Statistical and Machine Learning Approaches to Assessing Foreign Aid Effectiveness in Kenya: An ARDL Framework

Noel Jepkosgei Rutto
137090

Submitted in total fulfilment of the requirements for the degree of Master of Science in Statistical Sciences of Strathmore University

Institute of Mathematical Sciences
Strathmore University
Nairobi, Kenya

June 2025

Declaration and Approval

Declaration

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made in the thesis itself.

Name of Candidate: **Noel Rutto**

Signature: ____________________

Date: May 26, 2025

Approval

The thesis of **Noel Jepkosgei Rutto** was reviewed and approved by the following:

Prof. Samuel Mwalili
Supervisor
Strathmore Institute of Mathematical Sciences

Dr. Geoffrey Madigu
Dean
Strathmore Institute of Mathematical Sciences

Prof. Bernard Shibwabo
Director
Graduate Studies

Abstract

This study investigates the impact of foreign aid and other macroeconomic factors on household consumption in Kenya, using household consumption as a proxy for poverty. It adopts a hybrid methodological approach, combining traditional econometric modelling with modern machine modelling techniques to balance causal inference with predictive accuracy.

The analysis is anchored in the Autoregressive Distributed Lag (ARDL) framework, which is well-suited to small samples and mixed integration orders. After establishing the presence of cointegration among variables, the model is reparameterised into an Error Correction Model (ECM) to distinguish between short-run and long-run effects. Diagnostic tests confirm the model's robustness. A Granger causality test reveals no temporal precedence from foreign aid to household consumption, while GDP per capita consistently emerges as a significant long-run driver.

To complement the explanatory power of ARDL, three machine learning models, LASSO regression, Random Forest, and XGBoost, are implemented to assess their ability to predict changes in household consumption. The LASSO model demonstrates the best performance across all evaluation metrics (MAE, RMSE, R^2), outperforming traditional ARDL and more complex ML models. Feature importance analyses using permutation importance and SHAP values reinforce the dominance of GDP per capita and lagged effects of foreign aid as key predictors.

Findings indicate that while econometric methods offer nuanced insight into short- and long-term dynamics, machine learning provides superior predictive power. The study underscores the potential of a hybrid modelling approach in low-frequency macroeconomic contexts, where data constraints limit the application of purely data-hungry methods. Ultimately, the results contribute to the discussion on how aid and macroeconomic variables influence poverty outcomes in developing economies.

Acknowledgement

I am eternally thankful to God Almighty for His grace, wisdom, and strength, which have sustained me through this research.

To my beloved parents, I express my deep appreciation for their unconditional support and countless sacrifices. Their emotional and financial encouragement has been the foundation for my academic pursuits. Their belief in my potential has not wavered, and I am forever indebted to them for their endless faith in me.

I am profoundly grateful to my supervisor, Prof. Samuel Mwalili, for his exceptional guidance, expertise, and mentorship throughout this thesis. His patience, especially during the extended timeline of this study, and his unwavering commitment to my research have been instrumental in bringing this work to fruition. I am truly fortunate to have the privilege of learning under his guidance.

I also want to thank my dear friend, Christine Kawira, for constantly pushing me to complete my studies. As my accountability partner, she has been a source of motivation and inspiration, helping me remain focused and determined to complete this study.

Thank you, one and all, for standing by me. This journey would not have been the same without you.

Dedication

To my parents, Mr & Mrs Rutto, and my loving husband, Stephen Ogugo, who have had my back when I was swamped.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ARDL	Autoregressive Distributed Lag
ECM	Error Correction Model
FGLS	Feasible Generalised Least Squares
FDI	Foreign Direct Investment
GDP	Gross Domestic Product
GCF	Gross Capital Formation
GMM	Generalised Method of Moments
HDI	Human Development Index
Infl	Inflation
LASSO	Least Absolute Shrinkage and Selection Operator
ML	Machine Learning
NPISHs	Non-Profit Institutions Serving Households
ODA	Official Development Assistance
OLS	Ordinary Least Squares
PP	Phillips-Perron (Test)
SSA	Sub-Saharan Africa
TRADE	Foreign Trade
VAR	Vector Autoregression
VECM	Vector Error Correction Model
WDI	World Development Indicators

Chapter 1: Introduction

1.1 Background of the Study

The role of foreign aid in addressing poverty has been debated in development economics. While aid is intended to foster economic growth and improve living standards, empirical findings on its effectiveness remain inconclusive. Some studies highlight its positive impact on poverty reduction, while others argue that it has negligible or adverse effects. These inconsistencies arise due to variations in methodological approaches, data sources, and country-specific factors, including governance, institutional frameworks, and economic structures (Baltagi and Baltagi, 2008).

Kenya has received foreign aid for over five decades, with funds allocated to infrastructure development, poverty alleviation, and human capital enhancement (Mahembe and Odhiambo, 2019). However, despite substantial aid inflows, poverty remains a persistent challenge, raising concerns about the effectiveness of foreign assistance in achieving its intended developmental outcomes. Recently, the situation has been further complicated by geopolitical changes, particularly the reduction in USAID funding under the "America First" policy (Kohnert, 2025). These funding cuts are disrupting key development programs in Sub-Saharan African countries, including Kenya, exposing the vulnerabilities of aid dependency.

To assess the impact of foreign aid on poverty reduction in Kenya, this study employs an Autoregressive Distributed Lag (ARDL) model, which was proposed by Pesaran et al. (2001), integrated with Machine Learning (ML) techniques such as LASSO regression and Random Forests. While ARDL models help analyse both short-run and long-run relationships, ML methods enhance model selection, improve robustness, and increase predictive accuracy (Hastie et al., 2009). By integrating these techniques, this study aims to provide a novel approach to assessing the aid-poverty relationship, offering better policy insights than traditional econometric models.

1.2 Statement of the Problem

The challenge in evaluating aid effectiveness on poverty reduction lies in modelling a dynamic, time-dependent system where multiple macroeconomic variables interact over time. Traditional approaches such as OLS regression often assume stationarity and fail to account for lag structures and non-stationary time series (Hamilton, 2020). Moreover, Kenya's economic conditions are influenced by various exogenous shocks, such as the COVID-19 pandemic and political crises, which introduce structural breaks and nonlinearities that conventional models struggle to capture.

A key limitation in existing studies is their reliance on panel regressions, which treat aid effectiveness as a cross-sectional problem, ignoring the longitudinal dynamics that evolve. Additionally, multicollinearity among macroeconomic indicators such as Gross Domestic Product (GDP), Foreign Direct Investment (FDI), and trade openness complicates the selection of relevant predictor variables, leading to potential model misspecification.

To overcome these challenges, this study will employ an ARDL model, which allows for different orders of integration among variables and provides a robust framework for analysing both short-term and long-term impacts (Pesaran et al., 2001). Additionally, Machine Learning (ML) methods, including LASSO regression and Random Forests, will be incorporated to enhance variable selection, improve model accuracy, and optimise predictive performance for out-of-sample data.

This research aims to explore the following key questions:

i. In what ways can statistical models be designed to assess both the immediate and long-term impacts of foreign aid on poverty in Kenya?

ii. What statistical techniques can improve feature selection and predictive accuracy in aid-effectiveness modelling?

iii. How do statistical evaluation metrics compare between classical ARDL and modern ML models in modelling aid-poverty dynamics?

iv. Can Machine Learning techniques enhance the identification of structural changes and nonlinearities in aid-effectiveness models?

By addressing these questions, this study will contribute to applied statistical modelling in development research and provide data-driven insights for improving aid allocation strategies.

