

6. Coupler FMKL-GLD quantile regression method.

3.8 Concluding Remarks

The exploratory data analysis carried out in this chapter focused on investigating the properties of the RGDP and its drivers. In summary, RGDP is related to all the economic growth drivers identified in this study. Thus, a linear relationship exists between the response variable (RGDP) and the economic growth drivers (explanatory variables) under investigation. RGDP and its drivers under consideration are significantly correlated. The RGDP drivers are correlated and, as such, show the presence of multicollinearity. It is also found that the RGDP and its identified drivers in this study contain outliers. Therefore, the statistically accepted models for RGDP and its identified economic growth drivers should be robust for predicting RGDP in the presence of multicollinearity and outliers. Hence, we need to propose the suggested models, which are the ridge regression model, robust principal component regression model, partial least square regression model, average centred penalized regression model, gaussian process regression model, a machine learning approach and Coupler Freimer–Mudholkar–Kollia–Lin-generalized lambda distributed (FMKL-GLD) quantile regression model. These robust methods are sequentially used to explore, model and predict RGDP in the next ten (10) quarters for each of the remaining chapters of this study.

CHAPTER 4

DETERMINATION AND ESTIMATION OF ECONOMIC GROWTH USING RIDGE REGRESSION TECHNIQUE

4.1 Introduction

In order to predict economic growth (RGDP) using the identified economic growth drivers, the OLS estimation technique is often used (Gujarati, 2003; Ebiwonjumi *et al*., 2022). Lukman *et al*, (2015) further emphasized that, in the presence of multicollinearity and outliers, OLS estimators are unstable and rugged to interpret, while outliers can significantly influence the coefficients of the model, leading to unreliable results. Thus, we propose a ridge regression model to estimate and predict reliable economic growth (RGDP) in the presence of multicollinearity and outliers. This involves deliberately using a biased estimating method or procedure to improve the accuracy of the estimated economic growth parameters. According to Hoerl and Kennard (1970) and cited by Ebiwojumi *et al*. (2022), it can be stressed that if the mean square error (MSE) or root mean square error (RMSE) performance metric criterion is to be used to assess the of the accuracy of the model, ridge regression model generates better and accurate estimates than the unbiased OLS estimates. Also, William and Bruce (2015) opined that the ridge regression method generates better and more promising results than deliberately deleting relevant variable(s) to avoid multicollinearity problem, and it also provides smaller mean square error (MSE) or root mean square error (RMSE) estimates than unbiased OLS method.

Simionescu, *et al*. (2017) analyzed factors that determine stable economic growth in the V4 countries (Czech Republic, Slovak Republic, Hungary, Poland) and Romania country between the period (2003-2016). A ridge regression method was employed for the study, and the results revealed that foreign direct investment promoted economic growth in all countries except the Slovak Republic. It was also revealed that the expenditure on education influenced economic growth in the Czech Republic, while the expenditure on human capital positively affected economic growth in Romania, Hungary and the Czech Republic. Zhang and Yuan (2018) analyzed factors that affect the proportion of tertiary industry in achieving sustainable and stable national economic growth in Beijing (2000-2015). In the study, six factors included the proportion of labour productivity, the proportion of employed people, the proportion of fixed-asset investment, the proportion of the actual utilization of foreign capital, the proportion of total energy consumption, and the proportion of the resident population. The ridge regression method used for analysis showed that the identified six factors significantly affected the development of tertiary industry. It was stressed that energy, population and investment were the most influential factors affecting national economic growth.

Wang *et al*. (2023) investigated the impact of technological standardization on economic growth by examining the magnitude effect of technological innovation as an intermediary variable in this pathway and formulating corresponding hypotheses. Considering the specific context of China and the evolution of standardization, the study focused on encompassing technological standardization, technological innovation, and economic growth within a unified framework, utilizing China's macroeconomic data between the period of (1989-2019). A ridge regression method was applied due to the presence of multicollinearity among the explanatory and mediation variables. The results revealed that more stable parameter values are obtained. Also, a comparative international analysis revealed that technological standardization as an intermediate variable pathway enhanced economic growth. Wang and Zhang (2023) carried out a study on modern economic management analysis of commercial insurance by investigating factors that affect households' willingness to purchase commercial insurance from the perspective of household expenditure using China family panel study survey data. A ridge regression method was employed for the study due to a strong correlation that showed evidence of a multicollinearity problem among the factors such as the annual expenditure of a family on food, beauty, tourism, education and training and medical healthcare. Thus, it was revealed from the results that the aforementioned factors significantly affect the households' willingness to purchase commercial insurance. Hence, there is a need for the development of commercial insurance based on the use of big data to achieve private customization, increase innovation, and face the development gap and rational control.

However, in the various studies and literature reviewed for this study, we are unable to find a study that estimates the parameters of the identified economic growth drivers such as internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN) and efficiently predict economic growth (RGDP) in Nigeria using ridge regression method. This is a gap worth filling to enhance the estimates' efficiency and accuracy and predict stable and reliable values for economic growth (RGDP) in Nigeria.

4.2 Research Methodology

This section explains the methodology of ridge regression for handling multicollinearity.

4.2.1 Ridge Regression Technique

Consider a general ridge regression model (Hoerl and Kennard, 1970) expressed as:

$$\hat{\beta}_{RR} = (X'X + KI_P)^{-1}\, X'Y, \qquad k \geq 0 \;, \tag{4.1}$$

In this case, 0 < K < 1. This is a biased technique adopted to reduce the estimated variances of the parameters. According to Swindel (1976) a ridge regression (RR) technique for parameter estimation can be stated by:

$$\hat{\beta}_{RR,b} = (X'X + KI_P)^{-1}\, X'Y + Kb, \qquad k \geq 0 \;, \tag{4.2}$$

where b is a prior estimate of β. As K increases to one, the RR estimator approaches b. Furthermore, an unbiased ridge regression (URR) estimation method defined by Crouse *et al.* (1995) is given as:

$$\hat{\beta}_{UR,J} = (X'X + KI_P)^{-1}\, X'Y + KJ, \qquad k \geq 0 \,, \tag{4.3}$$

where $J \sim N(\beta, \frac{\sigma^2}{K} I_P)$. for K > 0.

To consider the spectral decomposition of $X'X$, the model stated in (4.1) can be transformed and express in (4.4) using $X'X = T\Lambda T'$, where $TT' = T'T = I$.

$$Y = XTT'\beta + \epsilon \;,$$

$$Y = Z\alpha + \epsilon, \;\; , \tag{4.4}$$

with $Z = XT$, $\alpha = T'\beta$ where $Z'Z = T'X'XT = \Lambda = \text{diag}\,(D_1, D_2, D_3, \dots, D_P)$. The diagonal elements of Λ are the eigenvalues of $X'X$ and T consist of corresponding eigenvectors of $X'X$. Hence, the transformed OLS, ORR and URR with parameter α based on the spectral decomposition of $X'X$ are written as y

$$\hat{\alpha}_{OLS} = \Lambda^{-1}Z'Y \;\; , \tag{4.5}$$

$$\hat{\alpha}_{RR} = (\Lambda + KI_P)^{-1}Z'Y \;\; , \tag{4.6}$$

$$\hat{\alpha}_{UR,J} = (\Lambda + KI_P)^{-1}(Z'Y + KJ). \tag{4.7}$$

According to Batah and Gore (2009), a modified unbiased ridge regression estimation technique can be derived from unbiased ridge regression (URR). Similarly, ordinary ridge regression (ORR) is derived from the ordinary least squares (OLS) regression technique. This modification aims to reduce the inflated variances that arise when eliminating bias using URR. The underlying logic involves pre-multiplying the URR by the matrix $[I - K(X'X + KI_P)^{-1}]$ in order to obtain the estimator of β. Recalled (4.3) given as:

$$\hat{\beta}_J(K) = [I - K(X'X + KI_P)^{-1}]\,\hat{\beta}_{UR,J} \quad ,$$

$$\hat{\beta}_J(K) = [I - K(X'X + KI_P)^{-1}]\,((X'X + KI_P)^{-1}(X'Y + KJ) \;, \tag{4.8}$$

where $J \sim N(\beta, \frac{\sigma^2}{K} I_P)$ and K > 0. This estimator is called modified unbiased ridge regression because it is obtained from URR. The modified unbiased ridge estimation technique, through spectral decomposition and transformation, can be written as:

$$\hat{\alpha}_J(K) = [I - K(X'X + KI_P)^{-1}]\,\hat{\alpha}_{UR,J}, \tag{4.9}$$

The following properties of techniques can be obtained:

bias of $\hat{\alpha}_J(K)$

$$bias\left(\hat{\alpha}_J(K)\right) = E\left(\hat{\alpha}_J(K)\right) - \alpha \,,$$

$$bias\left(\hat{\alpha}_J(K)\right) = -KG_K^{-1}\beta \quad , \tag{4.10}$$

where $G = X'X \; and \; G_K = (G + KI)$,

Variance-Covariance Matrix of $\hat{\alpha}_J(K)$

$$Var\left(\hat{\alpha}_J(K)\right) = E\left[\left(\hat{\alpha}_J(K)\right) - E\left(\hat{\alpha}_J(K)\right)\right)\left(\hat{\alpha}_J(K) - E\left(\hat{\alpha}_J(K)\right)\right)' ,$$

$$Var\left(\hat{\alpha}_J(K)\right) = \sigma^2 W G_K^{-1} W', \tag{4.11}$$

where $W = [I - KG_K^{-1}]$.

Matrix Mean Squared Error (MMSE) of $\hat{\alpha}_J(K)$

$$MMSE\left(\hat{\alpha}_J(K)\right) = Var\left(\hat{\alpha}_J(K)\right) + \left[Bias\left(\hat{\alpha}_J(K)\right)\right]\left[Bias\left(\hat{\alpha}_J(K)\right)\right]' ,$$

$$MMSE\left(\hat{\alpha}_J(K)\right) = \sigma^2 W G_K^{-1} W' + K^2 G_K^{-1}\alpha\alpha' G_K^{-1} \quad , \tag{4.12}$$

Scalar Mean Squared Error (SMSE) of $\hat{\alpha}_J(K)$

$$SMSE\left(\hat{\alpha}_J(K)\right) = E\left(\hat{\alpha}_J(K) - \alpha\right)'\left(\hat{\alpha}_J(K)\right) - \alpha) \quad ,$$

$$SMSE\left(\hat{\alpha}_J(K)\right) = tr\left(MMSE\left(\hat{\alpha}_J(K)\right)\right) \,,$$

then

$$SSME\left(\hat{\alpha}_J(K)\right) = \sigma^2 \sum_{i=1}^{P} \frac{{D_i}^2}{(D_i + K)^3} + K^2 \sum_{i=1}^{P} \frac{(D_i + K){\alpha_i}^2}{(D_i + K)^3} \quad , \tag{4.13}$$

where $\{D_i\}$ are eigenvalues of $X'X$

$\hat{\alpha}_J(K = 0) = \hat{\alpha}_{LS} = (X'X)^{-1} X'Y$ is the OLS estimator

$lim_{K \to 0}\, \alpha(K) = \hat{\alpha}_{LS}$.

Estimation of Ridge Constant K

There are several methods for estimating ridge constant**.** According to Hoerl and Kennard (1970), Hoerl*et al*. (1975) and Crouse *et al.* (1995), this can be obtained based on the harmonic mean of optimal ridge constant values, which depends on the unknown parameters α and σ^2. Thus, the unknown parameters can be obtained using ordinary least square estimated values and, as such, the operational ridge constant given by Hoerl and Kennard (1970), that is, the value of K hat which minimizes the MSE can be expressed as:

$$\mathrm{K} = \frac{\hat{\sigma}^2}{\hat{\alpha}_{\max}} \quad , \tag{4.14}$$

where $\hat{\sigma}^2$ represents the error variance of the model, $\hat{\alpha}_{max}$ is the maximum among elements of α and it is defined as $\alpha = D'\hat{\alpha}$ with D being an orthogonal matrix. The estimated ridge constant called the (FG) ridge constant K can be expressed in the form given by

$$\hat{K}_{FG} = \frac{p\hat{\sigma}^2}{\sum_{i=1}^{p}\left[\hat{\alpha}_i^2 \Big/ \left[\left(\frac{\hat{\alpha}_i^4 D_i^2}{4\hat{\sigma}^4} + \frac{D_i \hat{\alpha}_i^2}{\hat{\sigma}^2}\right)^{1/2} - \frac{D_i \hat{\alpha}_i^2}{2\hat{\sigma}^2}\right]\right]} \quad , \tag{4.15}$$

Also, the HKB ridge constant by Hoerl *et al*. (1975) is given as

$$\hat{K}_{HKB} = \frac{p\hat{\sigma}^2}{(\hat{\alpha}'_{LS}\}'(\hat{\alpha}_{LS})} \quad , \tag{4.16}$$

Crouse et al. (1995) ridge constant K denoted by CJH ridge constant is expressed in the form given by

$$\hat{K}_{CJH} = \begin{cases} \dfrac{P\hat{\sigma}^2}{(\hat{\alpha}_{LS} - J)'\,(\hat{\alpha}_{LS} - J) - \hat{\sigma}^2 tr\ (X'X)^{-1}} & \text{if } (\hat{\alpha}_{LS} - J)'\,(\hat{\alpha}_{LS} - J) > \hat{\sigma}^2 tr\ (X'X)^{-1}, \\ \dfrac{P\hat{\sigma}^2}{(\hat{\alpha}_{LS} - J)'\,(\hat{\alpha}_{LS} - J)} & otherwise \end{cases} \quad , \tag{4.17}$$

where $\hat{\sigma}^2 = \frac{(Y - X\hat{\alpha}_{LS})'\,(Y - X\hat{\alpha}_{LS})}{n-p}$ is an unbiased estimator of σ^2 and $\hat{K}_{CJH}$ is a generalization of Hoerl *et al*. (1975) ridge regression constant K given in (4.16).

4.3 Empirical Results

In this section, we present the empirical results of the fitted ridge regression model discussed in Section 4.2 to capture multicollinearity among the explanatory variables. Table 4.1 shows the maximum likelihood parameter estimate of the OLS regression model.

Table 4.1: ML OLS Regression Results

	Parameter estimate	*p*-value
Intercept	7.4666	<0.0001
INDT	0.2817	<0.0001
EXDT	-0.1128	<0.0001
RINR	0.3885	<0.0001
REXR	0.0998	<0.0001
OPEN	0.5078	<0.0200
R-squared	0.9147	
Adjusted R-squared	0.9116	
F-statistic	289.6333	<0.0001

In Table 4.1, we present the results of the parameter estimation for the explanatory variables such as internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR), and trade openness (OPEN) used in assessing dependent variable (economic growth (RGDP)) in Nigeria. The OLS results reveal that INDT, EXDT, RINR, REXR, and OPEN are positively related to RGDP while, EXDT negatively related with RGDP. Specifically, INDT, RINR, REXR and OPEN increased economic growth RGDP by 28.2%, 38.8%, 9.9%, and 50.8%, respectively, while EXDT reduced the growth of the economy (RGDP) by 11.3%. The p-value < 0.05 shows the statistical significance of the estimated parameters for INDT, EXDT, RINR, REXR and OPEN. Furthermore, the adjusted R-square of 0.912 reveals that INDT, EXDT, RINR, REXR and OPEN as predictors can explain 91% of variations in RGDP. The F-statistic of 289.633 with p-value $<$ 0.05, which measured the overall significance of the model, revealed that the method was statistically significance in assessing the impact of INDT, EXDT, RINR, REXR and OPEN on RGDP in Nigeria. Thus, the high value of the coefficient of determination of 91%t from the Adjusted R-square, according to Granger (2003), is another evidence for the presence of multicollinearity. Also, since exploratory data analysis has revealed the presence of multicollinearity among the economic growth drivers identified for this study, we fit a ridge regression model to account for the multicollinearity with the appropriate ridge constant k. Thus, a ridge trace constant is plotted, as shown in Figure 4.1.

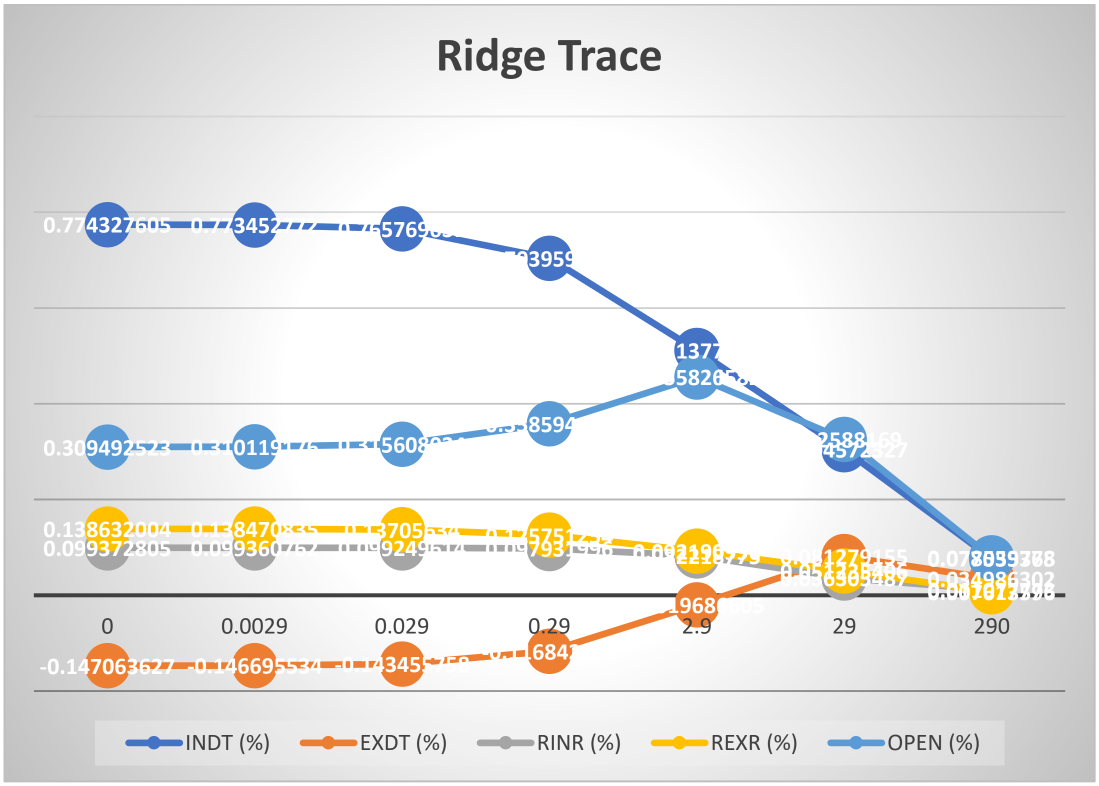


Figure 4.1: Ridge Trace Plot

In Figure 4.1, the ridge constant k that produces the most efficient estimate for the ridge regression model is chosen for the parameters estimate. Thus, we fit a ridge regression model with a ridge constant $k = 0.29$, as shown in Figure 4.1. it is observed from the ridge trace plot that each economic growth driver either contrasts or expands after the chosen ridge constant $k = 0.29$ to show its efficiency in fitting the ridge regression model for the results in Table 4.2.

Table 4.2: Ridge Regression Model Results

RGDP	Parameter estimate	*p*-value
INDT	0.7040	0.0000
EXDT	-0.1168	0.0621
RINR	0.0979	0.0301
REXR	0.1258	0.0096
OPEN	0.3586	0.0006

R-squared	0.8973	
Adjusted R-squared	0.8910	
F-statistic	147.7685	0.0000

In Table 4.2, we present results of the parameter estimate for the economic growth drivers such as internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN) that can be used to explore the economic growth rate (RGDP) in Nigeria. The parameters are estimated using the ridge regression technique to address the multicollinearity problem. From Table 4.2, it is revealed that INDT, RINR, REXR and OPEN are positively and significantly related to RGDP at 5% level of significance as earlier observed from the ordinary least square results presented in Table 4.1 except EXDT, which is negative and significant in relation to the RGDP. Specifically, the results indicate that INDT, RINR, REXR and OPEN increase Nigeria's economic growth rate (RGDP) by 70.4%, 9.8%, 12.6% and 35.9%respectively. In comparison, EXDT reduces Nigeria's economic growth rate (RGDP) by 11.7% during the period under consideration. The respective p-values < 0.05 for INDT, RINR, REXR and OPEN show the statistical significance of the estimated parameters of the ridge regression model while, p-value $<$ 0.10 shows the statistical significance of estimated parameter for EXDT using ridge regression method.

Also, the adjusted R-square of 0.891, a measure for the coefficient of determination, reveals that 89% variations in the economic growth rate (RGDP) can be explained by changes in internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN) respectively in Nigeria. The F-statistic value of 147.769 with p-values < 0.05, which determines the overall significance of the model, reveals that the ridge regression technique is statistically significance in examining the impact of INDT, EXDT, RINR, REXR and OPEN on RGDP in Nigeria. However, to examine whether the multicollinearity problem observed using the ordinary least square method has been addressed, the test for collinearity among the identified economic growth drivers (explanatory variables) for this study is conducted using variance inflation factor (VIF) and the results are presented in Table 4.3.

Table 4.3: Variance Inflation Factor for the Ridge Regression Technique

Variable	VIF for Ridge Regression
INDT	9.1430
EXDT	2.7334
RINR	1.3858
REXR	1.5347
OPEN	6.2876

In Table 4.3, we present the variance inflation factor (VIF) results to examine the presence of multicollinearity. The results reveal the VIF for INDT, EXDT, RINR, REXR and OPEN to be 9.1431, 2.7335, 1.3859, 1.5347 and 6.2877 respectively. Thus, it can be emphasized that the multicollinearity problem has been sorted out since all the VIF statistics < 10.00 for all the identified economic growth drivers (explanatory variables) are under investigation. Hence, the ridge regression technique and the selected ridge constant k = 0.29 optimally addressed the problem of multicollinearity in the data used for this study. It is also important to emphasize that this model efficiently predicts or forecasts the economic growth rate (RGDP) for the next quarters based on the parameter estimates obtained. The predictive performance metrics and the forecasting plot showing the forecast for the next ten (10) quarters are presented in Table 4.4 and Figure 4.2, respectively.

Table 4.4: Performance Metrics Evaluation for the Fitted Ridge Regression Model

Model Metrics for Ridge Regression model	Estimates
Root Mean Squared Error (RMSE)	0.2910
Mean Absolute Error (MAE)	0.2093
Mean Absolute Percentage Error (MAPE)	71.9538
Theil Inequality Coefficient (TIC)	0.1493
Bias Proportion (BiasP)	0.0000
Variance Proportion (VarP)	0.0223

Table 4.4 presents the criterion for selecting the optimal and efficient ridge regression technique employed in this study. This is determined by root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). Thus, from the results in Table 4.4, the

RMSE for the fitted ridge regression model is 0.2910. The MAE for the fitted ridge regression models is 0.2093, and the MAPE is 71.9538.

4.4 Forecast for RGDP using Ridge Regression Model

Figure 4.2 shows the plot of the ridge regression model's predictive efficiency in generating stable and reliable values for economic growth rate (RGDP) based on the data under consideration in this study.

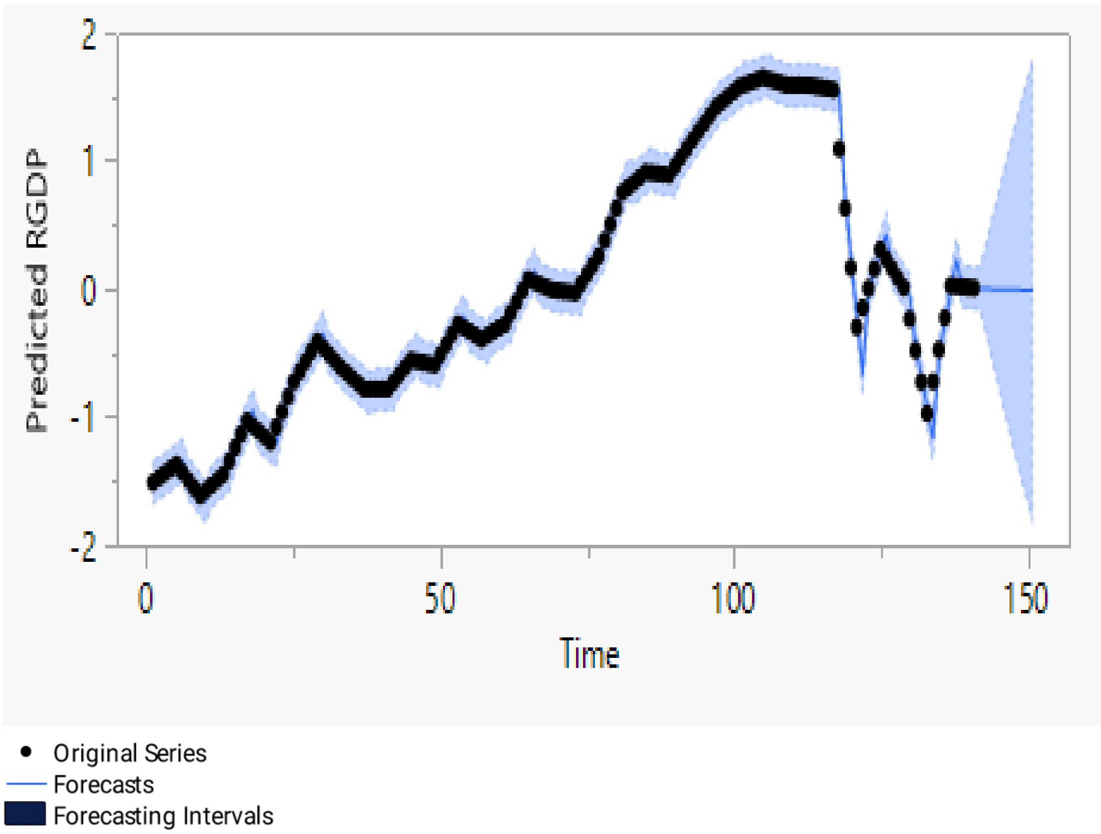


Figure 4.2: Forecast Plot for RGDP using Ridge Regression Technique

4.5 Concluding Remarks

An examination of an efficient estimator for the parameters of the economic growth (RGDP) and its determinants, such as INDT, EXDT, RINR, REXR and OPEN in the presence of multicollinearity, was thoroughly investigated using the ordinary least square and ridge regression techniques. The results show that multicollinearity is caused by the INDT, as revealed by the variance inflation factor. The study also observed that the ordinary least square estimation method underestimated the standard error of the parameters due to the multicollinearity problem among the identified economic growth drivers (explanatory variables) for this study. However, it is evident from the results obtained using the ridge regression technique with appropriate ridge constant *k* that the multicollinearity problem has been addressed, as the variance inflation factor for the economic growth drivers (explanatory variables) is all less than 10.0000. Also, the standard error of the parameters or the determinants of the RGDP is efficient and better estimated when the ridge regression constant k is 0.29, as indicated by the ridge trace. Hoerl and Kennard (1970) emphasized that a ridge regression model with an appropriate ridge constant chosen from the ridge trace statistic produces a minimum mean square of error. When the ridge regression constant k is zero, it is observed that the parameters estimated are the same as the ordinary least square regression method.

Therefore, based on the findings from the method, it can be emphasized that the ridge regression technique is efficient in estimating the parameters of the economic growth drivers under investigation in the presence of multicollinearity. Hence, this serves as a great benefit to the policymakers as the study provides a better understanding of the contributions of each of the identified or aforementioned economic growth drivers to the growth rate of the economy (RGDP), the aforementioned economic growth drivers and in particular the negative influence of external debt on the growth of the economy in Nigeria. Furthermore, it will significantly benefit the researchers by strengthening their understanding of the appropriate estimation technique when multicollinearity is established in the data under consideration. However, we aim to explore and model and suggest a better predictive model for economic growth (RGDP); given this, a robust principal component regression method shall be examined and discussed in Chapter 5 of this study.

CHAPTER 5

ROBUST PRINCIPAL COMPONENT METHOD FOR ESTIMATION OF ECONOMIC GROWTH PARAMETERS

5.1 Introduction

In this chapter, we propose a robust principal component regression method to simultaneously address the problem of multicollinearity and outliers in the data set, as established in Chapter 3 of this study. According to Draper and Smith (1981) and Myers (1986) as cited in Ebiwonjumi *et al.*(2023), a principal component can be used to address the multicollinearity problem associated with a predictive model. Thus, it can generate estimates and predictions better than ordinary least squares. Empirically. Adongo et al. (2018) studied macroeconomic variables, which many economists and researchers have long neglected for preliminary investigation to ascertain their similarities and differences. In the study, a principal component analysis was employed to assess nine macroeconomic variables, and the results revealed the level of redundancy among the variables in the correlation matrix. Also, Scree plots of principal components revealed the grouping of variables into three factors or components.

Alphonsus and Raji (2019) applied principal component analysis to reduce the dimension of Sixteen morphological variables with very high severe multicollinearity measured from 50 multiparous Bunaji cows. Thus, it was revealed that the collinear variables were combined to form composite scores, effectively reducing the morphological variables' dimension to four uncorrelated components. Therefore, principal component analysis can be used to address multicollinearity problems and dimension reduction of morphology. Nugrahadi, Maipita and Situmeang (2020) carried out a study to analyze the dominant factors affecting the optimal performance of the economic policy of the small business in the food, beverage and tobacco sector development in North Sumatra, Indonesia. Maipita and Situmeang (2020) employed the principal component analysis to examine investment, wage, inflation, exchange rate, population, total workers, industrial output growth, interest rate and total credit. The result revealed that population and credit are the two dominant factors.

Mbaluka et al. (2022) applied principal component analysis and hierarchical regression model on Kenyan Macroeconomic variables. The study considered 18 macroeconomic variables data that were extracted from Kenya National Bureau of Statistics and World Bank for the period (1970-2019). The principal component analysis was used to reduce the dimensionality of the data, and a hierarchical regression model was fitted on the extracted components to determine and predict economic growth. However, this study aims to estimate and predict RGDP based on its identified economic drivers in the presence of multicollinearity and outliers using a robust principal component regression method. To our knowledge, this technique has not been used within the Nigerian economic context. Hence, this study aims to fill a gap that exists in the literature.

5.2 Research Methodology

In this section, we discuss a robust principal component regression to address the problems of multicollinearity and outliers. This study uses three robust estimators to extract robust principal components The robust estimators, namely the M-estimator, S-estimator and the MM-estimator, are discussed.

5.2.1 M-Estimators

The most common general method of robust regression is M-estimation, introduced by Huber (1973), which is nearly as efficient as OLS. Instead of minimizing the sum of squared errors, M-estimation minimizes a function of the errors. The objective function for the M-estimate is given as:

$$min \sum_{i=1}^{n} p \frac{e_i}{S} = \min \sum_{i=1}^{n} \rho \left(\frac{Y_i - X' \hat{\beta}_i}{S} \right), \tag{5.1}$$

S is an estimate of scale often formed from a linear combination of the residuals. The function gives the contribution of each residual to the objective function. A reasonable ρ should have the following properties:

$\rho(e) \geq 0,\ \rho(0) = 0, \rho(e) = \rho(-e))$, and $(\rho(e_i) \geq\ \rho\left(e_i^{;}\right) for\ |e_i| \geq\ \left|e_i^{;}\right|$.

The system of normal equations to solve this minimization problem is found by taking partial derivatives with respect to β and setting them equal to 0, yielding the expression given by:

$$\sum_{i=1}^{n} \varphi \left(\frac{Y_i - X' \hat{\beta}_i}{S} \right) X_i = 0\ , \tag{5.2}$$

where φ is a derivative of ρ. The choice of the φ function is based on the preference of how much weight to assign outliers. Newton-Raphson and Iteratively Reweighted Least Squares (IRLS) are the two methods to solve the M-estimates in nonlinear normal equations. IRLS expresses the normal equations given in the form expressed as:

$$X'\varphi X\hat{\beta}_i = X'\varphi Y \,. \tag{5.3}$$

5.2.2 S-Estimator

According to Rousseeuw and Yohai (1984), S estimators are derived from a scale statistic in an implicit way, corresponding to $s(\theta)$ where $s(\theta)$ is a certain type of robust M-estimate of the scale of the residuals $e_1(\theta),\ e_2(\theta), \dots ,\ e_n(\theta)$. They are defined by minimization of the dispersion of the residuals:

minimize $S(e_1(\theta),\ e_2(\theta), \dots ,\ e_n(\theta))$ with final scale estimate

$\hat{\sigma} = S(e_1(\theta),\ e_2(\theta), \dots ,\ e_n(\hat{\theta}))$.

The dispersion $e_1(\theta),\ e_2(\theta), \dots ,\ e_n(\theta)$ is defined as the solution of the expression given by:

$$\frac{1}{n}\sum_{i=1}^{n} P\left(\frac{e_i}{S}\right) = K \,, \tag{5.4}$$

where K is a constant and $P\left(\frac{e_i}{S}\right)$ is the residual function. Based on Rousseeuw and Yohai (1984), Tukey's bi-weight function was suggested and is defined by Setting c = 1.5476 and K = 0.1995 to give a 50% breakdown point as expressed (Rousseeuw and Leroy, 1987).

$$p(x) = \begin{cases} \dfrac{x^2}{2} - \dfrac{x^4}{2c^2} + \dfrac{x^6}{6c^4} & for\ |x| \le c \\ \dfrac{c^2}{6} & for\ |x| > c \end{cases} \,. \tag{5.5}$$

5.2.3 MM-Estimator

MM-estimation is a special type of M-estimation (Yohai, 1987), which combines the high asymptotic relative efficiency of M-estimators with the high breakdown of a class of estimators called S-estimators. The MM-estimator was among the first robust estimators to have these two properties simultaneously. It refers to the fact that multiple M-estimation procedures are carried out in the computation of the estimator. MM-estimator was described in three stages as follows:

Stage 1. A high breakdown estimator is used to find an initial estimate, which we denote $\tilde{\beta}$. The estimator needs to be efficient. Using this estimate, the residuals are computed by:

$$r_i(\beta) = y_i - x_i^T \tilde{\beta} \,, \tag{5.6}$$

Stage 2. Using these residuals from the robust fit and where K is a constant and the objective is given as:

$$\frac{1}{n}\sum_{i=1}^{n} \rho\left(\frac{r_i}{s}\right) = K \,, \tag{5.7}$$

function ρ, an M-estimate of scale with 50% breakdown point is computed. This $s(r_1(\tilde{\beta}), r_2(\tilde{\beta}),$ $\ldots r_n(\tilde{\beta}))$ is denoted s_n

The objective function used in this stage is labeled ρ_0.

Stage 3. The MM-estimator can now be defined as an M-estimator of β using a re-descending score function expressed as:

$$\emptyset_1(u) = \frac{\partial \rho_1(u)}{\partial u} \,, \tag{5.8}$$

and the scale estimate s_n obtained from stage 2. Thus, an MM-estimator $\hat{\beta}$ is defined as a solution to expression given by:

$$\sum_{i=1}^{n} x_{ij} \, \emptyset_1\left(\frac{y_i - x_i^T \tilde{\beta}}{s_n}\right) = 0; \;\; j = 1, 2 \ldots, p \;\; . \tag{5.9}$$

5.3 Robust Principal Component Regression Technique

This technique provides a unified way to handle multicollinearity, requiring some calculations that are not usually included in standard regression analysis (Ebiwonjumi *et al*. 2023). The principal component analysis follows from the fact that every linear regression model can be restated in terms of a set of orthogonal explanatory variables (Oyewole and Agunbiade (2020). These new variables are obtained as linear combinations of the original explanatory variables, which are referred to as the principal components (Oyewole and Agunbiade, 2020; Ebiwonjumi *et al*., 2023). Consider Y to be an $n \times 1$ matrix of response variable, X is an $n \times p$ matrix of the independent variables, β is a $p \times 1$ vector of unknown constants and ε is an $n \times 1$ vector of random errors. Thus, there exists a matrix V, satisfying the expression given as:

$$V'(X'X)V = \Lambda \text{ and } V'V = VV' = 1 \,, \tag{5.10}$$

where Λ is a diagonal matrix with ordered characteristics roots of $X'X$ on the diagonal. The characteristic roots are denoted by $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p$. V may be used to calculate a new set of explanatory variables as given and expressed as:

$$Z = \left(Z_{(1)}, Z_{(2)}, \dots, Z_{(p)}\right) , \tag{5.11}$$

$$XV = \left(X_{(1)}, X_{(2)}, \dots, X_{(p)}\right) , \tag{5.12}$$

which are linear functions of the original explanatory variables. The $Z's$ are referred to as principal components. Hence, the regression model can be restated in terms of the principal components as given by the expression:

$$Y = Z\alpha + \varepsilon , \tag{5.13}$$

$$Z'Z = V'X'XV = V'V\Lambda V'V, \tag{5.14}$$

where $Z = XV$ and $\alpha = V\beta$. The least square estimator of α and the variance covariance matrix of $\hat{\alpha}$ are given as:

$$\hat{\alpha} = (Z'Z)^{-1} Z'Y = \Lambda^{-1} Z'Y , \tag{5.15}$$

$$Var(\hat{\alpha}) = \sigma^2 (Z'Z)^{-1} = \sigma^2 \Lambda^{-1} . \tag{5.16}$$

Thus, a small eigenvalue of $X'X$ implies that the variance of the corresponding regression coefficient will be large. From expression in (5.10) and (5.14) which can be put as:

$$Z'Z = V'X'XV = V'V\Lambda V'V = \Lambda , \tag{5.17}$$

is often referred to as the eigenvalues λ_j as the variance of the jth principal component. If all λ_j equal to unity, the original regressors are orthogonal, while if a λ_j is exactly equal to zero, then it implies a perfect linear relationship between the original regressors. The principal component regression approach combats multicollinearity by using less than the full set of principal components in the model. To obtain the principal components estimators, assume that the regressors are arranged in order of decreasing eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p > 0$. In principal components regression the principal components corresponding to near zero eigenvalues are removed from the analysis and least squares method is applied to the remaining components as in (5.13) that can be expressed as:

$$Y = XVV'\beta + \varepsilon , \tag{5.18}$$

$$Y = Z\alpha + \varepsilon , \tag{5.19}$$

where $Z = XV, \alpha = V'\beta$ and $V = \left(v_1, v_2, v_3, \dots, v_p\right) = (V_r, V_{p-r})$ is a $p \times p$ orthogonal matrix with the expression given by:

$$\left(V_r, V_{p-r}\right)' X'X\left(V_r, V_{p-r}\right) = A = \begin{pmatrix} A_r & 0 \\ 0 & A_{p-r} \end{pmatrix}, \quad (5.20)$$

where $0 < r \leq p$, $\mathrm{A} = diag\left(v_1, v_2, v_3, \dots, v_p\right)$, $A_r = diag(v_1, v_2, v_3, \dots, v_r)$, $A_{p-r} = diag\left(v_{r+1}, v_{r+2}, v_{r+3}, \dots, v_p\right)$ and $v_1 \geq v_2 \geq v_3 \geq \dots \geq v_p > 0$ are the ordered eigenvalues of $X'X$. By definition, expression obtained from (5.18) can be put as:

$$Z = XV = \left(Z_r, Z_{p-r}\right). \quad (5.21)$$

is the $n \times p$ matrix of the principal components, where $Z_i = XV_i$ is the ith principal component. Considered Z_{p-r} that contains principal components corresponding to near zero eigenvalues, which implies that Z can be partitioning into Z_r and Z_{p-r}, such that Z_{p-r} are eliminated. Thus, expression in (5.19) can further be written and expressed as:

$$Y = Z_r \alpha_r + Z_{p-r} \alpha_{p-r} \alpha + \varepsilon, \quad (5.22)$$

The least square estimator of α is given as:

$$\hat{\alpha} = (A)^{-1} Z'Y, \quad (5.23)$$

From expression in (5.3) and (5.23), the M-estimator of α can be given as:

$$\hat{\alpha}_M = \left(A_\varphi\right)^{-1} Z' \hat{\varphi} Y \quad (5.24)$$

where $A_\varphi = Z' \hat{\varphi} Y, \hat{\varphi}$

is the derivative of p. Based on (5.22) and (5.23), the principal component estimator of α can be expressed in the form given by:

$$\hat{\alpha}_{PC} = (A_r)^{-1} Z_r{}' Y. \quad (5.25)$$

To address multicollinearity among the economic growth drivers in this study, the principal components with the highest cumulative proportion of variance explained are selected to replace the original variables. The results and discussions of various analyses, including a robust principal component regression model, its associated diagnostic tests, and the evaluations conducted, are presented in the next section.

5.4 Empirical Results

This section presents the empirical results of the fitted models discussed in Section 5.2. To address the multicollinearity among the explanatory variables, principal component analysis was performed to extract the principal components. Table 5.1 shows the results of the principal component analysis conducted for this study.

Table 5.1: Principal Components Analysis

Principal Components Analysis					
Eigenvalues: (Sum = 5, Average = 1)					
Number	Value	Difference	Variance Proportion	Cumulative Variance Explained	
PC1	2.2456	0.6149	0.4491	0.4491	
PC2	1.6307	1.0742	0.3261	0.7753	
PC3	0.5564	0.0544	0.1113	0.8866	
PC4	0.5020	0.4369	0.1004	0.9870	
PC5	0.0650	---	0.0130	1.0000	
Ordinary correlations:					
	INDT	EXDT	RINR	REXR	OPEN
INDT	1.0000				
EXDT	0.6073	1.0000			
RINR	0.1275	0.3875	1.0000		
REXR	-0.0923	-0.2898	-0.4594	1.0000	
OPEN	0.8200	0.2534	-0.0089	0.2416	1.0000

Table 5.1 reports the results of principal component analysis used to deal with multicollinearity in the dataset and extract non-correlated principal components (PCs). This reduction is achieved by using a subset of the principal components to explain the variation in economic growth (RGDP). Specifically, Table 5.1 provides the principal components for the explanatory variables INDT, EXDT, RINR, REXR, and OPEN, which are as follows:

$$\mathrm{PC1} = 0.6134\mathrm{INDT} + 0.5292\mathrm{EXDT} + 0.2921\mathrm{RINR} - 0.1844\mathrm{REXR} + 0.4736\mathrm{OPEN}, \quad (5.26)$$

$$\mathrm{PC2} = 0.2382\mathrm{INDT} - 0.1930\mathrm{EXDT} - 0.5331\mathrm{RINR} + 0.6258\mathrm{REXR} + 0.4796\mathrm{OPEN}, \quad (5.27)$$

$$\mathrm{PC3} = -0.1676\mathrm{INDT} - 0.2802\mathrm{EXDT} + 0.7863\mathrm{RINR} + 0.4716\mathrm{REXR} + 0.2290\mathrm{OPEN}, \quad (5.28)$$

$$\mathrm{PC4} = -0.1638\mathrm{INDT} + 0.7211\mathrm{EXDT} - 0.0072\mathrm{RINR} + 0.5622\mathrm{REXR} - 0.3701\mathrm{OPEN}, \quad (5.29)$$

$$\mathrm{PC5} = 0.7155\mathrm{INDT} - 0.2899\mathrm{EXDT} + 0.1097\mathrm{RINR} + 0.1889\mathrm{REXR} - 0.5968\mathrm{OPEN}. \quad (5.30)$$

Also, from Table 5.1 and based on the Kaiser's rule of thumb that says the principal component with the eigenvalues greater than 1 must be considered for fitting an appropriate model, it is found from the results that two components which are PC1 and second PC2 components account for 44.9% and 32.6% of the variance and as such the remaining components are insignificance. The scree plot presents in Figure 5.1 also establishes the significance of selecting PC1 and PC2 as the two components for fitting appropriate of robust principal component models for the prediction of RGDP for this study.