1.3 Significance of the Study

This research is driven by the essential need to enhance the methodological rigour and practical relevance of statistical analysis in evaluating the impact of foreign aid. Although traditional economic research has been predominantly based on theoretical frameworks and policy discussions, this study seeks to bridge the gap between theory and practice using a robust statistical framework to measure the causal relationship between foreign aid and poverty reduction. In doing so, it aims to provide empirically grounded insights that can inform both academic discourse and policy formulation.

From a policy perspective, the study provides empirical insights into how aid compares to foreign direct investment, trade, inflation, and GDP growth in influencing poverty reduction. It highlights the risks associated with aid volatility, particularly in light of USAID funding cuts under the Trump administration, offering recommendations for mitigating the effects of external funding disruptions on Kenya's development programs. Additionally, by leveraging ML-based forecasting models, this research equips policymakers with better decision-making tools to optimise aid allocation and development planning.

The findings of this study are expected to guide donors, development agencies, and policymakers in structuring aid programs for greater effectiveness and sustainability. By bridging traditional econometrics and modern data science, the research provides a more comprehensive and adaptable framework for evaluating aid effectiveness, with potential applications in other developing economies facing similar challenges.

1.4 Objectives of the Study

The primary objective of this study is to develop a statistically rigorous framework for analysing the relationship between foreign aid and poverty reduction in Kenya by integrating autoregressive modelling and machine learning techniques.

The specific objectives are:

i. To formulate an Autoregressive Distributed Lag (ARDL) model that estimates the short-term and long-term effects of foreign aid on poverty reduction in Kenya.

ii. To assess the predictive accuracy and robustness of machine learning techniques in modelling aid-poverty relationships in Kenya.

iii. To compare the effectiveness of ARDL-based econometric models with ML-enhanced forecasting approaches in capturing aid dynamics and poverty trends in Kenya.

1.5 Scope and Limitations

This study focuses on the statistical relationship between foreign aid and poverty reduction in Kenya from 1974 to 2024, using data sourced from the World Development Indicators (WDI) database. Unlike cross-country panel studies, this research is confined to Kenya, ensuring a country-specific analysis.

There are several limitations to this approach. First, due to data availability, the study uses Households and NPISHs Final Consumption Expenditure per Capita (constant 2015 US$) as a proxy for poverty. While this may not fully capture the multidimensional nature of poverty, it is widely regarded as a reliable and welfare-consistent measure of material well-being in developing countries. Ravallion and Chen (2017) argue that consumption is not only more observable than income but also more reflective of welfare in low-income settings, especially where informal employment and income underreporting are prevalent. Alternative indicators such as the Human Development Index (HDI), the Multidimensional

Poverty Index (MPI), and the Gini coefficient were initially considered. However, HDI data for Kenya has only been available since 1990, Gini coefficients are reported sporadically (typically every decade), and MPI lacks annual coverage, rendering them unsuitable for dynamic time-series analysis over five decades.

Second, reliance on secondary data sources introduces potential measurement inconsistencies. Third, while ARDL models integrated with machine learning techniques offer methodological strengths, they also have inherent limitations. ARDL assumes linearity and stable relationships, whereas ML models may suffer from overfitting and reduced interpretability without appropriate tuning and validation.

Additionally, Kenya's economy has experienced structural breaks due to events such as the COVID-19 pandemic and political crises, which may introduce volatility in the results. However, integrating machine learning approaches helps mitigate some of these challenges by enhancing variable selection and improving predictive robustness.

As the analysis is country-specific, findings may not be directly generalizable to other contexts, though the methodological approach can inform similar studies elsewhere. Despite these constraints, this study offers a methodologically rigorous and statistically innovative framework for evaluating aid effectiveness. The findings are expected to contribute meaningful insights to both academic debates and the design of evidence-informed development policy.

Chapter 2: Literature Review

2.1 Introduction

The relationship between foreign aid and poverty reduction in Sub-Saharan Africa (SSA) and Kenya is complex and contested. While some studies report positive impacts, others highlight persistent challenges and mixed outcomes. This chapter reviews the literature from three perspectives: (1) empirical findings on aid and poverty, (2) methodological advances in econometric and machine learning (ML) approaches, and (3) critical reflections on the limitations and opportunities for hybrid frameworks, particularly in the context of Kenya and SSA.

2.2 Theoretical Foundations of Foreign Aid and Poverty Reduction

Foreign aid has long been a subject of intense theoretical and empirical debate in development economics. Over the years, economists and social scientists have developed several theories to explain the role of foreign aid in economic growth and poverty reduction. These theories include Modernization Theory, Dependency Theory, Two-Gap Model, Dutch Disease Hypothesis, and Public Choice Theory. Each presents a unique perspective on the effectiveness of foreign aid, its potential benefits, and unintended consequences. Among these, Modernization Theory and Dependency Theory are the two most influential classical frameworks that have shaped international development discourse and informed debates on the effectiveness of foreign aid.

Modernization Theory

Modernisation Theory, strongly influenced by the work of Rostow (1990)Rostow (1960) and the Harvard sociologist Parsons (1960), posits that foreign aid accelerates economic growth by facilitating modernisation in recipient countries. According to this theory, underdeveloped nations can achieve economic progress by following the historical development trajectories of industrialised Western nations. Foreign aid serves as a catalyst for technological progress, infrastructure development, and institutional reforms, ultimately fostering industrialisation and poverty reduction.

Rostow (1990) argued that all societies progress through five stages of economic development: traditional society, preconditions for take-off, take-off, drive to maturity, and high mass consumption. Foreign aid helps countries transition through these stages by financing critical investments in infrastructure, education, and industrialisation. In this theory, aid is also seen as a tool to bridge the resource gap in developing countries, enabling them to achieve self-sustaining growth.

However, critics argue that Modernization Theory adopts a Eurocentric, one-size-fits-all approach, assuming that Western development models are universally applicable. Banda (2020), using Tanzania and Malawi as case studies, contends that existing policies based on this theory have been insufficient in addressing poverty in Africa. Instead, they advocate for homegrown development strategies considering each country's unique historical, cultural, and economic contexts. Chiamogu et al. (2012) argues that while aid has some benefits, its effectiveness is severely limited by donor-driven policies, poor governance, and systemic inequities. Actual poverty reduction requires structural reforms beyond aid alone.

Dependency Theory

Dependency Theory emerged in the 1950s and 1960s as a counterargument to Modernization Theory. It critiques the global economic structure, arguing that underdevelopment in the Global South is not merely a result of internal factors but is a product of historical exploitation and systemic global inequalities. Prominent scholars such as Prebisch (1962), Baran (1973), Gunder (1989), Amin (1972), Dos et al. (1970), and Wallerstein (1974) have contributed to this framework.

Dependency Theory posits a structural division of the global economy into an unequal core-periphery dichotomy, where developed nations (the core) dominate capital, technology, and trade systems. In contrast, developing countries (the periphery) remain economically subordinated as suppliers of raw materials and cheap labour (Gunder, 1989; Amin, 1972). Rejecting the notion of underdevelopment as an intrinsic condition, the theory emphasises historical exploitation through colonial and neo-colonial systems that systematically extract wealth from the periphery to the core, perpetuating cyclical dependency. Crucially, it critiques foreign aid as a mechanism reinforcing this imbalance. It argues that development assistance often entrenches reliance on donor nations through conditionalities, market asymmetries, and political leverage rather than fostering autonomous growth.

Empirical studies suggest that aid-dependent economies struggle to achieve self-sustaining growth, as external financial support often perpetuates cycles of poverty rather than alleviating them. Bienefeld (1988); Durokifa and Ijeoma (2018); Bräutigam and Knack (2004) all agree that aid perpetuates poverty, but Bienefeld, Durokifa and Ijeoma frame it as systemic exploitation while Bräutigam and Knack rejects the deterministic dependency claims and acknowledge aid's potential to reinforce dysfunctional systems in poorly governed states.

Dependency theorists like (Kabonga, 2016) and (Lopes, 2024)) argue that donor-imposed policies and aid conditionalities systematically restrict economic autonomy across Africa, reinforcing cycles of external reliance. Kabonga 's broad critique aligns with Lopes 's analysis of how Europe's ”charity dependency” framework, exemplified in cases like Angola's oil-backed loans and Senegal's exploitative fishing agreements, undermines local agency.

While other theories offer valuable critiques of aid, their scope remains narrower compared to the overarching frameworks of Modernisation and Dependency Theory. The Two-Gap Model (Chenery, 1967) frames aid as a tool to address savings and trade deficits but neglects institutional barriers. Public Choice Theory (Boone, 1996) highlights elite capture and rent-seeking, while the Dutch Disease Hypothesis warns of aid-induced economic distortions. The modernisation and dependency theories remain the dominant lenses for analysing aid's paradoxical role in development, bridging both its aspirational potential and structural constraints.

Complementing these classical perspectives, emerging literature has increasingly underscored the importance of institutional quality and governance in determining aid effectiveness, particularly in fiscal allocations to sectors such as health and education (Iwegbu and Dauda, 2022). This shift reflects a growing consensus that the institutional context plays a critical role in mediating the impact of aid interventions beyond economic metrics.

2.3 Empirical Evidence from Aid Effectiveness Studies

The empirical literature presents divergent findings on foreign aid's effectiveness in poverty reduction, with results varying significantly across methodological approaches and country contexts. Studies employing time-series techniques like the Autoregressive Distributed Lag (ARDL) model have produced particularly insightful yet mixed results. Ugwuanyi et al. (2017) applied the ARDL approach to Nigeria (1981-2014) and found significant long-term poverty reduction effects from Official Development Assistance (ODA), particularly when aid was directed toward infrastructure and human capital development. Similarly, Edward and Karamuriro (2020) used ARDL modelling for Uganda and demonstrated positive economic growth impacts from aid inflows, which indirectly contributed to poverty alleviation.

However, other ARDL-based studies present more nuanced findings. Ogbodo and Attamah (2019) found no significant poverty reduction impact from foreign aid in their Nigerian study (1981-2017), attributing this to short-term aid volatility and implementation challenges. Haruna et al. (2023) employed linear and nonlinear ARDL approaches for Nigeria, revealing complex dynamics where aid initially showed negative short-term effects (potentially due to Dutch Disease effects) before exhibiting positive long-term impacts when institutional factors were accounted for.

Cross-country studies using panel data methodologies have similarly produced conflicting evidence. For example, Iwegbu and Dauda (2022) found that when foreign aid is combined with effective fiscal policy in education and health, poverty headcount reductions are significant in many SSA countries. However, others, such as Tang et al. (2017) and Moyo (2010), emphasized that aid effectiveness hinges on macroeconomic policy discipline and institutional integrity. The micro-macro paradox also emerges here, despite strong project-level outcomes, national poverty or GDP growth often remain unaffected (Ndikumana and Pickbourn, 2017).

Mahembe and Odhiambo (2019) and Mahembe and Odhiambo (2021)applied the Generalized Method of Moments (GMM) estimators to Sub-Saharan African countries and found that foreign aid significantly reduced poverty when complemented by sound domestic policies. Their later work in 2021 highlighted education and health investments as key transmission channels. Conversely, SULE (2019) found minimal poverty reduction effects from aid in Nigeria, emphasising how governance failures and corruption can undermine potential benefits.

Quantile regression approaches have added further dimensionality to the debate. Asongu et al. (2019) demonstrated that aid's effectiveness varies across different poverty thresholds, with more pronounced effects in middle-income brackets compared to extreme poverty segments. Mohameda and Mzeea (2017) similarly found heterogeneous impacts across different quantiles of human development indicators, suggesting that aid's benefits may not always reach the poorest populations.

Overall, the conditional nature of aid effectiveness emerges as a consistent theme across methodologies and contexts. The findings of this study align closely with evidence from both Kenya (Khainga and Oduor, 2009) and SSA more broadly (Iwegbu and Dauda, 2022), which emphasize that aid can support poverty reduction when accompanied by appropriate targeting, governance, and complementary policy reforms. At the same time, this thesis diverges methodologically by employing ARDL alongside machine learning models, uncovering nonlinear effects and lag structures that conventional econometric methods may miss. By doing so, it contributes a more granular and context-sensitive understanding of how aid interacts with household welfare dynamics over time.

Akobeng (2020) showed through GMM estimation that aid only reduces poverty in democratic institutional settings, while Maruta et al. (2020) found that institutional quality mediates aid's growth impacts. These findings are particularly relevant for Kenya, where democratic institutions have evolved significantly in recent decades but still face governance challenges.

While Kenya is often subsumed in broader SSA panel studies (e.g., (Mahembe and Odhiambo, 2021);(Anetor et al., 2020)), a handful of country-specific studies offer insights into the aid-growth and aid-poverty dynamics. M'amanja and Morrissey (2006), for instance, found that although investment and imports positively influenced per capita income, aid, particularly in the form of external loans, had a statistically significant negative effect on Kenya's long-run growth. Similarly, Gitaru (2015) identified a negative relationship between aid grants and per capita output, attributing the results to weak monitoring and macroeconomic instability.

In contrast, Khainga and Oduor (2009) provided evidence that well-targeted Official Development Assistance (ODA) significantly improved poverty outcomes, especially among the poorest households, albeit with diminishing returns for near-poor groups. Sectoral analyses such as Borter (2017) and recent case studies highlight that while aid to health, education, and agriculture has increased, these inflows do not always correlate with sectoral GDP growth, suggesting inefficiencies in allocation, alignment with national priorities, or implementation. These divergent findings reflect the conditional nature of aid effectiveness and underscore the importance of sectoral targeting, institutional quality, and governance.

2.4 ARDL and Time-Series Analysis in Aid Effectiveness Research

Methodological Landscape in Aid-Poverty Studies

Empirical research on aid effectiveness has employed diverse econometric approaches with distinct strengths and limitations. Generalised Method of Moments (GMM) estimators (e.g., (Mahembe and Odhiambo, 2021)) are prevalent in panel studies for addressing endogeneity through instrumental variables but often mask country-specific dynamics. Fully Modified OLS (FMOLS), used by Farah et al. (2018) in Ethiopia and Maruta et al. (2020) in cross-country growth studies and the Feasible GLS (FGLS)—applied by Anetor et al. (2020) for SSA poverty analysis—are robust for non-stationary data and heteroskedasticity (Phillips and Hansen, 1990). However, these methods require strict cointegration assumptions and struggle with mixed-order integration, limitations particularly acute in single-country analyses like Kenya's, where structural breaks and policy shifts necessitate flexible modelling.

Rationale for ARDL Modeling

The Autoregressive Distributed Lag (ARDL) framework, pioneered by Pesaran and Smith (1995) and later refined in 2001, offers distinct advantages for studying Kenya's aid-poverty nexus. First, unlike Johansen's cointegration or FMOLS, ARDL accommodates variables with different integration orders [I(0) or I(1)], critical given the non-stationarity of macroeconomic data (Nkoro et al., 2016). This contrasts with other studies that discard potentially valuable I(0) variables. Secondly, the Pesaran bounds test avoids pretesting uncertainties associated with unit root tests, which are often inconclusive in small samples (Narayan, 2005). Finally, the error correction mechanism (ECM) within ARDL decomposes effects into transient dynamics and equilibrium relationships, aligning with theories of aid's delayed poverty impacts as suggested by Ugwuanyi et al. (2017).