The principal component scree plot and cumulative proportion plot for the selection of component is shown in Figure 5.1

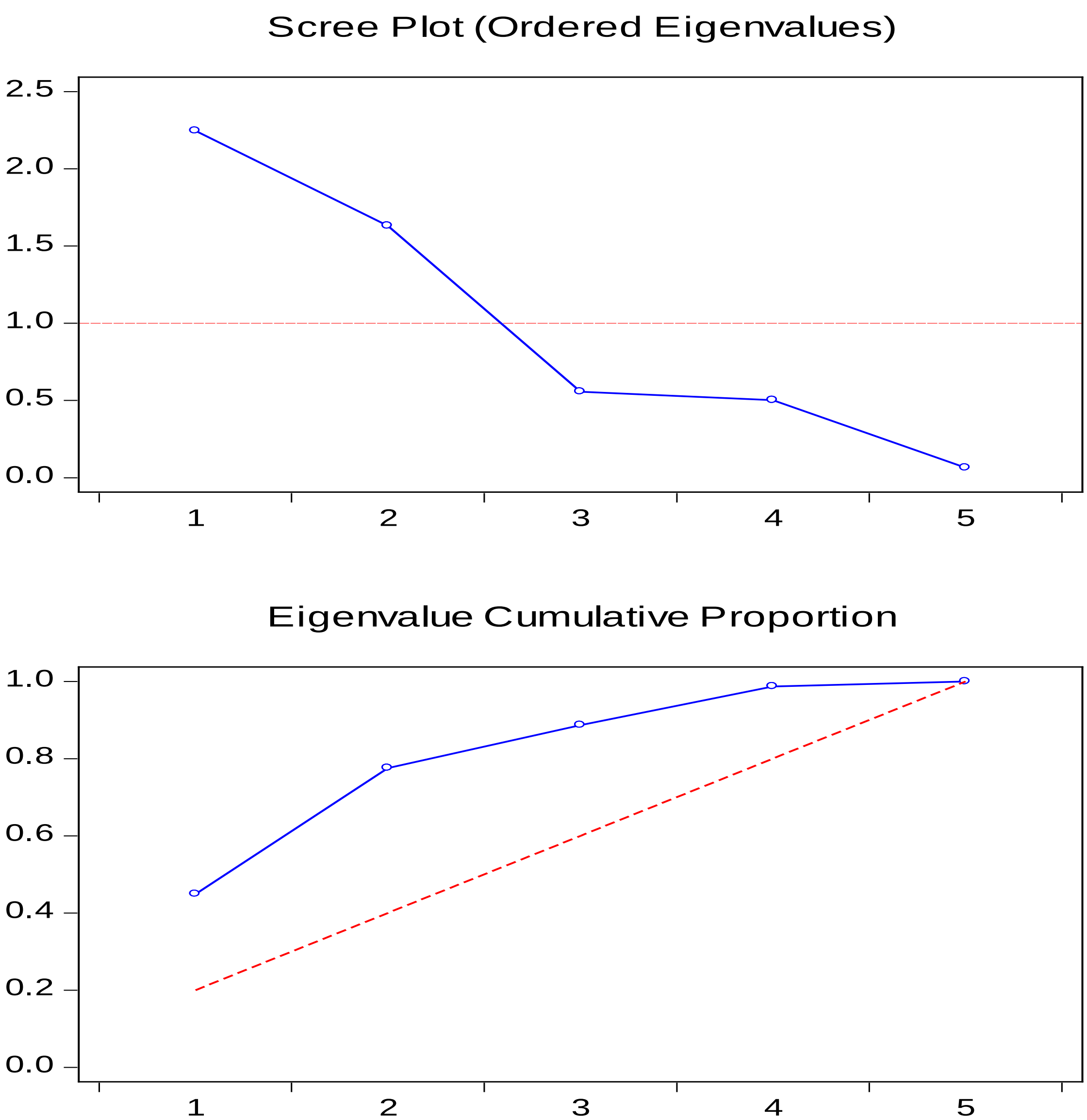


Figure 5.1: The Scree Plot and Cumulative Proportion Plot Showing the Number of Component of Economic Growth Drivers to be Selected

In Figure 5.1, we present the scree plot and the cumulative proportion of the eigenvalue of the components of the identified economic growth drivers. The results show that PC1 with eigenvalue

2.2456 and PC2 with the eigenvalue 1.6307 have the highest variance proportion estimates of 44.9% and 32.6% and with the cumulative variance proportion that accounts for 77.5% of the generated components of the identified economic growth drivers when compared with others estimated components such as PC3, PC4, and PC5 with eigenvalue that are less that one with their total variance cumulative proportion of 22.5%.

Thus, based on the principal component eigenvalues presented in Table 5.1 and the scree plot shown in Figure 5.1 the appropriate principal component analysis model to be fitted for the RGDP and the selected principal components PC1 and PC2 respectively can written in the form given by:

$$\mathrm{RGDP_i} = \propto_0 + \propto_1 \mathrm{PC1} + \propto_2 \mathrm{PC2} + \epsilon_i, \tag{5.31}$$

Hence, the results of the fitted robust principal components for determining RGDP in this study are presented in Table 5.2

Table 5.2: Robust Principal Component Regression (PCR) Results

	M-estimator		S-estimator		MM-estimator	
Variable	Estimate	*p*-value	Estimate	*p*-value	Estimate	*p*-value
C	10.2851	0.0000	10.3244	0.0000	10.2847	0.0000
PC1	0.3539	0.0000	0.3715	0.0000	0.3546	0.0000
PC2	0.2215	0.0000	0.2122	0.0000	0.2224	0.0000
R-squared	0.6617		0.7125		0.7031	
Adjusted R-squared	0.6568		0.7083		0.6988	
Rw-squared	0.9248				0.9172	
Adjust Rw-squared	0.9248				0.9172	
Rn-squared statistic	1048.933	0.0000	1316.421	0.0000	1076.843	0.0000

In Table 5.2, we present principal component regression results for the estimated parameters of the derived variables, principal component one (PC1) and principal component two (PC2), which are used to drive Nigeria's economic growth (RGDP). It must be noted that in this study, quarterly data that contained multicollinearity and outliers are used and based on this, a robust M-principal component regression (M-estimator) of weighted bisquare with 4.685 turning and median centred as scale, a robust S-principal component regression (S-estimator) with 1.5476 turning with 0.5 breakdown for 200 trials and a robust MM-principal component regression (MM-estimator) with

S turning of 1.5476 with 0.5 breakdown for 200 trials and M weighted bi-square with 4.685 turning designed to address the two identified problem of multicollinearity and outliers are fitted.

Table 5.2 presents robust results from the M-estimation method, indicating that PC1 and PC2 positively impact RGDP, increasing it by 35.39% and 22.15%, respectively, during the period under consideration. The p-values for these estimated principal components are less than 0.05, demonstrating their statistical significance as economic growth drivers in assessing RGDP in Nigeria. Similarly, the robust S-estimation method shows that PC1 and PC2 positively affect RGDP, with increases of 37.15% and 21.22%, respectively. These results are statistically significant, as indicated by p-values less than 0.05 for the estimated parameters of the principal components. Furthermore, the results from the robust MM-estimation method reveal a positive relationship between PC1 and PC2 and RGDP. Specifically, PC1 and PC2 lead to increases in RGDP of 35.46% and 22.24%, respectively, during the study period. The p-values for these principal components are also less than 0.05, confirming their statistical significance in determining RGDP in Nigeria.

Also, the adjusted R-square of 0.6568, 0.7083 and 0.6988 for the robust principal component regression M-estimator, S-estimator and MM-estimator reveal that 65.68%, 70.83% and 69.88% of proportional changes in RGDP can be explained by changes in PC1 and PC2 using the respective estimation methods above and as such emphasize the significance of the method in addressing the problem of multicollinearity and outlier in this study. The Rn-squared statistic of 1048.933, 1316.421 and 1076.843 with p-value < 0.05 are used to examine the overall significance of the fitted M-estimator, S-estimator and MM-estimator models and thus show the statistical significance of the robust principal component method in examining and predicting RGDP using PC1 and PC2 in the presence of multicollinearity and outliers. The test to confirm that the multicollinearity problem has been addressed is presented in Table 5.3.

Table 5.3: Test for the Absence of Multicollinearity

Variable	Coefficient Variance	VIF
PC1	0.00026	1.0000
PC2	0.00026	1.0000

Table 5.3 shows the result of the variance inflation factor (VIF) for examining the presence of multicollinearity after the extraction and selection of the appropriate principal components, such as PC1 and PC2, to serve as economic growth drivers (explanatory variables) to replace the original economic growth drivers that include INDT, EXDT, RINR, REXR and OPEN. In this study, PC1 and PC2 are the principal components with the highest proportion of eigenvalue that are appropriated and selected for the prediction of RGDP. Thus, there is a need to test for the presence of multicollinearity. Hence, it is found from Table 5.3 that VIF for PC1 and PC2 are 1.0000 and 1.0000 respectively. Thus, according to Allison (1999) and Freund and Littell (2000), as cited in Khalaf and Iguernane (2016), it was posited that the VIF < 10.00 indicates the absence of multicollinearity. Based on this evidence, it can be stressed that the multicollinearity problem has been addressed from the fitted robust principal component models, and as such, the estimated parameters of models are efficient and optimal for the prediction of RGDP in this study for Nigeria. Also, to determine the most efficient robust principal component estimator among the fitted robust estimators, the predictive power or performance evaluation metrics of the robust principal component estimators are considered, and the results are presented in Table 5.4.

Table 5.4: Performance Metrics Evaluation of the Robust Estimators

Forecast RGDP	M-estimator	S-estimator	MM-estimator
Root Mean Square Error	0.1927	0.1929	0.1928
Mean Absolute Error	0.1319	0.1321	0.1320
Mean Abs Percent Error	1.2672	1.2675	1.2686
Theil Inequality Coefficient	0.0093	0.0093	0.0093
Bias Proportion	0.0103	0.0104	0.0107
Variance Proportion	0.0284	0.0132	0.0265
Covariance Proportion	0.9612	0.9763	0.9627

In Table 5.4, we present the forecasting power or performance evaluation metrics of the fitted robust principal component models that simultaneously addressed the presence of multicollinearity and outliers in this study. From the results in Table 5.4, it is revealed that the root mean square error (RMSE) of the M-estimator with a value of 0.1927 is the smallest in comparison with the S-estimator and MM-estimator with values of 0.1929 and 0.1928 respectively. The mean absolute error (MAE) of the fitted robust principal estimation method reveals that the M-estimator has the smallest mean absolute error value of 0.1319 if compared with the S-estimator and MM-estimator with values of 0.1321 and 0.1320 respectively. Also, the mean absolute percentage error (MAPE) of the fitted robust principal component estimation technique shows that the M-estimator with the value of 1.2672 has the smallest value in comparison with the S-estimator and MM-estimator with a MAPE value of 1.2675and 1.2686 respectively. The same results are obtained using bias proportion that shows that the M-estimator with the value 0.0103 is the smallest in comparison with the S-estimator and MM-estimator with values 0.0104 and 0.0107, respectively. The Theil inequality coefficient and the variance proportion also reveal the same result, and as such, it can be asserted that the M-estimator is the most efficient robust principal component estimation model for addressing the presence of multicollinearity and outlier jointly in modelling and predicting RGDP using PC1 and PC2 extracted from the leading economic growth drivers that include internal debt, external debt, interest rate, exchange rate and trade openness in Nigeria. Thus, the forecast for the next ten (10) quarters of the RGDP, as shown in Figure 5.2, is carried out using the M-estimator as the most efficient and robust principal component model.

5.5 Forecast for the RGDP using Robust Principal Component Estimator

Figure 5.2 illustrates the predictive efficiency of the fitted M-estimator model in generating stable and reliable RGDP values based on the data considered in this study for the next ten quarters. The plot in Figure 5.2 represents this forecast.

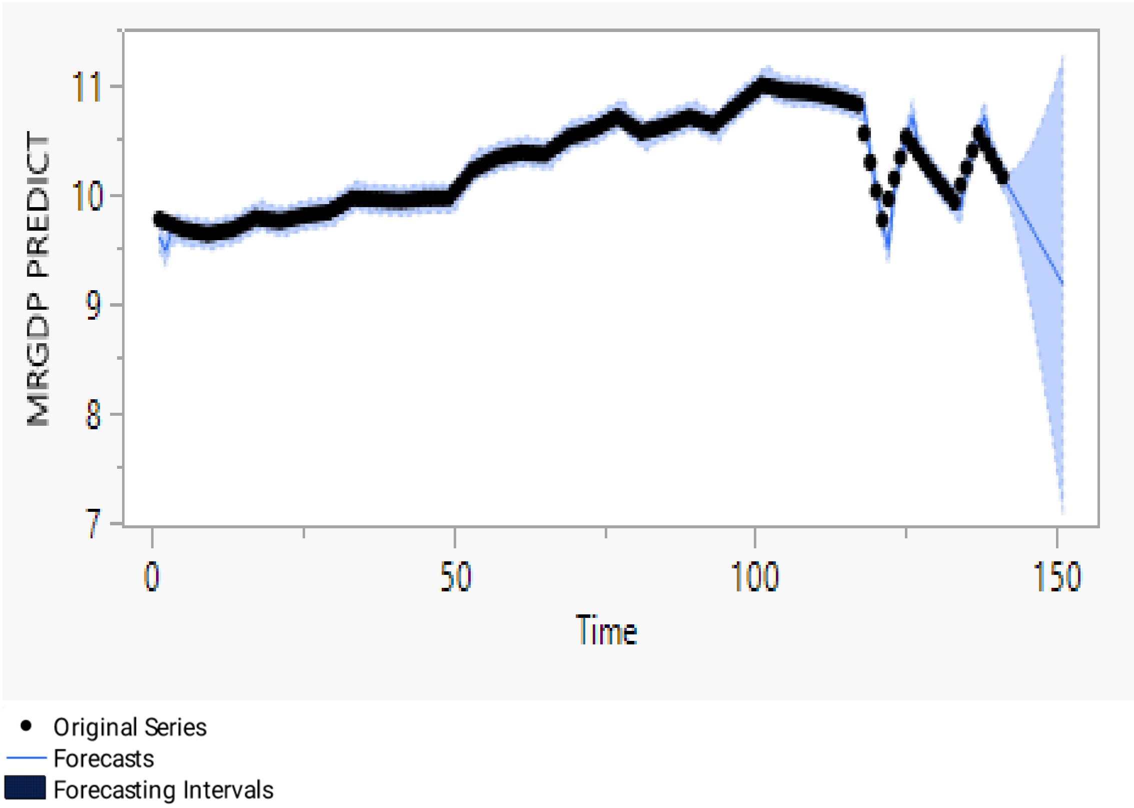


Figure 5.2: Forecast Plot using Robust Principal Component Regression Method

5.5 Concluding Remarks

An examination of an estimation of economic growth's parameters in Nigeria based on the identified determinants (INDT, EXDT, RINR, REXR, and OPEN) in the presence of multicollinearity and outliers as basic assumptions violation. An exploratory and diagnostic analysis establishes the relationship between the RGDP and the drivers above or determinants. The VIF and Grubb's test carried out on the macroeconomic data set under consideration established the presence of multicollinearity and outlier. Consequently, to jointly handle the problems and to obtain efficient parameter estimates for INDT, EXDT, RINR, REXR and OPEN on RGDP, a robust principal component analysis technique was employed to produce efficient estimates. The principal component analysis that addressed the multicollinearity problem generated two principal components, which are PC1 and PC2. Thus, a robust regression that includes M-estimator, S-estimator and MM-estimator models is fitted using PC1 and PC1 as explanatory variables on RGDP, which produce efficient parameter estimates for the prediction of RGDP in this study.

However, it is revealed that the M-estimator is the most efficient and optimal estimation technique for investigating the impact of PC1 and PC2 on RGDP in Nigeria during the period under study. This assertion is based on the root mean square error (RMSE) of the robust principal component regression model (M-estimator) value of 0.1927, which is smaller than the values obtained when compared with the S-estimator and M-estimator, respectively. The M-estimator's mean absolute error (MAE) also has the smallest value of 0.1319 when compared with the S-estimator and MM-estimator. Also, the M-estimator's mean absolute percentage error (MAPE) with the value of 1.2672 is the minimum when compared with both the S-estimator and MM-estimator. The results obtained using bias proportion show that principal component regression based on the M-estimator with the value of 0.0103 is the smallest in comparison with the S-estimator and MM-estimator, respectively. Theil inequality coefficient and the variance proportion also reveal the same result. As such, it can be asserted that the M-estimator is the most efficient robust fitted estimation model for addressing the multicollinearity and outlier problems in modelling and predicting RGDP in this study.

Moreover, the results of the M-estimator indicate that PC1 and PC2 have a positive impact on RGDP. Specifically, PC1 and PC2 increase RGDP by approximately 35.39% and 22.15%, respectively, during the period under consideration. The *p*-values, which are less than 0.05 for the estimated principal components serving as economic growth drivers, demonstrate the statistical significance of PC1 and PC2 in assessing RGDP in Nigeria. Additionally, the R-squared statistic with a p-value less than 0.05 confirms the overall significance of the fitted M-estimator model in examining the relationship between PC1 and PC2 (extracted as economic growth drivers) and RGDP in Nigeria, in the presence of multicollinearity and outliers.

Therefore, it can be emphasized that economic recession, crash in crude oil prices at the international market, insecurity, terrorist activities, an astronomical increase in the naira to dollar exchange rate and, above all, the Covid-19 pandemic have been heavily witnessed during the aforementioned period greatly affected the identified economic growth drivers or determinants. This impact is translated to the RGDP during the period under investigation. Hence, this study greatly benefits the policymakers as it provides a better understanding of the existing relationship between RGDP and its drivers above or determinants in this study. Also, a robust principal component regression based on a fitted M-estimator model remains robust statistically for modelling, estimating parameters and predicting the value of RGDP when multicollinearity and outliers are jointly present in the data set. However, there is a need to examine other estimation techniques to choose the best model appropriate for predicting or forecasting RGDP in Nigeria, thus leading to the discussion of the partial least square method being a component selection method in chapter 6 of this study.

CHAPTER 6

PARTIAL LEAST SQUARE (PLS) METHOD FOR ESTIMATION OF ECONOMIC GROWTH

6.1 Introduction

In order to account for variation in the dependent and explanatory variables, the partial least regression model is proposed. According to Helland (1990) as cited in Ebiwonjumi, *et al*. (2023), it is a method that can be used to construct a predictive model with many collinear explanatory variables. In this technique, the dependent variable Y is regressed against $x_1, x_2, x_3, \dots, x_p$ explanatory variables, in an effort to obtain new factors or components ($F_1, F_2, F_3, \dots, F_p$) that will satisfactorily replace the explanatory variables. The new factors or components ($F_1, F_2, F_3, \dots, F_p$) are called latent variables thus, each component serves as a linear combination of $x_1, x_2, x_3, \dots, x_p$. This technique is similar to principal component regression as the methods attempts to find new factors that can serve as explanatory variables in determining Y variables. The main difference between the methods is that the principal component focuses on the variance in the predictors. At the same time, partial least squares incorporate information about the covariance between the dependent variable and the predictors (explanatory variables). Consequently, Garthwaite (1994) emphasized that the basis of this is to reduce the dimension of the regression by using a few components rather than the whole explanatory variables (X). Sawatsky *et al*. (2015) emphasized that partial least square regression is a statistical modelling technique that extracts latent factors to explain predictor and response variation.

Yoon *et al*. (2015) compared generating composite indices weights using principal component analysis and partial least square methods. In the study, it was found that the partial least squares method generated weights that substantially increased the composite indices better than the principal component analysis method based on predictive performance for the outcome variables considered. Ghazari (2017) investigated factors that influence demand for banknotes in Malaysia using quarterly data obtained from Bank Negara Malaysia that spanned between (2007-2014). In the study, gross domestic product, interest rates, inflation rates, exchange rates and market sentiments were the factors considered. The partial least square regression method employed for the analysis revealed that gross domestic product and interest rates significantly influenced

Malaysia's banknote demand. Also, the banknote demand behaviour was accurately predicted for Bank Negara in Malaysia. Thus, there is a need to propose appropriate policy actions to manage the cash supply chains efficiently.

Dirsehan and Henseler (2022) conducted a study to determine the indices' optimum weights using the partial least square regression method. The study analysed the effects of the market potential index on foreign direct investment and gross domestic product by implementing different weighting schemes. Thus, the assessment showed that the partial least square model was a good predictive model for the study. Lazorec (2023) assessed the economic resilience process of EU countries using a multidimensional approach between (2008-2021). In the study, GDP growth rate and employment rate were considered as a measure of economic resilience. They identified the most critical determinants that influenced the economic resilience of the EU countries considering economic, social and institutional dimensions. The partial least squares regression employed for the analysis with three fitted distinct models for the resilience process, resistance, recovery, and transformation showed that the recent crisis was felt more in the economic and social dimensions. Thus, it can be emphasized that improving the institutional system of the EU countries would enhance their economic resilience.

However, in this chapter, we propose a partial least squares regression model for predicting economic growth (RGDP) in the presence of multicollinearity. The identified drivers include internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR), and trade openness (OPEN). This technique represents a robust approach that has been extensively used in various works and studies involving these variables in the literature.

6.2 Research Methodology

In this section, we discuss in detail the use of partial least squares regression to address the issues of multicollinearity and outliers in the dataset. This process involves transforming the original explanatory variables and the response variable to obtain latent variables or factors. These factors are used to fit a predictive model that efficiently predicts economic growth (RGDP).

6.2.1 Partial Least Squares Regression Model

Consider a general linear model given as:

$$Y = X^{'}\beta + \varepsilon, \tag{6.1}$$

The ordinary least square estimator of β is given by:

$$\hat{\beta} = (X'X)^{-1}X'Y, \tag{6.2}$$

The problem often arises if X is or likely to be singular, either because of the number of variables exceeds the number of observations or because of the presence of multicollinearity. To address this problem, PLS decomposes X into orthogonal scores T and loadings P as $X = TP$. Thus, to derive the partial least square estimators of β, the matrix X assume a bilinear decomposition that can be expressed in the form:

$$X = t_1P_1' + t_2P_2' + t_3P_3' + \cdots + t_rP_r', \tag{6.3}$$

$$X = \sum_{i=1}^{r} t_iP_i' = TP', \tag{6.4}$$

where t_i are the linear combinations of X, which can be written as X_{ri}. The $P \times 1$ vectors denoted by P_i are usually known as loadings. Unlike the weights in principal component regression that is the eigenvectors (j_i), the r_i are not orthogonal. However, t_i like the principal components Z_i, are orthogonal. According to Wold (1966) and De Jong (1993), two popular algorithms can be used to obtain partial least square estimators. These are non-linear iterative partial least squares (NIPALS) and simple partial least squares (SIMPLS). In the first method, the orthogonality is imposed through the computation of t_i as the linear combination of the residual matrices denoted by E_i (Frank, 1987 and Geladi and Kowalski, 1986). In other words, it can be given and expressed as:

$$t_i = E_{i-1}w_iE_i, \tag{6.5}$$

such that, it can be expressed in the form:

$$t_r = X - \sum_{i=1}^{r} t_iP_i' \; and \; E_0 = X, \tag{6.6}$$

where w_i are orthogonal. Thus, making it two sets of weighted vectors w_i and $r_i, i = 1, 2, 3, \dots, m$. In most of the algorithms for multivariate and univariate partial least square, the first step is to compute either w_i or $r_i, i = 1, 2, 3, \dots, m$, that can help us in calculating the linear combination of the t_i. Thereafter, $\mathrm{P_i}$ are computed through the regression of X on $\mathrm{t_i}$. Putting m factors into consideration. Then, the following relationship can be expressed and obtained as:

$$\mathrm{T_m = XR_m} \,, \tag{6.7}$$

$$\mathrm{P_m = X'T_m(T_m'T_m)^{-1}} \,, \tag{6.8}$$

$$\mathrm{R_m = W_m(P_m'w_m)^{-1}} \,, \tag{6.9}$$

where the first m is the dominant factors, that capture most of the variance in X and it has the ability to maximize the efficiency of the estimated parameter of the model. In the expression given in (6.9), two sets of weighted vectors were connected through a linear transformation. Also, from equations (6.7) and (6.8), $\mathrm{P_m'R_m}$ is equivalent to $\mathrm{I_m}$. Thus, the existence of this transformation can be obtained in the expressions given as follows:

$$\mathrm{R_m'P_m = R_m'X'T_m(T_m'T_m)^{-1}} \,, \tag{6.10}$$

$$\mathrm{R_m'P_m = T_m'T_m(T_m'T_m)^{-1}} \,, \tag{6.11}$$

$$\mathrm{R_m'P_m = I_m} \,. \tag{6.12}$$

After the extraction of m dimensional vector, the vector of fitted values for the partial least square can be used to represent the first m partial least square linear combinations denoted by T_m. The derivation can be obtained for the univariate case by the expression given as:

$$\mathrm{\hat{Y}_{PLS}^m = T_m(T_m'T_m)^{-1}T_m'y} \,. \tag{6.13}$$

According to Huber *et al*. (2005), it can be opined that the multivariate case can be obtained by replacing the vector $\hat{y}_{PLS}^m$ with the matrix $\hat{y}_{PLS}^m$. Thus, substituting XR_m for T_m and $\hat{\beta}_{OLS}$ for y in (6.13) and the result can be expressed in the form given by:

$$\mathrm{\hat{Y}_{PLS}^m = XR_m(R_m'X'X)^{-1}R_m'X'X\hat{\beta}_{OLS}} \,. \tag{6.14}$$

Then, it is expressed in the expression given as:

$$\mathrm{\hat{\beta}_{PLS}^m = R_m(R_m'X'XR_m)^{-1}R_m'X'X\hat{\beta}_{OLS}^m} \,, \tag{6.15}$$

Thus, this can be somewhat made simple for $\hat{\beta}_{OLS}$ by first substituting the expression given in (6.7) into (6.8) to obtain the expression given as:

$$\mathrm{P_m = X'XR_m(R_m'X'XR_m)^{-1}} \,. \tag{6.16}$$

Hence, using the expression given in (6.15) and (6.16), the result obtained can be expressed in the forms given by:

$$\hat{\beta}_{PLS}^m = R_m P_m' \hat{\beta}_{OLS} \,, \tag{6.17}$$

$$\hat{\beta}_{PLS}^m = w_m(P_m'w_m)^{-1}p_m'\hat{\beta}_{OLS} \,. \tag{6.18}$$

Thus, in the next session, the results and discussions of various analysis on fitted partial least square regression model and its associated diagnostic test and evaluation carried out were presented.

6.3 Empirical Results

This section reports the empirical results of the fitted models discussed in Section 6.2. In order to capture the multicollinearity that exists among the explanatory variables, the partial least squares method is used for the analysis. Table 6.1 presents the results for the non-cross-validated partial least squares.

Table 6.1: Non-Cross Validated Partial Least Square Result

Components X	Variance	Error	R-Sq (X)	PRESS	R-Sq (Y)or(pred)
1	0.4294	4.6861	0.8349	4.8898	0.8278
2	0.7265	3.5911	0.8735	3.8674	0.8638
3	0.8756	3.0195	0.8936	3.3132	0.8833
4	0.9000	2.4628	0.9133	2.8428	0.8999
5	1.0000	2.4208	0.9147	2.7830	0.9020

Table 6.1 shows the results of partial least square selection, which reveals the variance of the estimated first, second, third, fourth, and fifth components for the economic growth drivers (explanatory variables), which are 0.4294, 0.7265, 0.8756, 0.8999 and 1.0000 respectively with the associated error of 4.6861, 3.5911, 3.0195, 2.4628 and 2.4208. As a result, component 1 contributes the largest variability in predicting economic growth, with the highest estimated error of 4.6861. Thus, based on the error associated with the various estimated components for predicting economic growth (RGDP), it can be emphasized that using the entire 5 estimated components is more efficient because of the small error of 2.4208 compared to the selection first or second and any other components. In Table 6.1, the R-square indicates the amount of variance that various components of X can explain. It shows that the first, second, third, fourth, and fifth components explained 83.5%, 87.4%, 89.4%, 91.3% and 91.5%, respectively, for the variance of various linear combinations of economic growth drivers under consideration. Furthermore, R-square (pred) reveals the amount of variance of economic growth (RGDP) that each estimated component explains, and it is found that the first, second, third, fourth, and fifth components can explain 82.8, 86.4, 88.3, 90.0 and 90.2 of the variances of Y. In examining the efficiency of the

estimate, a predictive residual sum of square (PRESS) is used, and it is found that the first, second, third, fourth, and fifth components have a predictive residual sum of the square of 4.8898, 3.8674, 3.3132, 2.8428 and 2.7830 respectively. The plot we present in Figure 6.1 also reveals the amount of variance explained by the linear combination of economic growth drivers that includes internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR), degree of economy openness (OPEN) and the economic growth (RGDP) in relation with estimated components that are obtained after the transformation of the economic growth drivers mentioned.

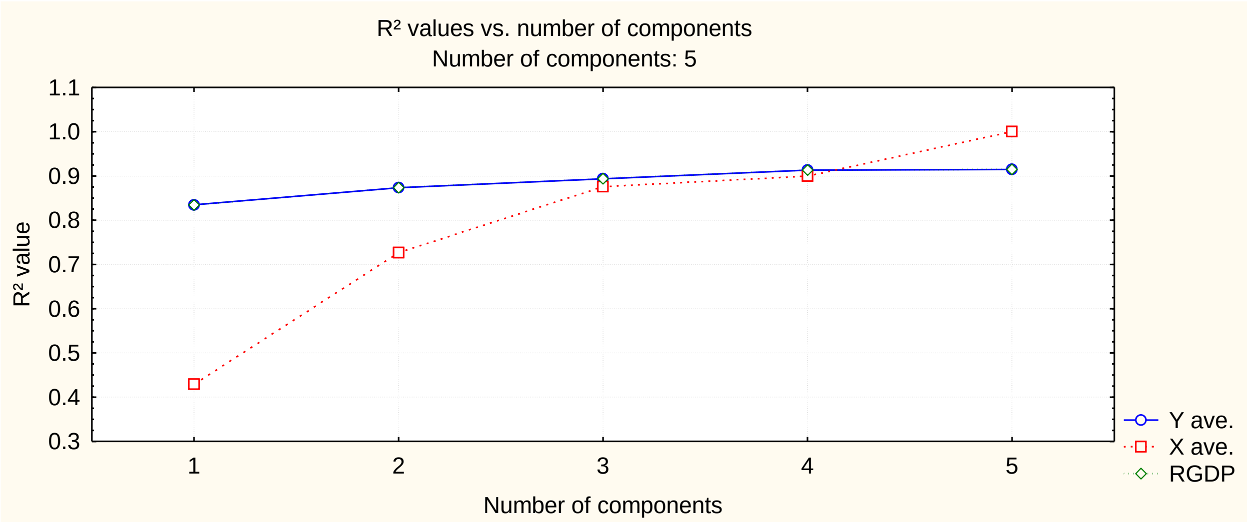


Figure 6.1: Plot showing the amount of variance explained by the components and RGDP

Table 6.2: Loading Weight for Economic growth and its drivers

Loading (X)	COMP 1	COMP 2	COMP 3	COMP 4	COMP 5
INDT	0.6683	-0.0627	0.0518	0.7515	-0.1493
EXDT	0.4133	-0.6981	-0.3554	-0.0430	0.0803
RINR	0.1117	-0.5461	0.7350	-0.6550	0.6920
REXR	0.0285	0.6306	-0.6693	-0.1894	0.7000
OPEN	0.6281	0.3423	0.0785	-0.6461	-0.0489
Loading (Y)					
RGDP	0.6314	0.1841	0.1741	0.4829	0.0544

Table 6.2 shows the loading weight for drivers of economic growth identified in the study that can be used to predict economic growth (RGDP). Thus, the loading weights are positive or negative,

indicating the contribution of economic growth drivers (explanatory variables) to various estimated components. The results in Table 6.2 show that INDT, EXDT, RINR, REXR, and OPEN contribution to the first component or factor is 66.8%, 41.3%, 11.2%, 2.9% and 62.8%, respectively. Also, INDT, EXDT, RINR, REXR, and OPEN contribution to the second component or factor is -6.3%, -69.8%, -54.6%, 63.1% and 34.2%, respectively. INDT, EXDT, RINR, REXR, and OPEN contributions to the third component or factor are 5.2%, -35.5%, 73.5%, -66.9% and 7.9%, respectively. For the fourth component or factor, the contributions of INDT, EXDT, RINR, REXR, and OPEN are 75.2%, -4.3%, -65.5%, -18.9% and -64.6% respectively and for the fifth component or factor, INDT, EXDT, RINR, REXR, and OPEN contributes -14.9%, 8.0%, 69.2%, 70.0% and -4.9% respectively. Thus, it can be emphasized that the results of the respective number of components that can be explained by the identified economic growth drives or determinants efficiently fit a predictive model for the prediction of economic growth rate (RGDP). Also, in Table 6.2, we present the loading results for economic growth (RGDP) and the components obtained from the linear combination of the economic growth drivers under consideration. Thus, it indicates that the first, second, third, fourth and fifth components are 63.1%, 18.4%, 17.4%, 48.3% and 5.4% efficient in predicting economic growth rate (RGDP) in this study. After obtaining the components, the next step is to estimate the model's coefficient; the results are shown in Table 6.3.

Table 6.3: ML Estimate of Partial Least Squared Model

Variables	Estimate	Standardized estimate
Constant	7.4666	0.0000
COMP 1	0.2817	0.9376
COMP 2	-0.1128	-0.2669
COMP 3	0.3885	0.1664
COMP 4	0.0998	0.1541
COMP 5	0.5078	0.1469
Sum of square regression		25.9683
Mean sum of square regression		5.1937
Sum of square residual		2.4208
Mean sum of square residual		0.0179
F-ratio		289.63
p-value		0.0000

Table 6.3 and Figure 6.2 show the coefficient of estimated parameters for predicting economic growth (RGDP) using the components of the economic growth drivers. Thus, from the results, it is found that the extracted first, second, third, fourth, and fifth components contribute 28.2%, -11.3%, 38.8%, 10.0% and 50.8%, respectively, to the prediction of economic growth rate (RGDP) in Nigeria after the linear combination and transformation of INDT, EXDT, RINR, REXR, and OPEN as the identified economic growth drivers. This result is obtained through the standardization of the estimated parameters. In Table 6.3, the partial least square model results reveal that the mean sum of square regression (MSSR) and mean square of error (MSE) are 5.1937 and 0.0179, respectively. The results indicate that the error margin of predicting economic growth using the model above is less than 5%, thus emphasizing the efficiency of the partial least square model in predicting economic growth rate (RGDP). The F-statistic value of 289.63 with associated p-value < 0.05 establishes the overall statistical significance of the partial least square regression model in predicting economic growth (RGDP).

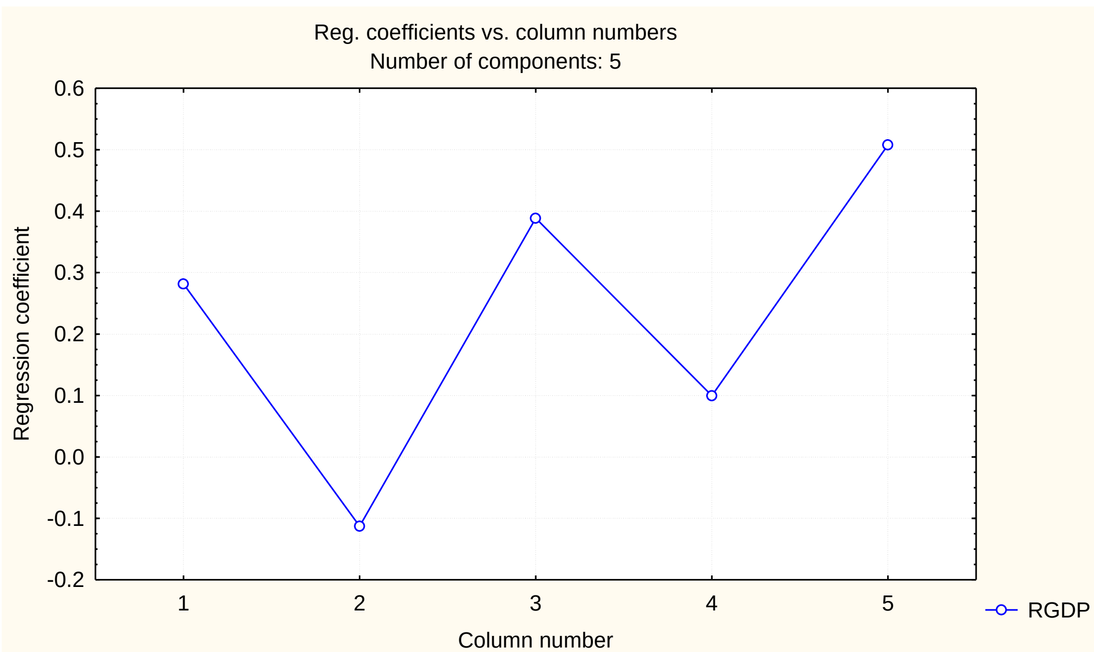


Figure 6.2: Graph showing the coefficient of estimated parameter for components

Figure 6.3: Scree plot showing the eigenvalue of estimated for the components

In Figure 6.3, we present the eigenvalues screen plot for the first, second, third, fourth, and fifth components that are generated from the linear combination and transformation of the identified economic growth drivers (explanatory variables) that include INDT, EXDT, RINR, REXR, and OPEN. Thus, from Figure 6.3, the eigenvalues screen plot shows that the first, second, third, and fourth component or factor contribute 41.9%, 22.8%, 13.3%, and 1.7% respectively, to the prediction of the economic growth (RGDP) in this study. Hence, it can be emphasized that even though the multicollinearity problem in the original data set has been addressed, the outliers may still be present in the estimates because of the non-cross-validation approach employed for this analysis. Therefore, it is imperative to carry out further analysis to address the outliers that might still be available in the estimates through the cross-validation and selection method of the components. The results of the cross-validation are shown in Table 6.4:

Table 6.4: Cross Validation Partial Least Square Results

Comp	R^2X	R^2X(Cumul.)	Eigenvalues	R^2Y	R^2Y(Cumul.)	Q^2	Q^2(Cumul.)	Sig.
1	0.4294	0.4294	2.0943	0.8349	0.8349	0.7824	0.7824	S
2	0.2971	0.7265	1.1387	0.0386	0.8735	0.0052	0.7835	S

Table 6.4 presents the cross-validated results of the partial least square regression model. The results reveal that two components or factors are generated from the linear combination of identified economic growth drivers (explanatory variables) under consideration to predict Economic growth rate (RGDP). The cross-validated partial least squares model results reveal that two components extracted explained 87.35% proportion of the economic growth RGDP (dependent variables) and, as such, efficient and appropriate for the prediction of the RGDP. It is also found that the variance denoted by the R-square of X for the first and second components extracted and selected from the identified economic growth drivers (explanatory variables) in this study are 0.4294 and 0.2971, respectively. This result shows that 42.9% and 29.7% of the variance of the first and second components with eigenvalues of 2.0943 and 1.1387 can be accounted for in predicting economic growth rate (RGDP) in this study. Thus, it emphasizes that a total of 72.6% of the variance has been accounted for or explained as a result of the linear combination of the identified economic growth drivers (explanatory variables) using the cross-validated partial least square model. In Table 6.4, the R-square Y(pred) shows the amount of variation in the economic growth (RGDP) that can be explained by the first and second extracted or generated components. Thus, it is revealed that 83.5% and 3.9% of the proportional variation in economic growth rate (RGDP) can be explained by the first and second components extracted through the cross-validated approach employed. The Q-square reveals that the cross-validated predictive value of 0.7834 and 0.0052 for the first and second extracted components indicates that a total of 87.4% proportional variation of the components can be efficiently and significantly explained in predicting economic growth rate (RGDP) with the cross-validated predictive value of 78.9%. In a quest to know the order of importance of the identified economic growth drivers that generate the two extracted components to predict economic growth rate (RGDP) through the cross-validation approach of the partial least square model, the variable importance for projection of economic growth result is presented in Table 6.5.

Table 6.5: Variable Importance for Projection (VIP) of Economic Growth

Variables	Variable number	VIP	Importance
INDT	1	0.6678	1
OPEN	5	0.6556	2
EXDT	2	0.3253	3
REXR	4	0.1014	4
RINR	3	0.0898	5

Table 6.5 reveals the result of the identified economic growth drivers (explanatory variables) in this study and the level of their importance or rank in predicting economic growth (RGDP). In Table 6.5, it is found that INDT, OPEN, EXDT, REXR and RINR with the variance importance projection (VIP) values of 66.8%, 65.6%, 32.5%, 10.1%, and 9.0% are ranked as first, second, third, fourth, and fifth respectively, as the order of importance in predicting the RGDP in Nigeria. This result emphasizes the relevance of internal debt (INDT) or domestic borrowing and a high degree of economic openness or trade openness (OPEN) in growing Nigeria's economy. This result is also represented in a bar chart, as shown in Figure 6.4, for a better description and understanding or visualization of the importance of identified economic growth drivers (explanatory variables) in predicting Nigeria's economic growth rate (RGDP).