The general form of an ARDL(p, q) model:

$$Y_t = \alpha + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=0}^{q} \beta_j X_{t-j} + \epsilon_t \tag{2.1}$$

Where:

Y_t is the dependent variable,

X_t represents the explanatory variables,

p and q denote the number of lags for the dependent and independent variables, respectively,

α is the intercept,

ϕ_i and β_j are coefficients to be estimated,

ϵ_t is a white-noise error term.

In the presence of cointegration, ARDL can also be parameterised to ECM in the form:

$$\Delta Y_t = \alpha + \sum_{i=1}^{p} \phi_i \Delta Y_{t-i} + \sum_{j=0}^{q} \beta_j \Delta X_{t-j} + \lambda_1 Y_{t-1} + \lambda_2 X_{t-1} + \epsilon_t. \tag{2.2}$$

where λ_1 and λ_2 represent long-run equilibrium relationships.

Enhancing ARDL with Machine Learning Techniques

While ARDL excels in parametric estimation, ML techniques address gaps left by conventional methods. These include nonlinearities, model uncertainties and structural breaks. The threshold effects in aid allocation (e.g., diminishing returns beyond 10% of GDP, implied by Asongu et al. 's quantile regression) are detectable via Random Forests but obscured in linear models. The LASSO regression (Tibshirani, 1996) penalizes spurious regressors (e.g., the lagged inflation in FGLS models like Anetor et al. (2020), mitigating overfitting, a known ARDL challenge (Lütkepohl, 2005). Regarding structural breaks, they can be automatically detected by neural networks, which OLS and GLS models, like Farah et al. (2018), treat as ad hoc dummy variables.

2.5 Gaps in Literature

The methodological limitations of existing approaches, particularly their inability to capture complex nonlinear relationships and interaction effects, suggest the need for innovative analytical frameworks.

This need is particularly acute in single-country contexts like Kenya, where existing studies either subsume it in SSA panels or use static OLS specifications ill-suited for its volatile aid inflows. This study aims to bridge this gap by deploying ARDL-ECM to quantify temporal aid-poverty linkages and avoid the rigid cointegration requirements. It will also enhance robustness through ML-based sensitivity checks and provide policy-relevant counterfactuals, which are left unaddressed by GLS and GMM frameworks.

Chapter 3: Research Methodology

This chapter outlines the methodological framework adopted to examine the dynamic relationship between foreign aid and poverty reduction in Kenya. The methodological approach combines the temporal causality analysis of autoregressive distributed lag (ARDL) modelling with machine learning ensembles' predictive power, addressing inferential and forecasting objectives.

3.1 Research Design

This study adopts a quantitative, time-series econometric research design to examine the dynamic relationship between foreign aid and poverty in Kenya. It employs a hybrid methodological approach, combining the Autoregressive Distributed Lag (ARDL) model for causal inference with machine learning techniques for predictive evaluation and model validation.

The ARDL framework is selected for its ability to accommodate small samples and variables of mixed integration orders, enabling analysis of both short-run and long-run effects. To complement the explanatory power of the ARDL model, machine learning methods, including LASSO regression, Random Forest, and XGBoost, are applied to enhance feature selection and assess out-of-sample predictive performance.

This integrated approach supports the inferential objectives (identifying causal effects) and predictive goals (evaluating future poverty trends), providing a more comprehensive understanding of aid effectiveness.

3.2 Data Sources and Variables

The dataset used in this study is sourced from the World Development Indicators (WDI) and spans 1974-2024. The dependent variable is household final consumption expenditure per capita, expressed in constant 2015 US dollars. Let $y_t \in \mathbb{R}$ denote this variable for year t; it is used as a welfare-consistent proxy for poverty. As argued by Ravallion and Chen (2017), consumption-based measures are more reliable than income in capturing relative deprivation, particularly in low-income economies where informal sectors dominate and income reporting is unreliable.

Let $\mathbf{x}_t \in \mathbb{R}^K$ denote the vector of macroeconomic predictors observed at time t, defined as:

$$\mathbf{x}_t = \begin{bmatrix} \text{AID}_t \\ \text{FDI}_t \\ \text{TRADE}_t \\ \text{GCF}_t \\ \text{GDPpc}_t \\ \text{INFL}_t \\ \text{POP}_t \end{bmatrix},$$

where each component represents a standard economic measure, as detailed in Table 3.1.

The annual frequency of the data further supports its suitability for econometric analysis using time-series frameworks. Given the macroeconomic scope of this study, seasonality is not considered a major concern at this level of temporal aggregation. Table 3.1 summarises the variables used in the analysis and their expected theoretical relationships.

3.3 Data Transformation and Stationarity Testing

All time series variables were tested for stationarity prior to model estimation to ensure valid inference and avoid spurious regressions. Stationarity is a key requirement in time series econometrics, particularly for the ARDL framework, which assumes that none of the series is integrated of order two or higher.

Stationarity was assessed using the Augmented Dickey-Fuller (ADF) test, which was corroborated by the Phillips-Perron (PP) test. Both tests evaluate the null hypothesis of a unit root at the 5% significance level. Lag lengths for the ADF test were selected using the Akaike Information Criterion (AIC). The tests were conducted on both levels and in different forms for each series.

Let $z_t \in \{y_t\} \cup \{x_{jt}\}_{j=1}^K$ denote any univariate time series drawn from the set of variables used in the analysis. For series identified as non-stationary at levels$I(0)$, the following transformations were applied to the corresponding z_t:

First-differencing was applied to variables to remove non-stationarity:

$$\Delta Z_t = Z_t - Z_{t-1}$$

Logarithmic transformation was applied to strictly positive variables with high positive skewness:

$$Z_t^* = \log(Z_t)$$

Table 3.1: Variable Description and Expected Relationships

Variable	**Symbol**	**Expected Sign**	**Justification**
Households and NPISHs Final Consumption Expenditure per Capita	HhC_t	Dependent Variable	Proxy for economic well-being and poverty reduction
Foreign Direct Investment, net outflows (% of GDP)	FDI_t	Positive	Contributes to economic growth and employment by increasing capital inflows and fostering business expansion
Net ODA received (% of GNI)	AID_t	Positive	Foreign aid finances development projects, improves welfare and alleviates poverty through targeted investments
Trade (% of GDP)	$TRADE_t$	Positive	Trade openness enhances economic activity, encourages specialisation, and contributes to higher income levels
Gross Capital Formation (% of GDP)	GCF_t	Positive	Capital investment boosts production capacity, drives economic growth, and supports long-term sustainability
Inflation, consumer prices (annual %)	$INFL_t$	Negative	High inflation reduces purchasing power, destabilises the economy, and discourages long-term investment
GDP per capita (constant 2015 US$)	$GDPpc_t$	Positive	Higher GDP per capita indicates economic prosperity, better living standards, and greater consumer purchasing power
Population, total	POP_t	Positive/Negative	Large populations can strain resources but also provide a labour force that drives economic expansion -*This was later excluded due to non-stationarity*

Square root transformation was used for non-negative variables with moderate skewness:

$$Z_t^* = \sqrt{Z_t}$$

Box-Cox transformation was applied to variables requiring more flexible normalisation and variance stabilisation:

$$Z_t^{(\lambda)} = \begin{cases} \dfrac{Z_t^\lambda - 1}{\lambda}, & \lambda \neq 0 \\ \log(Z_t), & \lambda = 0 \end{cases}$$

where the transformation parameter λ was estimated via maximum likelihood.

Table 3.2 presents the final results of the Augmented Dickey-Fuller (ADF) test. Variables such as Population, which failed to exhibit consistent stationarity even after multiple transformations, were excluded to preserve the integrity of the ARDL and machine learning models. Additionally, a Chow test for structural breaks identified a breakpoint at observation 27, indicating a significant shift in the underlying statistical properties of the series over time.

Table 3.2: ADF Test Results for Stationarity (5% Significance Level)

Variable	Test Statistic	Critical Value (5%)	p-value	Decision
Infl	-3.499	-2.93	< 0.05	Stationary
Hh_C_diff	-4.321	-2.93	< 0.05	Stationary
FDI_diff	-6.549	-2.93	< 0.05	Stationary
ODA_diff	-4.613	-2.93	< 0.05	Stationary
Trade_diff	-5.784	-2.93	< 0.05	Stationary
GCF_diff	-6.954	-2.93	< 0.05	Stationary
GDP_pc_diff	-3.410	-2.93	< 0.05	Stationary

Given this instability and to avoid the risk of spurious regression, the Population variable was excluded from the final analysis. Its redundancy further justifies its exclusion, as key variables such as GDP per capita inherently account for population size, while inflation indirectly reflects population-driven consumption patterns. Removing the variable helps maintain the parsimonious structure and reliability of the ARDL and machine learning models.

3.4 Estimation Theory and Methodological Foundations

Following the stationarity adjustments, the empirical analysis proceeds with the estimation of parametric and non-parametric models that capture both causal and predictive relationships. This section formalises the statistical properties of estimators used in the study and presents the estimation techniques employed for both the ARDL and machine learning frameworks.

Estimator Properties

Let $\hat{\theta}_T$ denote an estimator of a true scalar parameter θ, based on a sample of size T. Three classical properties of estimators are relevant in this context:

Unbiasedness: An estimator is unbiased if:

$$\mathbb{E}[\hat{\theta}_T] = \theta.$$

In the linear regression context, the OLS estimator satisfies this property under classical assumptions such as exogeneity and homoscedasticity. In contrast, regularised methods like LASSO introduce bias intentionally to control variance and improve prediction accuracy.

Consistency: An estimator is consistent if:

$$\hat{\theta}_T \xrightarrow{p} \theta \quad \text{as} \quad T \to \infty.$$

This property ensures convergence in probability to the true value, given that the underlying stochastic process satisfies conditions such as stationarity, weak dependence, and correct model specification.

Efficiency: Among all unbiased estimators, $\hat{\theta}_T$ is efficient if it has the smallest variance:

$$\text{Var}(\hat{\theta}_T) \leq \text{Var}(\tilde{\theta}_T), \quad \forall \tilde{\theta}_T \text{ unbiased.}$$

OLS is efficient under the Gauss–Markov conditions, while MLE is asymptotically efficient under standard regularity assumptions.

Estimation Techniques

The study utilises both classical and modern estimation techniques, selected based on their suitability for inference, robustness, and predictive validity.

Ordinary Least Squares (OLS). For linear models, the OLS estimator is defined as:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \sum_{t=1}^{T} \left(y_t - \mathbf{x}_t^\top \boldsymbol{\beta} \right)^2.$$

The closed-form solution is:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y},$$

where $\mathbf{X} \in \mathbb{R}^{T \times K}$ is the matrix of explanatory variables and $\mathbf{y} \in \mathbb{R}^T$ is the response vector. OLS is used in estimating the ARDL model coefficients, given its interpretability and desirable large-sample properties.

Maximum Likelihood Estimation (MLE). The MLE identifies parameters that maximise the log-likelihood function:

$$\ell(\theta) = \sum_{t=1}^{T} \log f(y_t \mid \mathbf{x}_t; \theta),$$

leading to the estimator:

$$\hat{\theta}_{\text{MLE}} = \arg\max_{\theta} \ell(\theta).$$

Under standard regularity conditions, MLE is consistent and asymptotically normal:

$$\sqrt{T}(\hat{\theta}_{\text{MLE}} - \theta) \xrightarrow{d} \mathcal{N}(0, \mathcal{I}^{-1}(\theta)),$$

where $\mathcal{I}(\theta)$ is the Fisher information. While not applied directly in this study, MLE provides the conceptual foundation for several time series diagnostics and likelihood-based selection criteria.

Cross-Validation (CV). Cross-validation is used to assess predictive performance and select tuning parameters in the machine learning models. In k-fold cross-validation, the dataset is split into k subsets; each fold is used once as a validation set, while the model is trained on the remaining $k-1$ folds. The cross-validation error is defined as:

$$\text{CV}(f) = \frac{1}{k} \sum_{i=1}^{k} \sum_{t \in \mathcal{D}^i_{\text{valid}}} \left(y_t - \hat{f}^{-i}(\mathbf{x}_t) \right)^2,$$

where $\hat{f}^{-i}$ denotes the function fitted without fold i. Cross-validation is essential in training the LASSO, Random Forest, and XGBoost models used in this study.

Bias–Variance Trade-off and Regularisation

The expected prediction error for a given input $\mathbf{x}_t$ can be decomposed as:

$$\mathbb{E}\left[(y_t - \hat{f}(\mathbf{x}_t))^2 \right] = \left(\mathbb{E}[\hat{f}(\mathbf{x}_t)] - f(\mathbf{x}_t) \right)^2 + \mathbb{E}\left[\left(\hat{f}(\mathbf{x}_t) - \mathbb{E}[\hat{f}(\mathbf{x}_t)] \right)^2 \right] + \sigma^2,$$

representing the squared bias, variance, and irreducible error, respectively. Regularisation techniques reduce the variance component by introducing bias, thereby improving generalisability.

LASSO regression addresses this trade-off by solving the following optimisation problem:

$$\hat{\boldsymbol{\beta}}_{\text{LASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{t=1}^{T} (y_t - \mathbf{x}_t^\top \boldsymbol{\beta})^2 + \lambda \|\boldsymbol{\beta}\|_1 \right\},$$

where $\lambda \geq 0$ is the regularisation parameter tuned via cross-validation. The ℓ_1 penalty induces sparsity in the coefficient vector, which is particularly useful for addressing multicollinearity and improving prediction performance in small samples.

3.5 ARDL Model Specification

The Autoregressive Distributed Lag (ARDL) model is employed to estimate the dynamic relationship between foreign aid and poverty, proxied by household consumption per capita. The ARDL approach is suitable for small samples and accommodates regressors that are a mix of $I(0)$ and $I(1)$, provided none are integrated of order two or higher. This flexibility makes it particularly appropriate for macroeconomic time series like those used in this study.

Let y_t denote the dependent variable and $\mathbf{x}_t = (x_{1t}, x_{2t}, \ldots, x_{Kt})^\top$ the K-dimensional vector of regressors. The general ARDL($p, q_1, q_2, \ldots, q_K$) model is specified as:

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{q_k} \theta_{jk} x_{k,t-j} + \varepsilon_t,$$

where:

- α is the intercept term,
- ϕ_i and θ_{jk} are the short-run coefficients for the lagged dependent and independent variables respectively,
- $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$ is a white noise error term.

Lag orders p and q_k were selected using the Akaike Information Criterion (AIC), implemented via the `auto_ardl()` function in R. This model allows for asymmetric lag structures across variables, reflecting the temporal heterogeneity in how macroeconomic shocks may influence household consumption.

Bounds Testing for Cointegration

The ARDL model supports bounds testing for cointegration as developed by Pesaran et al. (2001). This involves estimating the unrestricted error correction form of the ARDL model and testing the joint significance of the lagged level variables.

The error correction form is specified as:

$$\Delta y_t = \gamma_0 + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{q_k-1} \delta_{jk} \Delta x_{k,t-j} + \lambda \left(y_{t-1} - \boldsymbol{\beta}^\top \mathbf{x}_{t-1} \right) + \varepsilon_t,$$

where:

- γ_i and δ_{jk} capture short-run dynamics,
- λ is the error correction term, expected to be negative and statistically significant if a long-run relationship exists,
- $\boldsymbol{\beta} \in \mathbb{R}^K$ contains the long-run equilibrium coefficients.