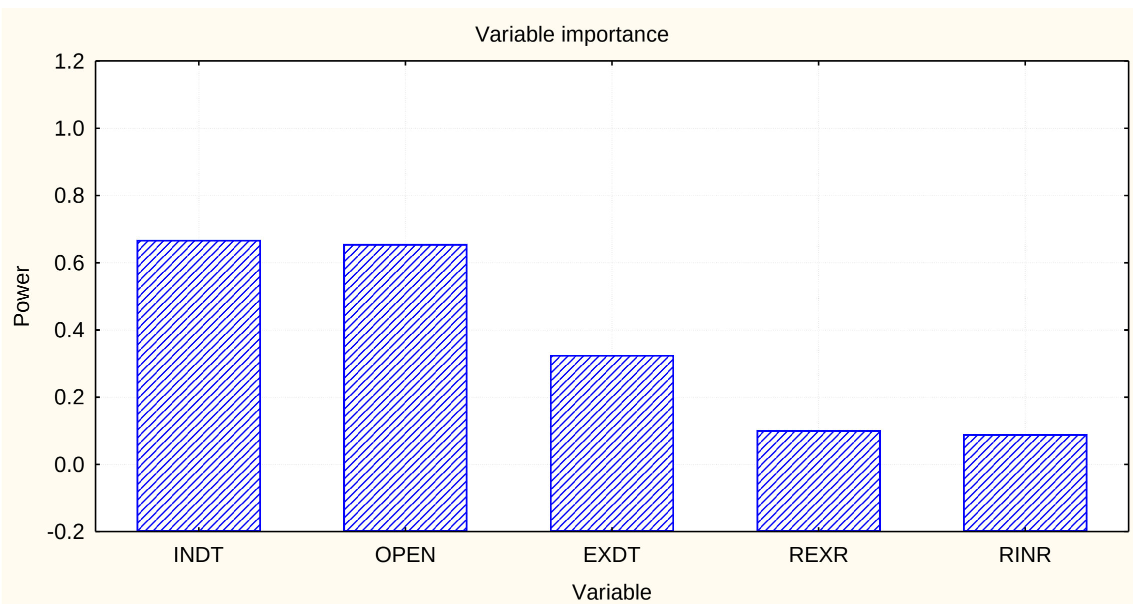


Figure 6.4: Variable Importance Representation for Projection of Economic Growth

Next is Table 6.6 and Table 6.7, where we present partial least square weights and loadings obtained from the identified economic growth drivers under consideration for extracting the cross-validated components and their required coefficients for efficient and optimal prediction of the economic growth rate (RGDP).

Table 6.6: Cross Validated Weight of the Economic Growth Drivers for the Components

Variable	Component 1	Component 2
INDT	0.6828	0.0920
EXDT	0.2719	-0.8921
RINR	0.0835	-0.1777
REXR	0.0785	0.3155
OPEN	0.6684	0.2541

Table 6.6 shows the weight of the identified economic growth drivers (explanatory variables) for the first and second components from the cross-validated method, dominated by the largest contribution from INDT and OPEN, as emphasized in Table 6.5 and Figure 6.4. In Table 6.6, it is observed that INDT and OPEN contributions are high and positive, representing 68.3% and 66.8% contributions to the extraction of the first component compared with others such as EXDT, RINR and REXR that contribute 27.2%, 8.4% and 7.9% respectively. In the second component, it is found that INDT, REXR and OPEN are positive and contribute 9.2%, 31.6%, and 25.4%, respectively, for the extraction of the component. Others, such as EXDT and RINR, have negative contributions toward the extraction of the second component, thus limiting the efficiency of the component by 89.2% and 17.8%, respectively.

Table 6.7: Economic Growth Drivers Loading for Cross Validated Components

X loading (Number of components is 2)			
Variable	Variable number	Component 1	Component 2
INDT	3	0.6682	-0.0627
EXDT	4	0.4133	-0.6981
RINR	5	0.1117	-0.5461
REXR	6	0.0286	0.6306
OPEN	7	0.6281	0.3423

Table 6.7 shows the result of the partial least square loading for the identified economic growth drivers to predict economic growth (RGDP) in Nigeria. Thus, the loadings are either positive or negative, as revealed in Table 6.7. This indicates the contribution of identified economic growth drivers (explanatory variables) to the extracted cross-validated components. Specifically, it is

revealed that INDT, EXDT, RINR, REXR, and OPEN contribution to the first component or factor are 66.8%, 41.3%, 11.2%, 2.9% and 62.8%, respectively. Also, INDT, EXDT, RINR, REXR, and OPEN contribution to the second component or factor is -6.3%, -69.8%, -54.6%, 63.1% and 34.2%, respectively. Thus, it implies that the respective variation of the components can be explained by the identified economic growth drivers or determinants that are required for efficient and optimal prediction of the RGDP. This result is also presented in graphical form in Figures 6.5a, 6.5b and 6.5c for the first and second extracted components and the combined components extracted. In Figure 6.5a, it is specifically revealed that all the identified economic growth drivers are in a positive position, thus reflecting their positive contribution to the first component. In contrast, in Figure 6.5b, INDT, EXDT and RINR are in a negative position of the second component. REXR and OPEN are in the positive position of the second component to show their respective positive and negative contributions to the extracted second component through the cross-validated partial least square model. The combined graphical representation and the position of the identified economic growth drivers indicate the contributions of the economic growth drivers to the extracted components through the cross-validated method, as shown in Figure 6.5c.

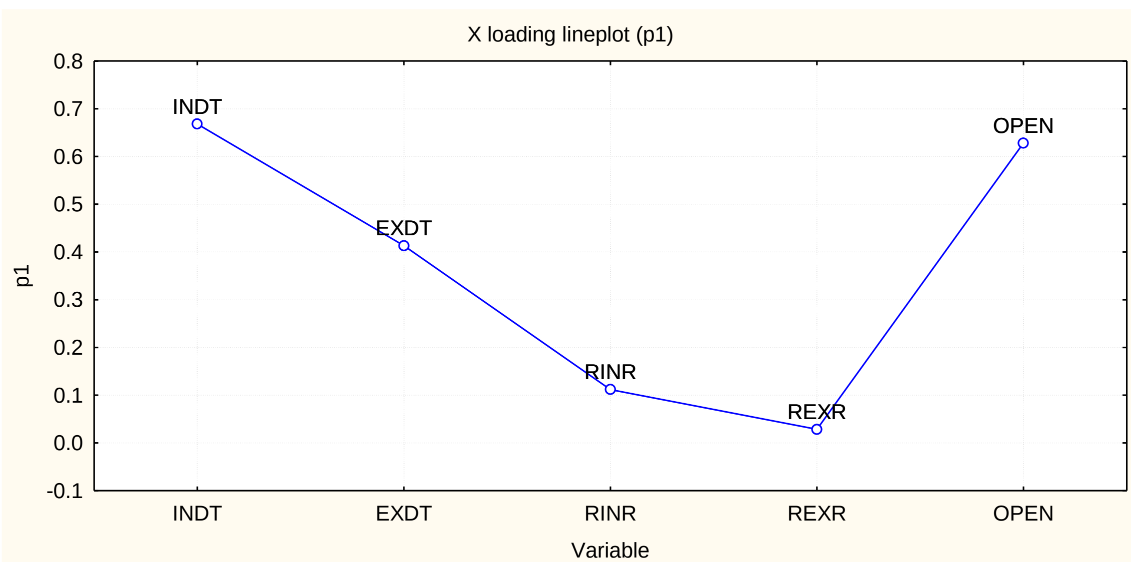


Figure 6.5a: loading plot the contribution of economic growth drivers for the first cross validated component

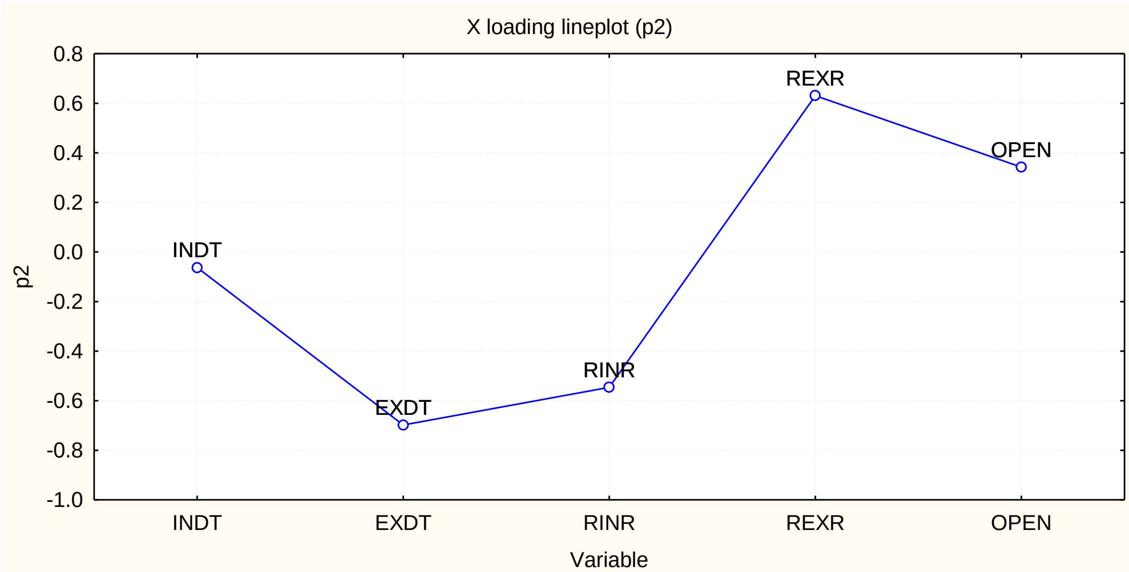


Figure 6.5b: loading plot and the contribution of economic growth drivers for the second cross validated component

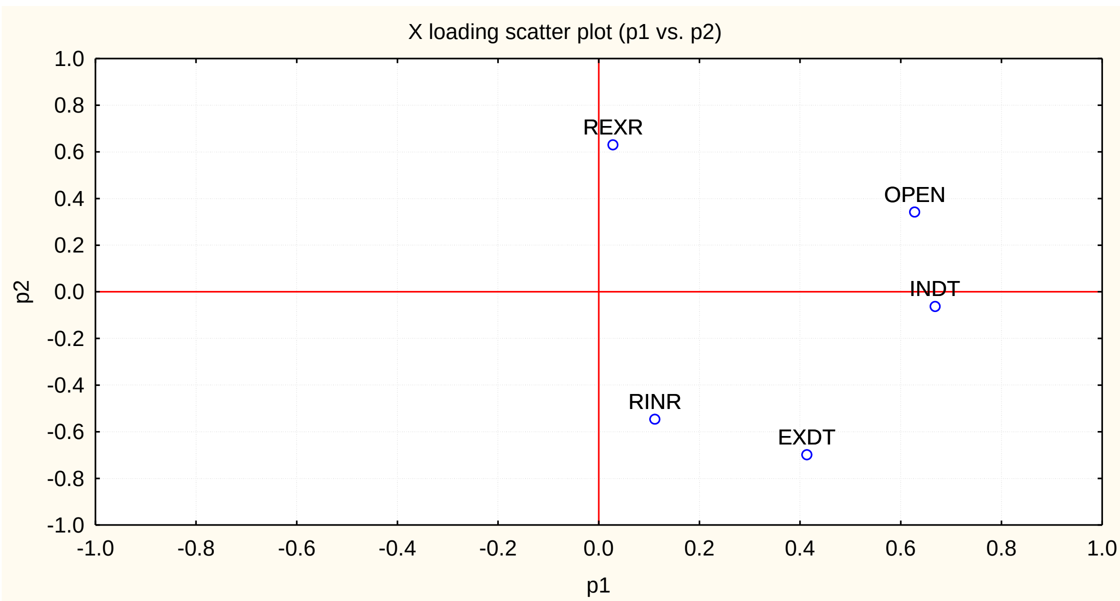


Figure 6.5c: loading plot and the contribution of economic growth drivers for the combined cross validated component

Table 6.8: The Economic Growth RGDP Loading (Y loading for components)

Variable	Variable number	Component 1	Component 2
RGDP	2	1.0000	1.0000

In Table 6.8, we present the result of the loading for the RGDP and the extracted components obtained from the linear combination and transformation of the identified economic growth drivers under consideration in this study through the cross-validation method. The results we present in Table 6.8 and Figure 6.6 indicate that the first and second extracted components through the cross-validated method are 100% efficient and optimal in predicting economic growth in Nigeria compared to the non-cross-validated method, where it was revealed that the first and second component is 63.1% and 18.4% respectively efficient and optimal in predicting economic growth. Thus, Table 6.2 indicated the presence of outliers even when the multicollinearity problem has been sorted out. This is also stressed, as shown in the eigenvalues plot of the extracted components through the non-cross-validated technique shown in Figure 6.1. Thus, the extracted components from the linear combination and transformation of the identified economic growth drivers through the cross-validated method can efficiently and optimally predict the RGDP.

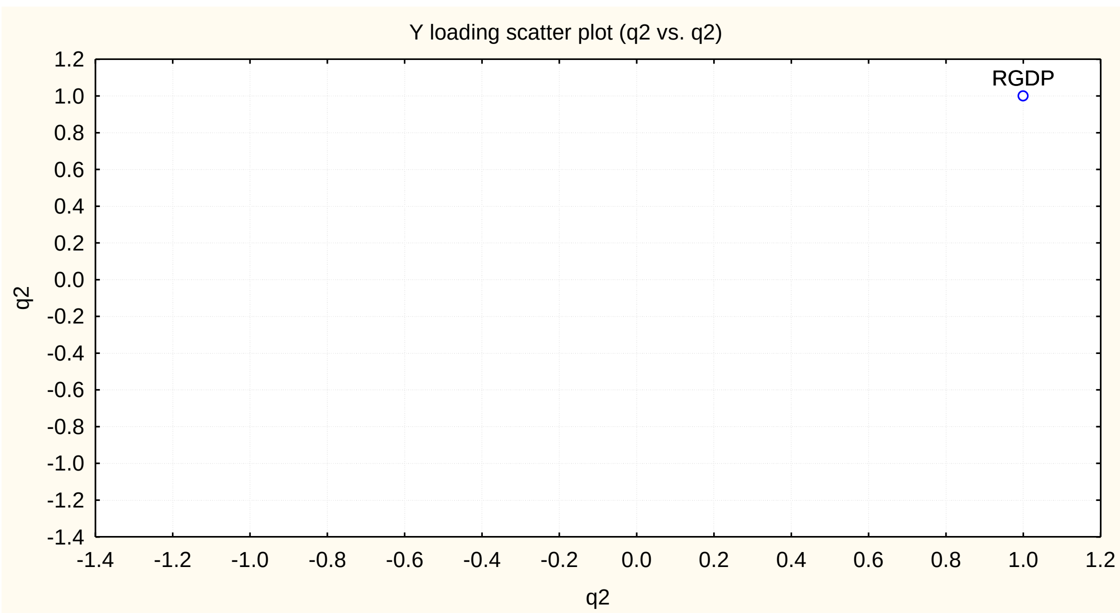


Figure 6.6: Graph showing the contribution of components to economic growth (RGDP)

Table 6.9: Cross Validated Partial Least Square Estimates for Extracted Components

Comp.	Estimate	Eigenvalues	Total variance (%)	Cum eigenvalue	Cum. (%)
1	0.6314	2.0943	41.8867	2.0943	41.8867
2	0.1841	1.1387	22.7733	3.2330	64.6600

Table 6.9 presents the cross-validated partial least square estimates for the extracted components to predict RGDP. It is found from the results that the extracted first and second components are 63.1% and 18.4% respectively efficient and optimum in predicting economic growth (RGDP) in Nigeria with the eigenvalue of 2.0943 and 1.1387 respectively, using cross-validation partial least square method. Also, Table 6.9 and Figure 6.7 show that the variance of the first and second extracted components estimates are 41.9% and 22. 8%, respectively, using the cross-validated method. This result is more efficient than the variance obtained using a non-cross-validated method, which is 42.9% and 29.7%, respectively. In cumulative, the total variance obtained for the two extracted components through cross-validation is 64.7%, showing its efficiency in predicting economic growth rate (RGDP) over the value of 72.7% obtained through the non-cross-validation method.

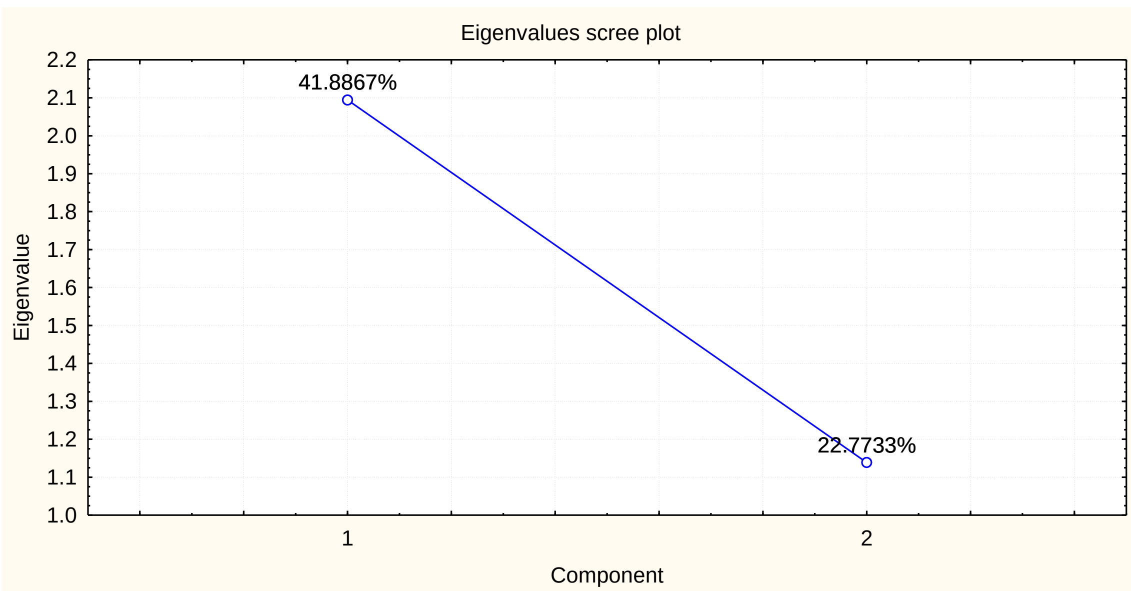


Figure 6.7: Eigenvalues scree plot of the generated components

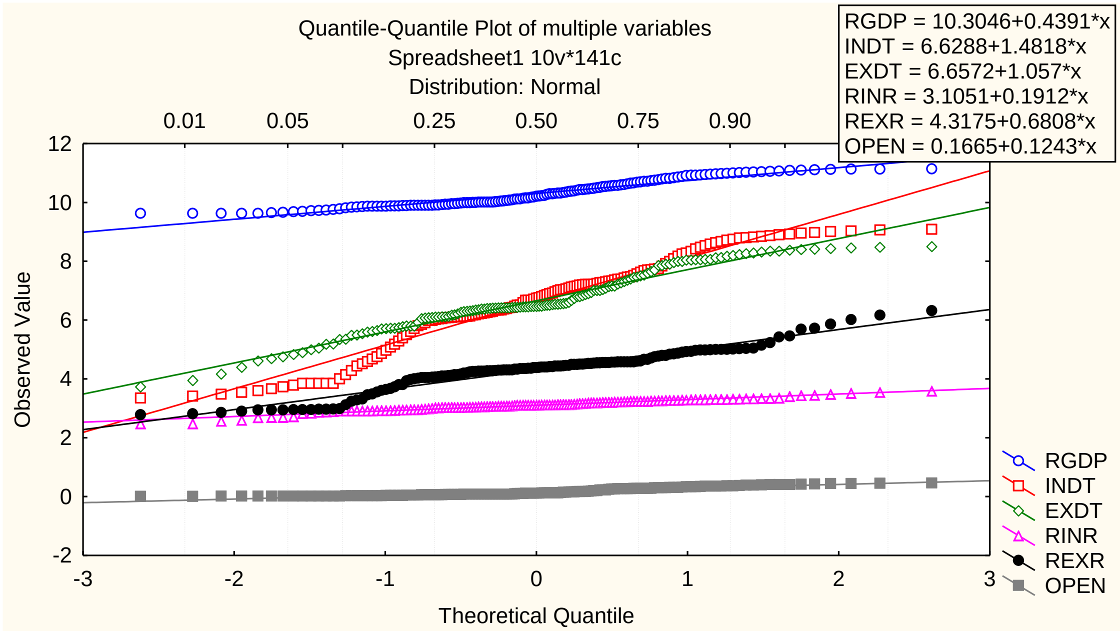


Figure 6.8: The Q-Q Plot showing the normality of economic growth and its drivers

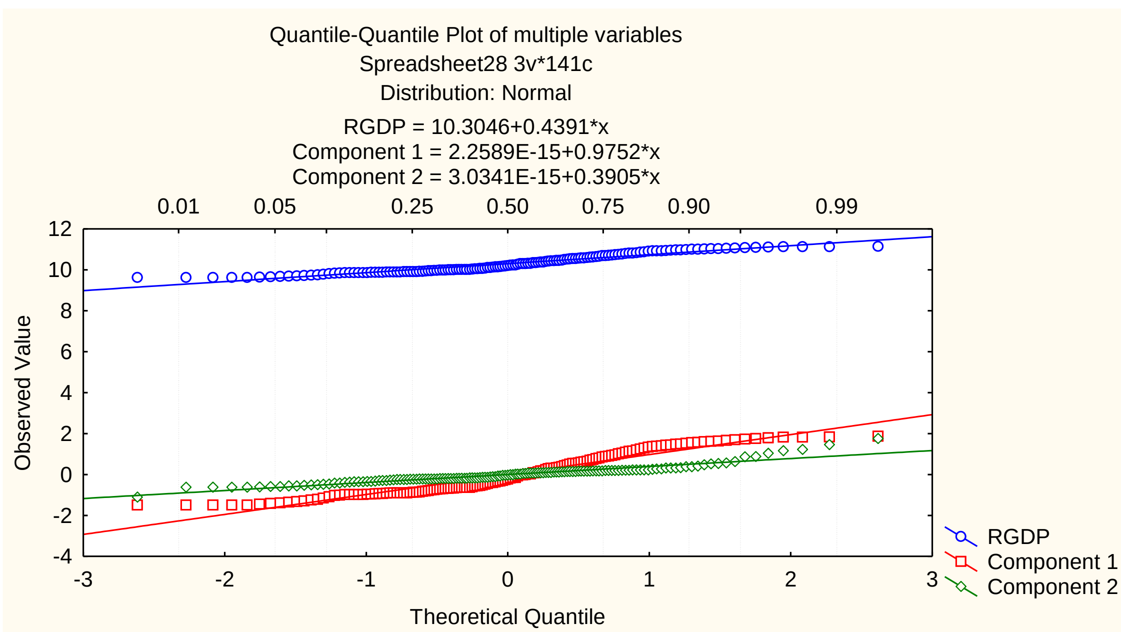


Figure 6.9: Q-Q Plot showing the normality of economic growth and the contribution of components

In Figure 6.8 and 6.9 we present Q-Q plots that show economic growth (RGDP) and its identified drivers, including: internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN) are from normal population that contained outliers as shown by the plotted values in Figure 6.8. However, from the extracted first and second components through the cross-validated partial least square method, as shown in Figure 6.9, it is revealed that the extracted first and second component estimates from the linear combination and transformation of the identified economic growth drivers that the outliers have been addressed based on the values of the components with approximate mean and standard error of 2.2589E-15 + 0.9752 and 3.0341E-15 + 0.3905 respectively. These are less than the ones obtained for identified economic growth drivers shown in Figure 6.8, making them suitable and appropriate for predicting economic growth (RGDP).

In addition, in Figure 6.10, Shapiro-Wilk test statistics, which is the most efficient test for normality, shows that the values of 0.9290, 0.9290, and 0.9146 for RGDP, first component and second component respectively that are obtained from the linear combination and transformation of INDT, EXDT, RINR, REXR OPEN respectively and the probability value of the Shapiro-Wilk statistics with *p*-value < 0.05 indicates that all identified economic growth drivers under investigation are normally distributed.

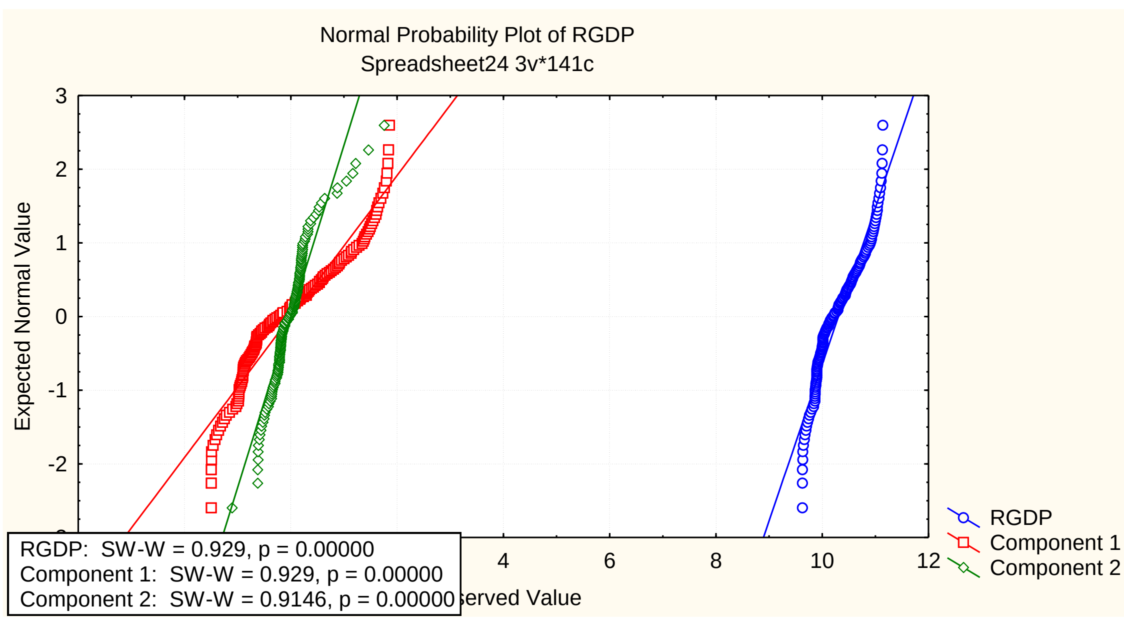


Figure 6.10: Normal Probability Plot showing the contribution of components and RGDP

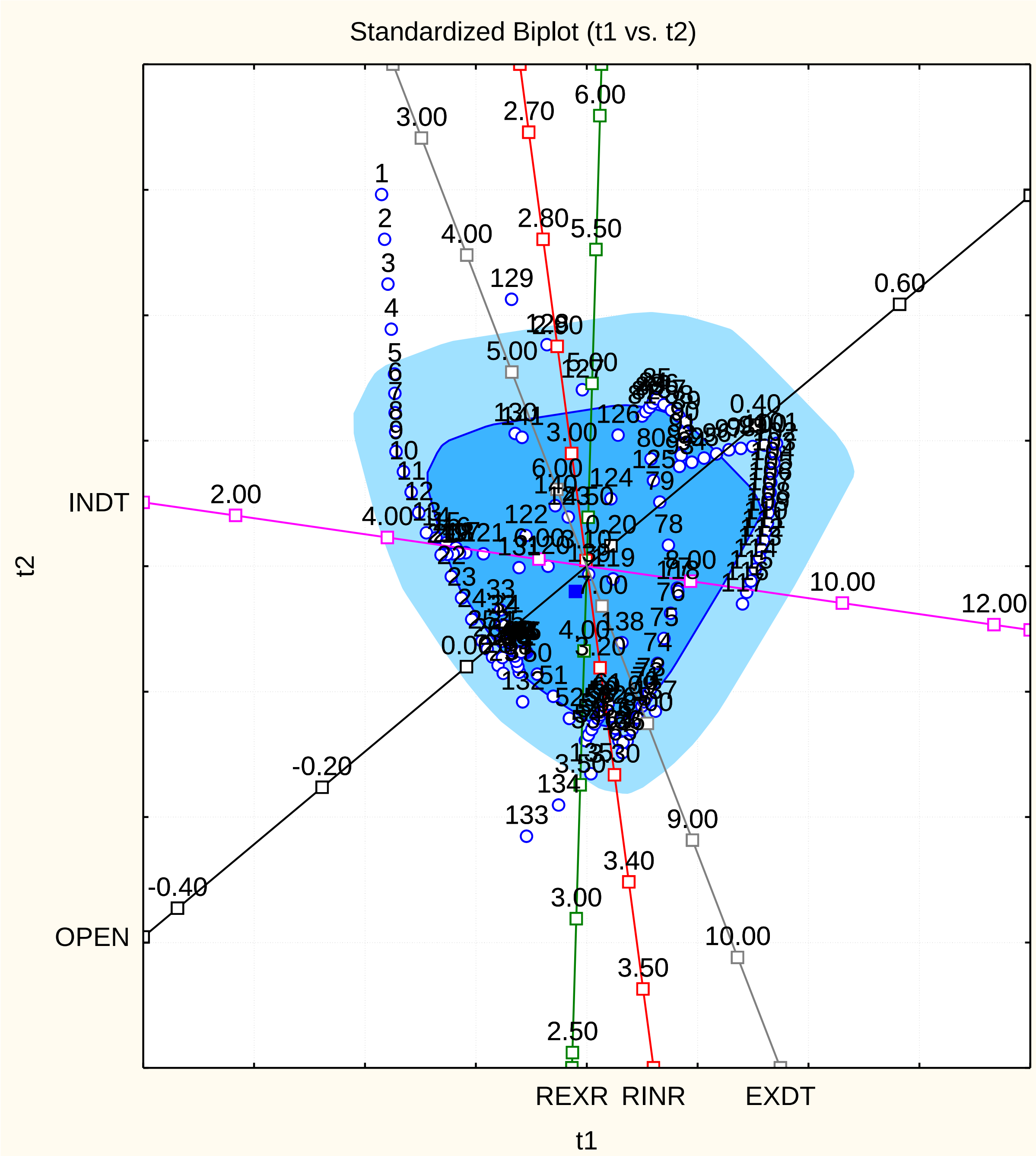


Figure 6.11: Standardized biplot showing the efficiency of the two Cross Validated Components

In Figure 6.11, we present a standardized biplot that shows and validates the extracted two components from the linear combination and transformation of the identified economic growth drivers that include INDT, EXDT, RINR, REXR and OPEN and the contributions of each extracted component estimates in predicting the economic growth rate (RGDP). The two colours shown in the biplot represent the first and the second extracted components with respective inner deep blue colours and the light blue colour that is obtained through the cross-validated partial least square method. It further shows that the two extracted components are this study's major drivers or determinants for predicting economic growth (RGDP). In the inner deep blue colour region of the biplot, it is found that all the identified economic growth drivers such as INDT, EXDT, RINR, REXR, and OPEN values are concentrated in the region to form the first extracted component for predicting economic growth rate (RGDP). This also informs why it is revealed in Table 6.4, the positive contributions of the linear combination of the identified economic growth drivers in obtaining the first component to predict the economic growth rate (RGDP). In the light blue colour region, the major values concentrated in the region are the ones from REXR and OPEN as the identified economic growth drivers and the reason for their positive contributions toward the extraction of the second component through the linear combination and transformation of the identified economic growth drivers. The values of the identified economic growth drivers outside the two identified regions prove why the five components extracted using the non-cross-validated partial least square method are inefficient in predicting the economic growth rate (RGDP) in Nigeria.

Table 6.10: Performance Metric Evaluation for PLS Models

Forecast RGDP	Non-Cross Validated PLS	Cross Validated PLS
Root Mean Square Error	0.2316	0.1388
Mean Absolute Error	0.2062	0.1094
Mean Abs Percent Error	1.9926	1.0561
Theil Inequality Coefficient	0.0112	0.0067
Bias Proportion	0.0000	0.0000
Variance Proportion	0.0858	0.0276
Covariance Proportion	0.9141	0.9723

In Table 6.10, we present predictive power or performance metrics evaluation for the partial least square model that simultaneously addressed the presence of multicollinearity and outliers in this study. From Table 6.10, it is revealed that the root mean square error (RMSE) of the validated partial least square with a value of 0.1388 is the smallest in comparison with the non-cross-validated partial least square with a value of 0.2316. The mean absolute error (MAE) of the fitted for the cross-validated partial least square is observed to be the smallest with a value of 0.1094 when it is compared with the non-cross-validated partial least square with a value of 0.2062. Also, the mean absolute percentage error (MAPE) of the fitted cross-validated partial least square model with the value of 1.0561 is the smallest MAPE when it is compared with the non-cross-validated partial least square with a value of 1.9926. The Theil inequality coefficient and the variance proportion also reveal the same result, and as such, it can be emphasized that the cross-validate partial least method is more efficient and optimal for predicting economic growth rate (RGDP) in the presence of multicollinearity and outliers. Thus, the next ten (10) quarters' forecast of the economic growth rate (RGDP) shown in Figure 6.10 is carried out using a cross-validated partial least square model.

6.4 Forecast Plot for Cross Validated Partial Least Square Model

In Figure 6.10, we present a forecast plot using a validated partial least square to show the predictive efficiency of the model in predicting stable and reliable values for economic growth rate (RGDP) based on the data under consideration in this study.

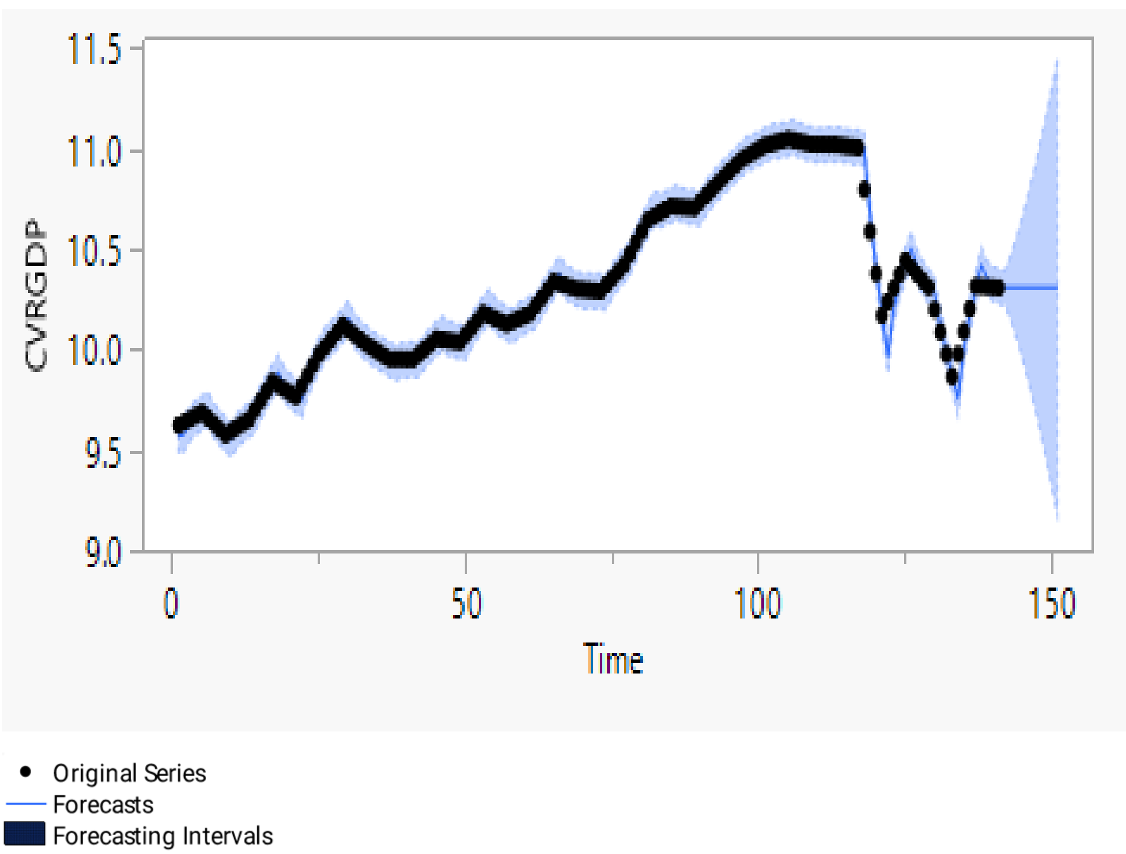


Figure 6.12: Forecast Plot for the Cross Validated Partial Least Square Method

6.5 Concluding Remarks

This chapter thoroughly examines the estimation of economic growth parameters in Nigeria based on identified drivers or determinants (INDT, EXDT, RINR, REXR, and OPEN), considering the presence of multicollinearity and outliers. Exploratory and diagnostic analysis tests establish the relationship between economic growth (RGDP) and these economic drivers. Multicollinearity and outliers are identified using VIF and Grubbs' test, respectively, on the dataset under consideration. To address these issues and obtain efficient parameter estimates for INDT, EXDT, RINR, REXR, and OPEN, both non-cross-validated and cross-validated partial least squares regression methods are employed. These methods aim to provide an optimal estimate for the model to predict the economic growth rate (RGDP) in Nigeria effectively.

Thus, from the result, a non-cross-validated partial least square method reveals the extraction of five components to be selected to predict economic growth rate (RGDP) based on the linear combination and transformation of the identified economic growth drivers under consideration for this study. While a cross-validated partial least square method reveals the extraction of only two components to be selected based on the same linear combination and transformation process for the identified economic growth drivers under consideration to predict the economic growth (RGDP). From the results, the R-square indicates that 91.5% and 72.6% of the variance in economic growth drivers can be explained by the five and two components extracted and selected from the non-cross-validated and cross-validated partial least square method, respectively. Also, R-square (pred) shows that 90.2% of variabilities in economic growth (RGDP) can be explained by the five extracted and selected components, with the predictive residual sum of square (PRESS) efficiency value of 2.7830. However, from the validated partial least method result, it is found that 87.4% variations in economic growth (RGDP) can be explained by the two extracted and selected components with cross-validated predictive value of 78.4%, thus a more efficient estimate as well as improved significance prediction of the two components, over the non-cross validated five components.

In a non-cross validated partial least square selection method, the result of loading for the economic growth (RGDP) and the extracted components from the linear combination of the identified economic growth drivers under consideration reveals that the extracted and selected first, second, third, fourth, and fifth components contributions are 63.1%, 18.4%, 17.4%, 48.3% and 5.4% respectively in predicting economic growth rate (RGDP) in Nigeria. However, the cross-validation results show that the first and second extracted and selected components are 100% in predicting economic growth. As such, the two extracted and selected components are more efficient than the five components extracted and selected using the non-cross-validated method. It is also found that after the extraction of the components through the linear combination and transformation, the first, second, third, fourth, and fifth components influenced the growth of the economy by 28.2%, -11.3%, 38.8%, 10.0% and 50.8% respectively while, cross-validated method revealed that the extracted and selected first and second components are 63.1% and 18.4% respectively efficient and optimum in predicting economic growth in Nigeria. Thus, from the eigenvalue scree plot, the variance for the first and second extracted and selected components are 41.9% and 22. 8%, respectively, forming a perfectly straight line between the two extracted components to emphasize that the data set's outlier and multicollinearity problems have been addressed. This situation is not observed in the case of the five extracted and selected components using a non-cross-validated method.

To further validate these results, the quantile plot and normal probability plot showing Shapiro-Wilk statistic value with p-value < 0.05 show that the extracted and selected components are normally distributed. Also, the biplot reveals that all the identified economic growth drivers: INDT, EXDT, RINR, REXR, and OPEN values are concentrated in the region of extracted and selected first and second components, showing the positive contributions of the linear combination and transformation of the identified economic growth drivers to the extracted components in efficiently predicting economic growth (RGDP). Also, from the variable importance projection, it is found that the economic growth in Nigeria depends largely on internal borrowing and economic openness to engender international patronage.

Therefore, it can be stressed that economic recession, crash in crude oil prices at the international market, insecurity, and terrorist activities all led to insufficient availability of funds and inadequate internal funding through borrowing to grow the economy. The production level, particularly agricultural and manufacturing products, which are the alternative sources through which the economy can grow, is also hampered. The closure of all international borders during this period affected the openness of the economy for exportation and importation activities, and as such, international patronage was very low, which hindered the growth of the economy. An astronomical increase in the naira to dollar exchange rate and, above all, the Covid-19 pandemic that was heavily witnessed during the aforementioned period greatly affected the identified economic growth drivers; thus, its impact is translated to Nigeria's economic growth (RGDP) during the period under investigation. Hence, this study serves as a great benefit to policymakers as it establishes that internal borrowing and economic openness are essential for the growth of the economy and the need for policy direction in these areas. Also, a cross-validated partial least square method of extracting and selecting the components required for predicting the growth of the economy is established as an efficient technique to be used in dealing with multicollinearity and outlier problems in a data set. However, there is a need to examine other robust statistical techniques in order to choose the best model appropriate for predicting or forecasting the economic growth rate RGDP in Nigeria and, as such, the need to explore and discuss the robustness of the average penalized least square method in Chapter 7.

CHAPTER 7

ESTIMATION OF ECONOMIC GROWTH USING AVERAGE CENTERED PENALIZED REGRESSION METHOD

7.1 Introduction

In order to ensure an efficient estimate of the parameters and stable predictive value for economic growth (RGDP), we introduce the average centered penalized least square regression method. According to Tibshirani *et al.* (2015), the average penalised regression method's high prediction accuracy and computational efficiency make this technique appropriate for this study. James *et al.*, (2013) opined that in similarity with ordinary least squares method, penalized least square regression methods estimate the regression coefficients by minimizing the residual sum of squares by placing a constraint on the size of the regression coefficients. This constraint or penalty on the size of the regression causes coefficient estimates to be biased, but it improves the overall prediction error of the model by decreasing the variance of the coefficient estimates (Hastie *et al.*, 2001; Yuzbasi *et al.*, 2017). James *et al.* (2013) posited that penalized least square involves the selection of model subsets in an attempt to improve prediction accuracy by sacrificing a small quantity of bias for preserving or discarding variable(s). Thus, a continuous process of variable or model selection approach known as regularization or shrinkage methods includes ridge least squares method, least absolute shrinkage and selection operator (LASSO) and elastic net least squares method.

Lukman *et al.* (2018) study focused on identifying and retaining factors that contributed immensely to economic growth in Nigeria based on some variable selection methods such as stepwise regressions, partial least square regression, and LASSO. In the study, twelve factors were available to predict economic growth. The results revealed that stepwise regressions could have been more efficient in the presence of multicollinearity. Also, it was observed that partial least squares and the LASSO model efficiently and significantly identified the positive impact of oil revenue, non-oil revenue and capital expenditure on transfers as Nigeria’s economic growth predictors and, as such, retained them. Tuyet (2020) assessed the impact of tariff reductions on fluctuations in customs revenues in Vietnam using the data collected from the government’s web portal and the World Bank’s website between (2002-2017). The LASSO regression method was employed to estimate and predict the relationship between tariff reductions and customs revenue.

The results revealed that tariff reductions positively contributed to customs revenues. It was also found that a reduction in tariffs led to an increase in import turnover, the level of compliance with tax laws by import-export enterprises increased, and smuggling and trade fraud decreased.

Agu *et al*. (2022) conducted a study to fit predictive models for a typical dataset using the ordinary least squares method and three other methods: ridge regression, LASSO regression, and principal component regression. Their study compared the results using the estimated mean square error of these methods and found that principal component regression outperformed the others with minimal mean square error. Therefore, they suggested that principal component regression would be optimal for building predictive models for related datasets. In a separate study, Bayrakcı (2022) examined the factors influencing the profitability indicators of the banking sector in Turkey. The study considered variables such as securities portfolio, equity, NPL ratio, asset share, deposit, inflation, and gross domestic product as factors affecting the return on assets (ROA) and return on equity (ROE) ratios for the top 10 deposit banks in 2020, based on their asset size. Using the LASSO regression method, Bayrakcı found that all variables considered as factors, except deposit, did not significantly affect the ROA and ROE ratios, which are measures of profitability.