The null hypothesis $H_0 : \lambda = 0$ corresponds to the absence of a long-run relationship. Rejection of the null implies cointegration, validating the use of an Error Correction Model (ECM).

Error Correction Representation

Upon confirmation of cointegration, the ARDL model is reparameterised into an Error Correction Model (ECM) to explicitly capture both short-run dynamics and the speed of adjustment toward the long-run equilibrium. The ECM form is:

$$\Delta y_t = \gamma_0 + \sum_{i=1}^{p-1} \gamma_i \Delta y_{t-i} + \sum_{k=1}^{K} \sum_{j=0}^{q_k-1} \delta_{jk} \Delta x_{k,t-j} + \lambda \, \mathrm{EC}_{t-1} + \varepsilon_t,$$

where the error correction term is defined as:

$$\mathrm{EC}_{t-1} = y_{t-1} - \boldsymbol{\beta}^\top \mathbf{x}_{t-1}.$$

The parameter λ measures the speed at which deviations from the long-run equilibrium are corrected. A significant and negative λ indicates a stable adjustment process, wherein past disequilibria gradually diminish over time.

Model Diagnostics

To ensure the validity of the ARDL model, several diagnostic tests are conducted to assess the robustness of the estimated results. The Breusch-Godfrey LM test is applied to detect the presence of autocorrelation in the residuals, ensuring that the model does not suffer from serial correlation, which could distort inference. The Breusch-Pagan test is employed to examine heteroskedasticity, verifying whether the variance of the residuals remains constant across observations. Finally, the Jarque-Bera test assesses the normality of residuals, confirming whether they follow a Gaussian distribution, a key assumption for valid hypothesis testing and inference in time-series econometric models. These diagnostic procedures collectively enhance the reliability of the model's estimates and ensure the robustness of statistical conclusions.

3.6 Machine Learning Framework

To complement the ARDL model and benchmark predictive performance, this study implements three supervised learning models: LASSO regression, Random Forest, and XGBoost. These models are non-parametric or regularised extensions of classical approaches, suited for high-dimensional, nonlinear, or weakly structured data settings. They are evaluated based on their ability to accurately predict the response variable y_t using transformed predictor vectors $\mathbf{x}_t$, post-stationarity treatment.

Each model is trained on a training subset of the data and evaluated on a holdout test set. Predictive performance is assessed using three standard loss metrics:

$$\text{MAE} = \frac{1}{n}\sum_{t=1}^{n} |y_t - \hat{y}_t|,$$
$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (y_t - \hat{y}_t)^2},$$
$$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2},$$

where $\hat{y}_t$ denotes the predicted value and $\bar{y}$ is the mean of the observed values in the test set.

LASSO Regression

The Least Absolute Shrinkage and Selection Operator (LASSO) is a regularised linear regression model that penalises the absolute magnitude of regression coefficients. It is defined by the optimisation problem:

$$\hat{\boldsymbol{\beta}}_{\text{LASSO}} = \arg\min_{\boldsymbol{\beta}} \left\{ \sum_{t=1}^{T} (y_t - \mathbf{x}_t^\top \boldsymbol{\beta})^2 + \lambda \|\boldsymbol{\beta}\|_1 \right\},$$

where $\|\boldsymbol{\beta}\|_1 = \sum_{j=1}^{K} |\beta_j|$ is the ℓ_1-norm and $\lambda \geq 0$ is a regularisation parameter. The penalty induces sparsity by shrinking some coefficients exactly to zero, effectively performing variable selection.

The tuning parameter λ is selected using k-fold cross-validation, minimising:

$$\mathrm{CV}_{\mathrm{LASSO}}(\lambda) = \frac{1}{k} \sum_{i=1}^{k} \sum_{t \in \mathcal{D}^i_{\text{valid}}} \left(y_t - \hat{f}_\lambda^{-i}(\mathbf{x}_t) \right)^2 .$$

Random Forest

Random Forest is an ensemble method that aggregates predictions from multiple decision trees. Each tree is trained on a bootstrap sample of the data and at each split, a random subset of predictors is considered. The model prediction is the average of the outputs from all trees:

$$\hat{y}_t = \frac{1}{M} \sum_{m=1}^{M} f_m(\mathbf{x}_t),$$

where $f_m(\cdot)$ is the prediction from the m-th tree and M is the total number of trees.

Random Forest reduces variance by decorrelating the trees, and it is robust to multicollinearity and outliers. Hyperparameters such as the number of trees M and the number of predictors sampled at each split are tuned via cross-validation.

XGBoost

Extreme Gradient Boosting (XGBoost) is a scalable implementation of gradient-boosted decision trees. It builds trees sequentially, where each subsequent tree attempts to minimise the residual errors of the previous ensemble. The objective function includes both a loss term and a regularisation term:

$$\mathcal{L} = \sum_{t=1}^{T} \ell(y_t, \hat{y}_t) + \sum_{m=1}^{M} \Omega(f_m),$$

where $\ell(y_t, \hat{y}_t)$ is the squared loss and $\Omega(f_m)$ penalises model complexity. Specifically,

$$\hat{y}_t^{(m)} = \hat{y}_t^{(m-1)} + \eta f_m(\mathbf{x}_t), \quad r_t^{(m)} = y_t - \hat{y}_t^{(m-1)},$$

where $\eta \in (0, 1]$ is the learning rate and $r_t^{(m)}$ are the residuals fitted at iteration m.

Hyperparameters such as learning rate, maximum tree depth, and subsample ratio are tuned using cross-validation.

Feature Importance and Interpretability

To interpret the contribution of features in ML models, Permutation Importance is used to measure the change in model performance when a feature's values are randomly permuted and SHAP Values, based on cooperative game theory, quantifies the marginal contribution of each feature to the model's prediction for a given observation.

These methods enhance the transparency of machine learning models, offering insight into which variables most influence predicted household consumption and aligning with the study's goal of policy-relevant analysis.

3.7 Summary

This hybrid methodology offers a comprehensive framework for investigating both causality and predictive accuracy. The ARDL-ECM approach enables interpretable estimation of short- and long-run macroeconomic relationships, while the machine learning models are used to assess out-of-sample predictive performance and uncover nonlinear dynamics. To further validate the robustness of these approaches, a simulation study was incorporated to test model behaviour under varying sample sizes, noise levels, and functional forms. This layered strategy enhances both inferential credibility and generalisability, supporting the development of rigorous, evidence-informed policy insights.

3.8 Software

All data analysis and modelling were conducted using R version 4.4.1. Core modelling tasks were performed using the **ARDL**, **glmnet**, **randomForest**, and **xgboost** packages. Data manipulation and visualisation were supported by **dplyr** and **ggplot2** from the **tidyverse** ecosystem, while summary tables were generated using **kableExtra**.

Chapter 4: Results

4.1 Introduction

This chapter presents the empirical findings of the study, based on both traditional econometric modeling and modern machine learning approaches. The aim is to evaluate the short-run and long-run impact of foreign aid and other macroeconomic variables on poverty, proxied by household consumption. In addition, the chapter assesses the predictive power of machine learning techniques and compares their performance against the traditional Autoregressive Distributed Lag (ARDL) framework.

4.2 Descriptive Statistics and Exploratory Analysis

The descriptive statistics Table 4.1 highlight key trends in the transformed variables. Household consumption (*Hh_C_diff*) shows moderate variability (SD = 13.71), with values ranging from 1,976 to 2,022, reflecting substantial fluctuations in poverty-related expenditure despite differencing. Inflation (*Infl*) exhibits positive skewness (mean = 11.54, SD = 8.04), suggesting episodic price instability, likely tied to macroeconomic shocks or policy shifts.

Foreign direct investment (*FDI_diff*) is the most volatile variable (SD = 40.87, range: -83.66 to 109.77), indicating erratic capital inflows that necessitate careful lag specification in modeling. In contrast, official development assistance (*ODA_diff*) displays minimal variation (mean = 0.01, range: -0.36 to 0.35), implying limited year-to-year changes and potentially weak standalone predictive power unless lagged effects are accounted for.

Table 4.1: Descriptive Statistics

Statistic	N	Mean	St. Dev.	Min	Max
Hh_C_diff	47	1,999.00	13.71	1,976	2,022
Infl	47	11.54	8.04	1.55	45.98
FDI_diff	47	8.53	40.87	−83.66	109.77
ODA_diff	47	0.01	0.14	−0.36	0.35
Trade_diff	47	−0.03	1.80	−6.99	5.72
GCF_diff	47	−0.65	5.87	−14.43	19.93
GDP_pc_diff	47	0.02	3.07	−11.66	6.37

Trade and gross capital formation (*Trade_diff*) and *GCF_diff*) exhibit symmetric distributions and low standard deviations, signaling stability in external sector indicators. Differenced GDP per capita *GCF_diff*) appears balanced, though its time series plot (Figure 4.1) reveals intermittent spikes, likely corresponding to economic shocks or growth surges.

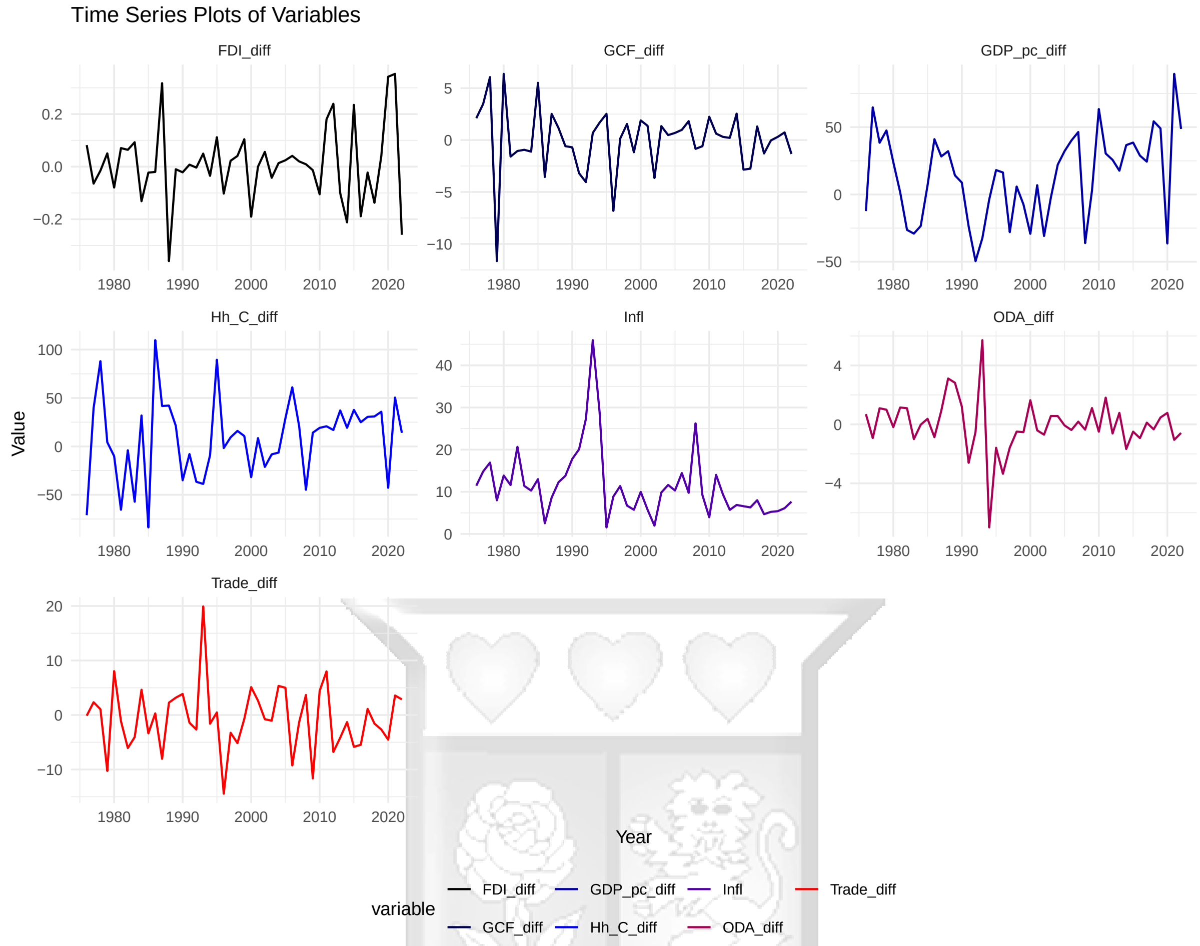

Figure 4.1: Time series of the variables.

The correlation matrix (Figure 4.2) underscores key relationships: a moderate positive correlation between *GDP_pc_diff* and *Hh_C_diff* (≈ 0.63) aligns with expectations that rising income boosts consumption. Conversely, inflation correlates negatively with household consumption (≈ -0.38), implying that price pressures erode real expenditure. No extreme correlations among predictors are observed, mitigating multicollinearity concerns for ARDL and machine learning models.

Density plots (Figure 4.3) further clarify distributions. *Infl*, *FDI*, and *GDP_pc_diff* exhibit skewness and non-normality, justifying transformations or lag adjustments. *Trade_diff* and *GCF_diff* show stable, symmetric distributions, while *GDP_pc_diff*'s occasional outliers align with macroeconomic disruptions. These patterns reinforce the need for robust time series specifications and support the variables' inclusion in subsequent analyses.

4.3 ARDL Model Analysis

The Autoregressive Distributed Lag (ARDL) model was employed to evaluate short- and long-run relationships between household consumption and key macroeconomic variables. To determine the optimal