Therefore, in this chapter, we propose variable(s) or model selection method known as regularized or shrinkage technique to explore and fit stable and reliable model via continuous process selection methods due to presence of multicollinearity and outliers to predict economic growth (RGDP) for Nigeria using the identified drivers that include internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN). Thus, this approach will enrich the body of knowledge as there remains to be more knowledge in various works and studies on the aforementioned macroeconomic variables in the literature.

7.2 Research Methodology

In this study, penalized least square regression methods comprised of least absolute shrinkage and selection operator (LASSO), ridge regression and elastic net method, which is an extension of OLS by adding the penalty term, are considered in this section. This chapter shall use this approach to explore and predict economic growth based on the earlier identified drivers.

7.2.1 Average Centered OLS Panelized Regression Model

Yuzbasi et al. (2017) argued that response variables such as RGDP in this study can be influenced by other predictors. According to Hoerl and Kennard (1970), when there is dependency among predictors, using the OLS method can lead to estimation errors. Therefore, ridge regression is proposed as an alternative method for parameter estimation, introducing a penalty (ℓ) that reduces the variance of the estimates (Zhang, 2011; Hoerl and Kennard, 1970). This technique effectively shrinks the correlation coefficients of the explanatory variables (Friedman et al., 2001; Hastie et al., 2001). Without restrictions on the parameter β(s), extreme values with wide variances may result. To mitigate this, it is necessary to regularize the estimated parameters β(s) of the model (James et al., 2013).

Least Absolute Shrinkage and Selection Operator (LASSO) is a penalized technique that restricts the model's parameters restricting all the absolute values to be less than a given value. This can be done by regulating the process where the estimated parameters of the model shrink or part of the parameters are reduced to zero. In the process of selecting the parameters, variables associated with the estimated parameters that are not zero after regularization are made to be part of the variables for model specification. In this case, the forecast error will be reduced (Tibshirani *et al.*, 2015).

According to Zou and Hastie (2005), an elastic net can be used jointly for regulating and selecting variables and a group of variables where correlation coefficients exist. Elastic net, as a penalized regression technique, is a single linear function that combines Ridge and LASSO features. The linear equation can be estimated using a two-step method that involves determining parameter estimates for a given regression model as the first step and using the LASSO technique for regularization as a second step to reduce the unnecessary estimated values. This process is more efficient than the LASSO or ridge as an individual penalized technique. Thus, the performance of the penalized regression techniques such as Ridge, LASSO and Elastic Net are considered.

7.2.2 Shrinkage Model

The model specification to examine the efficiency of the aforementioned penalized regression technique is expressed as:

$$Y = X\beta + \mu, \quad (7.1)$$
$$\mu \sim N(0, \sigma^2).$$

In this case, $Y = (y_1, y_2, y_3, \ldots, y_n)'$ is a vector of response variable observations, X denoted an nxr matrix for the predictors, $\beta = (\beta_1, \beta_2, \ldots, \beta_r)$ is a vector of estimated parameters of the model and μ denoted the error terms of the model and its variance denoted by $var(\mu) = \sigma^2 I$ (Doreswamy *et al.* 2013). The coefficients of the regression model can be estimated using ordinary least square through the minimization of the expression given by:

$$\underset{\beta_0, \beta}{argmin} \{\textstyle\sum_{i=1}^{r}(y_i - \beta_0 - X_{i1}\beta_1 - X_{I2}\beta_2 - \ldots - X_{ir}\beta_r)^2\}, \quad (7.2)$$

It must be noted that, the estimated parameters $\beta_1, \beta_2, \ldots, \beta_r$ from the expression given in (7.2) are not zeros thus, a large *r* makes the interpretation of the estimated parameters to become a challenge. That is, n < r leads to inefficient estimate of the parameters using OLS. Thus, the need for alternative estimation techniques that equate the expression under consideration to be zero. Therefore, the estimated parameters are restricted or regulated (penalty).

7.2.3 Ridge Penalized Regression

This method is appropriate when many endogenous variables are independent. Specifically, ridge regression is an appropriate technique when there are many explanatory variables with small estimated parameter values and regulates the estimated variance of the correlated independent variables from obtaining a large variance that can lead to the insignificance of the estimated parameters. Hence, the estimated parameters of the linear model using the ridge penalized regression method are given as follows:

$$\hat{\beta}_{ridge} = \underset{\beta}{argmin} \parallel y - \beta' X \parallel^2 + \lambda \parallel \beta \parallel \frac{2}{2}, \quad (7.3)$$

where $\parallel y - \beta' X \parallel^2 = \sum_{i=1}^{r}(y_i - \beta' X)^2$ denoted by ℓ_2, an average loss function, $X_i{}'$ is the ith value of vector X, $\parallel \beta \parallel \frac{2}{2} = \sum_{j=1}^{r} \beta_j{}^2$ the ℓ_2 model penalty on β. Then, λ which is greater than or equal to zero, is a parameter which normalizes penalty in the model. Thus, a comparative influence of data that are not error independent can be examined by λ using cross-validated approach that allows the substitution of λ in equation (7.3), then, the next expression given can be obtained:

$$\underset{\beta_0, \beta}{argmin} \{\textstyle\sum_{i=1}^{r}(y_i - \beta_0 - X_{i1}\beta_1 - X_{I2}\beta_2 - \ldots - X_{ir}\beta_r)^2\} \; s.t \; \lambda \textstyle\sum_{j=1}^{r} \beta_j{}^2 \le t, \quad (7.4)$$

where, parameter t can be defined by the user.

7.2.4 Least Absolute Shrinkage and Selection Operator (LASSO)

This technique can regulate the sum of squares for the loss of absolute error and constraints of the sum of the absolute value of coefficients. According to Tibshirani (1996) as cited in Zari *et al*. (2019), the constraint has a regularizing effect on the parameters estimated. Also, it sets some to zero to provide a proper interpretation for the fitted regression model. In LASSO, the penalized technique is obtained as a sparse solution using the optimization problem given by:

$$\hat{\beta}_{lasso} = \underset{\beta}{argmin} \parallel y - \beta' X \parallel^2 + \lambda \parallel \beta \parallel_1, \quad (7.5)$$

where $\parallel \beta \parallel_1 = \sum_{j=1}^{r} \beta_j$, the norm penalty of ℓ_2 under β that represented the sparsity level and $\lambda \geq 0$ serve as shrinking parameter for the estimated model. Thus, expression (7.5) can be written as:

$$\underset{\beta_0, \beta}{argmin} \{\sum_{i=1}^{r} (y_i - \beta_0 - \beta_{i1}\beta_1 - X_{I2}\beta_2 - \ldots . - X_{ir}\beta_r)^2\} \; s.t \; \lambda \sum_{j=1}^{r} |\beta_j| \leq t, \quad (7.6)$$

ℓ_1 is the penalty on β that helps in reducing the number of components to zero, with excellence selection that can be carried out simultaneously; and as such, it is a good technique used to shrink the estimates of linear model.

7.2.5 Elastic Net

This is a penalized least square and variable selection technique with turning parameter $\alpha \geq 0$. It is a technique that combined the features of LASSO and Ridge models. Elastic net is designed to address the correlation and selection problem from ridge and LASSO penalized least square methods, and penalties value denoted by ℓ_1 and ℓ_2 respectively. Thus, elastic net penalized model reduced to ridge model if α = 1. According to Zhao and Yu (2006) and Doreswamy *et al*., (2013), it can be emphasized that Elastic net penalized technique consist of LASSO and ridge penalized techniques properties respectively. Thus, if α = 0, the penalty function is undefined and if α > 0 the penalty function will be convex. Elastic net regression technique with penalized parameters ℓ_1 and ℓ_2 can be expressed in the form given as:

$$\hat{\beta}_{elastic} = \left(1 + \frac{\lambda_1}{n}\right)\{argmin \parallel G \parallel^2 + \lambda_2 \parallel \beta \parallel^2 + \lambda_1 \parallel \beta \parallel_1\} \quad (7.7)$$

setting $\alpha = \frac{\lambda_2}{\lambda_1 + \lambda_2}$, the function expressed in (7.7) is similar to minimization of expression given as:

$$\hat{\beta}_{elastic} = \left(1 + \frac{\lambda_1}{n}\right)\{argmin \parallel G \parallel^2 + \lambda_2 \parallel \beta \parallel^2 + \lambda_1 \parallel \beta \parallel_1 \text{ s.t } (1-\alpha) \parallel \beta_1 \parallel + \alpha \parallel \beta \parallel^2 \leq t \ , \qquad (7.8)$$

Where $G = y - \beta' X$ and the Elastic net penalty parameters are given as $(1-\alpha) \parallel \beta_1 \parallel + \alpha \parallel \beta \parallel^2$.

Evaluation Techniques

It is important to select penalized regression models by adopting suitable evaluation techniques. Thus, in this study, the evaluation techniques to be adopted are root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and lambda penalty (λ), with the least adopted evaluation techniques selected.

7.3 Empirical Results

This section presents the empirical results of the fitted average penalized least squares regression models discussed in Section 7.2 in order to address the problem of multicollinearity among the explanatory variables; thus, in Table 7.1, we present the maximum likelihood estimate for LASSO, ridge, and elastic net as an average penalized regression.

Table 7.1: ML Estimates for the Average Centered Penalized Techniques

LASSO Technique						
Variables	Estimate	Std Error	Wald Chi-Sqr.	Prob > Chi-Sqr	Lower 95%	Upper 95%
Constant	10.3297	0.0131	614660.99	<.0001*	10.3039	10.3556
INDT	4.2717	0.5561	59.0021	<.0001*	3.1817	5.3617
EXDT	-0.9760	0.2818	11.9886	0.0005*	-1.5284	-0.4235
RINR	0.4007	0.1721	5.4163	0.0199*	0.0632	0.7382
REXR	0.4948	0.1852	7.1348	0.0076*	0.1317	0.8580
OPEN	0.5262	0.4691	1.2581	0.2620	-0.3932	1.4457
Ridge Technique						
Constant	10.2914	0.0135	574342.53	<.0001*	10.2648	10.3180
INDT	4.0771	0.6663	37.4351	<.0001*	2.7710	5.3832
EXDT	-1.1178	0.3826	8.5316	0.0035*	-1.8679	-0.3677
RINR	0.7247	0.2305	9.8784	0.0017*	0.2728	1.1767
REXR	0.6629	0.1548	18.3298	<.0001*	0.3594	0.9663
OPEN	0.7843	0.5565	1.9863	0.1587	-0.3063	1.8750
Elastic Net Technique						
Constant	10.2857	0.0123	694138.25	<.0001*	10.2615	10.3099

INDT	3.4575	0.5829	35.1736	<.0001*	2.3148	4.6001
EXDT	-0.9222	0.2935	9.8721	0.0017*	-1.4976	-0.3469
RINR	0.6414	0.2114	9.2038	0.0024*	0.2270	1.0557
REXR	0.5809	0.1335	18.9355	<.0001*	0.3193	0.8426
OPEN	1.1485	0.4905	5.4811	0.0192*	0.1870	2.1100

In Table 7.1, we present the results of an average centered penalized regression parameter estimates for each identified economic growth drivers. The economic growth drivers include internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR) and trade openness (OPEN). These are used as drivers to determine and predict economic growth RGDP (Ebiwonjumi *et al*. 2022 and 2023). This study employs the average penalised regression methods of LASSO, ridge and elastic net techniques. In Table 7.1, the results of the LASSO technique show that the identified economic growth drivers or determinants are linearly related to RGDP. The result further shows that INDT, RINR, REXR and OPEN are positive, and their contributions to RGDP are 4.27%, 0.40%, 0.49% and 0.52%, respectively. While EXDT is negative, it causes a decline in the RGDP by 0.97% during the period under study. The standard error values of 0.6663, 0.3826, 0.2305 and 0.1548 with *p*-value < 0.05 for the estimated parameters of the economic growth drivers consisting of INDT, EXDT, RINR and REXR show the statistical significance of the identified economic growth drivers in determining and predicting economic growth (RGDP) in Nigeria.

Also, in Table 7.1, we present the results of the ridge estimation technique, which reveals that the identified economic growth drivers have a linear relationship with RGDP. Specifically, the results indicate that INDT, RINR, REXR, and OPEN are positive and contribute 4.07%. 0.72%, 0.66%and 0.78%, respectively, to the economic growth rate (RGDP). The results also reveal that EXDT is negative and causes a decline in economic growth (RGDP) by 1.11% in the period under investigation. The standard error values of 0.6663, 0.3826, 0.2305 and 0.1548 and the *p*-value < 0.05 for the estimated parameters such as INDT, EXDT, RINR and REXR evidently establish the statistical significance of the identified economic growth drivers in determining and predicting economic growth (RGDP) in Nigeria.

The elastic net estimation method reveals that economic growth drivers are linearly related to economic growth (RGDP). Thus, in Table 7.1, the results show that INDT, RINR, REXR and

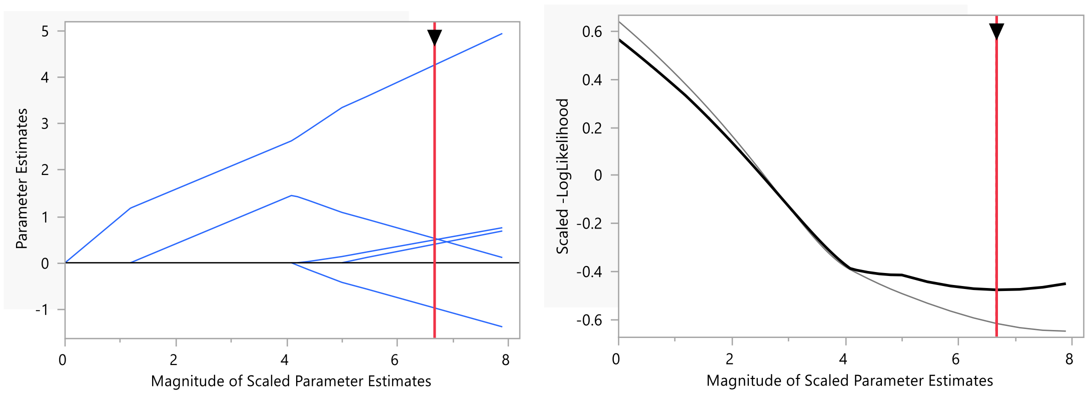
OPEN contribute positively to the RGDP in Nigeria to the turn of 3.45%. 0.64%, 0.58% and 1.14%, respectively. It also reveals that the contribution of EXDT is negative and, as such, leads to a decline in the economic growth rate (RGDP) by 0.92%. The standard error values of 0.5829, 0.2935, 0.2114, 0.1335 and 0.4905 and p-value < 0.05 for the estimated parameters of the aforementioned economic growth drivers (INDT, EXDT, RINR and REXR), evidently show the statistical significance of the identified economic growth drivers in examining and predicting economic growth rate (RGDP) in Nigeria.

Table 7.2: Distribution and Scaled Parameter for Average Centered Penalized Techniques

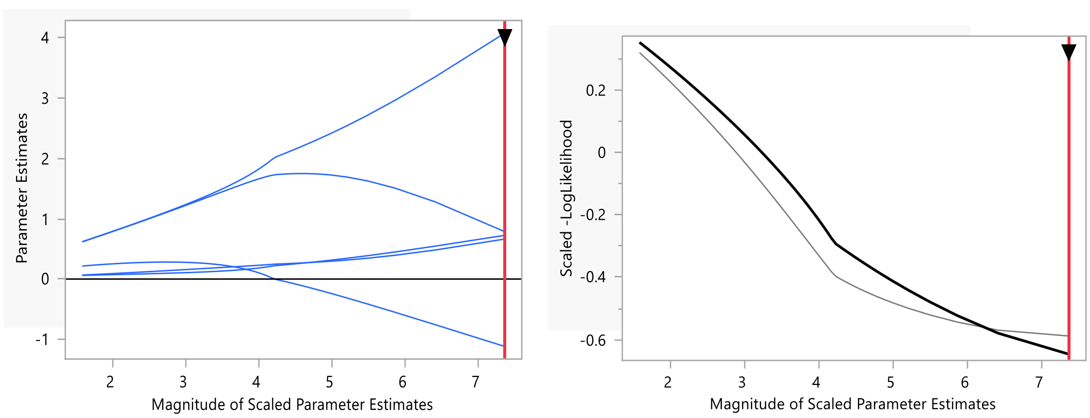
Normal Distribution Parameters	Estimate	Std Error	Wald ChiSquare	Prob > ChiSquare	Lower 95%	Upper 95%
LASSO Scale	0.1304	0.0149	76.6142	<.0001*	0.1012	0.1596
Ridge Scale	0.1344	0.0160	70.4006	<.0001*	0.1030	0.1658
Elastic Net Scale	0.1222	0.0119	104.0088	<.0001*	0.0987	0.1457

In Table 7.2, we present the results of penalty parameter lambda for the penalized regression models: LASSO, ridge and elastic net technique employed in this study. Thus, from the results, the distribution parameter lambda or estimated scale for the LASSO, ridge and elastic net are 0.1304, 0.1344 and 0.1222, respectively. The standard error values of the techniques are 0.0149, 0.0160 and 0.0119 with p-value < 0.05. As such, it emphasizes the significance of the scale parameter establish the identified economic growth drivers and the RGDP under consideration need penalization. Thus, in Figure 7.1, the solution path for the employed average penalized least square regression techniques is shown, and it reveals the trends for the 98 training and 43 validated data used for this study. The solution path describes the magnitude of the scaled parameters and efficiently maximizes the likelihood function of the estimates. The vertical line in the above plot indicates the optimal value for lambda. The following two plots show the display of the coefficient 'path' and the sorted magnitude of the coefficients at the optimal lambda.

LASSO Technique



Ridge Technique



Elastic Net Technique

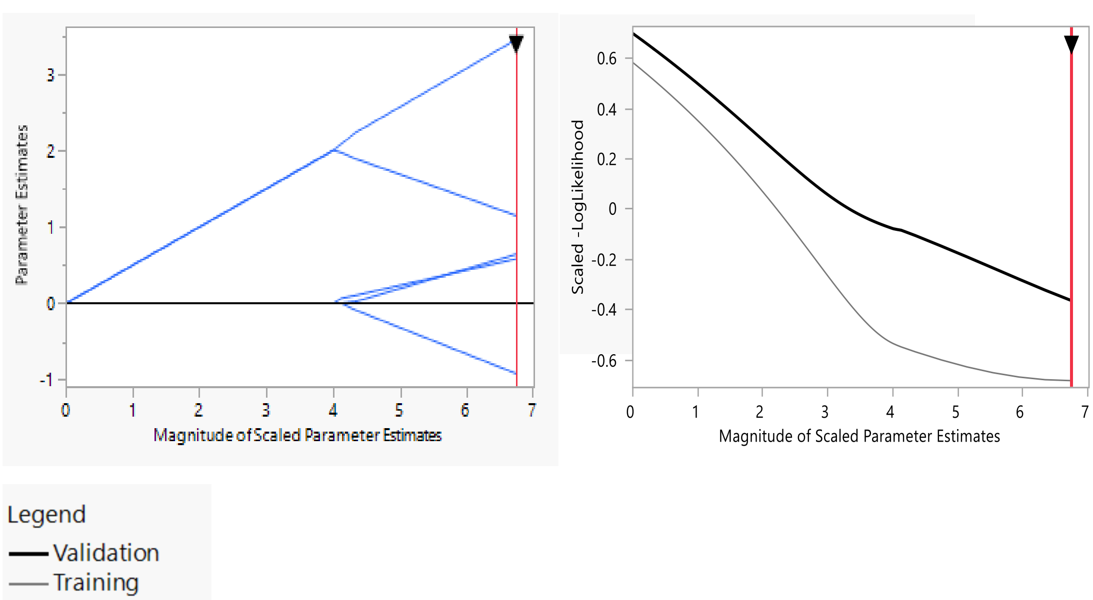
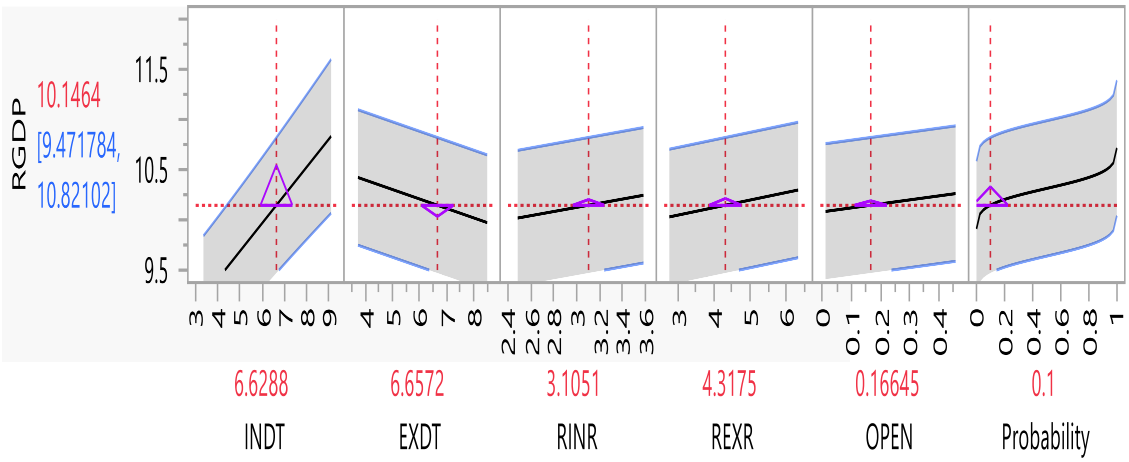


Figure 7.1: Solution Path for the Average Centered Penalized Techniques

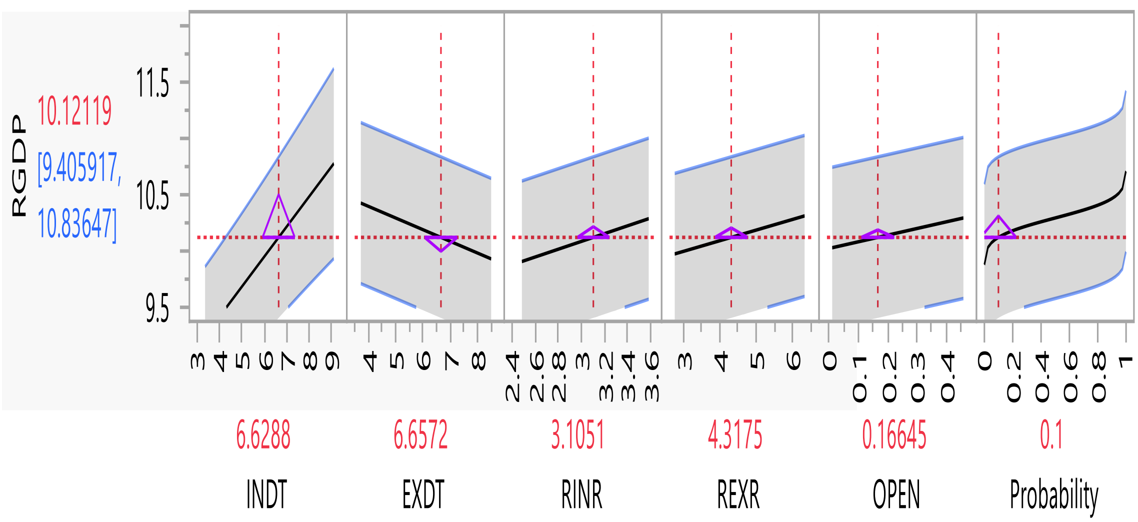
Quantile Prediction for Average Centered Penalized Techniques

The quantile prediction plot is shown in Figure 7.2 for the penalized regression techniques used for this study that include LASSO, ridge and elastic net method.

LASSO Technique



Ridge Technique



Elastic Net Technique

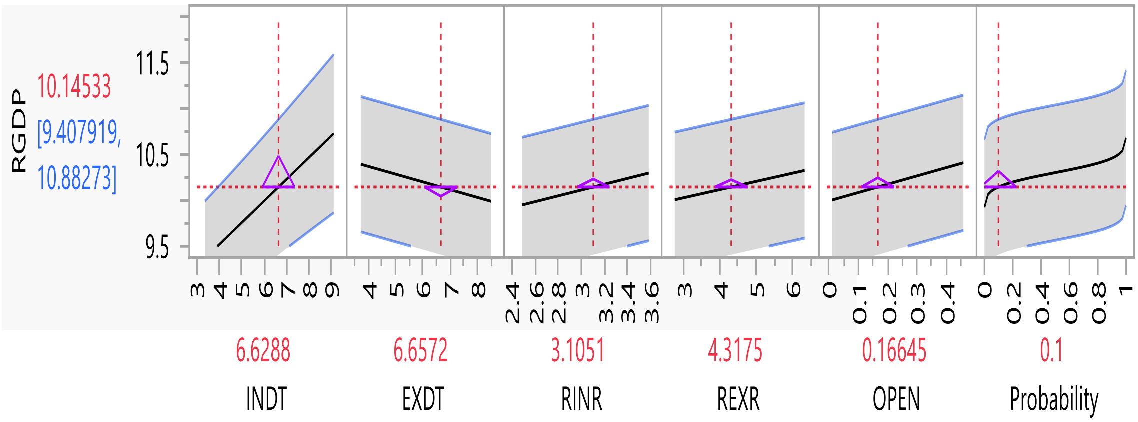
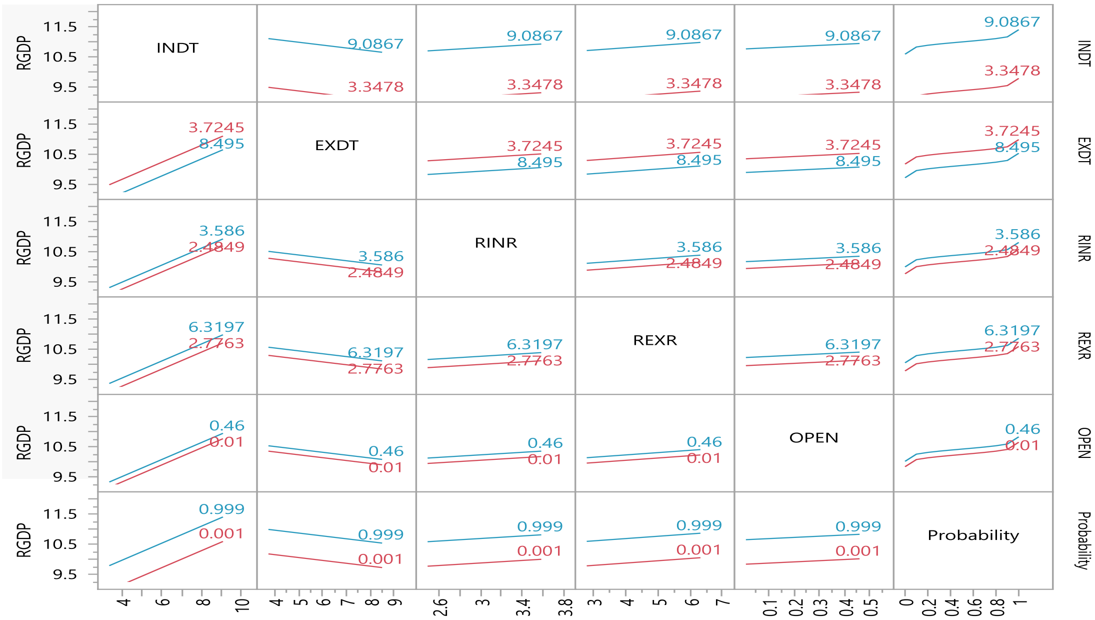


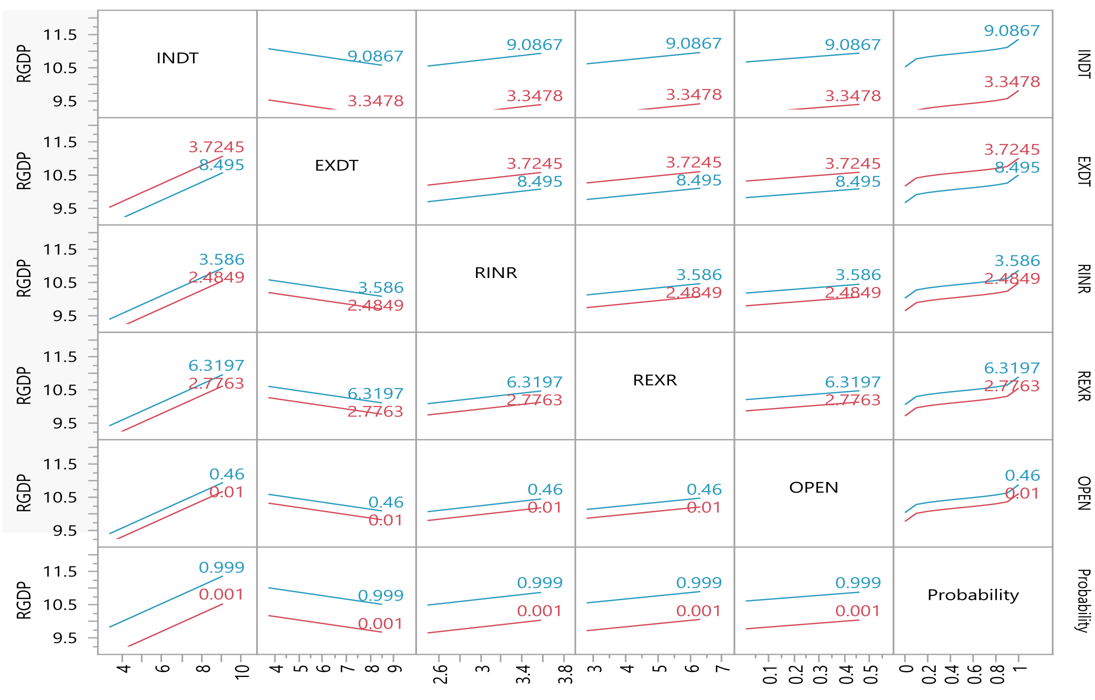
Figure 7.2: Quantile Prediction Plot for Average Centered Penalized Techniques

In Figure 7.2, we present plots that reveal the sensitivity of the LASSO, ridge and elastic net as the penalized regression technique employed for predicting economic growth (RGDP). In the graph, the dotted red line represents expected distribution of the p-value while, the blue line represents the observed distribution. Thus, it is indicated that 6.6288% of INDT, 6.6572% of EXDT, 3.1051% of RINR, 4.3175% of REXR and 0.16645% of OPEN predicted 10.1464% of the economic growth in RGDP at the error margin of 0.1 or 10% error margin, when LASSO penalized regression technique is employed. The 95% confidence interval reveals that the predicted economic growth (RGDP) value is 10.1464%. This is good and precision because the value is within the confidence interval, as shown in the LASSO plot. In Figure 7.2, the plot reveals that 6.6288% of INDT, 6.6572% of EXDT, 3.1051% of RINR, 4.3175 of REXR and 0.16645% of OPEN efficiently predict 10.1212% of economic growth rate (RGDP) at 10% error margin when ridge penalized regression technique is used. The 95% confidence interval reveals that the predicted value for economic growth (RGDP) is 10.1212%. This is good because the value is within the confidence interval, as the ridge plot shows. Also, the plot shows that 6.6288% of INDT, 6.6572% of EXDT, 3.1051% of RINR, 4.3175 of REXR and 0.16645% of OPEN predict economic growth RGDP to be 10.1453% at 10% error margin, when elastic net penalized regression technique is adopted. The 95% confidence interval reveals that the predicted economic growth (RGDP) value is 10.1453%. This is good and signifies precision because the value is within the confidence interval, as shown in the elastic net plot. Thus, it can be emphasized that the LASSO penalized regression technique has the highest predictive value of 10.1464% for economic growth (RGDP) in Nigeria compared to the predicted ridge and elastic net values. This is also evidence in the precision of the solution path plots for the penalized techniques as shown in Figure 7.1 . The plot among the economic growth drivers is considered, and it is shown in Figure 7.3 for each of the average penalized regressions under consideration.

LASSO Technique



Ridge Technique



Elastic Net Technique

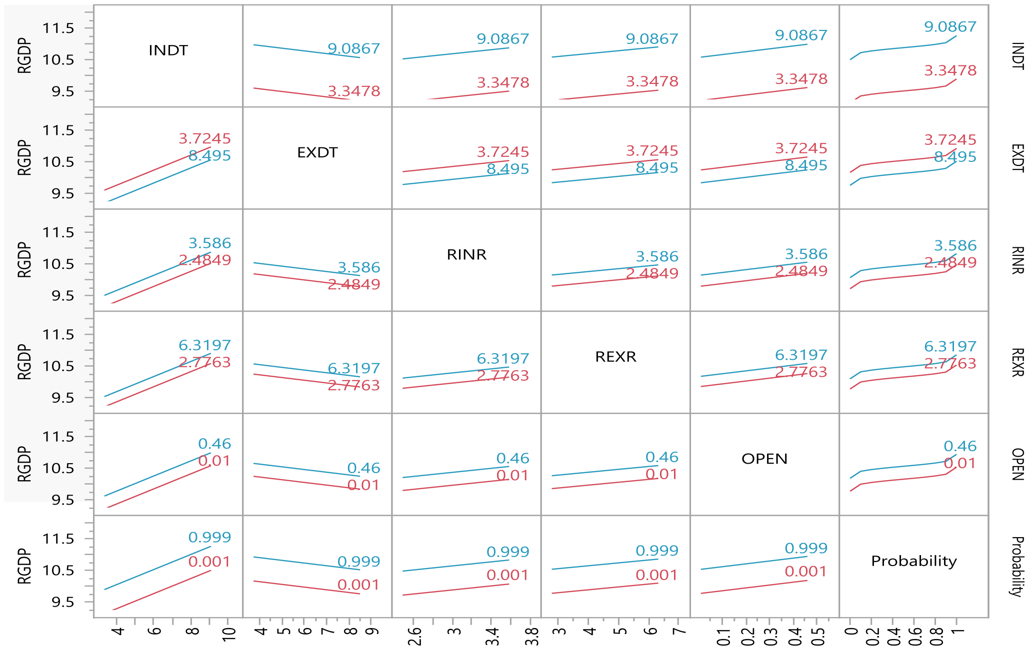
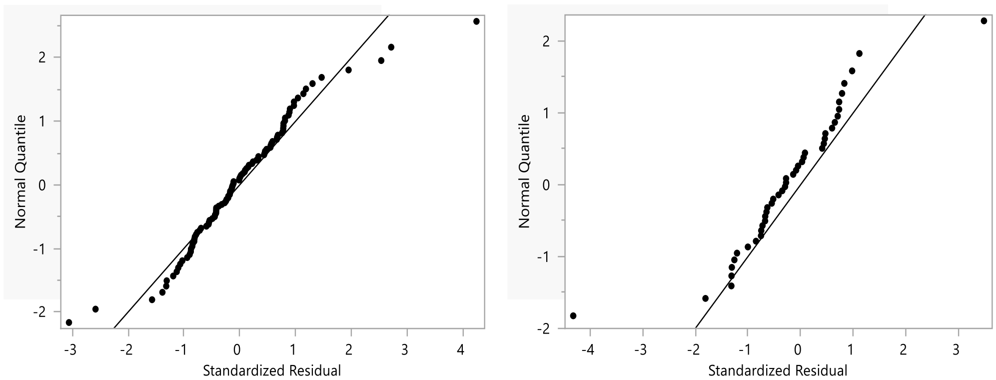


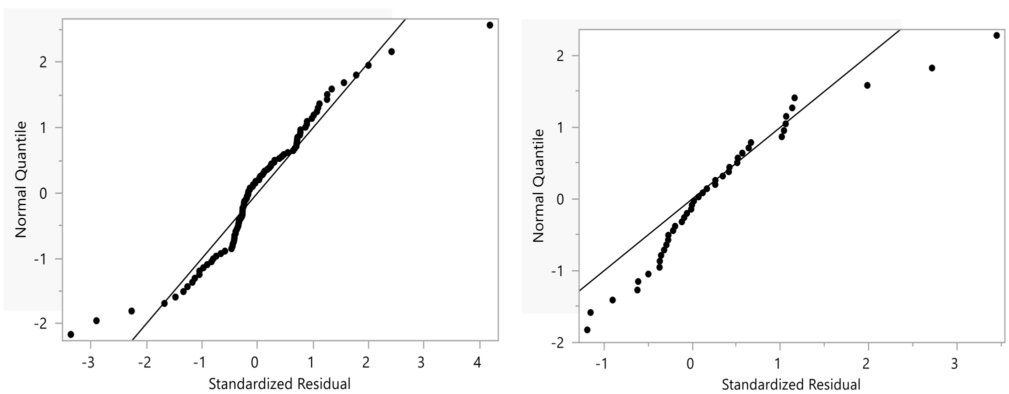
Figure 7.3: Interaction Plot for the Average Centered Penalized Techniques

In Figure 7.3, we present the plots for the interaction profile of the average centered penalized method such as lasso, ridge and elastic net employed in this study. In Figure 7.3, the blue and the red line show the relationship between two predictors such as INDT and EXDT, INDT and RINR, INDT and REXR and INDT and OPEN and so on in determining their joint effect on RGDP. Thus, the interaction plots show evidence of parallel lines between INDT and EXDT, INDT and RINR, INDT and REXR, and INDT and OPEN for all the three average penalized regression methods employed for this study. Thus, it indicates the independence of the identified economic growth drivers and such do not influence each other in determining and predicting economic growth rate (RGDP). Following the interaction plots, we also consider the normal quantile plots for each method's average panelized methods under study, as shown in Figure 7.4.

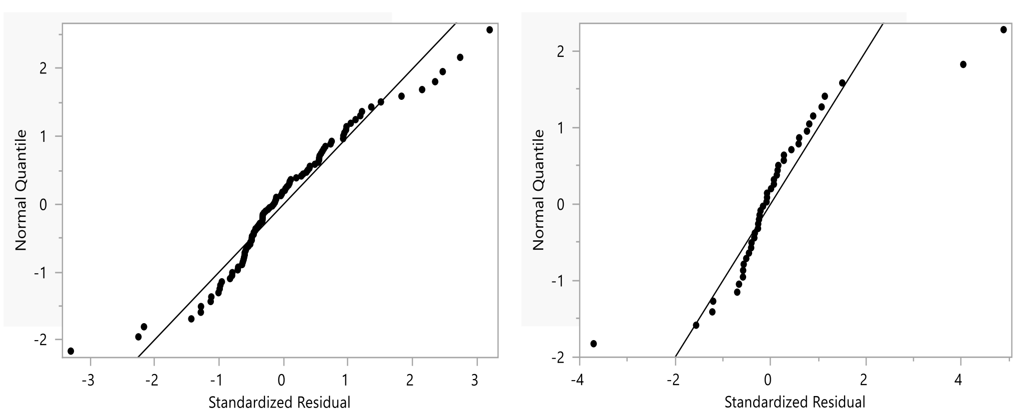
LASSO Technique



Ridge Technique



Elastic Net Technique

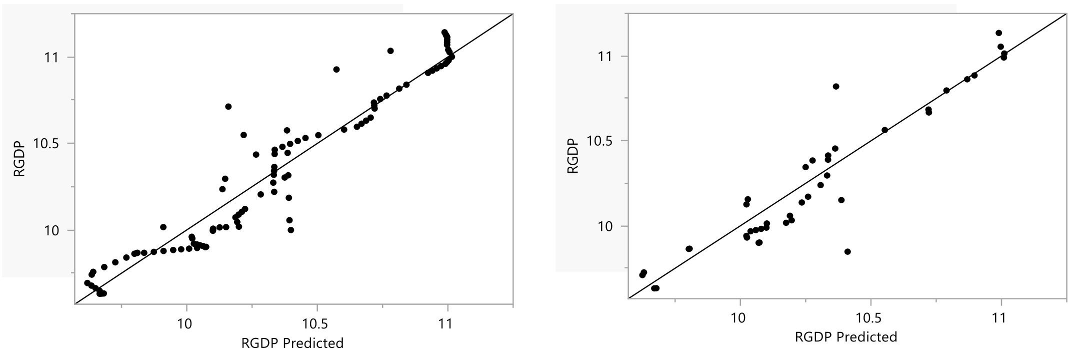


Training Validation

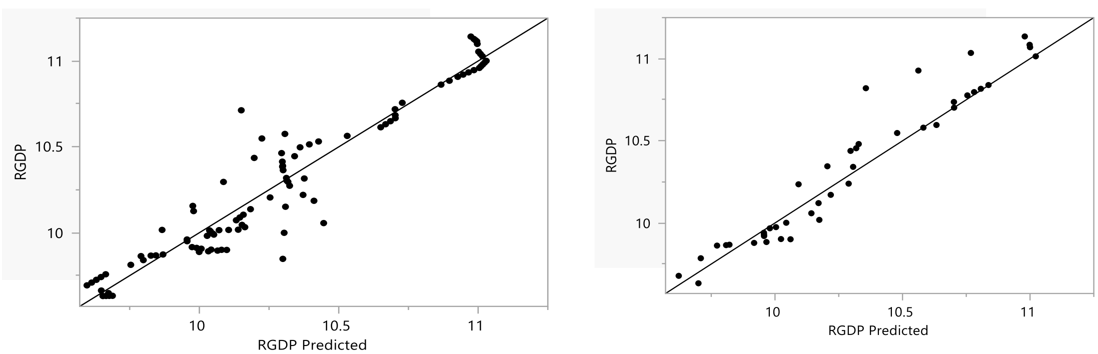
Figure 7.4: Normal Quantile Plot for the Average Centered Penalized Techniques

The actual predicted plot for the average centered penalized techniques used in this study is shown in Figure 7.5.

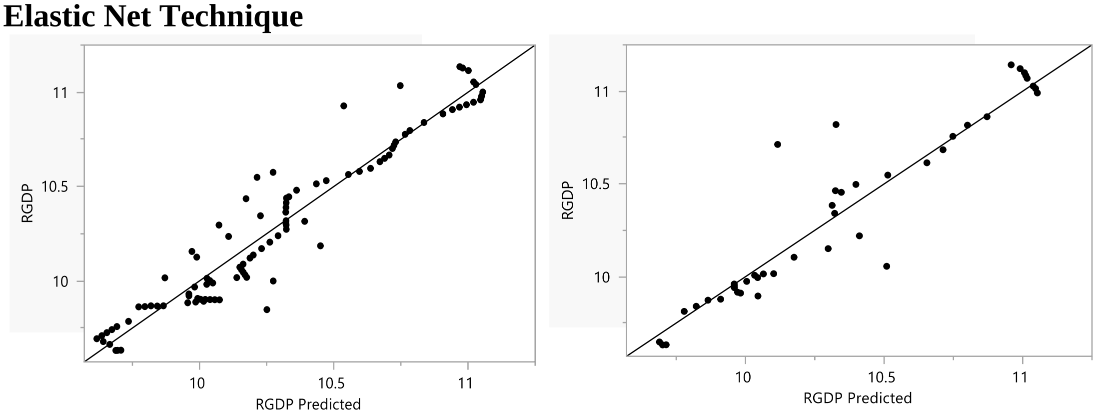
LASSO Technique



Ridge Technique



Elastic Net Technique



Training Validation

Figure 7.5: Predicted Plot for the Average Centered Penalized Techniques

In Figure 7.4 and Figure 7.5, we present normal quantile plots and the predicted plots for the average centered penalized regression techniques that include LASSO, ridge and elastic net employed for this study. The plots show that the data set under consideration is normally distributed, as revealed by all the techniques. Thus, the dotted lines represent the individual predicted observations during the period under consideration using the average centered penalized least methods.

Having determined the contributions of the identified economic growth drivers and established their independence using various fitted average centered penalized least squares techniques, it is also imperative to investigate the predictive power of these methods. Evaluating their ability to make stable and reliable predictions or forecasts for the economic growth rate (RGDP) is crucial. Therefore, performance metrics must be considered to assess the predictive strength of the methods discussed in this chapter.

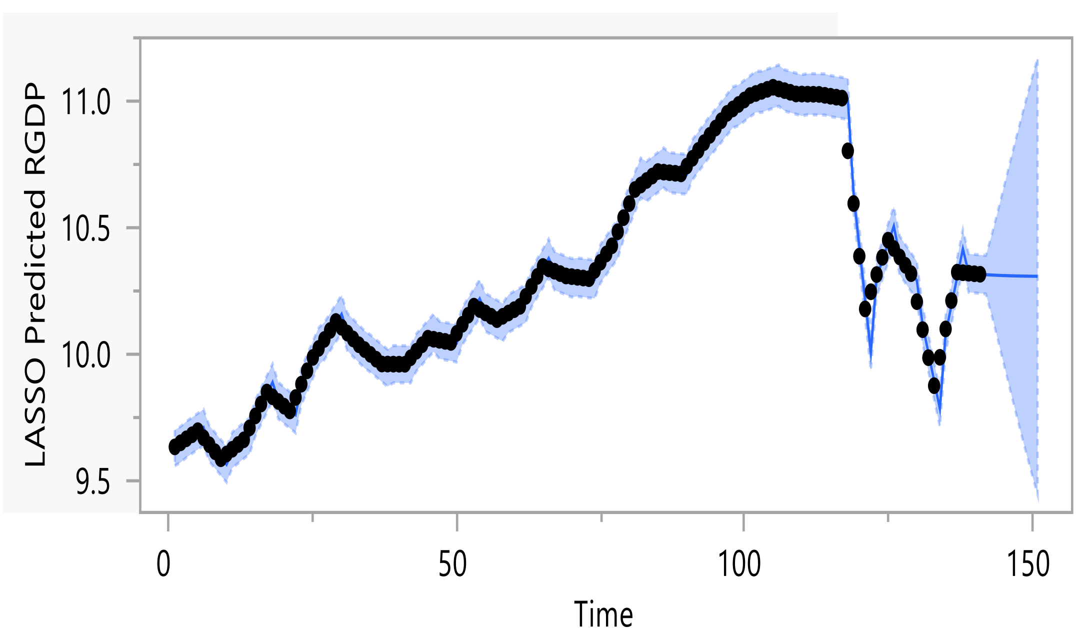
Table 7.3: Performance Metric of the Average Centered Penalized Regression Models

	LASSO Technique		Ridge Technique		Elastic Net Technique	
Evaluation M.	Training	Validation	Training	Validation	Training	Validation
Rows	98	43	98	43	98	43
Sum of Freq	98	43	98	43	98	43
-LogLikelihood	-60.5600	-20.5270	-57.5990	-27.7730	-66.9370	-15.6620
BIC	-89.0250	-14.7260	-83.1040	-29.2170	-101.7790	-4.9960
AICc	-105.875	-23.855	-99.954	-38.346	-118.630	-14.1250
R-Square	0.9190	0.8730	0.9100	0.9190	0.9200	0.8920
Lambda Penalty	0.0840		-		0.0000	
RMSE	0.2895		0.3197		0.2979	
MAE	0.2174		0.2400		0.2237	
MAPE	2.1298		2.3489		2.1985	

In Table 7.3, we present the results of various performance metric evaluations to be considered for selecting the best-fit model among the average-centered penalized least square techniques employed for this study. Here, we use root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error and lambda penalty. Thus, from the results, the RMSE for the fitted LASSO, ridge and elastic net penalized least square techniques are 0.2895, 0.3197 and 0.2979, respectively. The MAE for the fitted LASSO, ridge and elastic net penalized least square techniques are 0.2174, 0.2400 and 0.2237, respectively. The MAPE for the fitted LASSO, ridge and elastic net penalized least square techniques are 2.1298, 2.3489 and 2.1985, respectively. The R-Square for the fitted LASSO, ridge and elastic net penalized least square techniques were 0.919, 0.910 and 0.920, respectively, indicating that identified economic growth drivers explained 91.9%, 91.0% and 92.0% variations or changes in the economic growth in Nigeria using LASSO, ridge and elastic net as penalized least square regression techniques respectively. The lambda penalty for the fitted model using the LASSO technique is 0.0840, which is higher than that of the elastic net technique value of 0.0000. Thus, it can be emphasized, based on these performance metrics that LASSO technique technique is the most efficient average centered penalized regression model for this study.

7.4 Forecast for RGDP using LASSO Penalized Regression Method

Based on the RMSE, MAE, MAPE and lambda penalty, LASSO is selected as the most efficient average-centered penalized least regression method to be used in predicting a stable and reliable economic growth rate (RGDP) for this study. Therefore, in the next ten (10) quarters, the economic growth rate (RGDP) is predicted or forecasted using the LASSO penalized regression model. The result is shown as forecast in the plot presented in Figure 7.6.



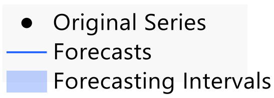


Figure 7.6: Forecast Plot for the adopted Average Centered Penalized Techniques

7.5 Concluding Remarks

In this chapter, we explore and estimate the parameters of the identified economic growth drivers (INDT, EXDT, RINR, REXR, and OPEN) to predict the economic growth rate (RGDP) for Nigeria. An exploratory and diagnostic analysis reveals the existing relationships between economic growth (RGDP) and its identified drivers or determinants. However, due to the presence of multicollinearity and outliers in the dataset, ensuring robust estimates for predicting economic growth is essential. To address these issues simultaneously and obtain efficient parameter estimates, we employ an average cantered penalized regression technique. This technique includes LASSO, ridge, and elastic net regressions, which are fitted and discussed to select the optimal method for predicting economic growth using the identified drivers.

From the LASSO results, it is revealed that the effects of ‘INDT, RINR, REXR and OPEN’ are positive and their contributions to economic growth rate RGDP are 4.27%, 0.40%, 0.49% and 0.52% respectively. while, the effect of EXDT is negative and as such causes a decline in the economic growth rate (RGDP) to the turn 0.97% during the period under study, Also, in ridge estimation technique, the results reveal that INDT, RINR, REXR and OPEN are positive and contributes 4.07%. 0.72%, 0.66%and 0.78%, respectively, to the economic growth rate (RGDP); the results also show that EXDT is negative and, as such, causes a decline in economic growth (RGDP) by 1.11% in the period under investigation. The elastic net method results show a positive contribution of ‘INDT, RINR, REXR and OPEN’ to economic growth (RGDP), amounting to 3.45%. 0.64%, 0.58% and 1.14%, respectively. The elastic net method results further reveal the contribution of EXDT to RGDP to be negative, leading to a decline in RGDP by 0.92% during the period under consideration. The results show the statistical significance of the identified economic growth drivers in determining and predicting economic growth RGDP using average penalized regression methods discussed in this study.

Also, the RMSE for the fitted LASSO, ridge and elastic net penalized least square techniques are 0.2895, 0.3197, and 0.2979, respectively, and MAE for the fitted models are 0.2174, 0.2400 and 0.2237, respectively, and the MAPE for the fitted models are 2.1298, 2.3489 and 2.1985 respectively. The R-Square for the fitted LASSO, ridge and elastic net penalized least square techniques indicate that the identified economic growth drivers explained the 91.9%, 91.0% and 92.0% variations or changes in the economic growth rate in Nigeria when LASSO, ridge and elastic net, as penalized least square regression techniques are respectively employed. The lambda penalty for the fitted model using the LASSO technique is 0.084, which is higher than that of the elastic net technique value of 0.000. However, the ridge technique must be more efficient in generating a lambda penalty.

Thus, it can be emphasized that the sensitivity of LASSO, ridge, and elastic net as penalized regression techniques is crucial for predicting the economic growth rate in Nigeria. Among these, LASSO stands out as the most sensitive and efficient technique for examining and predicting RGDP in Nigeria, due to its minimal values of RMSE, MAE, and MAPE. LASSO is shown to provide stable and reliable forecasts even in the presence of multicollinearity and outliers, outperforming the other penalized regression methods discussed in this study.

Therefore, based on the findings, LASSO as an average centered penalized regression method is recommended as the most efficient and appropriate technique for predicting the state of the economy in Nigeria. Additionally, there is an urgent need to address the retrogressive economic situation through policy formulation that focuses on the identified economic growth drivers or determinants.

To further examine and predict stable and reliable economic growth (RGDP) for Nigeria, another technique called the Gaussian process regression method is proposed and discussed in detail in Chapter 8 of this study.

CHAPTER 8

GAUSSIAN PROCESS REGRESSION METHOD FOR ESTIMATING AND PREDICTING ECONOMIC GROWTH

8.1 Introduction

In this chapter, we propose the Gaussian process regression method to enhance the estimated parameters' stability for efficient and reliable economic growth prediction (RGDP). According to Ria and Yogo (2021), the Gaussian Process regression method is a flexible statistical method that provides a good fit for various types of data, structures, and distributions and produces better predictions. Recently, these methods have been predominantly used, including non-parametric analysis and latent stochastic processes (stationarity and non-stationarity methods) such as the Gaussian process regression technique. Qamar and Tokdar (2014) asserted that Gaussian process regression is one of the important techniques to be considered for data analysis in time series data. Interestingly, this technique and others have been embedded within the machine learning community to address nonlinear effects and non-stationary signals for accurate, appropriate inference from the fitted model. Durrande *et al*. (2010) emphasized that the Gaussian process regression model is beneficially useful for prediction, optimization, or Monte Carlo-based quantification of uncertainty.

Dejan *et al*. (2011) applied the Gaussian process regression method as an emerging non-parametric Bayesian model to model and estimate daily financial data from the U.S. commodity market. The results obtained from the fitted Gaussian process regression model were compared with the Bayesian vector autoregression model and benchmark model that was commonly used in financial mathematics. The results revealed that the Gaussian process regression method was appropriate in modelling and estimating financial data because it provided stable and reliable results like the Bayesian vector autoregression model. Sameh and Moh'd (2021) conducted a study to forecast monthly gasoline prices in Jordan using Gaussian process regression. The monthly prices of two types of gasoline (octane-90 and octane-95) were collected between January 2008 and December 2019. The performance prediction metrics considered for the study were RMSE and MAPE. Thus, the results showed that the Gaussian process regression model could predict gasoline prices accurately. Hence, this helped the policymaker formulate and implement fuel subsidies and tax policies. Ria and Yogo (2021) applied Gaussian process regression to model and estimate future claims for motor vehicle insurance. The data used for the study were historical data on the motor vehicle insurance business line of PT XYZ from January 2017 to December 2019. The Gaussian process regression method employed for analysis revealed that the method accurately predicted future claims from PT XYZ insurance.

Kamonrat *et al*. (2022) examined the Gaussian process regression method's predictive power on Thailand's stock price. The stock price data in Thailand were divided into 2 data sets: the data in 2015-2020 and the data in 2020 due to the massive change during the COVID-19 pandemic. Gaussian process regression was employed for the prediction, and the results were compared with the artificial neural network and recurrent neural network. The evaluation performance metrics such as the RMSE, MAE and MAPE were used, and the results revealed that Gaussian process regression performed better than the other methods for both data sets with high predictive accuracy. Therefore, the Gaussian process regression method was more appropriate and reliable for the prediction of stock price. Alexandra and Vasiliy (2022) assessed the possibility of using Gaussian process regression as a surrogate modelling method to replace time-consuming calculations related to the modelling of COVID-19 dynamics. The Gaussian process regression was used as a surrogate to replace detailed simulations with a COVID-19 multiagent model. Experiments were conducted with various kernels, and in accordance with the quality metrics of the model, kernels were identified in which the Gaussian process regression method gave the most accurate result. It was further revealed that the results showed the potential and possibility of using Gaussian process regression to model and conduct an uncertainty qualification of the multiagent model of COVID-19 propagation. Therefore, to avoid spurious estimation and ensure efficient prediction, a Gaussian process regression method, a machine learning technique, is used to explore, estimate and predict economic growth based on its identified drivers under consideration in this study.

8.2 Research Methodology

This section discusses the Gaussian process regression method, marginal functions, kernel functions, and stationarity and non-stationarity kernel in detail.

8.2.1 Gaussian Process Regression Model

Consider economic growth rate (RGDP) as the target variable denoted by y as presented in column vector $y = (y_1, y_2, \dots, y_N)'$ and the explanatory variables comprising of identifying economic growth drivers denoted by $X = (x_1, x_2, \dots, x_N)$, where $x_i = (x_{i1}, x_{i2}, \dots, x_{id})'$is a *d*-dimensional column vector and *d* is the number of explanatory variables under consideration. The domain of the $\mathrm{j^{th}}$ variable is denoted by $\mathrm{X_j}$ and the joint domain of all explanatory variables is denoted by $\mathrm{X} = \mathrm{X}_1 \times \mathrm{X}_2 \times \dots \times \mathrm{X_d}$. A Gaussian process (GP) stated according to Rasmussen and Williams (2006) is an expression for the distribution of nonlinear functions. Thus, for $x,\ x' \in X$, GP is defined as:

$$f(x) \sim GP\big(\mu(x), k(x, x')\big)\,, \tag{8.1}$$

where $\mu(x)$ is the mean and $k(x, x')$ is a positive-semi-definite kernel function that defines the covariance between any two realizations of $f(x)$ and $f(x')$ which can be expressed by

$$k(x, x') = cov\big(f(x), f(x')\big)\,. \tag{8.2}$$

The mean is often assumed to be zero, i.e., $\mu(x) = 0$, and the kernel has parameter θ, $i.e., k(x, x')|\theta)$. For any finite collection of inputs $X = (x_1, x_2, \dots, x_N)$, the function values $f(X) = (f(x_1), f(x_2), \dots, f(x_N))'$ have joint multivariate Gaussian distribution

$$f(X) \sim \mathrm{N}(0, \mathrm{K}x, x(\theta))\,, \tag{8.3}$$

where elements of the NxN covariance matrix are defined by the kernel $[\mathrm{K}x, x(\theta)\,]_{i,j} = k\big(x_i, x_j\big)|\theta)$. Thus, the use of the following hierarchical Gaussian process model given by

$$\left.\begin{aligned} \theta &\sim \tau(\varphi) \\ f &\sim \mathrm{N}(0, \mathrm{K}x, x(\theta)) \\ y &\sim \mathrm{N}(f, \sigma_\varepsilon^2 I) \end{aligned}\right\}\,, \tag{8.4}$$

where $\tau(\varphi)$ is defined as the prior for the kernel parameters (including σ_ε^2), σ_ε^2 is the noise variance and *I* is the NxN identity matrix. According to Rasmussen and Williams (2006), a Gaussian noise model, the marginal of f can be analytically determined by the expression given in the form:

$$p(y|X, \theta) = \int p(y|f, X, \theta)p(f|X, \theta)df \tag{8.5}$$

$$p(y|X, \theta) \sim \mathrm{N}(0, \mathrm{K}x, x(\theta) + \sigma_\varepsilon^2 I)\,, \tag{8.6}$$

In the expression given in (8.5) and (8.6), a flexible model is the additive GP model with *D* kernels which is defined as

$$f(x) = f^{(1)}(x) + f^{(1)}(x) + \dots + f^{(D)}(x)\,, \tag{8.7}$$

$$y = f(x) + \varepsilon \quad , \tag{8.8}$$

where each

$$f^{(j)}(x) \sim GP\left(0, k^{(j)}\left(x, x' \mid \theta^{(j)}\right)\right)$$

is a distinct GP with kernel specific parameters $\theta^{(j)}$ and ε is the additive Gaussian noise. By definition, it must be noted that for any number of explanatory variables $X = (x_1, x_2, \dots, x_d)$, each $\text{GP} = f^{(j)}(X)$ follows a multivariate Gaussian distribution. Since a sum of multivariate Gaussian random variables is still Gaussian, the latent function f also follows a multivariate Gaussian distribution denoted by the expression given by

$$\Phi = \theta^{(1)}, \theta^{(2)}, \dots, \theta^{(j)}, \sigma_\varepsilon^2 I). \tag{8.9}$$

The marginal likelihood for the target variable y can be given in the form express as:

$$p(y|X, \Phi) = \mathrm{N}\left(0, \sum_{j=1}^{D} \mathrm{K}_{X,X}^{(j)}\left(\theta^{(j)}\right) + \sigma_\varepsilon^2 I\right), \tag{8.10}$$

where the latent function *f* has been marginalized out as given in (8.5) and (8.6). To simplify the notation, it can be expressed as:

$$\mathrm{K}_y(\,\Phi) = \sum_{j=1}^{D} \mathrm{K}_{X,X}^{(j)}\left(\theta^{(j)}\right) + \sigma_\varepsilon^2 I. \;, \tag{8.11}$$

This is done for the purpose of identifying explanatory variables that are important in predicting the target (response) variable. Thus, it is assumed that each Gaussian process depend only on small important explanatory variables that can be selected as subset of the identified explanatory variables.

$$f^{(j)}(\boldsymbol{x}) \colon \mathrm{X}^{(j)} \longrightarrow y$$

where

$$\mathrm{X}^{(j)} = \prod \mathrm{X}_i \,, \mathrm{i} \in \mathrm{I}_j \subseteq \{1, 2, \dots, d\}\,, \tag{8.12}$$

and y is the domain for target (response) variable. I_j are indices of the explanatory variables associated with the j^{th} kernel.

8.2.2 Kernel Functions for Explanatory Variables

In this study, the explanatory variables to be used are the identified economic growth drivers, which are continuous in nature. Thus, it is essential that when fitting an additive Gaussian process regression model, there is need to determine appropriate kernels for different explanatory variables under consideration and their interactions.

8.2.3 Stationary kernels

A kernel involving one or two explanatory variables is specified to fit a stationary Gaussian process regression (GP) model. Thus, a squared exponential (SE) kernel for continuous explanatory variables is given in the form expressed as

$$\mathrm{K}_{se}\left(x_i, x_j \,|\, \theta_{se}\right) = \sigma_{se}^2 \, exp\left(-\frac{(x_i - x_j)^2}{2\rho_{se}^2}\right), \tag{8.13}$$

where ρ_{se} is the length-scale parameter, σ_{se}^2 is the magnitude parameter and $\theta_{se} = (\rho_{se}, \sigma_{se}^2)$. Length-scale ρ_{se} controls the smoothness and magnitude parameter, σ_{se}^2controls the magnitude of the kernel. Periodic kernel for continuous explanatory variables is given as:

$$\mathrm{K}_{pe}\left(x_i, x_j \,|\, \theta_{pe}\right) = \sigma_{pe}^2 \, exp\left(-\frac{2sin^2(\tau(x_i - x_j)/\alpha)}{\rho_{\mathrm{pe}}^2}\right), \tag{8.14}$$

where ρ_{pe} is the length-scale parameter, σ_{pe}^2 is the magnitude parameter, α is the period parameter and $\theta_{pe} = (\rho_{pe}, \sigma_{pe}^2, \alpha)$ Length-scale ρ_{pe} controls the smoothness, σ_{pe}^2 controls the magnitude and α is the period of the kernel. In this model, α corresponds to a year. Constant kernel is given as:

$$\mathrm{K}_{co}\left(x_i, x_j \,|\, \theta\right) = \sigma_{co,}^2 , \tag{8.15}$$

where $\theta = (\sigma_{co}^2\,)$ is the magnitude parameter of the constant signal.

In categorical form, the kernel for explanatory variables or discrete-valued covariates is given by

$$\mathrm{K}_{ca}\left(x_i, x_j\right) = \begin{cases} 1, & if\ x_i = x_j \\ 0, & otherwise \end{cases}, \tag{8.16}$$

Also, for explanatory variables that assume binary form, its kernel can be expressed as

$$\mathrm{K}_{ca}\left(x_i, x_j\right) = \begin{cases} 1, & if\ x_i = 1\ and\ x_j = 1 \\ 0, & otherwise \end{cases} \tag{8.17}$$

The product of two kernels can also be obtained between any two valid kernels such as the pairs of kernels represented by $\mathrm{K}_{bi}(.)\ and\ \mathrm{K}_{se}(.)$. Thus, a product kernel can be expressed as

$$\mathrm{K}_{bixse}(.) = \mathrm{K}_{bi}\left(x_{ip}, x_{jp} \,|\, \theta_{bi}^{(p')}\right) \ \mathrm{K}_{se}\left(x_{iq}, x_{jq} \,|\, \theta_{se}^{(q')}\right), \tag{8.18}$$

where $\theta_{bi}^{(p')}$ and $\theta_{se}^{(q')}$ are kernel parameters for the $\mathrm{p^{th}}$ and $\mathrm{q^{th}}$ covariates, respectively.

8.2.4 Non-stationary Kernel

It is important and reliable to emphasize that the target (response) variable, which in this study, involves changes in an economic growth rate that the identified economic growth drivers can determine. This makes fitting Gaussian process regression with a squared exponential kernel challenging since the kernel is stationary, so changes are constant over the period under consideration. It can be emphasized, according to Heinonen *et al*. (2016), Tolvanen *et al*. (2014) and Saul *et al*. (2016), that a non-stationary Gaussian process regression can be fitted by using the special non-stationary kernel, such as the neural network kernel that can be defined by the kernel parameters depending on explanatory variables to be considered or via input or output warpings (Snelson *et al*., 2004). This study adopted the input warping approach and define a bijective mapping $\psi : (-\infty, +\infty) \rightarrow (-c, c)$ for a continuous explanatory variables identified for this study at period t as given by

$$\psi(t) = 2c, \left(-0.5 + \frac{1}{1 + \exp\left(-a(t-b)\right)} \right), \tag{8.19}$$

where $\mathrm{a, b,}$ and c are predefined parameters: a controls the size of the effective time window, b controls its location, and c controls the maximum range. The non-stationary kernel can then be defined as:

$$\mathrm{K}_{ns}(\, t, t' \,|\, \theta_{se}) = \sigma_{se}^2 \, exp\left(-\frac{(\psi(t) - \psi(t'))^2}{2\rho_{se}^2} \right), \tag{8.20}$$

where θ_{se} are the parameters of the SE kernel.

8.3 Empirical Results

This section presents the empirical results of the fitted Gaussian process regression model discussed in Section 8.2 to address the problem of multicollinearity and outliers in this study. Thus, Table 8.1 presents the maximum likelihood estimates for the Gaussian process regression method.

Table 8.1: ML Estimates for Gaussian Process Regression Model

Panel A			
Variables	Theta	Total Sensitivity	Main Effect
INDT	0.4969	0.5903	0.3853
EXDT	0.1140	0.1012	0.0021

RINR	0.7424	0.0756	0.0137
REXR	3.7881	0.3671	0.1498
OPEN	3.1143	0.0245	0.0049
Panel B			
RGDP	σ^2	Nugget (θ)	-2*LogLikelihood
10.2973	0.2837	0	-879.9303

In Table 8.1, we present maximum likelihood estimate results of the Gaussian process regression model for the identified economic growth drivers such as INDT, EXDT, RINR, REXR and OPEN under investigation in this study. The theta represents the estimated value for the magnitude of the parameters required for determining and predicting economic growth rate (RGDP). Thus, it is revealed from Table 8.1 that the magnitude of the INDT, EXDT, RINR, REXR and OPEN are 0.4969, 0.1140, 0.7424, 3.7881 and 3.1143 respectively. The sensitivity, the expected contribution of the aforementioned economic growth drivers, is also revealed to be 59.03%, 10.12%, 7.56%, 36.71% and 2.45%, respectively, in determining and predicting economic growth in Nigeria. However, the main effects or contributions of the economic growth drivers to the RGDP are 38.53%, 0.21%, 1.37% and 0.49% for INDT, EXDT, RINR, REXR and OPEN, respectively. Thus, there is a need for better improvement of the identified economic growth drivers to ensure and enhance efficient and optimal prediction of the fitted Gaussian process regression model. The nugget is 0, and it is the kernel parameter value on the likelihood of minimizing the model error to obtain the best fit Gaussian process regression model that predicts an RGDP value of 10.29% with an associated variance of 0.28%. Also, in Table 8.2, we present the Gaussian process regression results for variables interaction required to fit the predictive Gaussian process model for economic growth.

Table 8.2: ML Gaussian Process Regression Result for the Variables Interaction

Variables	INDT Interaction	EXDT Interaction	RINR Interaction	REXR Interaction	OPEN Interaction
INDT		0.0709	0.0074	0.1234	0.0032
EXDT			0.0005	0.0271	0.0007
RINR				0.0525	0.0015
REXR					0.0142
OPEN					

Table 8.2 shows the Gaussian process regression results for the interaction of identified economic growth drivers under consideration in this study. The results reveal that the contribution of the interaction of INDT and EXDT, INDT and RINR, INDT and REXR, INDT and OPEN, EXDT and RINR, EXDT and REXR, EXDT and OPEN, RINR and REXR, RINR and OPEN and REXR and OPEN to obtain the gaussian process regression model for the prediction of economic growth (RGDP) are 7.1%, 0.74%, 12.34%, 0.32%, 0.05%, 2.71%, 0.07%, 5.25%, 0.15%, 1.42% respectively. Based on the fitted Gaussian process regression model and the identified economic

growth drivers' interaction, the plot for the actual predicted values of economic growth (RGDP) is shown in Figure 8.1.

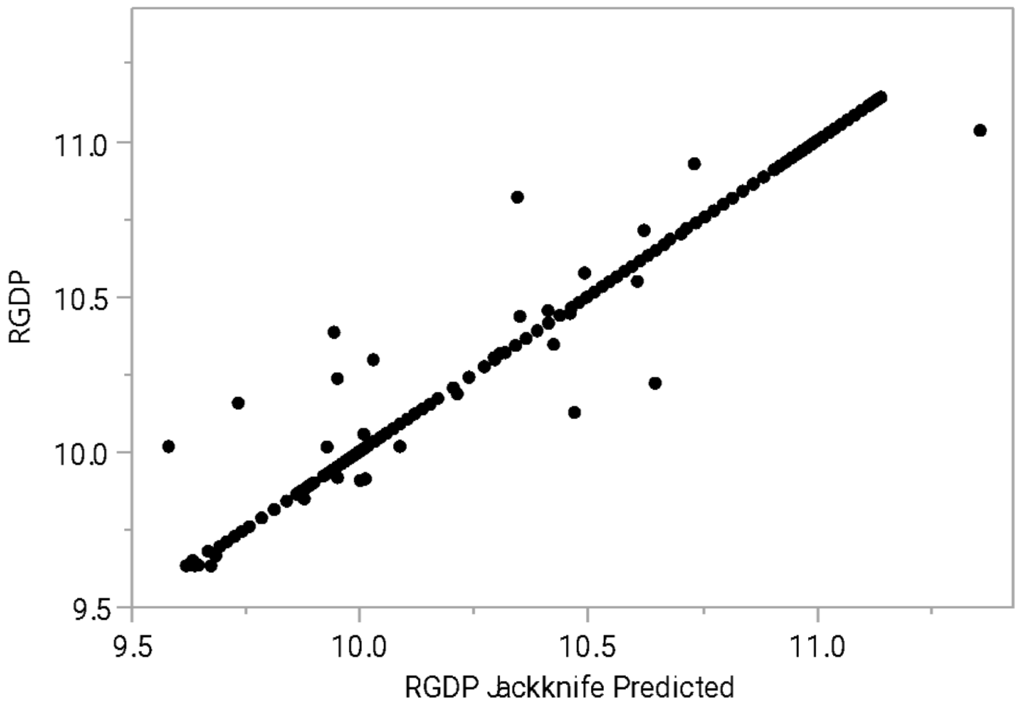
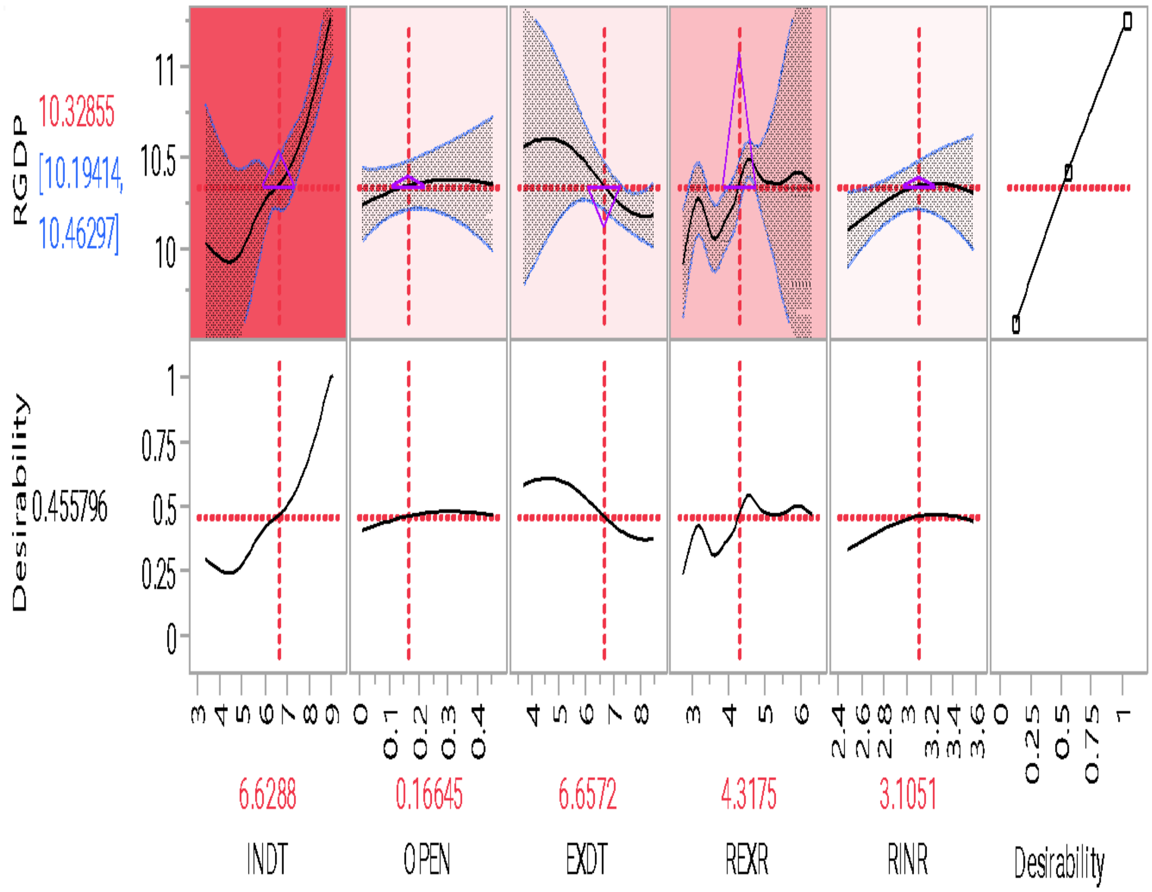


Figure 8.1: Gaussian Process Model's Plot for predicted values of RGDP

From Figure 8.1, having obtained the contributions of the identified economic growth drivers in this study that can be used to examine economic growth using the Gaussian process regression technique, it is also imperative to consider the validation and the efficiency of the Gaussian process regression method in predicting economic growth rate (RGDP) based on the available dataset. The dot shows individual predicted values for RGDP during the period under study using gaussian process model. Thus, in Figure 8.1, the Gaussian process regression model predicted plot for the RGDP, and it shows that during the period under investigation, there is a high proportional increase in the economic growth (RGDP) as most of the values concentrated on a straight line with small deviated values scatter below and above the main line. In addition, the Gaussian process quantile prediction plot showing the sensitivity of the economic growth drivers and desirability level for the RGDP is shown in Figure 8.2.



.

Figure 8.2: Quantile Prediction Plot for the Gaussian Process Regression Method showing Variables Sensitivity

In Figure 8.2, we present a Gaussian process regression model quantile plot that shows the sensitivity of the economic growth drivers in predicting economic growth (RGDP) and its desirability level. The vertical line in the above plot indicates the optimal value for lambda. The following two plots show the display of the coefficient 'path' and the sorted magnitude of the coefficients at the optimal lambda. Thus, in the plot, it is revealed that 6.63% of INDT, 0.17% of OPEN, 6.66% of EXDT, 4.3175 of REXR and 3.1051% of RINR predict a 10.33% economic growth rate (RGDP) at the desirability value of 0.46 using the gaussian process regression model. However, the Gaussian process interaction plot for the identified economic drivers in predicting economic growth (RGDP) is shown in Figure 8.3.

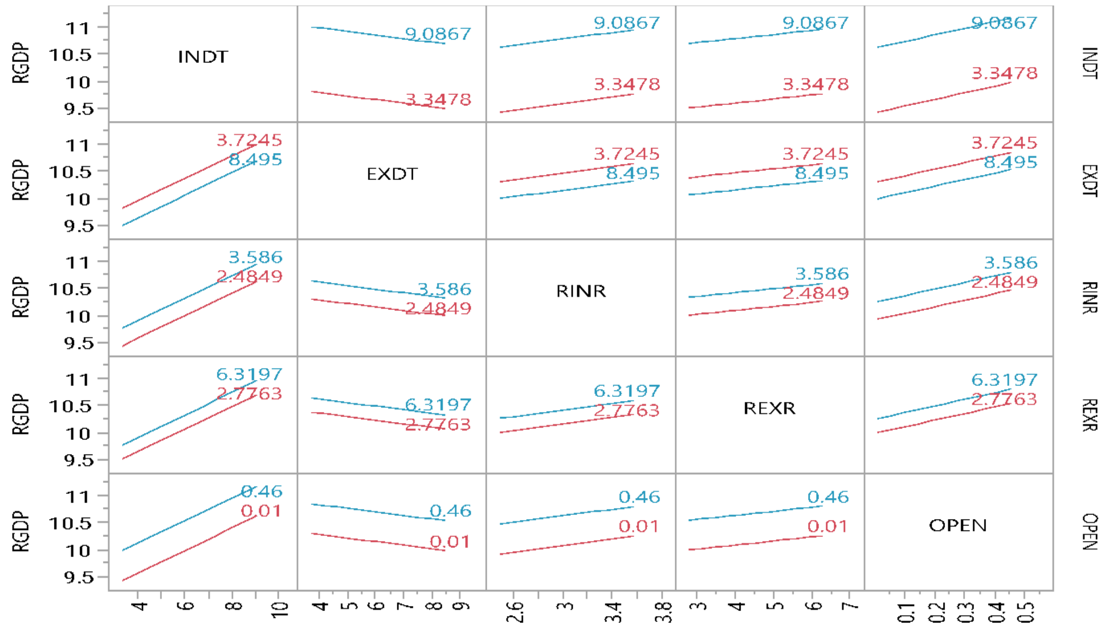


Figure 8.3: Variables Interaction Plot for the Gaussian Process Regression Model

Figure 8.3 presents the interaction plots for the identified economic growth drivers using a Gaussian process regression method. The blue and the red line show the relationship between two predictors such as INDT and EXDT, INDT and RINR, INDT and REXR and INDT and OPEN and so on in determining their joint effect on RGDP. In Figure 8.3, the interaction plots show evidence of parallel lines. This is observed between INDT and EXDT, INDT and RINR, INDT and REXR, and INDT and OPEN in the plots using the Gaussian process regression technique. Thus, it reveals that the identified economic growth drivers do not significantly influence each other in determining their various contributions to economic growth rate (RGDP) prediction, indicating the independence of the identified economic growth drivers in examining and predicting Nigeria's economic growth rate (RGDP). Following this, the variable (economic growth drivers) importance results for the fitted Gaussian process regression model are presented in Table 8.3.

Table 8.3: Variable Importance for Gaussian Process Regression Model Result

	Independent Uniform			**Independent Resampled**		
Variables	**Main Effect**	**Total Effect**		**Main Effect**	**Total Effect**	
INDT	0.3830	0.7440		0.5630	0.8520	

EXDT	0.1220	0.5110		0.0690	0.3240	
RINR	0.0110	0.1730		0.0310	0.1030	
REXR	0.0200	0.1340		0.0280	0.0920	
OPEN	0.0120	0.0520		0.0210	0.0500	

Table 8.3 presents the results for the variable importance of the identified economic growth drivers based on the fitted Gaussian process regression model. Table 8.3 reveals the main effect and total effect contributions of the identified economic growth drivers for independent uniform and independent resampled by the Gaussian process regression method. Thus, it is revealed that the main effect contributions of INDT, EXDT, RINR, REXR and OPEN for predicting economic growth rate (RGDP) based on independent uniform Gaussian process regression method are 38.30%, 12.20%, 1.10%, 2.00%, and 1.20% respectively. In the independent resampled Gaussian process regression method, the main effect contributions have improved and were found to be 56.30%, 6.90%, 3.10%, 2.80%, and 2.10%, respectively, for the INDT, EXDT, RINR, REXR and OPEN in predicting economic growth (RGDP) in this study. The marginal Gaussian process regression model plots for independent uniform and independent resampled for the identified economic growth are shown in Figure 8.4a and 8.4b, respectively.

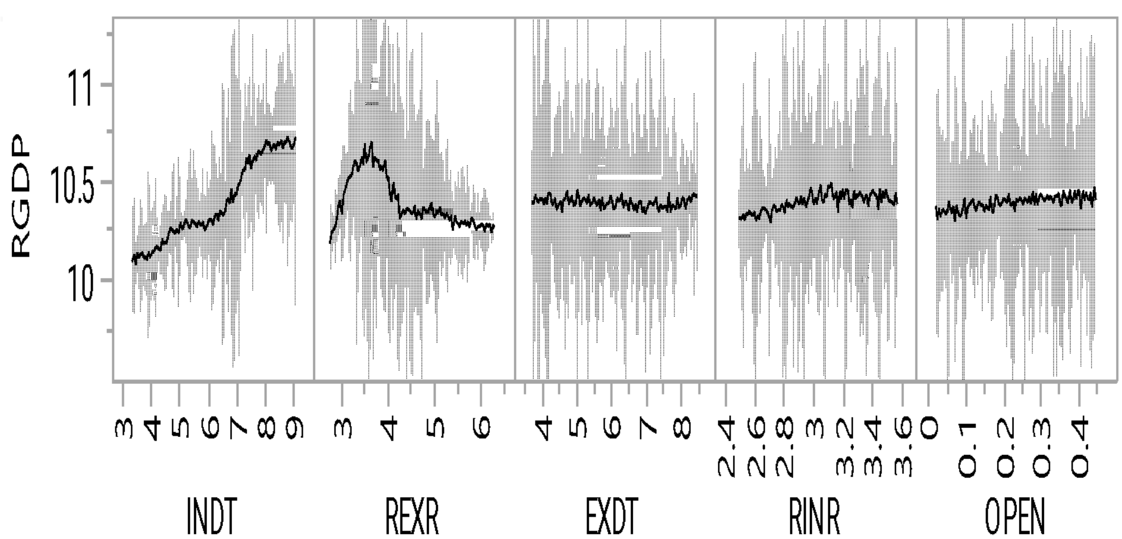


Figure 8.4a: Marginal Plot for Independent Uniform using Gaussian Process Regression Method

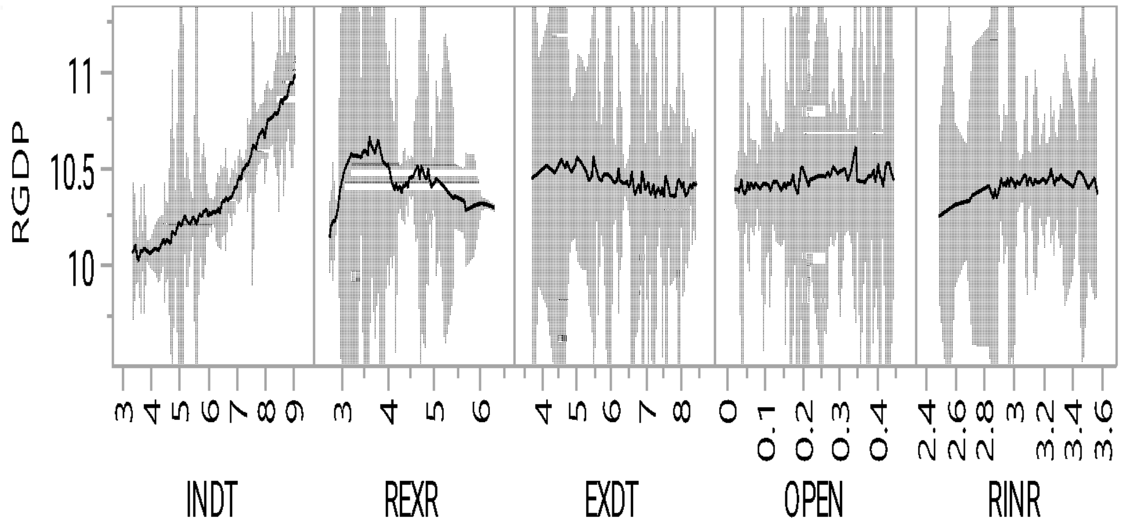


Figure 8.4b: Marginal Plot for Independent Resampled using Gaussian Process Regression Method

Figures 8.4a and 8.4b display marginal Gaussian process regression plots for independent uniform and independent resampled data, respectively. These plots illustrate the relationships between individual-identified economic growth drivers and RGDP. From these figures, it is evident that an increase in INDT leads to a significant proportionate increase in economic growth (RGDP), with the effect being smoother when using the independent resampling method. This significant increase is not observed for other identified economic growth drivers such as EXDT, RINR, REXR, and OPEN under both independent uniform and resampled approaches. The plots highlight the impact of fluctuations in REXR on economic growth. Initially, REXR shows a positive contribution for a certain period, but it then sharply declines and continues to negatively influence economic growth (RGDP). Other economic growth drivers such as EXDT, RINR, and OPEN exhibit both positive and negative relationships with RGDP; however, their impacts are not as strong as those of INDT and REXR. This analysis underscores the importance of addressing exchange rate fluctuations to enhance economic growth. The Gaussian process regression model indicates that the acquisition of foreign or external debt does not contribute positively to the economy's growth.

Table 8.4: 2-Dimension Estimate of Identified Economic Growth Drivers and RGDP

Economic Growth Drivers	Estimates	2-Dimensional (2D)	Predicted RGDP
INDT	6.2172	-	-
EXDT	6.1097	INDT and EXDT	10.5373
RINR	3.0354	INDT and RINR	10.5179
REXR	4.5480	INDT and REXR	10.5958
OPEN	0.2350	INDT and OPEN	10.4674

Table 8.4 shows the results of the estimated value for the identified economic growth drivers such as INDT, EXDT, RINR, REXR and OPEN under consideration with their respective estimated values of 6.2172%, 6.1097%, 3.0354%, 4.5480% and 0.2350% for predicting for economic growth RGDP using a Gaussian process regression method. The 2-dimensional (2D) for the identified economic growth drivers are considered, and the predicted economic growth (RGDP) values are 10.5373% for INDT and EXDT, 10.5179% for INDT and RINR, 10.5958% for INDT and REXR 10.4674% for INDT and OPEN respectively. Thus, 2D results for the identified economic growth drivers reveal that INDT and REXR have 10.5958%, the highest predictive value for economic growth (RGDP) during the period under investigation. This result affirmed the significant contribution of INDT and REXR to economic growth, as shown in Figures 8.4a and 8.4b. This result is represented in graphical form and presented in Figure 8.5.

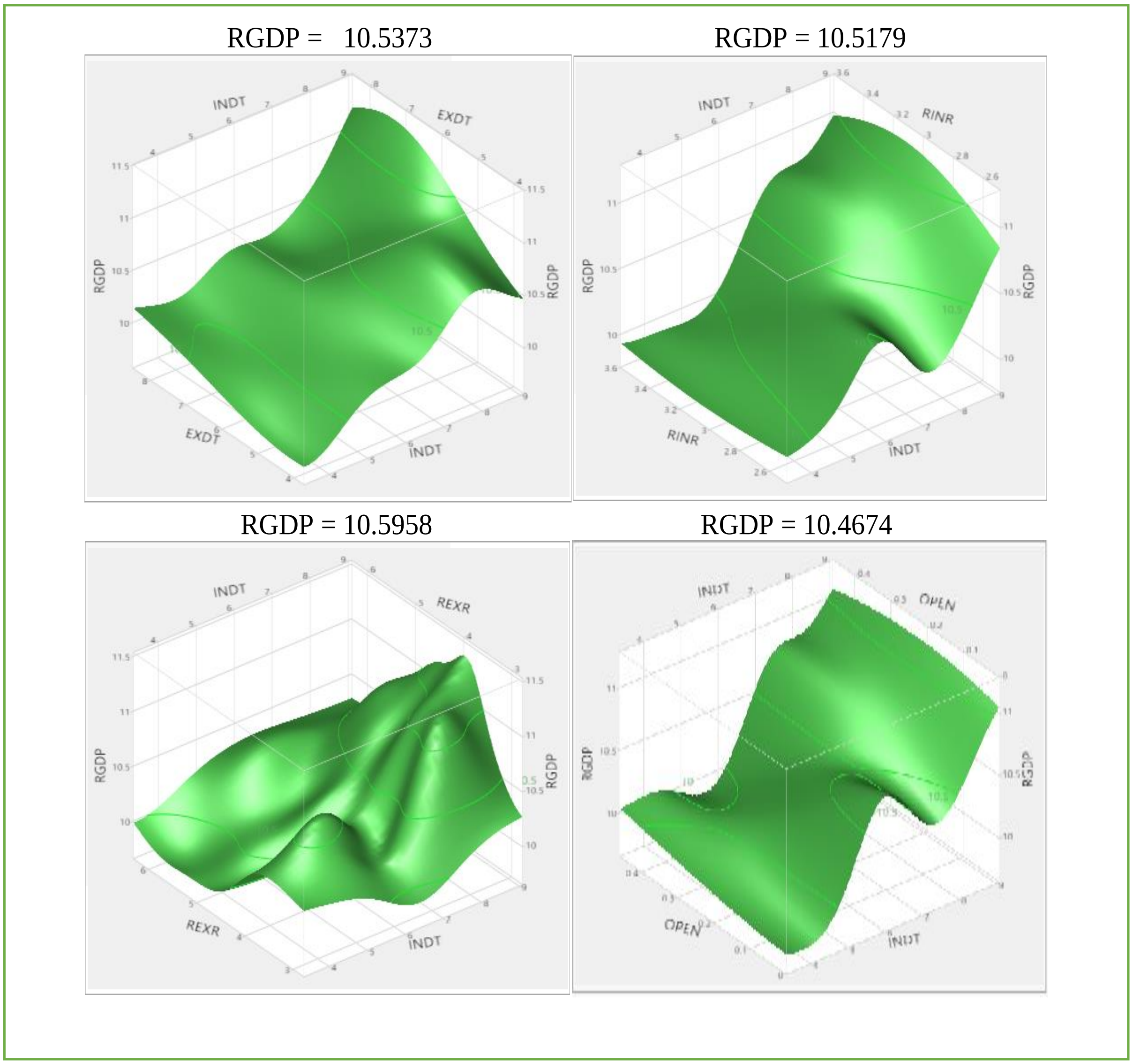


Figure 8.5: Predicted RGDP using Gaussian Process Regression Model 2D Plot

Table 8.5 presents the estimated values for the identified economic growth drivers that optimally predict economic growth using the Gaussian process regression method. The green colours indicate clusters of the actual values of two predictors such as INDT and EXDT, INDT and RINR, INDT and REXR and INDT and OPEN that majorly predict RGDP during the period under investigation.

Table 8.5: Identified Economic Growth Drivers and Optimal Predicted RGDP

Economic Growth Drivers	Estimates
INDT	6.6288
EXDT	6.6572
RINR	3.1051
REXR	4.3175
OPEN	0.1664
Response	
RGDP	10.3174

Table 8.4 and Figure 8.5 show the estimated value for the identified economic growth drivers that optimally predict stable and reliable economic growth (RGDP). Thus, in Table 8.4, the contributions of INDT, EXDT, RINR, REXR, and OPEN in predicting economic growth (RGDP) are 6.6288%, 6.6572%, 3.1051%, 4.3175%, and 0.1664%, respectively. Thus, the combined estimated values for the identified economic growth drivers optimally predict 10.3174% for economic growth (RGDP). This result is also shown in the plot in Figure 8.5, where the contour lines and corresponding optimal surface for the predicted value economic growth (RGDP) are shown in cube form.

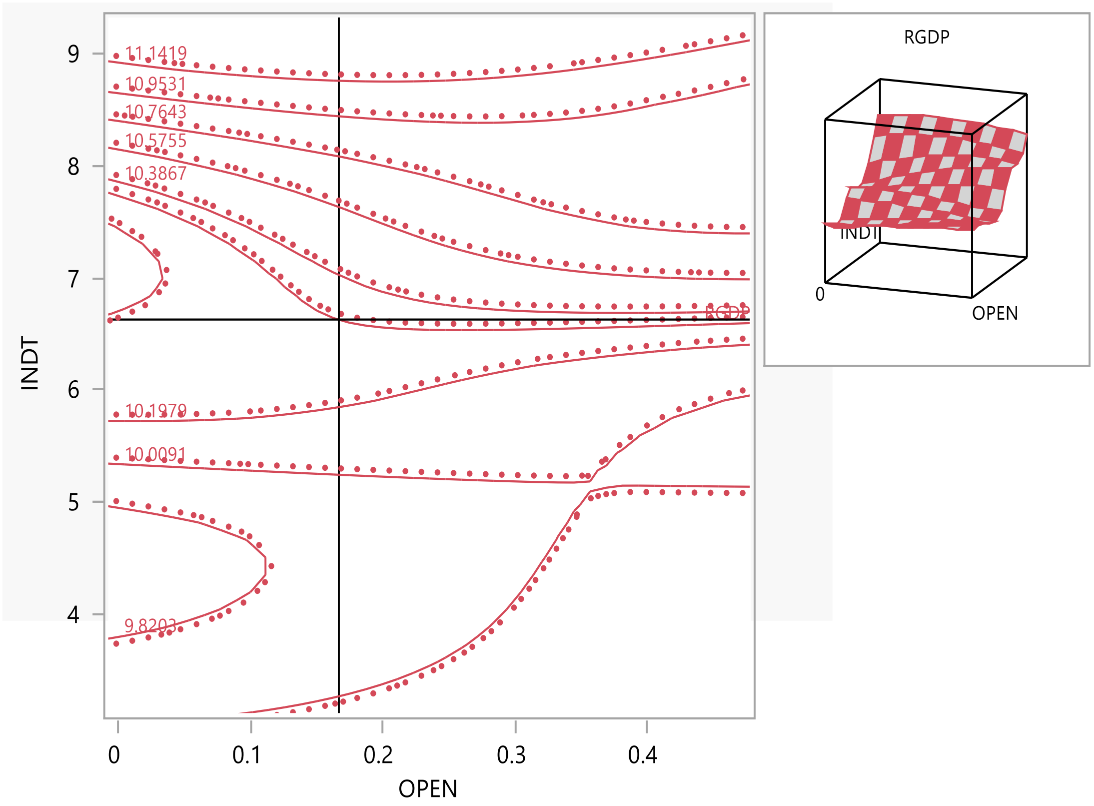


Figure 8.6: Optimal Predicted RGDP Plot Showing Contour and corresponding Surface

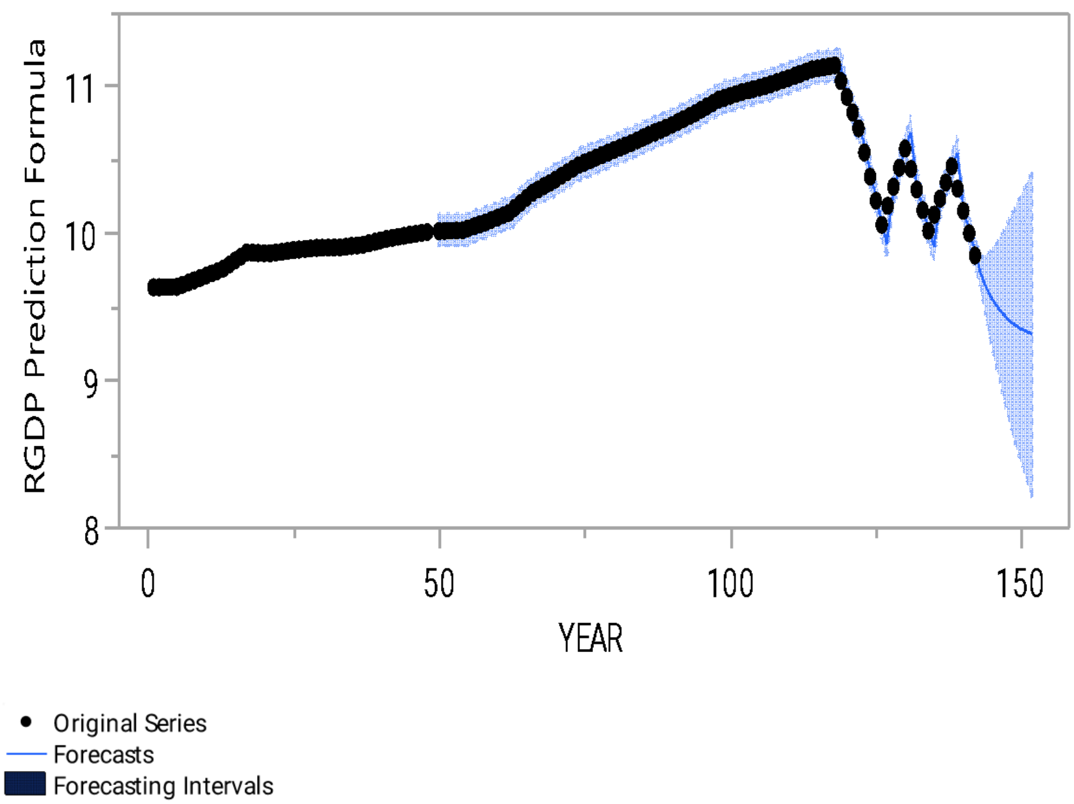
Table 8.6: Performance Metrics for Gaussian Process Regression Model

Measure of Efficiency	
AIC	-109.554
BIC	-94.3587
-2LogLikelihood	-121.554
RMSE	0.1310
MAE	0.0942
MAPE	0.9164

In Table 8.6, we present the performance metric to evaluate the Gaussian process regression method and its predictive power to select optimal and efficient predict stable and reliable values for the economic growth RGDP using a Gaussian process regression method. This result is determined using root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). Thus, from Table 8.5, the RMSE for the fitted Gaussian process regression technique is 0.1310. The MAE for the fitted Gaussian process regression technique is 0.0942, and the MAPE is 0.9184.

8.4 Forecast for RGDP using Gaussian Process Regression Method

In Figure 8.6, we present a plot that shows the forecasts or predicted values of the economic growth rate (RGDP) for the next ten (10) quarters based on the efficiency of the Gaussian process regression method in generating stable and reliable predictive values of economic growth rate using the data set under consideration.



Source: Researcher's Computation, 2024

Figure 8.7: Forecast Plot for RGDP using Gaussian Process Regression Method

8.5 Concluding Remarks

In this chapter, we thoroughly investigate the estimation of economic growth parameters in Nigeria concerning the identified drivers of economic growth (INDT, EXDT, RINR, REXR, and OPEN) using the Gaussian process regression technique. This machine learning-based approach is proposed to explore and predict the economic growth rate (RGDP) while addressing issues of multicollinearity and outliers, which violate the assumptions of the linear model. An exploratory and diagnostic analysis of the collected data reveals the relationships between the identified economic drivers and RGDP. Consequently, we propose the Gaussian process regression technique to obtain optimal, efficient, and stable parameters that can effectively predict economic growth (RGDP) in Nigeria. The results indicate that the Gaussian process estimated parameters for INDT, EXDT, RINR, REXR, and OPEN are 0.4969, 0.1140, 0.7424, 3.7881, and 3.1143, respectively.

The sensitivity, which serves as the expected contributions of the aforementioned economic growth drivers, was revealed to be 59.03%, 10.12%, 7.56%, 36.71% and 2.45%, respectively, in determining and predicting economic growth rate. However, from the Gaussian process model, the main effect or contributions of the economic growth drivers to the RGDP are 38.53%, 0.21%, 1.37% and 0.49% for INDT, EXDT, RINR, REXR and OPEN, respectively. Also, in this study, the results indicate that 6.63% of INDT, 0.17% of OPEN, 6.66% of EXDT, 4.3175 of REXR and 3.1051% of RINR reliably predicted 10.33% as an average economic growth in RGDP at the desirability value of 0.46. Also, the interaction plots show parallel lines indicating that independent variables (economic growth drivers) under consideration do not influence each other in determining their contributions to RGDP. The marginal output of the fitted Gaussian process regression model shows the main effect contributions of INDT, EXDT, RINR, REXR and OPEN in predicting RGDP based on independent uniform and improved contributions using independent resampled Gaussian process regression method for predicting RGDP.

The RMSE, MAE and MAPE with the associated values 0.1310, 0.0942 and 0.9184, respectively, establish the efficiency of the Gaussian process regression technique in predicting economic growth. Therefore, based on the findings, in enhancing economic growth's stability and efficient predictive values in the presence of multicollinearity and outliers, the proposed Gaussian process regression method remains versatile in modelling and predicting economic growth. Also, governments and policymakers must properly harness the benefit of trade openness to grow the economy. The development of infrastructure and the economy's growth through internal or external borrowing is not sustainable enough due to unstable exchange rates and high interest rates and, as such, the need for policy direction to checkmate the negative impact of the identified monetary policy instrument toward economic growth enhancement. However, to further explore, model and predict economic growth in Nigeria, another robust statistical method, a coupler FMKL-GLD quantile regression method, is proposed, and this is discussed in detail in Chapter 9 of this research work to predict RGDP in Nigeria.

CHAPTER 9

COUPLER FMKL-GLD AND QUANTILE METHOD FOR ESTIMATION OF ECONOMIC GROWTH

9.1 Introduction

In this chapter, we propose the Freimer–Mudholkar–Kollia–Lin generalized lambda distribution (FMKL-GLD) quantile regression method to explore, estimate, and predict economic growth. This method accounts for the distribution properties—location, scale, and shape parameters of the variables, which can be influenced by outliers in the data. According to Ramberg and Schmeiser (1974) and Freimer *et al*. (1988), generalized lambda distributions (GLDs) are appropriate for such situations due to their quantile function nature. Additionally, Su (2007b, 2010b) emphasized that advancements in GLDs and computer technology have made fitting these distributions to data feasible. Koenker (2005) and Koenker and Bassett (1978) noted that the rich shapes generated by GLDs provide an attractive means of formulating accurate and flexible quantile regression models, offering smoother regression coefficients compared to conventional quantile regression. Ramberg and Schmeiser (1974) and King and MacGillivray (1999), as cited in Darare *et al*. (2021), showed that the six regions in which the shape parameters can lie correspond to the shapes of the GLD. Ramberg et al. (1979) and Chalabi *et al*. (2010, 2012) introduced the notation for the generalized lambda distribution. The modern FMKL GLD, proposed by Freimer *et al*. (1988), places only one restriction: $\lambda_4 > 0$. Su (2007b) emphasized that the primary motivation for developing the FMKL GLD was to ensure the distribution is defined over all λ_3 and λ_4. values.

Su (2005) used GLD quantile regression to investigate three datasets: the Belgian Engel dataset, pipeline repair cost data, and the simulated motorcycle acceleration data. The FMKL-GLD quantile regression used for the analysis revealed the simplicity of GLD quantile regression compared to standard quantile regression. The results also showed that FMKL-GLD quantile regression provided a robust reference line to outliers and produced zero mean residuals. Besides, it provided a reference line with smooth regression coefficients across different quantiles. Chalabi *et al*. (2012) introduced new parameterizations of the GLD where the median and interquartile range were found to be interchangeable with the location and shape parameters of the distribution. Thus, it showed the advantage of GLD in modelling financial returns, as the GLD family can accommodate a great range of distribution shapes. The robustness of the GLD thus allowed for a single distribution to be used to model the data from various asset classes.

Canan et *al*. (2016) investigated the ability of five alternative distributions to the normal distribution that represented the behaviour of daily equity index returns between the period (1979–2014). The study considered the skewed Student-t distribution, the generalized lambda distribution, the Johnson system of distributions, the normal inverse Gaussian distribution, and the g-and-h distribution. Thus, it was revealed that the generalized lambda distribution was a proper and good alternative for modelling the behaviour of daily equity index returns. However, considering the presence of multicollinearity and outliers in the datasets used for this study can influence the distribution properties such as location, scale and shape parameters, which is lacking in various work and studies previously carried out in relation to the economic growth drivers under investigation. Therefore, it is imperative to consider a predictive model that can properly address this situation, and as such, we propose a coupler FMKL-GLD quantile regression method is employed as a robust method for efficient, stable and reliable prediction of economic growth based on its identified drivers in this study.

9.2 Research Methodology

This section outlines the methodology of the FMKL-GLD quantile regression method.

9.2.1 The Coupler FMKL-GLD Quantile Regression Model

Consider a quantile probability density function of the GLD, which is known as the Ramberg-Schmeiser Generalized Lambda Distribution (RS GLD), and its inverse distribution function, known as Tukey's lambda distribution (TLD) (Darare *et al*., 2021). This can be expressed as:

$$F^{-1}(P|\lambda) = F^{-1}(P|\lambda_i) \; \forall \; i = 1, 2, ..4$$

$$F^{-1}(P|\lambda_i) = \lambda_1 + \frac{P^{\lambda_3} - (1-P)^{\lambda_4}}{\lambda_2} \tag{9.1}$$

where P are the probabilities, $P \in [0,1]$, λ_1 and λ_2 are the location and scale parameters, while λ_3 and λ_4 are the shape parameters that define the strengths of the lower and upper tails, respectively.

According to Chalabi *et al*. (2010), it can be stated that the original one-parameter Tukey's Lambda Distribution (TLD) results is the limiting case $\lambda_1 = 0$ and $\lambda_2 = \lambda_3 = \lambda_4 = \lambda$.. This model is divided into two parts and is adopted to explore and predict the economic growth (RGDP) in Nigeria in relation to the macroeconomic variables such as internal debt (INDT), external debt (EXDT), interest rate (RINR), exchange rate (REXR), and trade openness (OPEN). The first part deals with the FMKL-GLD regression reference line generation and the second part concentrates on finding the quantile regression coefficients based on the reference line that can be obtained from the FMKL GLD as discussed below.

Freimer *et al*. (1988) introduced this parameterization to improve the RS (Ramberg and Schmeiser 1974) parameterization given in (9.1). Hence, the quantile function for the distribution of FMKL-GLD is defined as stated:

$$Q(P) = \lambda_1 + \frac{\frac{P^{\lambda_3 - 1}}{\lambda_3} - \frac{(1-P)^{\lambda_4 - 1}}{\lambda_4}}{\lambda_2} , \qquad (9.2)$$

where λ_1 and λ_2 are the location and scale parameters whereas λ_3 and λ_4 are the shape parameters. This distribution is most favorable because it is valid for all values of λ_3 and λ_4 where $\lambda_2 > 2$. If $\lambda_3 = \lambda_4 = 0$ the FMKL parameterization has the following quantile function given by:

$$F^{-1}(P) = \lambda_1 + \frac{\ln(p) - \ln(1-p)}{\lambda_2} . \qquad (9.3)$$

The FMKL-GLD takes different quantile forms if either λ_3 and λ_4 or both and are equal to zero. These forms can be represented as follows:

if $\lambda_3 = 0, \ \lambda_4 \neq 0,$

$$Q(P) = \lambda_1 + \frac{1}{\lambda_2}\left(\ln(p) - \frac{(1-P)^{\lambda_4 - 1}}{\lambda_4}\right), \quad 0 \leq p \leq 1 \qquad (9.4)$$

if $\lambda_3 \neq 0, \lambda_4 = 0$

$$Q(P) = \lambda_1 + \frac{1}{\lambda_2}\left(\frac{(P)^{\lambda_3 - 1}}{\lambda_3} - \ln(1-p)\right), \quad 0 \leq p \leq 1 \qquad (9.5)$$

if $\lambda_3 = 0, \lambda_4 = 0$

$$Q(P) = \lambda_1 + \frac{1}{\lambda_2}\left(\ln(p) - \ln(1-p)\right), \quad 0 \leq p \leq 1 \ , \qquad (9.6)$$

if $\lambda_3 \neq 0, \lambda_4 \neq 0$

$$Q(P) = \frac{\lambda_2}{P^{\lambda_3 - 1} + (1-P)^{\lambda_4 - 1}} \qquad 0 \le p \le 1 \;, \tag{9.7}$$

The probability density function of the FMKL-GLD can be obtained by using the relationship given by:

$$y = F^{-1}(y) = Q(P) \quad , \tag{9.8}$$

where $F(y) = p$ and $y = Q(P)$.

By differentiating $Q(P)$ with respect to y, the density function of FMKL-GLD can be obtained as expressed by:

$$\frac{dp}{dy} = f(y) \; and \; dy = d\big(Q(p)\big) \,, \tag{9.9}$$

These two relationships give $f(y) = \frac{dp}{d(Q(p))}$. From (9.2), the expression given as:

$$\frac{d(Q(p))}{dp} = \frac{\lambda_3 P^{\lambda_3 - 1} - \lambda_4 (1-P)^{\lambda_4 - 1}}{\lambda_2} \quad , \tag{9.10}$$

can be obtained. Substituting the result in (9.10) into $f(y) = \frac{dp}{d(Q(p))}$ the density function obtained can be given as:

$$f(y) = \frac{\lambda_2}{\lambda_3 P^{\lambda_3 - 1} - \lambda_4 (1-P)^{\lambda_4 - 1}} \quad , \tag{9.11}$$

where $0 \le p \le 1$. Thus, the probability density function of the FMKL-GLD when λ_3, λ_4 or both are equal to zero. The probability functions of the other forms of the FMKL-GLD can be in a similar fashion as:

$$f(y) = \frac{\lambda_2}{\lambda_3 P^{\lambda_3 - 1} - (1-P)^{\lambda_4 - 1}} \qquad 0 \le p \le 1. \tag{9.12}$$

The probability functions of the FMKL-GLD if $\lambda_3 = 0, \; \lambda_4 \neq 0$ can be obtained as:

$$f(y) = \frac{\lambda_2}{P^{\lambda_3 - 1} - \lambda_4 (1-P)^{\lambda_4 - 1}} \qquad 0 \le p \le 1 \;, \tag{9.13}$$

If $\lambda_3 = 0$ and $\lambda_4 = 0$, probability functions of the FMKL-GLD can be given as

$$f(y) = \frac{\lambda_2}{P^{\lambda_3 - 1} - (1-P)^{\lambda_4 - 1}} \qquad 0 \le p \le 1 \;, \tag{9.14}$$

If $\lambda_3 \neq 0, \; \lambda_4 \neq 0,$ probability functions of the FMKL-GLD can be expressed in the form:

$$f(y) = \frac{\lambda_2}{\lambda_3 P^{\lambda_3 - 1} - \lambda_4 (1-P)^{\lambda_4 - 1}} \qquad 0 \le p \le 1 \,, \tag{9.15}$$

The Maximum Likelihood Estimation

To use maximum likelihood estimation method, the quantiles for every observation pi for i = 1, 2, 3, ... n observations, need to be calculated under a set of initial values. This involves solving the expression in (9.9) numerically. This can be done by using the Newton-Raphson method. Having obtained the $Q(p_i)$, by substituting in the log-likelihood, the expression can be written as:

$$ML = \prod_{i=1}^{n} \ln\left(\frac{\lambda_2}{\lambda_3 {p_i}^{\lambda_3 - 1} - \lambda_4(1-p_i)^{\lambda_4 - 1}}\right) , \qquad (9.16)$$

The transformation of the expression in (9.16) can be stated and given as:

$$ML = n\ln(\lambda_2) - \prod_{i=1}^{n} \ln\left(\lambda_3 {p_i}^{\lambda_3 - 1} - \lambda_4(1-p_i)^{\lambda_4 - 1}\right) , \qquad (9.17)$$

This idea is valid and appropriated to maximize the likelihood in (9.16) and (9.17) using the Nelder-Mead also known as simplex search algorithm.

The Empirical Likelihood Goodness of Fit Test

The hypothesis to test the goodness of fit for the GLD distribution is given as:

$$H_0: f = f_0 \sim GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4) , \qquad (9.18)$$

$$H_1: f = f_1 \sim GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4) , \qquad (9.19)$$

The definition of the likelihood ratio test statistic for this hypothesis is given as:

$$LR = \frac{\prod_{i=1}^{n} fH_1(y_i)}{\prod_{i=1}^{n} fH_0(y_i)} , \qquad (9.20)$$

$$LR = \frac{\prod_{i=1}^{n} fH_1(y_i)}{\prod_{i=1}^{n} f(y_i|\lambda)} , \qquad (9.21)$$

where $y_1, y_2, y_3, \ldots, y_n$ follows a GLD distribution with the parameter $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ under the null hypothesis. Since f_0 and f_1 are unknown, the maximum likelihood method estimates $\boldsymbol{\lambda}$ of a GLD under the null hypothesis. Ning (2014) used the maximum empirical likelihood method to estimate the numerator can be written as:

$$L_f = \prod_{i=1}^{n} fH_1(y_i) = \prod_{i=1}^{n} fH_1(y_{(i)}) = \prod_{i=1}^{n} f_i , \qquad (9.22)$$

where $y_{(1)} \leq y_{(2)} \leq y_{(3)} \leq \dots \leq y_{(n)}$ are the order statistics of the observations $y_1, y_2, y_3, \dots, y_n$. The values of f_i to maximize L_f were obtained by using the constraint $f(s)ds = 1$ corresponding to the alternative hypothesis. The value can be obtained using:

$$f_j = \frac{2m}{n(Y_{j+m} - Y_{j-m})} \quad , \qquad (9.23)$$

where $\mathrm{Y_j = Y_1, if\ j \leq 1\ and\ Y_j = Y_n\ if\ j \geq n}$,

The likelihood ratio statistic based on the maximum likelihood empirical method is given as:

$$GLD_{mn} = \frac{\prod_{j=1}^{n} \frac{2m}{n(Y_{j+m} - Y_{j-m})}}{max \prod_{j=1}^{n} f(y_j|\lambda)} \quad , \qquad (9.24)$$

where $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)\ and\ 0 < \lambda < 1$,

Thus, in the next session, data exploration, descriptive analysis, test for outliers, multicollinearity, and granger causality as well as fitting FMKL-GLD quantile regression models and its associated diagnostics to determine the most efficient FMKL-GLD quantile regression models for estimating parameters economic growth in Nigeria based on the identified macroeconomic variables in this study.

9.3 Empirical Results

This section reports the empirical results of the fitted FMKL-GLD quantile regression models discussed in Section 9.2 to predict RGDP in the presence of multicollinearity and outliers in this study. Table 9.1 presents the maximum likelihood results for the estimated GLD parameters for the variables to fit the FKML-GLD model.

Table 9.1: ML Estimate of GLD parameters to Fit FKML-GLD Model

Returns	λ_1	λ_2	λ_3	λ_4	the AD statistic
RGDP	10.2499	1.2035	1.3437	0.9304	0.9999
INDT	6.8491	0.4215	0.6635	1.0603	0.9995
EXDT	6.9516	0.8196	0.3336	0.7856	0.9997
RINR	3.1222	9.7941	-0.0920	0.0713	0.9954
REXR	4.3947	4.4019	-0.4068	-0.2142	0.9999

OPEN	5.0623	5.0622	2.0101	0.5312	0.9999

Table 9.1 shows maximum likelihood estimates for the generalized lambda distribution (GLD) to fit the FMKL-GLD quantile model for investigating RGDP in relation to INDT, EXDT, RINR, REXR and OPEN as economic growth drivers in Nigeria. λ_1 and λ_2 represents the location and scale parameters while λ_3 and λ_4 representing the space parameters generated from the GLD fit of the identified economic growth drivers under investigation. The AD statistic > 0.05 shows the normality of the identified economic growth drivers under consideration or the identified economic growth drivers under consideration are from the normally distributed population. Thus, in Table 9.2 and Figure 9.1, we present FMKL-GLD quantile regression results and FMKL-GLD Q-Q plots for the FMKL-GLD quantile regression model.

Table 9.2: ML parameter estimates for FMKL-GLD Quantile Regression Model

Model	Estimate	*p*-value
Constant	8.1492	0.0000
INDT	0.1514	0.0000
EXDT	-0.0247	0.0000
RINR	0.2449	0.0000
REXR	0.0608	0.0000
OPEN	1.7597	0.0192
$\hat{\lambda}_1$	-0.0253	
$\hat{\lambda}_2$	34.1488	
$\hat{\lambda}_3$	-0.3303	
$\hat{\lambda}_4$	-0.5758	
KS test *p*-value	0.7949	
data driven smooth test *p*-value	0.6850	
Resample KS test> 0.05	91.4000	

Table 9.2 shows the estimated intercept and coefficients of the identified economic growth drivers to fit the FMKL-GLD model and the estimated lambda parameters $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$ and $\hat{\lambda}_4$. Specifically,

in Table 9.2, it is revealed that INDT, RINR, REXR and OPEN positively contribute 0.15%, 0.24%, 0.06% and 1.76% respectively to the economic growth. Also, it is found that EXDT contributes negatively to the economic growth in Nigeria thus, reduces economic growth by 0.02% and, as such, shows the tendency and potential of the identified economic growth drivers under study as veritable determinants for economic growth in Nigeria.. The location, scale and space parameters $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\hat{\lambda}_3$, and $\hat{\lambda}_4$ for the fitted FMKL-GLD model are -0.0253, 34.1488 -0.3303, and -0.5758, respectively.

Also, in Table 9.2, we present the results of Kolmogorov-Smirnov (KS) goodness of fit tests for the FMKL-GLD quantile regression model. The results show the *p*-values of the KS test, data-driven smooth test and Resample KS test. All the *p*-values are above the 5% significance level, which indicates that the FMKL-GLD fits the data moderately well. The GLD Q-Q plot in Figure 9.1 reveals this since most data points align with the data distribution line. Therefore, it can be emphasized that the FMKL-GLD regression model performs well in estimating Nigeria's economic growth parameters in the presence of multicollinearity and outliers considered in this study. Thus, based on the scale parameter, it can be emphasized that the FMKL-GLD quantile model is an efficient model that better describes the economic situation in Nigeria. However, Table 9.3 uses the maximum likelihood estimation method to obtain the FKML-GLD parameter values for the 25%, 50% and 75% quantile models. Thus, the GLD plot for FMKL-GLD is shown in Figure 9.1.

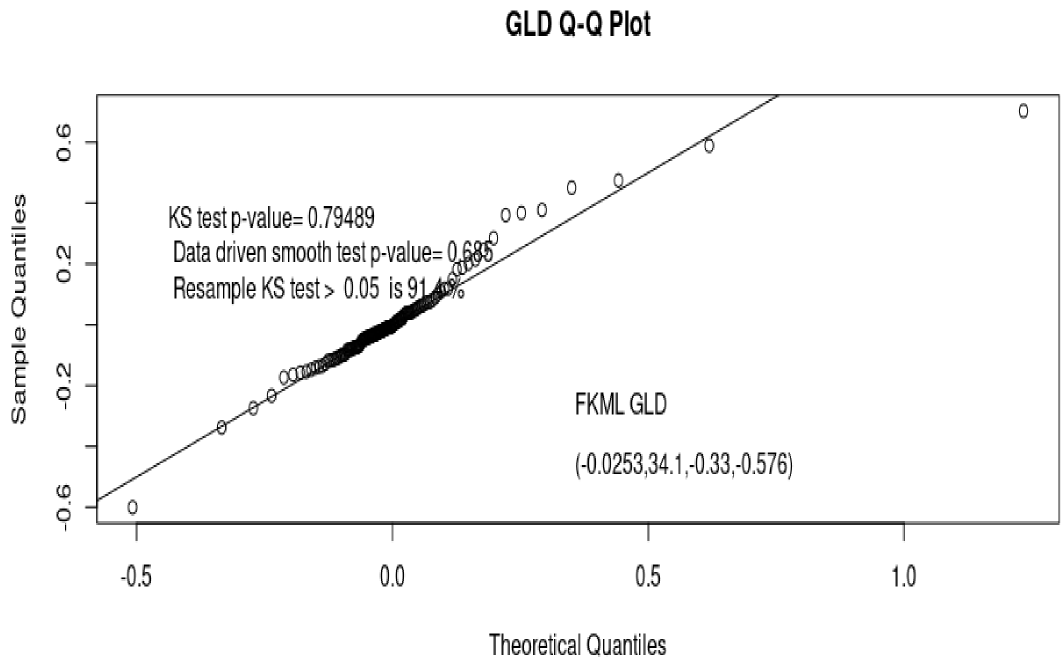


Figure 9.1: GLD Q-Q Plot for FMKL-FGLD Model

Table 9.3: ML parameter estimates of FKML-GLD Regression Model at different Quantile

Quantile	Parameters	Estimate	*p*-value	Lower 95% Interval	Upper 95% Interval
	C	8.1465	0.0000	7.7753	8.5177
	INDT	0.2035	0.0005	0.0912	0.3158
25th Percent	EXDT	-0.0679	0.0433	-0.1337	-0.0021
	RINR	0.2294	0.0005	0.1016	0.3573
	REXR	0.0599	0.0043	0.0191	0.1006
	OPEN	1.2848	0.0082	0.3377	2.2318
	C	7.8604	0.0000	7.2521	8.4686
	INDT	0.1794	0.0000	0.1017	0.2571
50th Percent	EXDT	-0.0392	0.1380	-0.0911	0.0127
	RINR	0.2941	0.0014	0.1161	0.4722
	REXR	0.0799	0.0003	0.0369	0.1230
	OPEN	1.4510	0.0007	0.6201	2.2818
	C	6.7134	0.0000	5.7058	7.7210

	INDT	0.3180	0.0000	0.2379	0.3981
75th Percent	EXDT	-0.0922	0.0007	-0.1446	-0.0399
	RINR	0.4901	0.0002	0.2337	0.7465
	REXR	0.1628	0.0000	0.0961	0.2296
	OPEN	-0.2183	0.6320	-1.1176	0.6810

Table 9.3 shows the results of FMKL-GLD quantile estimated parameters for 25%, 50% and 75% of the models fitted. Thus, from the results in Table 9.3, the coefficients of the 25% FMKL-GLD (FMKL-GLD 25Q) quantile regression model reveal that INDT, RINR, REXR and OPEN are positively related to RGDP. Specifically, the INDT, RINR, REXR and OPEN contributions to RGDP are 20.35%, 22.94%, 5.99% and 128.48%, respectively. It is also revealed that EXDT is negatively related to the RGDP and, as such, causes a decline of 6.79% in RGDP during the period under investigation. The *p*-value < 0.05 shows that the estimated parameters at a 5% level influence the economic growth rate. In Table 9.3, the coefficients of the 50% FMKL-GLD (FMKL-GLD 50Q) quantile regression model are revealed, and it shows that INDT, RINR, REXR and OPEN have a positive relationship with RGDP. Specifically, the INDT, RINR, REXR and OPEN contributions to RGDP are 17.94%, 29.41%, 7.99% and 145.10% respectively. The results also reveal that EXDT has a negative relationship with the RGDP, leading to a decline of 3.92% in RGDP during the period under investigation. The *p*-value < 0.05 shows that the estimated parameters at a 5% level affect the economic growth rate. Also, the results reveal the 75% FMKL-GLD (FMKL-GLD 75Q) quantile regression model coefficients, showing that INDT, RINR, and REXR are positively related to RGDP. Specifically, the INDT, RINR and REXR influence RGDP to the turn of 31.80%, 49.01%, and 16.28% respectively. At the same time, it is also revealed that EXDT and OPEN are negatively related to the RGDP and, as such, reduce RGDP by 9.22% and 21.83% during the period under investigation. The *p*-value < 0.05 shows that the estimated parameters at a 5% level influence the economic growth rate. The GLD Q-Q plots for 25%, FMKL-GLD quantile regression is shown in Figure 9.2, and it reveals the distribution of the data points.

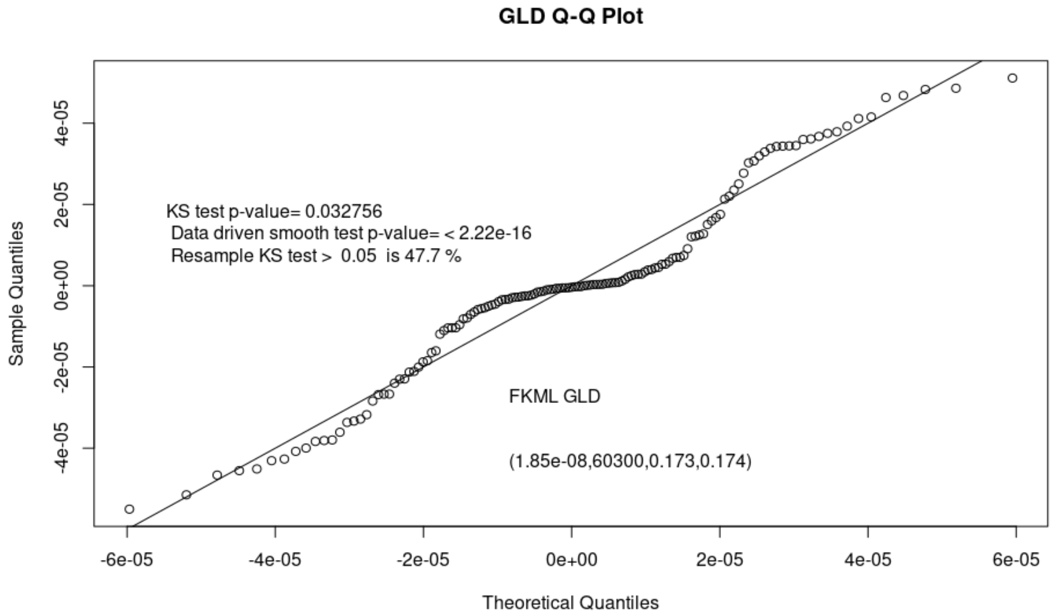


Figure 9.2: GLD Q-Q Plot for FMKL-GLD 25Q Quantile Model

The GLD Q-Q plots for 50% FMKL-GLD quantile regression is shown in Figure 9.3 and it reveals the distribution of the data points.

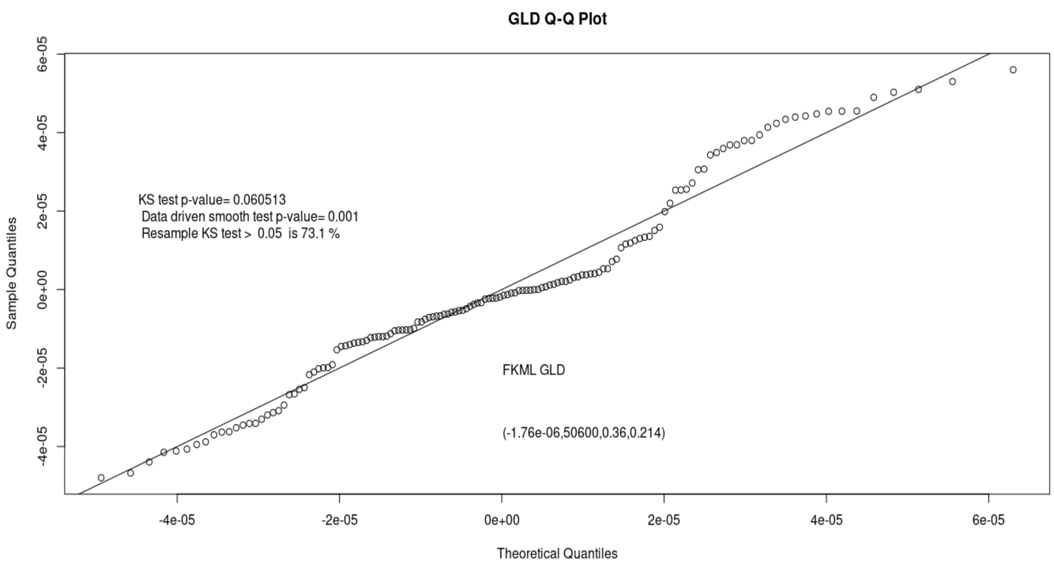


Figure 9.3: GLD Q-Q Plot for FMKL-GLD 50Q Quantile Model

The GLD Q-Q plot for 75% FMKL-GLD quantile regression is shown in Figure 9.4 and it reveals the distribution of the data points.

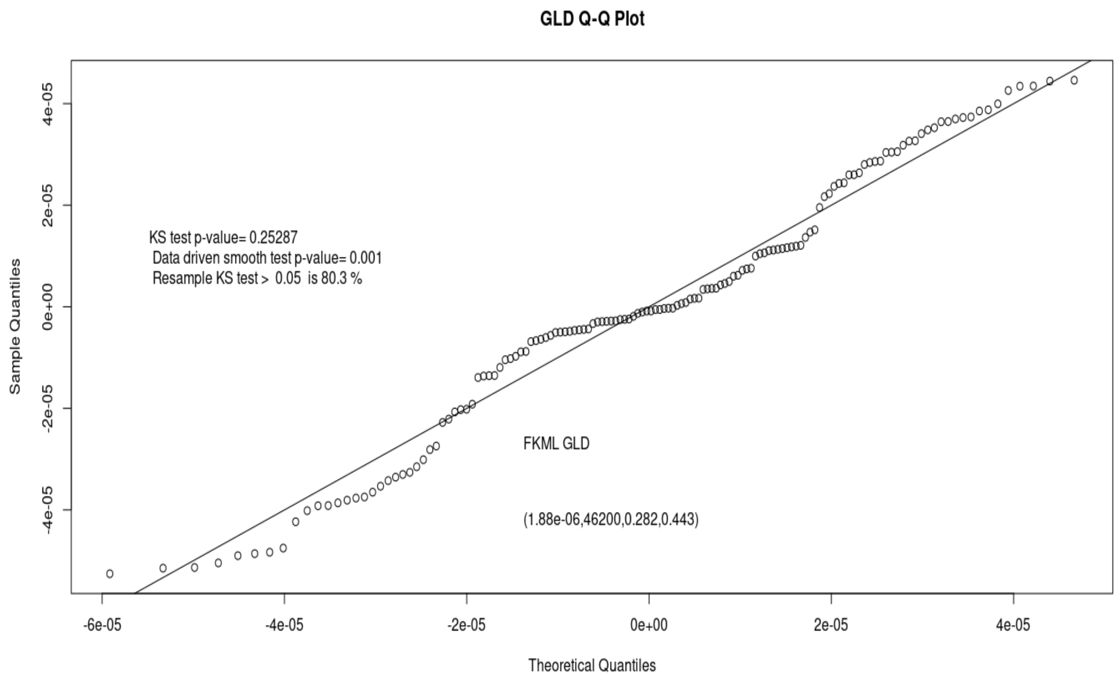


Figure 9.4: GLD Q-Q Plot for FMKL-GLD 75Q Quantile Model

The results presented in Table 9.4 are the *p*-values of the Kolmogorov-Smirnov (KS) test statistic and resample KS test shown in Figure 9.2, Figure 9.3 and Figure 9.4, respectively.

Table 9.4: KS and data driven smooth test results for fitted FMKL-GLD Quantile Models

Model	KS test *p*-value	data driven smooth test *p*-value	Resample KS test> 0.05
FMKL-GDP Q25	0.0328	0.0000	47.7
FMKL-GDP Q50	0.2529	0.0010	80.3
FMKL-GDP Q75	0.0605	0.0010	73.1

Figure 9.2, Figure 9.3 and Figure 9.4 show the Kolmogorov-Smirnov (KS) goodness of fit tests and the FMKL-GLD Q-Q plots for the FMKL-GLD quantile regression models such as FMKL-GLD 25Q, FMKL-GLD 50Q and FMKL-GLD 75Q. Table 9.4 shows the *p*-values of the K-S test, data-driven smooth test and Resample KS test. Thus, all the *p*-values < 0.05 level of significance except FMKL-GLD 50Q indicate that the FMKL-GLD quantiles fit the data moderately. Therefore, from the GLD Q-Q plot, it can be emphasized that the FMKL-GLD quantile regression

model, particularly FMKL-GLD 50Q, performs well in estimating the parameters of the identified economic growth for Nigeria. Having obtained the contributions of the identified economic growth drivers in this study by the various fitted FMKL-GLD quantile models that include FMKL-GLD 25Q, FMKL-GLD 50Q and FMKL-GLD 75Q, respectively, it is imperative to examine the goodness of fit of the fitted FMKL-GLD quantile models, and we present the results in Table 9.5.

Table 9.5: Performance Metrics for Fitted FMKL-GLD Quantile Models

Model	FMKL-GLD 25Q	FMKL-GLD 50Q	FMKL-GLD 75Q
Adj. R-Square	0.7313	0.7659	0.7539
Quasi-LR Stat	582.5285	751.7937	581.2553
Prob (Quasi-LR stat)	0.0000	0.0000	0.0000
RMSE	0.1581	0.1445	0.1701
MAE	0.1034	0.0875	0.1212
MAPE	0.9911	0.8417	`0`999
Bias P	0.2319	0.0118	0.2871

Table 9.4 shows the goodness of fit of the FMKL-GLDquantile models, including FMKL-GLD 25Q, FMKL-GLD 50Q and FMKL-GLD 75Q, respectively. This is done to determine the predictive power or forecasting efficiency of the FMKL-GLD 25Q, FMKL-GLD 50Q and FMKL-GLD 75Q, respectively. It is determined by using the model performance metrics such as root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and bias proportion (Bias P). Thus, from the results in Table 9.5, the RMSE for the fitted FMKL-GLD 25Q, FMKL-GLD 50Q, and FMKL-GLD 75Q quantile models are 0.1581, 0.1445 and 0.1701 respectively. The MAE for the fitted FMKL-GLD 25Q, FMKL-GLD 50Q, and FMKL-GLD 75Q quantile models are 0.1034, 0.0875 and 0.1212, respectively. The same results are also obtained for other model performance measures used for this study. The adjusted R-square of the fitted FMKL-GLD 25Q, FMKL-GLD 50Q and FMKL-GLD 75Q quantile models are revealed to be 73.13%, 76.59% and 75.39%, respectively, indicating the variations or changes in economic growth (RGDP) that can be explained by the identified economic growth drivers under consideration in fitting the FMKL-GLD quantile models. Thus, it can be emphasized that the

FMKL-GLD 50Q regression model is a more efficient FMKL-GLD quantile model for examining and predicting economic growth based on its smallest values of the model performance metrics such as RMSE, MAE and MAPE when it is compared with the remaining other FMKL-GLD quantile models considered for this study. Hence, the forecast for RGDP using FMKL-GLD 50Q quantile regression shall be considered the most efficient and reliable in this chapter.

9.4 Forecast for RGDP using FMKL-GLD 50Q Regression Model

Figure 9.5 shows the next ten quarter forecasts of RGDP using the selected robust FMKL-GLD 50Q regression model.

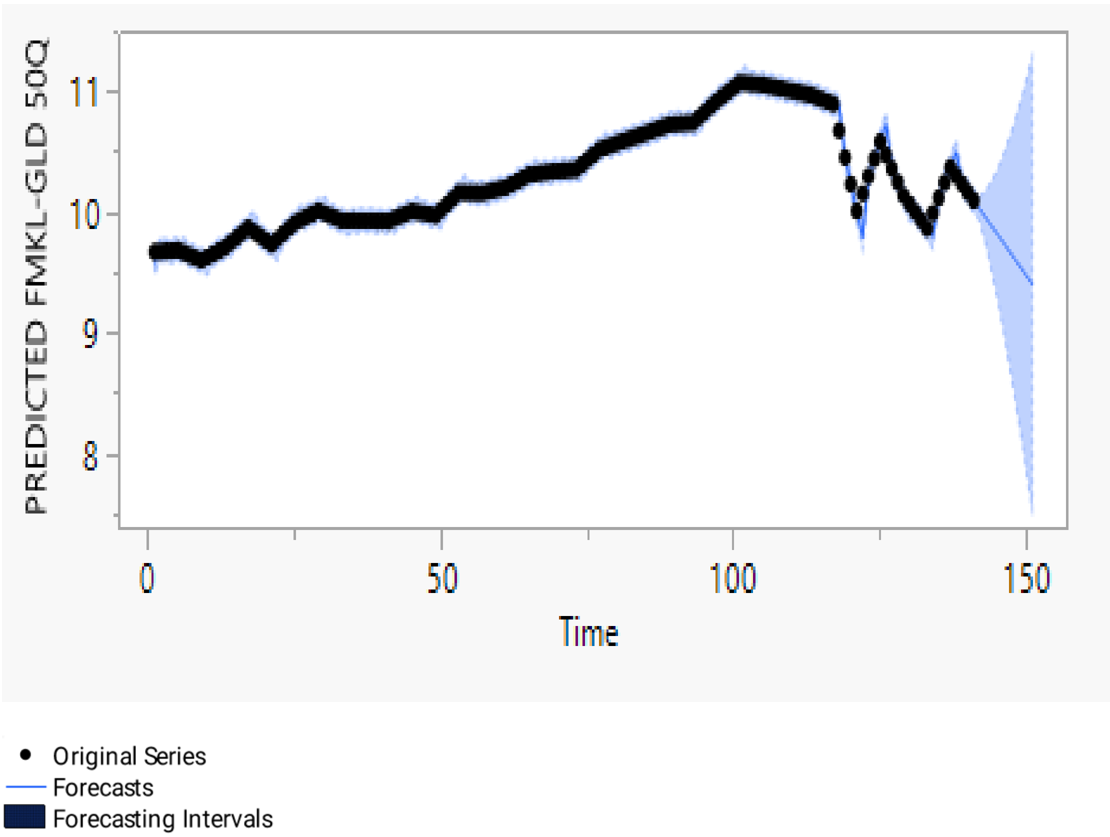


Figure 9.5: Forecast Plot for the adopted Gaussian Process Regression Technique

In Figure 9.5, we present a plot that shows the forecasts or predicted values of the economic growth rate (RGDP) for the next ten (10) quarters based on the efficiency of the FMKL-GLD 50Q quantile regression model in generating stable and reliable predictive values of economic growth rate using the data set under consideration.

9.5 Concluding Remarks

This chapter examines an estimation of economic growth parameters in Nigeria based on the identified drivers of economic growth, such as INDT, EXDT, RINR, REXR, and OPEN, in the presence of multicollinearity and outliers as an assumption violation. An exploratory and diagnostic analysis was carried out to establish the relationship between economic growth (RGDP) and the aforementioned economic drivers. Consequently, to simultaneously address these problems and obtain efficient parameter estimates, a coupler FMKL-GLD quantile regression model is fitted for the available data set for this study.

The results revealed that INDT, RINR, REXR, and OPEN contribute positively to economic growth, with increases of 0.15%, 0.24%, 0.06%, and 1.76%, respectively. Conversely, EXDT negatively impacts economic growth in Nigeria, reducing it by 0.02%. These findings indicate the potential of these factors as significant determinants of economic growth in Nigeria. The estimated location, scale, and shape parameters for the fitted FMKL-GLD model are -0.0253, 34.1488, -0.3303, and -0.5758, respectively. INDT, RINR, REXR, and OPEN positively and significantly influence RGDP, as revealed in the FMKL-GLD 25Q and FMKL-GLD 50Q quantile regression models while, INDT, RINR, and REXR positively and significantly influence RGDP, as revealed in the FMKL-GLD quantile regression model. The predictive performance of the fitted FMKL-GLD quantile models is evaluated using RMSE, MAE, MAPE, and bias proportion. The RMSE values for the FMKL-GLD 25Q, FMKL-GLD 50Q, and FMKL-GLD 75Q models are 0.1581, 0.1445, and 0.1701, respectively. The MAE values for the FMKL-GLD 25Q, FMKL-GLD 50Q, and FMKL-GLD 75Q models are 0.1034, 0.0875, and 0.1212, respectively.

Therefore, based on the performance metrics, it can be emphasized that the FMKL-GLD 50Q regression model is the most efficient FMKL-GLD quantile model for examining and predicting economic growth in Nigeria, based on its smallest RMSE and MAE values. The FMKL-GLD quantile regression technique produces efficient and optimal parameter estimates even in the presence of multicollinearity and outliers in the dataset. Based on the location, scale, and shape parameters from the FMKL-GLD and GLD Q-Q plots, as well as the estimated parameters for RGDP, it can be stressed that the FMKL-GLD is an efficient model that accurately describes the economic situation in Nigeria. The model indicates that the economy is experiencing inverse growth, highlighting the need for significant efforts and strong will to address the situation. To achieve this, adopting economic openness as a policy direction is essential for the development and growth of Nigeria's economy. In Chapter 10, the conclusion and recommendations for this study are discussed.

CHAPTER 10

CONCLUSION AND RECOMMENDATIONS

10.1 Conclusion

A classical and ordinary least squares (OLS) regression technique is a widely accepted method for modelling, estimating model parameters, and making predictions based on the fitted model. However, several conditions, known as assumptions, must be satisfied for the appropriate use of this technique. These correlated assumptions include linearity, independence of the variables, adequate sample size, no correlation of error terms, normality, absence of outliers, and stationarity of the variables. It must be noted that when dealing with modelling and prediction involving macroeconomic variables, proper data exploration and diagnostic evaluation are essential to ascertain the nature and structure of the dataset. This study focuses on the exploration, modelling, and prediction of economic growth in Nigeria based on identified macroeconomic variables. These variables include internal and external debt, interest and exchange rates (as monetary policy instruments), and trade openness (as a fiscal policy instrument), all considered in the presence of multicollinearity and outliers.

Therefore, an ordinary least square regression, a classical and efficient method required for the modelling and prediction, is rendered inefficient and, as such, the need for robust statistical methods for this purpose. Hence, the general aim of this study is to model, estimate and predict economic growth using internal debt, external debt, interest rate, exchange rate and degree of economic openness as economic growth drivers in Nigeria in the presence of multicollinearity and outliers as an assumption violation. Specifically, this is examined through a thorough data exploratory analysis and diagnostic evaluation coupled with various robust statistical techniques ranging from ridge regression method, robust principal component regression technique such as M-estimator, S-estimator and M-estimator methods, partial least square regression method; average centered penalized least square regression technique, gaussian process regression method based on machine learning and a coupler FMKL-GLD quantile regression method. These are structured and discussed in detail in various chapters in this study, as stated earlier in Chapter 1 of this thesis.

A number of conclusions have been drawn from the findings of this research. The conclusions are based on the robustness of the alternative statistical methods over ordinary least square regression in the presence of multicollinearity and outliers. Also, a conclusion is made on the estimated parameters of the predictors under consideration for the benefits of government and policymakers for policy formulation and implementation. Therefore, on the basis of the various adopted data exploratory and diagnostic evaluation as well as the estimation, model and predictive methods, the stationarity test establishes that KPSS is effective and optimal in determining the stationarity of the variable(s) at level [(0)] when the model(s) satisfied intercept but no trend condition and for both trend and intercept condition.

Findings based on Granger causality establish that those variables that Granger caused RGDP can be used to predict economic growth, and those explanatory variables that Granger caused any of the other explanatory variable(s) are not independent and, as such, are suspected or suggested variables causing multicollinearity problem in the study. Also, it can be stressed that those explanatory variables that do not Granger cause any other explanatory variable(s) are independent and, as such, cannot cause any multicollinearity problem. The presence and statistical significance of multicollinearity caused by internal debt (INDT) outliers in the data set under consideration can be attributed to the period of severe economic challenges in Nigeria. This study also establishes the relevance and appropriateness of all the robust statistical techniques and predictive models used as an alternative to the ordinary least square estimation method in examining and forecasting economic growth based on the identified economic drivers used in carrying out this study in Nigeria.

However, in order to avoid spurious results and missing out on vital information in modelling, estimation of the model parameters and prediction, the identified robust statistical methods earlier mentioned in the thesis structure are thoroughly used to explore and estimate the model economic growth (RGDP) and its determinants such as INDT, EXDT, RINR, REXR and OPEN based on the detection of multicollinearity and outliers. In this study, it is apparent that the ridge regression method addressed the multicollinearity problem in the explanatory variables or determinants of the RGDP when the ridge regression constant k is 0.29 as obtained through ridge trace statistics.

When the ridge regression constant k is zero, the result is the same as the ordinary least square regression method.

A robust principal component analysis regression technique, incorporating M-estimator, S-estimator, and MM-estimator methods, was adopted in this study. The findings establish that the robust principal component regression analysis based on the M-estimation technique outperformed those based on the S-estimator and MM-estimator. The M-estimation technique optimally and efficiently estimated and predicted RGDP using the principal components PC1 and PC2 generated from the identified predictors—INDT, EXDT, RINR, REXR, and OPEN—in Nigeria during the study period.

The non-cross-validated partial least squares method reveals that five components are extracted and selected. In contrast, the cross-validated partial least squares method identifies and selects two components based on the linear combination and transformation of the identified economic growth drivers or predictors to predict economic growth (RGDP) in Nigeria. In the non-cross-validated partial least squares method, 91.5% of the variance in the economic growth drivers is explained by the five extracted components. Meanwhile, in the cross-validated method, two components explain 72.6% of the variance in the predictors. Additionally, the cross-validated method shows that 87.4% of the variation in economic growth is explained by the two components, with a cross-validated predictive value of 78.4%. This result is better and more efficient, improving the prediction accuracy over the non-cross-validated method, which extracted and selected five components to predict economic growth.

Among the average centered penalized least squares regression techniques employed—LASSO, ridge, and elastic net—it is established that the LASSO penalized least squares method is the most efficient and appropriate for examining and forecasting economic growth in Nigeria. Consequently, it can be emphasized that LASSO, as an average centered penalized regression technique, is more sensitive and effective than other penalized regression models in this category for predicting RGDP in Nigeria.

Another essential technique this study considered is the Gaussian process regression, a machine learning-based method. In obtaining optimal, efficient and stable parameters that efficiently predict economic growth (RGDP) in Nigeria, this technique shows the sensitivity and the contributions of the identified predictors or economic growth drivers on economic growth (RGDP) to be positive. This technique also establishes that identified economic growth drivers or explanatory variables predicted 10.33% economic growth (RGDP) based on the available data during the period under investigation. The multicollinearity problem is perfectly addressed as the interaction plot result evidently shows the independence of the identified economic growth drivers in efficiently predicting a stable and reliable economic growth rate.

Additionally, using the FMKL-GLD quantile regression method, the median quantile regression method (FMKL-GLD 50Q) provides the best estimate and most efficient predictive model for investigating economic growth and its identified predictors in Nigeria, even in the presence of multicollinearity and outliers. Overall, the performance of various robust statistical techniques for estimating, modelling, and predicting economic growth based on the identified explanatory variables (economic growth drivers) in this study is evaluated using root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). A method with the minimum values of these evaluation metrics is considered to have better performance and is preferred over others. Table 10.1 presents the performance and efficiency of the robust statistical methods under consideration for modelling and predicting economic growth in Nigeria, offering alternatives to the ordinary least squares regression method in the presence of multicollinearity and outliers.

Table 10.1: Overall Performance and Efficiency of the Robust Statistical Methods

Methods	RMSE	MAE	MAPE
Ridge Regression Model	0.2909	0.2092	71.9537
Robust Principal Component Regression (M-Estimator)	0.1927	0.1319	1.2672
Partial Least Square Regression (Cross Validated) Method	0.1388	0.1094	1.0561
Average Centered Penalized Regression (LASSO) Method	0.2895	0.2174	2.1298
Gaussian Process Regression Method	0.1310	0.0942	0.9184

FMKL-GLD-50Q Quantile Regression Method	0.1445	0.0875	0.8417

Table 10.1 shows the results to assess the general performance and the efficiency of the robust statistical methods used in this study. This is examined using performance metrics evaluation that include root mean square error (RMSE), mean absolute error (MAE) and mean absolute percentage error (MAPE). Thus, Table 10.1 shows that RMSE, MAE and MAPE for the fitted ridge regression technique are 0.2909, 0.2092 and 71.9537, respectively. The RMSE, MAE, and MAPE results for the robust principal component regression technique (M-estimator) values are 0.1927, 0.1319, and 1.2672, respectively. The RMSE, MAE and MAPE results for partial least square regression (cross validate) method values are 0.1388, 0.1094 and 1.0561, respectively. The RMSE, MAE and MAPE for the fitted average centered penalized regression least square regression LASSO technique are 0.2895, 0.2174 and 2.1298, respectively. The RMSE, MAE and MAPE with the associated values 0.1310, 0.0942 and 0.9184, respectively, also establish the efficiency of the Gaussian process regression technique in predicting economic growth.

The RMSE, MAE and MAPE results for the fitted FMKL-GLD 50Q quantile regression model with the values of 0.1445, 0.0875 and 0.8417, respectively. Thus, it can be emphasized based on the findings from this study that the proposed coupler FMKL-GLD 50Q quantile regression model followed by the Gaussian process regression method are the most efficient and optimal robust statistical methods that are effectively used to predict stable and reliable values for the economic growth rate in Nigeria. Hence, the statistical methods are ranked in the order of coupler FMKL-GLD 50Q quantile regression technique, Gaussian process regression method, partial least square (cross-validated) method, robust principal component regression method (M-estimator), average centered penalized least square regression based on least absolute shrinkage and selection operator (LASSO) regularization method and ridge regression method respectively.

On the other hand, a conclusion based on the estimated parameters or the predictors under consideration is beneficial to government and policymakers. This is because it can be used for policy formulation and implementation. The study provides a better understanding of the existing relationship between economic growth and the identified economic drivers or predictors. In this study, it is established that the coupler FMKL-GLD quantile regression model revealed that ‘INDT, RINR, REXR and OPEN’ increase the growth of the economy by 17.94%, 29.42%, 7.99% and 145.10% respectively. It was also found that EXDT led to a decline of 3.92% in economic growth in Nigeria during the period under consideration. Also, from the variable importance projection result, it can be emphasized that internal borrowing and trade openness are crucial for Nigeria's economic growth.

Also, the Gaussian process model, the main effect or contributions of the economic growth drivers to the RGDP are 38.53%, 0.21%, 1.37% and 0.49% for INDT, EXDT, RINR, REXR and OPEN, respectively. Also, in this study, the results indicate that 6.63% of INDT, 0.17% of OPEN, 6.66% of EXDT, 4.3175 of REXR and 3.1051% of RINR reliably predicted 10.33% as an average economic growth in RGDP at the desirability value of 0.46.

Therefore, it can be concluded that economic recession, crash in crude oil prices at the international market, insecurity, and terrorist activities all led to insufficient availability of funds and inadequate internal funding through borrowing to grow the economy. The production level, particularly agricultural and manufacturing products, which are the alternative sources through which the economy can grow, is also hampered due to lack of adequate finance either by high-interest rates and instability in the exchange rate as well as colossal debt servicing attached to foreign loans.

The closure of all international borders during the period under study affected the openness of the economy for exportation and importation activities. As such, international patronage could have been higher, hindering the economy's growth. An astronomical increase in the naira to dollar exchange rate and, above all, the Covid-19 pandemic that was heavily witnessed during the period under study greatly affected the identified economic drivers or predictors, and by implication, the impacts are translated to the economic growth (RGDP) during the period under investigation.

10.2 Limitations of the Study

The main objective of this study is to model, estimate, and predict economic growth in Nigeria in the presence of multicollinearity and outliers using robust statistical methods. While economic growth has many drivers, this study focuses on a select few due to the unavailability or inaccessibility of data on other potential explanatory variables. Consequently, the limited scope of identified economic growth drivers serves as a limitation to this study.

10.3 Recommendations

Based on the findings and conclusion from the exploratory analysis, modelling, estimate and prediction of economic growth in Nigeria under the violation of assumptions of the linear model, the following recommendations were made:

- Government and policymakers must effectively leverage the benefits of trade openness driven by exports to foster international patronage, thereby promoting the growth of the Nigerian economy.
- The adoption of economic openness is a crucial policy direction that Nigeria must rigorously pursue to foster economic development and growth.
- Infrastructure development and economic growth through borrowing, whether internal or external, are not sufficiently sustainable to foster economic growth. Therefore, generating more revenue is essential, which should be judiciously used to target development that will stimulate economic growth.
- Given the volatility in exchange rates and high interest rates, there is a need for policy directions aimed at mitigating and minimizing their negative impacts, while enhancing overall economic growth.
- All forms of insecurity plaguing the country must be genuinely addressed and reduced, if not completely eradicated, to enhance productivity, particularly in the agricultural and manufacturing sectors of the economy.
- Furthermore, greater priority must be placed on ensuring proper public spending on healthcare, education, and other essential social services to improve living conditions and develop the citizens, thereby fostering productivity that can lead to economic growth.

10.4 Suggestions for Further Study

In this study, various robust statistical methods have been employed to model, estimate and predict economic growth growth in Nigeria using internal debt, external debt, interest rate, exchange rate and degree of economic openness as the predictors. However, Nigeria's economy had gone through various regimes such as military and civilian regime and the economy had the run by different political party with the occurrence of outbreak of pandemics (Ebola and Covid 19). Thus, the need for further study to investigate the structural breaks and propose appropriate model(s) that can be efficiently used to stimate and predict stable and reliable values for the economic growth in Nigeria or other developing economy in the World. Also, a comparative study can be carried among some selected developing economy or developed economy that have experience same economic crisis and pandemic outbreaks during the period to be considered for investigation.

REFERENCES

Abiodun, A. and Uche, J.U. (2023). Domestic public debt and economic growth in Nigeria: Effect of political regimes, *International Journal of Finance and Banking Research*. *9*(3), pp.50-57.

Adeeb, B., Saeed, S. and Ali, F., (2013). *Do Fiscal, monetary and trade policies matter for growth? Empirical evidence from Pakistan* (Doctoral dissertation, Dissertation, Mohammad Ali Jinnah University, Islamabad).

Adejoke, S.I., Ibiyemi, O., Olusuyi, E.A. and. Babatunde, A.G., (2013). Trade openness and economic growth in Nigeria: An empirical analysis. *International Journal of Humanities and Social Science Invention, 2*(6), pp.2319-7722.

Ademola, I.S., Olusuyi, A.E., Ibiyemi, O. and Babatunde, G.A., (2013). Trade openness and economic growth in Nigeria (1981–2009): an empirical analysis. *International Journal of Humanities and Social Science Invention*, *2*(6), pp.101-113.

Ademola, S.S., Tajudeen, A.O. and Adewumi, Z.A., (2018). External debt and economic growth of Nigeria: An empirical investigation. *South Asian Journal of Social Studies and Economics*, *1*(2), pp.1-11.

Adigwe, P.K., Echekoba, F.N. and Onyeagba, J.B., (2015). Monetary policy and economic growth in Nigeria: A critical evaluation. *IOSR Journal of Business and Management (IOSR-JBM)*, *17*(2), pp.110-119.

Adofu, I. and Abula, M., (2010). Domestic debt and the Nigerian economy. *Current Research Journal of Economic Theory*, *2*(1), pp.22-26.

Afaha, J.S. and Njogo, B.O., (2012). Trade openness and its effect on the Nigerian economy (1970-2010). *Arabian Journal of Business and Management Review (OMAN Chapter)*, *2*(1), p.49.

Afolabi, B., Danladi, J.D. and Azeez, M.I., (2017). International trade and economic growth in Nigeria. *Global Journal of Human-Social Science: Economics*, *17*(5), pp.1-12.

Afolabi, A.Ì., Sotan, K.A. and Salahudeen, S.O., (2020). Trade policy and economic growth in Nigeria. *Economic Thought journal*, (5), pp.84-99.

Agbanobi, T., (2017). Looking inwards: Will Nigeria tread the path of economic recovery and growth in 2017. *Economic and Policy Review*, *16*(1), pp.3-11.

Agu, S.C., Ugah, J.O., Ajaero, G.N. and Ibegbulem, C.A., (2022). Comparative Study of Least Squares, Ridge, Lasso and Principal Component Regression Algorithms. *African Journal of Advances in Science and Technology Research*, *8*(1), pp.01-07.

Agunbiade, D.A., (2012). A note on the effect of the Multicollinearity Phenomenon of a Simultaneous Equation Model. *Journal of Mathematical Sciences: Advances and Applications*, *15*(1), pp.1-12.

Ajao, M.G., (2013). The determinants of real exchange rate volatility in Nigeria. *Ethiopian Journal of Economics, 24*(2), pp.43-62.

Ajayi, I. and Adewusi, D., (2020). Effect of public debt on economic growth in Nigeria: An Empirical Investigation' *International Journal of Business and Management Review*, 8 (1), pp.18-38.

Ajayi, M. A. and Aluko, O. A., (2017). Evaluating the relative impact of monetary and fiscal policy in Nigeria using the St. Louis equation. *Acta Universitatis Danibus Economica,* 13(1), pp.456–471.

Ajayi, L. B. and Oke, M. O.. (2012). Effect of external debt on economic growth and development of Nigeria. *International Journal of Business and Social Science,* 3(12), pp.297-304.

Ajayi, F. O. and Atanda, A. A., (2012). Globalization and macroeconomic stability in Nigeria: An autoregressive adjustment analysis, *European Journal of Globalization and Development Research,* 3(1), pp.175-193.

Ajie, H.A., Akekere, J. and Ewubare, D.B., (2014). Praxis of public sector economics and finance. *Port Harcourt: Pearl Publishers.[4]*.

Akhanolu, I.A., Babajide, A.A., Akinjare, V., Oladeji, T. and Osuma, G., (2018). The effect of public debt on economic growth in Nigeria: An empirical investigation. *International Business Management*, *12*(6), pp.436-441.

Akinmulegun Sunday, O., (2018). Exchange rate fluctuation and industrial output growth in Nigeria. *International Journal of Economics and Financial Research*, *4*(5), pp.145-158.

Akpan, E.O. and Atan, J.A., (2011). Effects of exchange rate movements on economic growth in Nigeria. *CBN Journal of Applied Statistics*, *2*(2), pp.1-14.

Akram, N., (2016). Public debt and pro-poor economic growth evidence from South Asian countries. *Economic research-Ekonomska istraživanja*, *29*(1), pp.746-757.

Alajekwu, U.B., Ezeabasili, V.N. and Nzotta, S.M., (2013). Trade openness, stock market development and economic growth of Nigeria: Empirical evidence. *Research Journal of Finance and Accounting*, *4*(3), pp.120-127.

Matveeva, A. and Leonenko, V., (2022). Application of gaussian process regression as a surrogate modeling method to assess the dynamics of COVID-19 propagation. *Procedia Computer Science*, *212*, pp.340-347.

AlBassam, B.A., (2013). The relationship between governance and economic growth during times of crisis. *European Journal of sustainable development*, *2*(2), pp.1-18.

Ahmad, A.U., Balakrishnan, U.V. and Jha, P.S., (2019). Detection of collinearity effects on Explanatory Variables and Error Term in Multiple Regression. *International Journal of Innovative Technology and Exploring Engineering (TM), IJITEE*.pp1599-1615.

Alkhateeb, T. T. Y. Mahmood, H. Z. and Sultan, Z. A. (2016). The relationship between exports and economic growth in Saudi Arabia, *Asian Social Science*, 12(4), pp.117-124

Kharusi, S.A. and Ada, M.S., (2018). External debt and economic growth: The case of emerging economy. *Journal of economic integration*, *33*(1), pp.1141-1157.

Allison, P. D. (1999). Logistic regression using the SAS System: Theory and application. Cary, NC: SAS Institute.

Alper, A.E., (2017). Exchange rate volatility and trade flows. *Fiscaoeconomia*, *1*(3), pp.14-39.

Alphonsus, C. and Raji, A.O., (2019). Application of Principal Component Analysis (PCA) for correcting multicollinearity and dimension reduction of morphological parameters in Bunaji Cows. *Nigerian Journal of Animal Science*, *21*(2), pp.1-8.

Altee, A.H., Adam, M.H., Esmaeel, S.E. and Saled, M.S., (2014). Financial development trade Openness and economic growth: Evidence from Sultanate of Oman. *Journal of economic and sustainable development*, *5*, p.23.

Amassoma, D., (2017). The nexus between exchange rate variation and economic growth in Nigeria, *Journal of Entrepreneurship, Business and Economics*, 5(1), pp.1-40.

Amirkhalkhali, S. and Dar, A., (2019). Trade openness, factor productivity, and economic growth: recent evidence from OECD countries (2000-2015). *Applied Econometrics and International Development*, *19*(1), pp.5-14.

Anderson, L.C. and Jordan, J.L., (1968), Monetary and fiscal actions: A test of their relative importance in economic stabilization. *Federal Reserve Bank of St. Louis Review*, *80*, pp.29-45.

Anoka, F. and Takon, N., (2014). Exchange rate depreciation and government policy is Nigeria: An empirical evidence. *The Business and Management Review*. *4*(3), pp.161-170.

Arowolo T.O., Adewale, F.L. and Kayode, A., (2016). A comparative study of some method of handling multicollinearity in an autocorrelated error, *African Journal of Science and Technology Science and Engineering Series, 13*(2), pp.68-72.

Arshad, M.U. and Usman, M., (2017) Impact of export on economic growth of Pakistan, *Journal of Economics and Sustainable Development*, *8*(16), pp.115-121.

Asghar, N. and Hussain, Z. (2014). Financial development, trade openness and economic growth in developing countries: Recent evidence from panel data, *Pakistan Economic and Social Review*, *52*(2), pp.99-126.

Asiama J.P. and Kumah, F.Y., (2010). Determinants of real exchange rate movements: Evidence from a panel of African countries. *West African Finance and Economic Review, 2*(2), pp.48-88.

Asinya, F. A. and Takon, N., (2014). Exchange rate depreciation and government policy is Nigeria: An Empirical Evidence. *The Business and Management Review*. *4*(3), pp.161-170.

Askin, G.R., and Montgomery, D.C., (1980). Augmented robust estimators. *Techonometrics*, *22*, pp.333-341.

Atique, R. and Malik, K., (2012). Impact of domestic and external debts on economic growth of Pakistan, *World Applied Science Journal*. *20*(1), pp.120-129.

Awujola, A., (2013). External trade and Nigeria economy: An impact analysis, *Journal of Poverty, Investment and Developmen*t, *An Open Access International Journal*, *1*, pp.157-163.

Ayinde, K., Lukman, A.F. and Arowolo, O.T., (2015). Combined parameters estimation methods of linear regression model with multicollinearity and autocorrelation, *Journal of Asian Scientific Research*, *5*(5), pp.243-250.

Ayomitunde, A.T., Olaniyi, O.O., Zannu, S.M. and Stephen, B.A., (2018). Monetary policy and economic growth in Nigeria: an ARDL-bound testing and ECM approach. *European Journal of Economics Law Politics*, *6*(4), pp.1-11.

Ayres, R.U. and Warr, B., (2010). The economic growth engine: How energy and work drive material prosperity, Edward Elgar Publishing, Cheltenham.

Ayunku, P.E. and Etale, L.M., (2016). Econometric analysis of external debt, exchange rate and economic growth in Nigeria, *Researchers World – Journal of Arts, Science & Commerce, 7*(2), pp.83-89.

Azu, N.P. and Nasiri, A., (2015) Exchange rate fluctuation and sustainable economic growth in Nigeria: VAR approach, *Journal of Economics and Sustainable Development*, *6*(13), pp.228-237.

Babalola, O.O., Danladi, J.D., Akomolafe, K.J. and Ajiboye, O.P., (2015). Inflation, interest rates and economic growth in Nigeria, *European Journal of Business and Management*, *7*(30), pp.91-102.

Babatunde, M.A., Bukula, H., Olodo, O.A. and Ibraheem, K., (2013). An empirical study of growth through trade: Nigerian evidence. *Arabian journal of business and management review (OMAN* Chapter), *3*(5).

Babu, J.O., Kiprop, S., Kalio, A.M. and Gisore, M., (2014). External debt and economic growth in the East Africa community. *African Journal of Business Management, 8*(21), pp.1011-1018.

Bağci, E. and Ergüven, E., (2016). Relations between interest rate, inflation, growth, and investment in Turkey, (2002-2015). *ISOR Journal of Economics and Finance*, 7(5), pp.43-49.

Balago, G.S., (2014). An empirical analysis of the relationship between government external borrowings and economic growth in Nigeria. *International Journal of Finance and Accounting*, *3*(4), pp.235-243.

Bamidele, T.B. and Joseph, A.I., (2013). Financial crisis and external debt management in Nigeria, *International Journal of Business and Behavioral Sciences*, *3*(4), pp.16-24.

Bakari, S. and Mabrouki, M., (2016) The relationship among exports, imports and economic growth in Turkey, *Munich Personal RePEc Archive*, MPRA Paper No. *76044*, pp.1-10.

Bala D.A. and Asemota J.O., (2013). Exchange rate volatility in Nigeria: Application of GARCH models with exogenous break. *CBN Journal of Applied Statistics*, *4*(1), pp.89-116

Barnett, V. and Lewis, T. (1994). Outliers in statistical data. New York, Wiley.

Basirat, M., Nasirpour, A. and Jorjorzadeh, A., (2014) The effect of exchange rate fluctuations on economic growth considering the level of development of financial markets in selected developing countries, *Asian Economic and Financial Review*, *4*(4), pp.517-528.

Bayar, Y., (2016). Public governance and economic growth in the transitional economies of the European Union. *Transylvanian Review of Administrative Sciences*, *48*, pp.5-18.

Bayrakcı, G.N., (2022). Analysis of the factors affecting the profitability of banks in Turkey by LASSO regression, *Ankara Yildirim Beyazit University Business Journal*, *2*(1), pp.46-57.

Belsley, D.A., Kuh, E. and Welsch, R.E., (1980). Regression diagnostics; identifying influence data and source of collinearity. Wiley, New York. http://dx.doi.org/10.1002/0471725153

Bodunrin, O.S., (2016). The impact of fiscal and monetary policy on Nigerian economic growth. *MPRA Paper No. 92811*, pp.1-19. https://mpra.ub.uni-muenchen.de/92811/.

Boakye, G.A. and Atuilik, W.A. (2024). Public debt, debt servicing and economic growth, *International Journal for Multidisciplinary Research (IJFMR)*, 6(3), pp.1-16.

Bryson, J.H. and Nelson, E., (2014). Which countries have external debt "issues"? Wells Fargo securities economics group. retrieved from http://www.realclearmarkets.com/docs/2014/12/External%20Debt%20_%20Dec%202014. Pdf

Buchinsky, M., (1998). The dynamics of changes in the female wage distribution in the USA: A quantile regression approach, *Journal of Applied Econometrics*, *13*, pp.1-30.

Bucur, R.D. and Harja, M., (2012). Homogeneous areas delimitation by considering the energy demand for plants growing in covered spaces *Environment Engineering Management Journal, 11*, pp.253-257.

Campbell. J.Y. and Peron, P., (1991). Pitfalls and opportunities: What macroeconomists should know about unit roots, *In O. J. Blanchard and S. Fisher (Eds.), NBER Macroeconomics Annual*, *6*, pp.141-220.

Canan, G.C., Melike, M. and Murat, T., (2016). Empirical distributions of daily equity index returns: A comparison, *Expert Systems with Applications*, *54*, pp.170–192.

Central Bank of Nigeria (CBN, 2013). *Understanding monetary policy series No 36*. Monetary department, 10th anniversary commemorative edition.

Central Bank of Nigeria (CBN, 2014). *Annual economic report. 31st December, 2014.*

Central Bank of Nigeria (CBN, 2015). *Annual economic report. 31st December, 2015.*

Central Bank of Nigeria (CBN, 2016). *Annual economic report. 31st December, 2016.*

Central Bank of Nigeria (CBN, 2017). *Annual economic report. 31st December, 2017.*

Central Bank of Nigeria (CBN, 2018). *Annual economic report. 31st December, 2018.*

Central Bank of Nigeria (CBN, 2019). *Annual economic report. 31st December, 2019.*

Central Bank of Nigeria (CBN, 2020). *Annual economic report. 31st December, 2020.*

Central Bank of Nigeria (CBN, 2021). *Annual economic report. 31st December, 2021.*

Central Bank of Nigeria (CBN, 2022). *Annual economic report. 31st December, 2022.*

Central Bank of Nigeria (CBN, 2023). *Annual economic report. 31st December, 2023.*

Chalabi, Y., Scott, D. J. and Wuertz, D., (2012). Flexible distribution modeling with the generalized lambda distribution. Retrieved from https://mpra.ub.uni-muenchen.de/43333/

Chalabi, Y., Scott, D. J. and Würtz, D., (2010). The generalized lambda distribution as an alternative to model financial returns. *Working paper, Eidgenössische Technische Hochschule and University of Auckland, Zurich and Auckland. Retrieved fro*m https://www.rmetrics.org/sites/default/files/glambda_0.pdf

Chatterjee, S. and Hadi, A.S., (1988). Sensitivity analysis in linear regression. Wiley *series in Probability and Mathematical Statistics*. Wiley, New York.

Chatterjee, S. and Hadi, A.S., (2006). *Regression analysis by Example*, 4th ed. NJ: John Wiley and Sons.

Chatterjee, S., Hadi, A.S. and Price, B., (2000). *Regression by example*, 3rd ed. New York: John Wiley and Sons.

Chenery, H.B. and Bruno, M., (1962). Development alternatives in an open economy: The case of Israel, *The Economic Journal*, *72*(285), pp.79-103.

Chenery, B. and Strout, A., (1966). Foreign assistance and economic development, *The American Economic Review*, *56*, pp.679-733.

Chidubem, B.C., (2023). Impact of public debt and debt servicing on economic growth in Nigeria (1980-2019), independent, pp.1-17. http://dx.doi.org/10.2139/ssrn.4447361

Chika, P.I., Uju, R.E. and Uche, C.N. (2023). Rising external debt and exchange rate: empirical evidence from Nigeria, *International Journal of Advanced Economics*, *5*(4), pp.90-106.

Christopher, E., (2014). Impact of Trade Openness on the output growth in the Nigerian. Economy *British journal of economics, management & trade, 4*(5), pp.755-768.

Chukwu, A.B., (2007). Panel data analysis of determinants of trade in Africa and Asia, *Nigerian Journal of Economics and Social Studies*, *49*(3), pp.160-172.

Claessens, S., (1996). The debt laffer curve: Some empirical estimates. *World Development*, *18*(12), pp.38-45.

Coccia, M., (2017). Asymmetric paths of public debt and of general government deficits across countries within and outside the European monetary unification and economic policy of debt dissolution. *The Journal of Economic Asymmetries*, *17*(2017), pp.17–31.

Crouse, R. Jin, C. and Hanumara, R., (1995). Unbiased ridge estimation with prior information and ridge trace, *Communication in Statistics-Theory and Methods*, *24***,** pp.2341-2354. MR1350665(96i:62073). Zbl 0937.62616

Danmola, R.A.,(2013). The impact of exchange rate volatility on the macro economic variables in Nigeria. *European scientific journal*, *9*(7), pp.152-165.

Darare, K., Chinhamu, K. and Chifurira, R., (2021). Investigating the causality relationship of South African financial data: An application of the FMKL-GLD quantile regression model, *M. Sc. thesis submitted to the School of Mathematics and Computer Science, University of KwaZulu-Natal, Durban South Africa, June, 2021.*

Da'Silva, C., (2014). Impact of trade openness on the output growth in Nigeria. *British Journal of Economics, Management and Trade*, *4*(5), pp.755-768.

Da'Silva, D. and Ehinomen, C., (2014). Impact of trade openness on the output growth in Nigeria. *British Journal of Economics, Management and Trade, 4*(5), pp.755-768.

Daumal, M. and Ozyurt, S., (2011). The impact of international trade flows on economic growth in Brazilian states, *Review of Economics and Institutions*, *2*(1), pp.1-25.

Dejan, P., Jan, S., Jan, P. and Jus, K., (2011). Financial modeling using gaussian process models. *A paper presented at the 6th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications*, pp.15-17. September 2011, Prague, Czech Republic.

Dirsehan, T. and Henseler, J., (2022). Modeling indices using partial least squares: How to determine the optimum weights? *Quality & Quantity* https://doi.org/10.1007/s11135-022-01515-5.

Dombi, A. and Dedak, I., (2019). Public debt and economic growth: what do neoclassical growth models teach us?. *Applied Economics*, *51*(29), pp.3104-3121.

Doreswamy, C. and Vastrad. M., (2013). Performance analysis of regularized linear regression models for oxazolines and oxazoles derivative descriptor dataset, *International Journal of Computational Science and Information Technology, 1*(4), pp.111-123.

Draper, N.R. and Smith, H., (1981). *Applied Regression Analysis*. 2nd ed. New York, NY: John Wiley & Sons.

Dritsakis, N., (2007). Exports and economic growth: An empirical investigation of EU, USA and Japan using causality tests, *International Review of Applied Economic Research, 1*(2), pp.79-92.

Dritsakis, N. and Stamatiou, P., (2016). Trade openness and economic growth: A panel co-integration and causality analysis of the newest EU countries, *Romanian Economic Journal*, *18*(59), pp.45-60.

Dwivedi, D.N., (2004). *Managerial Economics*, Sixth Edition, VIKAS Publishing House PVT Ltd, New Delhi, India

Dwivedi, D.N., (2006). *Principles of Economics*, New Delhi, Vikas Publishing House PVT, Ltd.

Ebiwonjumi, A., Chifurira, R. and Chinhamu, K., (2022). An efficient estimation technique for investigating economic growth and its determinants for Nigeria in the presence of multicollinearity, *International Journal of Finance & Banking Studies*, *11*(1), pp.107-119.

Ebiwonjumi, A., Chifurira, R. and Chinhamu, K., (2023). A robust principal component analysis for estimating the parameters of economic growth in Nigeria in the presence of multicollinearity and outlier, *Journal of Statistics Application and Probability*, *12*(2), pp.611-627.

Ebiwonjumi, A., Chifurira, R. and Chinhamu, K., (2023). New perspectives on predicting economic growth in the presence of multicollinearity and outlier, *Journal of Applied Mathematics and Information Sciences*, *17*(4), pp.599-613.

Egbetunde, T., (2012). Public debt and economic growth in Nigeria: Evidence from granger causality. *American Journal of Economics*, *2*(6), pp.101–106.

Ehigiamusoe, K.U. and Lean, H.H., (2019). Influence of real exchange rate on the finance-growth nexus in the West African Region, *Economies*, *7*(23), pp.1-21.

Ehinomen, C., (2014). Impact of trade openness on the output growth in Nigeria. *British Journal of Economics, Management and Trade*, *4*(5), pp.755-768. DOI: 10.9734/BJEMT/2014/6682

Ekor, M., Orekoya, T. Musa, P. and Damisah, O., (2021). Does external debt impair economic growth in Nigeria? *MPRA Paper No.107844*, pp.1-8. https://mpra.ub.uni-muenchen.de/107844/

Ekpo, A.H., (2017). The Nigerian economy: Current recession and beyond. *33rd Convocation lecture at Bayero University, Kano, March* 17, 2017.

Elenya, K. and Jude, O., (2013). Trade openness and economic growth: A comparative analysis of the pre and post-structural adjustment program (SAP) in Nigeria. *Asian Journal of Business and Economics*, *2*(4), pp.2231-3699.

Elom-Obed, F.O., Odo, S.I., Elom-Obed, O. and Anoke, C.I., (2017). Public debt and economic growth in Nigeria. *Asian Research Journal of Arts & Social Sciences*, *4*(3), pp.1–16. https://doi.org/10.9734/ARJASS/2017/36095

Emara, N. and Chiu, I.M., (2016). The impact of governance on economic growth: The case of Middle Eastern and North African Countries. *Topics in Middle Eastern and African Economics, 18*(1), pp.126-144.

Emara, N. and Jhonsa, E., (2014). Governance and economic growth: interpretations for MENA countries. *Topics in Middle Eastern and African Economies, 16*(2), pp.164-183.

Eme, O.A. and Johnson, A.A., (2012). Effect of exchange rate movement on economic growth in Nigerian**,** *CBN Journal of Applied Statistics, 2*(2), pp.1-28.

Eme, D.A. and Olugboyega, O.A., (2012). Exchange rate and macroeconomic Aggregate in Nigeria, *Journal of Economics and Sustainable Development, 3*(2), pp.93-102.

Etale, E L.M. and Etale, L.M., (2016). The relationship between exports, foreign direct investment and economic growth in Malaysia, *International Journal of Business, Management and Economic Research, 7*(2), pp.572-578.

Ezeuchenne, K. and Lawal, E.0., (2017). International trade and economic growth in Nigeria. *Journal of Humanities and Social Science, 22*(6), pp.33-43.

Fantessi, A.A., (2015). Foreign direct investment, domestic investment and economic growth in ECOWAS countries. *International Journal of Research in Economics and Social Sciences*, *5*(11), pp.50-60.

Farebrother, F., (1976). Further results on the mean squared error of the ridge regression, *Journal of the Royal Statistical Society B., 38***,** pp.248-250.

Fasanya, I.O., Onakoya, A.B.O. and Agboluaje, M.A., (2013). Does monetary policy influence economic growth in Nigeria? *Asian Economic Finance Review*, *3*(5), pp.635–646.

Fatbardha, M., Eglantina, H., U˘gur, E., Mirela, P. and Marian, C.V., (2020). The effect of exchange rate volatility on economic growth: case of the CEE Countries, *Journal of risk and financial management*, *13*(177), pp.1-13.

Fayissa, B. and Nsiah, C., (2013). The impact of governance on economic growth in Africa. *The Journal of Developing Areas, 47*(1), pp.91-108.

Freimer, M., Kollia, G., Mudholkar, G.S. and Lin, C.T., (1988). A study of the generalized Tukey lambda family. *Communications in Statistics-Theory and Methods*, *17*(10), pp.3547-3567.

Freund, R.J. and Littell, R.C., (2000). *SAS system for regression*, 3rd Edition

Friedman, J., Hastie, T. and Tibshirani, R., (2001), The elements of statistical learning, *1*, *Springer series in statistics* New York. *2*(16), pp.68, 112.

Garthwaite, P.H., (1994). An interpretation of partial least squares, *Journal of the American Statistical Association*, *89*(425), pp.122–127.

Ghazari, A.S., (2017). Determinants of the demand for banknotes in Malaysia: A partial least equation approach, *thesis submitted to Othman Yeop Abdullah Graduate School of Business, Universiti Utara Malaysia in partial fulfillment of the requirement for the degree of Master of Science*

Giannone, D., Lenza, M. and Reichlin, R., (2010). Market freedom and the global recession. *IMF Economic Review, Palgrave Macmillan; International Monetary Fund*, *59*(1), pp.111-135.

Gibbons, R.D., Hedeker, D. and DuToit, S., (2010). Advances in analysis of longitudinal data. *Annual Review of Clinical Psychology*, *6*, pp.79-107.

Gilchrist, W., (1997). Modelling with quantile distributions. J*ournal of Applied Statistics*, *24*(1), pp.113-122.

Gilchrist, W., (2000). *Statistical modelling with quantile functions*. CRC Press.

Gilchrist, W., (2008). Regression revisited, *International Statistical Review, 76*, pp.401–418.

Gbosi, N.G., (2012). *Modern labour economics and policy analysis.* 2nd edition, Pack publishers, Ebonyi State, Nigeria.

Gbosi, N.G., (2015). *Contemporary macroeconomic problems and stabilization policies*, 2nd edition. Spirit and truth publishers, Benin City, Nigeria.

Godwin, B K. and Thomas, J.G., (1994). Real interest rate equalization and the integration of international financial markets, *Journal of International Money and Finance, 13*(1), pp.107-124.

Gómez-Puig, M. and Sosvilla-Rivero, S., (2017). Public debt and economic growth: Further evidence for the Euro Area. *Research Institute of Applied Economics, Working Paper 2017/15*, pp.1–37.

Gordon, L.B. and Cosimo, M., (2018). Government debt in EMU countries. *The Journal of Economic Asymmetries*, *18*(C), pp.1–14.

Granger, C.W.J. and Newbold, P., (1974). Spurious regressions in econometrics, *Journal of Econometrics*, *2*, pp.111-120.

Greene, W H., (2003). *Econometric analysis*, 5th ed. New Jersey 07458: Prentice Hall Saddle River.

Guisan, M.C., (2011). Industry, foreign trade and employment in EU countries; comparison of France, Germany, Italy, Spain and the UK with the United States. *Applied Econometrics and International Development, 11*(2), pp.43-55.

Gujarati, D.N., (2003). *Basic econometrics*, 3rd edition. The McGraw Hill companies.

Gujarati, D.N., (2004). *Basic econometrics*, 4th edition. The McGraw Hill companies.

Gujarati, D.N., Porter, D.C. and Gunasekar, S., (2012). *Basic econometrics*, 5th edition. New Delhi, India: Tata McGraw Hill Education Private Limited.

Gunst, R. F. and Mason, R. L. (1980). Regression analysis and its application: A data-oriented approach, New York: Marcel Dekker.

Habib, M.M., Mileva, E. and Stracca, L., (2017). The real exchange rate and economic growth: revisiting the case using external instruments. *Journal of International Money and Finance*, *73*, pp.386-398.

Hadi, A.S., (1992). Identifying multiple outliers in multivariate data, *Journal of the Royal Statistical Society, Series B (Methodological)*, pp.761-771.

Haller, A.P., (2012). Concepts of economic growth and development, challenges of crisis and of knowledge, *Economy Transdisciplinarity Cognition*, *15*, pp.66-71.

Hamdan, B.S.S., (2016). The effects of exports and imports on economic growth in the Arab countries; A panel data approach, *Journal of Economics Bibliography*, *3*(1), pp.100-107.

Hameed, A., Ashraf, H. and Chandhary, M.A., (2008). External debt and its impact on economic business growth in Pakistan, *International Research Journal of Finance and Economics*, *20*, pp.132-140.

Hansen, L., (1982). Large sample properties of generalized method of moments estimators. *Econometrica*, *50*(4), pp.1029-1054.

Hao, L. and Naiman, D.Q., (2007). *Quantile regression* (Series/Number 07-149), CA: SAGE Publications.

Hastie, T., Tibshirani, R. and Friedman, J., (2001). The elements of statistical learning: data mining, inference and prediction. *Springer Series in Statistics*.

Heinonen, M., Mannerstrom, H., Rousu, J., Kaski, S. and Lahdesmaki, H., (2016). Non stationary gaussian process regression with hamiltonian monte carlo. In *Proceedings of the 19th*

International Conference on Artificial Intelligence and Statistics: Workshop and Conference Proceedings, pp.732–740.

Helland, I.S., (1988). On the structure of partial least squares regression, *Communication is Statistics, Simulations and Computations, 17*, pp.581–607.

Helland, I.S., (1990). Partial least squares regression and statistical methods, *Scandinavian Journal of Statistics*, *17*, pp.97-114.

Hocking, R R., Speed, F.M. and Lynn, M.J., (1976). A class of biased estimators in linear regression. *Technometrics*, *18*, pp.425-437.

Hoerl, A.E. and Kennard, R.W., (1970). Ridge regression biased estimation for non-orthogonal problems, *Technometrics, 8*, pp.27–51.

Hoerl, A.E., Kennard, R.W. and Baldwin, K.F., (1975). Ridge regression: Some simulation, *Communications in Statistics*, *4*, pp.105-123.

Holland, P.W., (1973). Weighted ridge regression: Combining ridge and robust regression methods, *NBER Working Paper Series*, *11*, pp.1-19.

Huber, P.J., (1964). Robust estimation of a location parameter. *Annals of Mathematical. Statistics.*, *35*, pp.73– 101.

Huber, P.J., (1973). Robust regression: Asymptotics, conjectures and monte carlo, *The Annals of Statistics*, *1*, pp.799-821.

Hussein, Y.A. and Abdalla, A.A., (2012). Generalized two stages ridge regression estimator for multicollinearity and autocorrelated errors, *Canadian Journal on Science and Engineering Mathematics*, *3*(3), pp.79-85.

Hussain, Z. and Asghar, N. (2014). Financial development, trade openness, and economic growth in developing countries: Recent evidence from panel data. *Pakistan journal of economic and social review, 52*(2), pp.99-126.

Ibi, E.E. and Aganyi, A., (2015). Impacts of external debt on economic growth in Nigeria: A vector autoregressive approach. *Journal of Business Management and Administration, 3*(1), pp.1-5.

Idris, M. and Ahmad, T.S.T., (2017). The productivity of public debt borrowing and economic growth in Sub-Saharan region: the Nigerian context, *International Journal of Economic, Commerce and Management, 5*(6), pp.340–373.

Idris, M. and Bakar, R., (2017). Fiscal operations and macroeconomic growth: the Nigerian experience. *International Journal Social, Administration and Management Sciences, 2*(1), pp.31–44.

Idris, M., Bakar R. and Ahmad, T.S.T., (2018). The effects of fiscal operations on economic growth and stability in Nigeria: empirical evidence based on time series data. *International Journal of Accounting and Economic Studies*, *6*(1), pp.36-47.

Igbodika, M.N., Jessie, I.C. and Andabai, P.W., (2016). Domestic debt and the performance of Nigerian economy. An empirical investigation. *European Journal of Research and Reflection in Management Sciences, 4*(3), pp.34-42.

Ijirshar, V.U., (2019). Impact of trade openness on economic growth among ECOWAS Countries: 1975-2017, *CBN Journal of Applied Statistics, 10*(1), pp.75-96.

Imimole, B. and Imuoghele, L E., (2012). Impact of public debt on an emerging economy: evidence from Nigeria (1980-2009), *International Journal of Innovative Research and Development*, *1*(8), pp.59-71.

Insah B., (2013). Modeling real exchange rate volatility. *Journal of economics and sustainable development, 4*(6), pp.61-69

Ishola, V. and Titiloye, T., (2020). Effect of fiscal policy and monetary policy on economic growth in Nigeria: A time series analysis. doi:https://doi.org/10.2139/ssrn.37061 57.

Ismail, B. and Manjula, S., (2016). Estimation of Linear Regression Model with Correlated Regressors in the Presence of Autocorrelation, *International Journal of Statistics and Applications*, *6*(2), pp.35-39.

Ismaila, M. and Imoughele, L.E., (2015). Macroeconomic determinants of economic growth in Nigeria: A co-integration approach, *International Journal of Academic Research in Economics and Management Sciences, 4*(1), pp.34-46.

Iya, I.B. and Aminu, U., (2013). An investigation into the impact of domestic investment and foreign direct investment on economic growth in Nigeria. *International Journal of Humanities Social Sciences and Education, 2*(7), pp.40-50.

Iyoha, M, and Okim, A., (2017). The impact of trade on economic growth in ECOWAS member countries: Evidence from panel data. *CBN Journal of Applied Statistics*, *8*(1), pp.23-49.

James, G., Witten, D., Hastie, T. and Tibshirani, R., (2013). An introduction to statistical learning with applications in R. Springer New York Heidelberg Dordrecht London.

Jelilov, G., (2016). The impact of interest rate on economic growth example of Nigeria. *African Journal of Social Sciences*, pp.51-64.

Jelilov, G., Waziri, F. and Isik, A., (2016). Interest rate behaviour and its relationship with economic growth in Nigeria: An error correction model approach. *The Empirical Economics Letters*, pp.245-255.

Jilenga, M.T., Xu, H. and Igor, M.G., (2016). The impact of external debt and foreign direct investment on economic growth: Empirical evidence from Tanzania. *International Journal of Financial Research*, *7*(2), pp.154-162.

John, I.J., Udoka, C.O., Okon, E.A. and Orok, A.B., (2022). The effect of external debt on economic growth in Nigeria, *Frontiers in Management Science*, *1*(2), pp.38-45.

Johnston, J., (1984). *Econometric methods*, 3rd edition. New York: Mc, Graw Hill

Johnny, N. and Johnnywalker, W., (2018). The relationship between external reserves and economic growth in Nigeria (1980-2016), *International Journal of Economics, Commerce and Management*, *6*(50), pp.213–241.

Joy, J. and Panda, P.K., (2020). Pattern of public debt and debt overhang among BRICS nations: An empirical analysis. *Journal of Financial Economic Policy*, *12*(3), pp.345–363.

Kabemba, C. and Kabwe, M. (2024). Effects of public debt on economic growth: an empirical evidence from Zambia (2011-2021). *American Journal of Industrial and Business Management*, *14*, pp.165-183.

Kamonrat, S., Ruethaichanok, K. and Kuntalee, C., (2022). A gaussian process regression model for forecasting stock exchange of Thailand, *Trends in Sciences*, *19*(6), pp.30-45.

Karvanen, J. and Nuutinen, A., (2008). Characterizing the generalized lambda distribution by L-moments, *Computational Statistics and Data Analysis*, *52*, pp.1971-1983.

Kasidi, F. and Said, A.M., (2013). Impact of external debt on economic growth: A case study of Tanzania, *Advances in Management and Applied Economics*, *3*(4), pp.59 – 82.

Kashif, M. and Sridharan, P., (2015). International reserves accumulation and economic growth: Evidence from India. *International Journal of Engineering and Management Research*, *5*(2), pp.1-18.

Kashif, M., Sridharan, P. and Thiyagarajan, S., (2017). Impact of economic growth on international reserve holdings in Brazil. *Brazilian Journal of Political Economy*, *37*(3), pp.605–614.

Kengdo, A.A.N., Ndeffo, L.N. and Avom, D., (2020). The effect of external debt on domestic investment in Sub-Saharan African regions. *The Economic Research Guardian*, *10*(2), pp.69–82.

Khalaf, G. and Iguernane, M., (2016). Multicollinearity and a ridge regression parameter estimation approach. *Journal of Modern Applied Statistical Methods*, *15*(2), pp.400-410.

Khalaf, G. and Shukur, G., (2005). Choosing ridge parameter for regression problems. *Communications in Statistics, Theory and Methods*, *34*, pp.1177-1182.

Khan, R.E.A. and Gill, A.A., (2014). Crowding-out effect of public borrowing: A case of Pakistan. *ResearchGate Publication*.

Khattry, B. and Rao, J.M., (2002). Fiscal faux past, An analysis of the revenue implications of trade liberalization. *World Development*, *30*, pp.1431–1444.

Kharusi, S.A. and Ada, M.S., (2018). External debt and economic growth: The case of emerging economy. *Journal of Economic Integration*, *33*(1), pp.1141–1157.

Koenker, R. and Bassett, G., (1978). Regression Quantiles, *Econometrica*, *46*(1), pp.33–50.

Koenker, R., (2005). Quantile Regression, Cambridge University Press, New York, USA, pp.138–141.

Kim, E., Ha, Y. and Kim, S., (2017). Public debt, corruption and sustainable economic growth. *Sustainability*, *9*, pp.1-30.

Kim, D.H., Lin, S.H. and Suen, Y.B., (2016). Trade, growth and growth volatility: new panel evidence. *International Review of Economics and Finance*, *45*, pp.384–399.

King, R.A.R. and MacGillivray, H.L., (1999). A starship estimation method for the generalized distributions. *Australian & New Zealand Journal of Statistics*, *41*(3), pp.353-374.

Lawali, B.Z., Machief, P.E. and Aliyu, T.K., (2020). Exchange rate and economic growth nexus: An impact analysis of the Nigerian economy, *Journal of Research in Emerging Markets*, *2*(4), pp.58-67.

Lazorec, M., (2023). Partial least squares regression applied for assessing the economic resilience in EU Countries, *Journal of Public Administration, Finance and Law*, https://doi.org/10.47743/jopafl-2023-28-15

Luana E.M.L., Natalie, P. and Ursula, W. (2023). The effect of governance quality on future economic growth: an analysis and comparison of emerging market and developed economies, *Springer Nature Business and Economic Journal*, *3*(108), pp.1-33.

Lukman, A.F, Arowolo, O. and Ayinde, K., (2014). Some robust ridge regression for handling multicollinearity and outliers, *International Journal of Sciences: Basic and Applied Research*, *16*(2), pp.192-202.

Lyndon, M.E. and Ikechukwu, S.O., (2019). The relationship between exchange rate volatility, trade balance and economic growth in Nigeria: An empirical analysis, *International Journal of Development and Economic Sustainability*,*7*(6), pp.1-14.

Madow, N., Nimonka, B., Brigitte, K.K. and Camarero, M., (2021). On the robust drivers of public debt in Africa: Fresh evidence from Bayesian model averaging approach. *Cogent Economics & Finance*, *9*(1), pp.1-23.

Madsen, H., (2008). *Time Series Analysis* (Boca Raton: Chapman and Hall/CRC).

Maduka, A.C., Obi, C. and Ogugua, O.J., (2016). *Economic development analysis, theories and practice.* Rex Charles and Patrick Limited Anambra: Nigeria.

Manwa, F. and Wijeweera, A., (2016). Trade liberalization and economic growth link: The case of Southern African custom union countries. *Economic Analysis and Policy*, *16*, pp. 1-28.

Marquardt, D.W., (1970). Generalized inverse, ridge regression, biased linear estimation and non–linear estimation, *Technometrics*, *12*, pp.591–612.

Martin, R., (2009). Are there lessons for Africa from China's success against poverty?, *World Development*, *37*(2), pp.303-313.

Matiti, C., (2013). The relationship between public debt and economic growth in Kenya, *International Journal of Social Sciences and Project Planning Management*, *1*(1), pp.65-86.

Mc Donald, G.C. and Galarneau, D.I., (1975). A monte carlo evaluation of some ridge type estimators, *Journal of American statistical Association*, *70*(350), pp.407-412.

Mehta, S., (2015). The dynamics of relationship between exports, imports and economic growth in India, *International Journal of Research in Humanities & Social Sciences*, *3*(7), pp.39-47.

Mhlaba, N., Phiri, A. and Nsiah, C., (2019). Is public debt harmful towards economic growth? New evidence from South Africa. *Cogent Economics & Finance*, *7*(1), pp.1-15.

Micheal, V.A. and Abiodun, A.A., (2018). Estimation of regression coefficients in the presence of multicollinearity, *Social and Basic Sciences Research Review*, *2*(10), pp.404-415.

Modigliani, F., (1961). Long-run implications of alternative fiscal policies and the burden of the national debt, *The Economic Journal*, *71*(284), pp.730-755.

Moh'd AL-Tamimi, K.A. and Jaradat, M.S., (2019). Impact of external debt on economic growth in Jordan for the period (2010-2017). *International Journal of Economics and Finance, 11*(4), pp.114-118.

Mokuolu, J.O., Adejayan, A.O. and Ariyo, C.O. (2024). The impact of public debt on economic growth, *International Journal of Humanities Social Science and Management*, *4*(2), pp.349-360.

Momodu, A.A., (2012). Effect of debt servicing on economic growth in Nigeria, *Reiko Journal of Business and Finance*, *4*(3), pp.116-124.

Mordi, M.C., (2006). The challenges of exchange rate volatility in economic management in Nigeria, *CBN Bulletin*, *30*(3), pp.17-25.

Moshe, B., (1997). The dynamics of changes in the female wage distribution in the USA: A quantile regression approach, *Journal of Applied Econometrics*, *13*(1), pp.1-30.

Mosteller, F. and Tukey, J.W., (1977). Data analysis and regression*: A second course in Statistics*, Reading MA: Addison-Wesley

Muhammad, I., Maria, J. and Muhammad, A., (2013). Comparison of shrinkage regression methods for Remedy of Multicollinearity Problem, *Middle-East Journal of Scientific Research*, *14*(4), pp.570-579.

Murray, L., Nguyen, H., Lee, Y., Remmenga, M.D. and Smith, D.W., (2012). Variance inflation factors in regression models with dummy variables, *Conference on Applied Statistics in Agriculture*. https://doi.org/10.4148/2475-7772.1034.

Mutoh, T., (1985). The equivalence of the burden of internal and external public debt, Economic Letters, *17*(4), pp.369-372.

Naes, T. and Marten, H., (1988). Principal component regression in NIR analysis: Viewpoints, background details selection of components," *Journal of Chemometrics, 2*, pp.155-167.

Nageri, K.I., Ajayi, O. and Olodo, H.B. (2013). An empirical study of growth through trade: Nigeria evidence. *34*(2342), pp.1-12.

Naiman D.Q. and Hao, L., (2007). Quantile regression. Newbury Park, CA: Sage Publications

National Bureau of Statistics (NBS) (2023). Poverty profile in Nigeria, 2023

Nauro, F.C., Paul, D.G. and Yuemei J., (2017). Structural reforms, growth and inequality: An overview of theory, measurement and evidence, *Institute of Labour Economics Discussion Paper Series IZA DP* 11159.

Ndubuisi, P., (2017). Analysis of the impact of external debt on economic growth in an emerging economy: Evidence from Nigeria. *African Research Review*, *11*(4), pp.156-173.

Nduka, K.E., (2013). Trade openness and economic growth in Nigeria, *Journal of Education and Practice*, *4*(1), pp.68-73.

Nelson, C.N., Nathaniel, E.U. and Fredrick, O.A., (2016). Trade openness and exchange rate fluctuations nexus in Nigeria, *European Journal of Scientific Research*, *138*(3), pp.139-144.

Nelson, C. and Plosser, C. (1982). Trends and random walks in macroeconomics time series: Some evidence and implications, *Journal of Monetary Economics*, *10*, pp.139-162.

Nkoro, E. and Uko A.K., (2016). Autoregressive distributed lag (ARDL) co-integration technique: application and interpretation. *Journal of statistical and econometric methods, 5*(4), pp.63-91.

Nuhu, M. and Olaposi, O.J., (2016). Trade liberalization, shocks and economic growth in Nigeria, *European Journal of Scientific Research*, *139*(3), pp.205-221.

Nwannebuike, U.S., Ike, U.J. and Onuka, O.I., (2016). External debt and economic growth: The Nigeria experience. *European Journal of Accounting Auditing and Finance Research*, *4*(2), pp.33-48.

Nwosa, P., Keji, S., Adegboyo, S. and Fasina, O., (2020). Trade openness and unemployment rate in Nigeria. *Oradea Journal Business Economic, 5*(2), pp.52-62.

Nyeche, E. (2024). Impact of exchange rate on economic growth in Nigeria, *International Journal of Advanced Economics*, *6*(6), pp.242-250.

Obansa, S.A.J., Okoroafor, O.K.D., Aluko, O.O. and Eze, M., (2013) Perceived relationship between exchange rate, interest rate and economic growth in Nigeria: 1970 - 2010, *American Journal of Humanities and Social Sciences*, *1*(3), pp.116-124.

Obayori, J.B., Krokeyi, W.S. and Kakain, S., (2019). External debt and economic growth in Nigeria. *International Journal of Science and Management Studies*, *2*(2), pp.1-6.

Ochalibe, A.I., Awoderu, B.K. and Onyia, C.C. (2017). External debt and economic development: Policy implications and poverty reduction in Nigeria. *International Journal of Academic Research and Reflection*, *5*(1), pp.1-15.

Odubuasi, A.C., Uzoka, P.U. and Anichebe, A.S., (2018). External debt and economic growth in Nigeria. *Journal of Accounting and Financial Management*, *4*(6), pp.98-108.

Ogbonna, K.S., Ibenta, S.N., Chris-Ejiogu, U.G. and Atsanan, A.N., (2019). Public debt services and Nigerian economic growth: 1970-2017. *European Academic Research*, *6*(10), pp.22-34.

Ogunjimi, J.A., (2019). The impact of public debt on investment: Evidence from development bank of Nigeria, *Journal of Economic and Sustainable Growth*, *3*(2), pp.1-28.

Ogunmuyiwa, M.S., (2011). Does public debt promote economic growth? *Current Research, Journal of Economic Theory*. *3*(1), pp.11-28.

Oguntegbe, K.F. and Alexander, V., (2019). Enhancing economic competitiveness in Nigeria: The roles of exchange rate and interest rate, *Global Scientific Journal*, *7*(6), pp.526-536.

Ojewumi, J.S. and Aremu, O.A., (2021). External debt and poverty incidence in Nigeria: error correction mechanism, *Nigerian Studies in Economics and Management Sciences*, *4*(2), pp.54-66.

Oke, M.O. and Sulaiman, L.A., (2012). External debt, economic growth and investment in Nigeria. *European Journal of Business and Management*, *4*(11), pp.67-75.

Okereafor, G., Ogungbangbe, B.M. and Anyanwu, A., (2015). Positioning Nigeria for global competitiveness in the 21st century: The policy imperatives. *International Journal of Management Science and Business Administration*, *1*(10), pp.56-70.

Okorontah, C.F. and Odoemena, I.U., (2016). Effect of exchange rate fluctuation on economic growth of Nigeria, *International Journal of Innovative Finance and Economic Research*, *4*(2), pp.1-7.

Okonkwo, C.S. and Odularu, G.O., (2013). External debt, debt burden and economic growth nexus: Empirical evidence and policy lessons from selected West African States, *Paper presented at the Conference of Centre for the Study of African Economies (CSAE), March 18-20, Oxford, UK: St. Catherine's College.*

Olaifa, F.G., Subair, K. and Biala, M.I., (2013). Trade liberalization and economic growth in Nigeria; A co-integration analysis. *Journal Business Economics and Finance*, *2*(3), pp.43-52.

Oluwatoyin, M.A. and Folasade, A.B., (2014). Trade openness, institutions and economic growth in sub-Saharan Africa (SSA), *Developing Country Studies*, *4*(8), pp.1-14

Oluwaseyi, A.A. and Adejoke, M., (2013). Trade openness, foreign investment and economic growth in Nigeria: A long-run analysis, *European Journal of Globalization and Development Research*, *7*(1), pp.446-458.

Omoju, O.E, Babatunde, K.A., Oladunjoye, N.O. and Adesanya, O.O., (2012). Foreign debt and economic growth in Nigeria, *The Indian Economic Journal: The Quarterly Journal of the Indian Economic Association*, *60*(1), pp.155-166.

Onyeiwu C., (2012). Domestic debt and the growth of Nigerian economy. *Research Journal of Finance and Accounting*, *3*(5), pp.45-56.

Opurk, O.C. and Edoumiekumo, G.S., (2013). Economic growth factors in Nigeria: The role of global trade. *American Journal of Humanities and Social Science, 1*(2), pp.51-55.

Owusu, E.L. and Odhiambo, N.M. (2014). Financial liberalization and economic growth in Nigeria: evidence based on the ARDL-bounds testing approach. *International Journal of Sustainable Economy, 6*(2), pp.130-141.

Oyewole, O. and Agunbiade, D.A., (2020). Regression techniques in the presence of multicollinearity and autocorrelation Phenomena: Monte carlo approach, *Anale. Seria Informatică*, *18*(1), pp.70-77.

Oyovwi, D.O., (2012). Exchange rate volatility and economic growth in Nigeria, *Mediterranean Journal of Social Sciences*, *3*(3), pp.399-407.

Ozkale, M.R. and Kaciranlar, S., (2007). Comparisons of the unbiased ridge estimation to the other estimations, *Communication in Statistics-Theory and Methods*, *36*, pp.707-723.

Pere, E., (2015). The impact of good governance in the economic development of Western Balkan countries. *European Journal of Government and Economics, 4(*1), pp.26-45.

Qamar, S. and Tokdar. S. T., (2014). Additive gaussian process regression, arXiv preprint arXiv:1411.7009.

Qayyum, U. and Haider, A., (2012). Foreign aid, external debt and economic growth nexus in low-income countries: *The Role of Institutional Quality*, *1*(22), pp.97-115.

Rahman, M. and Islam, A., (2020). Some dynamic macroeconomic perspectives for India's economic growth: Applications of linear ARDL bounds testing for cointegration and VECM. *Journal of Financial Economic Policy, 12*(4), pp.641–658.

Rais, S.I. and Anwar, T., (2012). Public debt and economic growth in Pakistan: A time series analysis (1972-2010), *Academic Research international*, *2*, pp.535-544.

Ramberg, J.S., Dudewicz, E.J., Tadikamalla, P.R. and Mykytka, E.F., (1979). A probability distribution and its uses in fitting data. *Technometrics*, *21*(2), 201-214

Ramberg, J.S. and Schmeister, B.W., (1974). An approximate method for generating asymmetric random variables. *Communication of the ACM, 17,* pp.78-82.

Rasmussen, C.E. and Williams, C.K.I. (2006). Gaussian processes for machine learning. The MIT Press.

Ria, N.S. and Yogo, P., (2021). Implementation of gaussian process regression in estimating reserve for motor vehicle insurance claims, *Journal of Asian Multicultural Research for Economy and Management Study*, *2*(1), pp.38-48.

Ring, T.S., Abdullah, M.A., Osman, W.S.M., Hamdan, R., Hwang, J.Y.T., Mohamad, A.A., Hassan, M. K. H. and Khalid, F.D. (2021). Impact of external debt on economic growth: The role of institutional quality. *International Journal of Academic Research in Economics and Management and Sciences*, *10*(3), pp.223–236.

Rodrik, D. (1992). The limits of trade policy reforms in developing countries. *Journal of economic perspectives*, *6*(1), pp.87-105.

Rousseeuw, P. and Yohai, V., (1984). Robust regression by means of s estimators in robust and nonlinear time series analysis. In: Franke, J., Härdle, W. and Martin, R.D., Eds., *Lecture Notes in Statistics*, 26, Springer-Verlag, New York, pp.256-274.

Rousseeuw, P.J. and Van Driessen, K., (2006). Computing least trimmed square regression for large data sets, *Technical Report, University of Antwerp*

Rousseeuw, P.J. and Leroy, A.M., (2005). Robust regression and outlier detection. Wiley Inter-science, New York (*Series in Applied Probability and Statistics*), 329 http://dx.doi.org/10.1002/0471725382

Saaed, J.A. and Hussain, M.A., (2015). The causal relationship among trade openness, financial development and economic growth: Evidence from Kuwait, *Journal of Emerging Issues in Economics, Finance and Banking*, *1*(4), pp.2306-3670.

Sameh, A.A. and Moh'd, T.A., (2021). Gaussian process regression for forecasting gasoline price in Jordan, *International Journal of Energy Economics and Policy*, *11*(3), pp.502-509.

Sarkar, N., (1996). Mean squared error matrix comparison of some estimators in linear regression with multicollinearity, *Statistics and Probability Letters*, *21*, pp.1987-2000.

Saul, A.D., Hensman, J., Vehtari, A. and Lawrence, N.D., (2016). Chained gaussian processes, *Journal of Machine Learning Research: Workshop and Conference Proceedings*, *51*, pp.1431–1440.

Saungweme, T., Odhiambo, N.M. and Camarero, M., (2019). Government debt, government debt service and economic growth nexus in Zambia: A multivariate analysis. *Cogent Economics & Finance*, *7*(1), 1622998. https://doi.org/10.1080/23322039.2019.1622998

Saxena, S.P. and Shanker, I., (2018). External debt and economic growth in India. *Social Sciences Asia*, *4*(1), pp.15-25.

Siddique, M., Selvanathan, E.A. and Selvanathan, S., (2015*)*. The impact of external debt on economic growth: Empirical evidence from highly indebted poor countries. *Discussion paper 15.10. University of Western Australia Department of Economics*

Snelson, E. Ghahramani, Z. and Rasmussen, C., (2003). Warped gaussian processes. In *Advances in Neural Information Processing Systems 16*, pp.337–344.

Spilioti, S. and Vamvoukas, G., (2015). The impact of government debt on economic growth: An empirical investigation of the Greek market. *The Journal of Economic Asymmetries, 12*, pp.34–40. https://doi.org/10.1016/j.jeca.2014.10.001

Stock, J.H. and Watson, M.W., (1988). Testing for Common Trends, *Journal of the American Statistical Association*, *83*, pp.1097-1107.

Su, S., (2005). A discretized approach to flexibly fit generalized lambda distributions to data. *Journal of Modern Applied Statistical Methods*, *4*(2), pp.407-424.

Su, S., (2007a). Fitting single and mixture of generalized lambda distributions to data via discretized and maximum likelihood methods: GLDEX in R. *Journal of Statistical Software*, *21*(9), pp.1-17.

Su, S., (2007b). Numerical maximum log likelihood estimation for generalized lambda distributions. *Computational Statistics and Data Analysis*, *51*(8), pp.3983-3998.

Su, S., (2007c). Distributional modelling of pipeline leakage repair costs for a water utility company, *American Journal of Mathematical and Management Sciences*, *27*, pp.369–400.

Su, S., (2010a). Handbook of distribution fitting methods with R. In E. Karian, & Z. Dudewicz (Eds.), *Fitting GLD to data Using the GLDEX 1.0.4 in R* (Chap. 15). CRC Press.

Su, C., (2010b). Application of EGARCH models to estimate financial volatility of daily returns: The empirical case of China, *Master Degree Project* No. 2010:142, University of Gothenburg, Gothenburg

Sulaiman, L.A. and Azeez, B.A., (2012). The effect of external debt on economic growth of Nigeria. *Journal of Economies and Sustainable Development*, *3*(8), pp.71-79.

Sulaiman, L.A. and Migiro, S.O., (2014). The nexus between monetary policy and economic growth in Nigeria: A causality test, *Public Municipal Finance, 3*(2), pp.35-40.

Elijah, S. and Musa, B.A., (2019). Dynamic impact of trade openness on the economic growth in Nigeria. *International Journal of Engineering and Advanced Technology*, *8*(5C), pp.606-616.

Swindel, B.F., (1976). Good ridge estimators based on prior information, *Communication in Statistics-Theory and Methods*, *11*, pp.1065-1075.

Thao, P.T.P., (2018). Impacts of public debt on economic growth in six ASEAN countries. *Retsumeikan Annual Review of International Studies*, *17*(1), pp.63-88.

Teusch, A., (2006). Introduction to the spectral and time series analysis with examples from Geodesy (München: Verlag der Bayerischen Akademie der Wissenschaften).

Tibshirani, R., Hastie, T. and Wainwright, M., (2015). Statistical learning with sparsity. The Lasso and generalizations, Monographs on Statistics and Applied Probability, *143*(143), pp.1-343.

Tibshirani, R., (1996). Regression shrinkage and selection via the LASSO, *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), pp.267-288.

Tinungki, G.M., (2019). The analysis of partial autocorrelation function in predicting maximum wind speed, *IOP Conference Series.: Earth Environmental Science*, *235*(1), pp.1-12.

Tolvanen, V. Jylanki, P. and Vehtari. A., (2014). Expectation propagation for nonstationary heteroscedastic gaussian process regression, *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*.

Tom-Ekine, N.T., (2014). Macroeconomics: Dimensions of competitive indicators and policy performance. Dominus printing Co, #7 Udi Street, Port Harcout, Rivers State, Nigeria.

Toutenburg, H and Heumann, C., (2008) Descriptive statistics: an introduction to methods and applications with R and SPSS (Berlin: Springer-Verlag Berlin Heidelberg).

Tuyet, L.A. (2020). Applying the LASSO method to predict the impact of tariff reductions on customs revenues in Vietnam. *Economics of Development*, 19(3), pp.19-31.

Ucan, O., Akyildiz, A. and Maimaitiaili, M., (2016). The relationship between export and economic growth in Turkey, *European Scientific Journal*, Special Edition (June), pp.61-70.

Ude, D.K and Agodi, J.E., (2015). Does trade openness make sense? Investigation of Nigeria trade policy. *International Journal of Academic Research in Economics and Management Sciences*, 4(1), pp.6-21.

Udeh, S.N., Ugwu, J.I. and Onwuka, I.O., (2016). External debt and economic growth: The Nigeria experience, *European Journal of Accounting Auditing and Finance Research*, *4*(2), pp.33-48.

Ufoeze, L.O., Okuma, C.N., Nwakoby, C. and Alajekwu, U.B., (2018). Effect of foreign exchange rate fluctuations on Nigerian economy, *Annals of Spiru Haret University Economic Series,* Issue *1*, pp.105-122.

Ugbor, I.K., (2014). Trade openness and economic growth in Nigeria: An impact. *Asian Journal of Business and Economics, 3*(10), pp.2241-3602.

Ugwuanyi, L. and Nwadiubu, A.O., (2022). Public debt and the Nigerian economy: 1987-2020, *International Institute of Academic Research and Development Journal of Business and African Economy*, *8*(1), pp.55-66.

Ujuju, L.E. and Etale, L M., (2016) Macroeconomic analysis of the relationship between monetary policy instruments and inflation in Nigeria, *International Journal of Business and Management Review*, *4*(6), pp.31-39.

Ulasan, B., (2012). Openness to international trade and economic growth: A cross country empirical investigation. *Economic Discussion Papers, No. 2012-25.* Pp.1-50.

Ullah, F. and Rauf, A., (2013). Impacts of macroeconomic variables on economic growth: A panel data analysis of selected Asian countries, *GE-International Journal of Engineering Research, 5*(2), pp.22-34.

Uzun, A., Karakoy, C., Kabadayi, B. and Emsen, O.S., (2012). The impacts of external debt on economic growth in transition economics. *Chinese Business Review*, *11*(5), pp.491-499.

Vaidogas, E.R. and Juocevicius, V., (2011). A critical estimation of data on extreme winds in Lithuania, *Journal Environment Engineering. Landscape Management*, *19*, pp.178-188.

Yaqub, J.O., (2011). Exchange rate changes and output performance in Nigeria: A sectorial analysis. *Pakistian Journal of Social Sciences*. *7*(5), pp.12-19.

Yohai, V.J., (1987). High breakdown points and high efficiency robust estimates for regression, *Annals of Statistics*, *15*, pp.642-650.

Yoon, J., Klasen, S., Dreher, A. and Krivobokova, T., (2015). Composite indices based on partial least squares, *Discussion Papers,* No. 171, Georg-August Universität Göttingen, Courant Research Centre - Poverty, Equity and Growth (CRC-PEG), Göttingen

Yusuf, A. and Mohd, S., (2021). The impact of government debt on economic growth in Nigeria, *Cogent Economics & Finance*, 9(1), pp.1-19.

Yuzbasi, B. Arashi, M. and Ahmed, S.E., (2017). Big data analysis using shrinkage strategies. arXiv preprint arXiv: 1704.05074v1, pp.1-17.

Zaman, R. and Muhammad, A., (2014). The role of external debt on economic growth: Evidence from Pakistan economy. *Journal of Economics and Sustainable Development, 5*(24), pp.140-147.

Zaren, F. and. Ari, A., (2013). Trade openness and economic growth: A panel causality test. *International Journal of Business and Social Science*, *4*(9), pp.50-53.

Zari, F., Reza, A.B. and Ozlem, G.A., (2019). Analysis of penalized regression methods in a simple linear model on the high dimensional data. *American Journal of Theoretical and Applied Statistics*, *8*(5), pp.185-192.

Zhang, F., (2011). Cross-validation and regression analysis in high dimensional sparse linear models. A PhD dissertation submitted to the department of Statistics, Stanford University, August, 2011.

Zhao, P. and Yu, B., (2006). On model selection consistency of LASSO, *Journal of Machine Learning Research*, *7*(11), pp.2541-2563.

Zou, H. and Hastie, T., (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B, 67*, pp.301-320.

Zouhaier, H. and Fatma, M. (2014). Debt and economic growth, *International Journal of Economics and Financial Issues, 4*(2), pp.440-448.