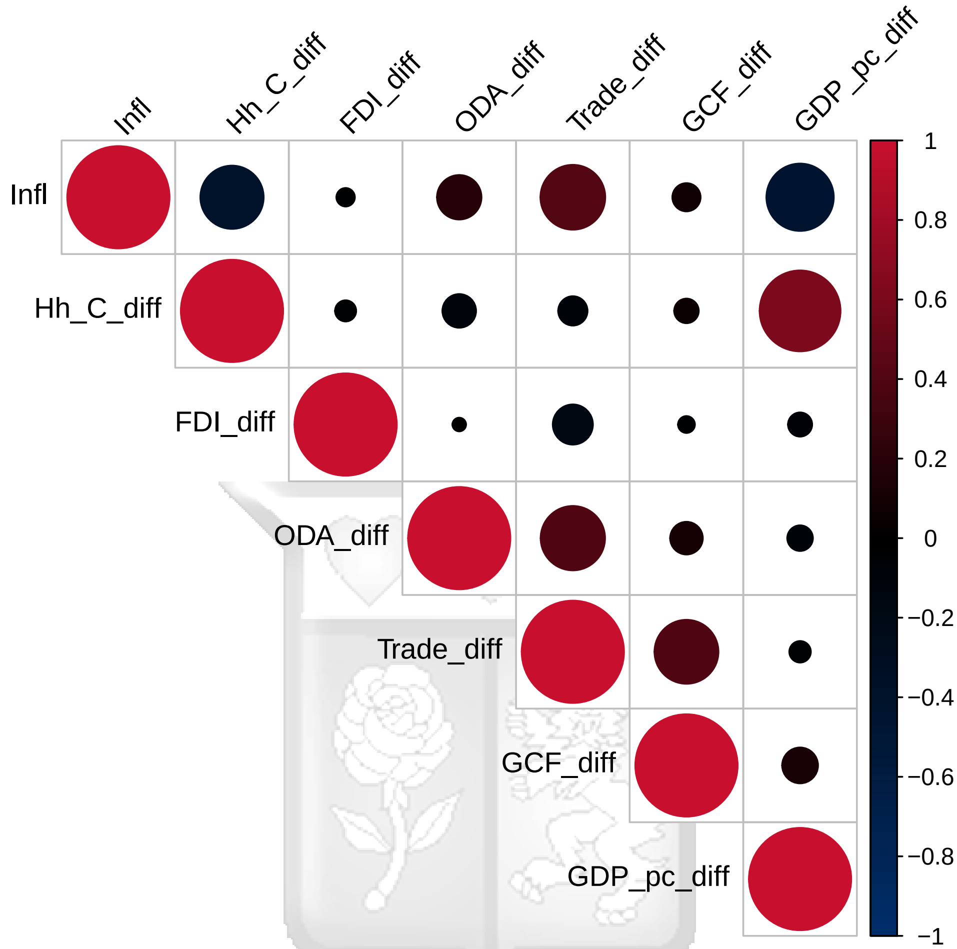

Figure 4.2: Time series of the variables.

lag structure, the *auto_ardl()* function was used with the Akaike Information Criterion (AIC) as the selection metric. A maximum of two lags was allowed for each variable, resulting in the optimal lag configuration shown in Table 4.2.

The ARDL model was then estimated using the selected lag order. Statistically significant predictors included the first lag of official development assistance (ODA), the second lag of gross capital formation (GCF), and GDP per capita. Notably, the lagged dependent variable was negative and significant, confirming short-run dynamic adjustment. The overall model fit was satisfactory, with an adjusted R-squared of 0.54 and an F-statistic that was significant at the 1% level, indicating strong explanatory power.

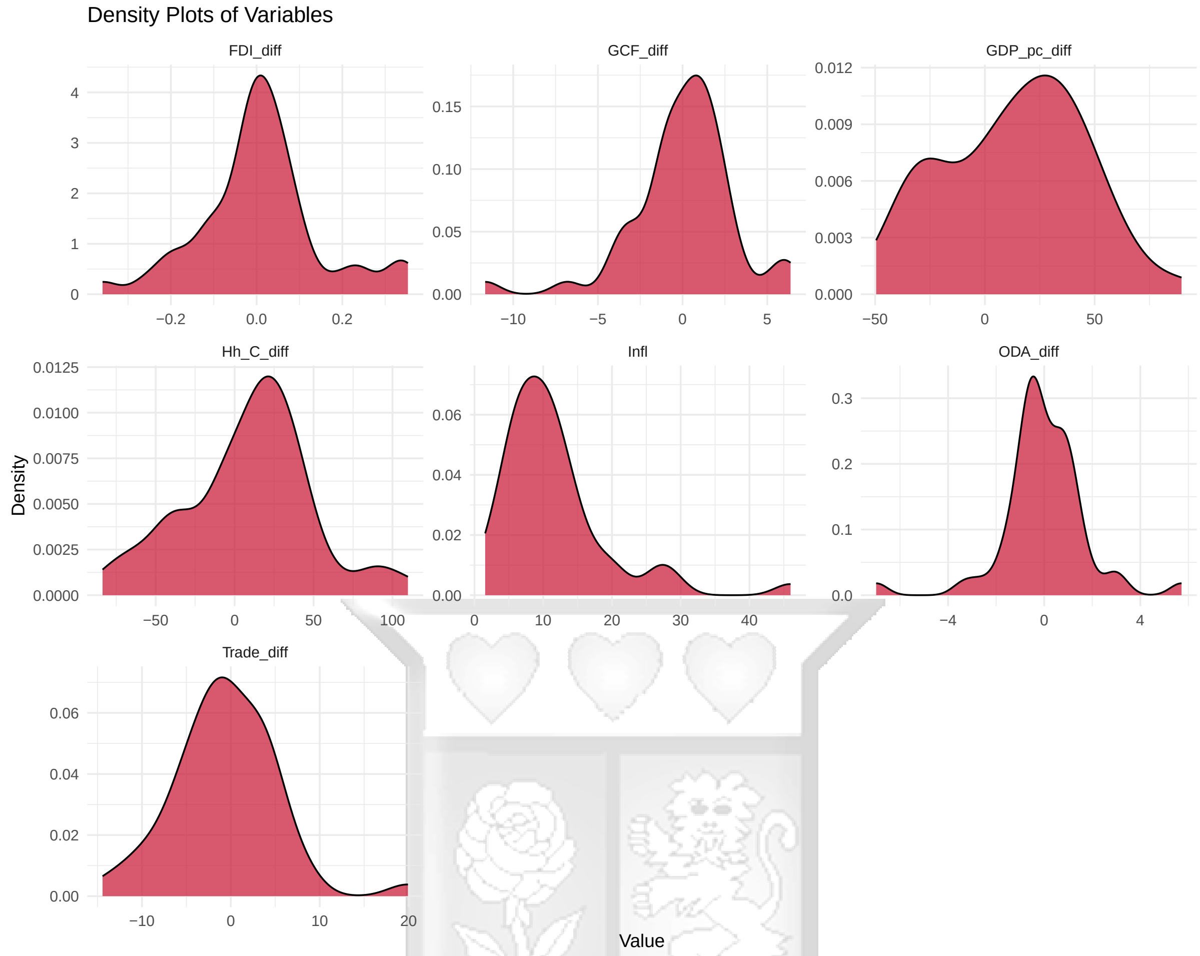

Figure 4.3: Time series of the variables.

4.3.1 Cointegration and Error Correction Model

The Bounds test for cointegration was conducted to assess whether a long-run equilibrium relationship exists among the variables. The computed F-statistic was 14.313 (p-value < 0.001), which exceeds the upper bound critical value at all conventional significance levels. This result strongly supports the presence of a cointegrating relationship.

The ARDL model was reparameterised into its Error Correction Model (ECM) form. The ECM estimates capture the short-run dynamics while incorporating the long-run adjustment process. The error correction term (ECT) was negative and statistically significant (coefficient = –1.44, p < 0.001), indicating rapid convergence to the long-run equilibrium. This suggests that deviations from equilibrium are corrected by more than 100% in a single period, implying potential volatility or overshooting in the adjustment process.

Table 4.2: ARDL Model Coefficients with Summary Statistics

Variable	Estimate	Std. Error	t-value	p-value
(Intercept)	-5.172	10.683	-0.484	0.632
Lag(Hh_C_diff, 1)	-0.441	0.155	-2.841	0.008
ODA_diff	0.035	2.777	0.012	0.990
Lag(ODA_diff, 1)	-7.829	2.433	-3.218	0.003
Lag(ODA_diff, 2)	-2.546	3.042	-0.837	0.409
Trade_diff	-0.120	1.042	-0.115	0.909
GCF_diff	1.708	1.677	1.019	0.316
Lag(GCF_diff, 1)	3.138	1.937	1.620	0.115
Lag(GCF_diff, 2)	4.418	1.698	2.602	0.014
Infl	-0.620	0.886	-0.700	0.489
Lag(Infl, 1)	1.086	0.845	1.285	0.208
GDP_pc_diff	0.738	0.171	4.314	0.000
Lag(GDP_pc_diff, 1)	0.299	0.201	1.488	0.147
FDI_diff	-15.488	29.612	-0.523	0.605
Summary Statistics				
Adjusted R-squared	0.537			
F-statistic	4.924			
p-value	0.0001			
Residual Standard Error	27.046			

Table 4.3: Error Correction Model (ECM) Coefficients and Summary Statistics

Variable	Estimate	Std. Error	t-value	p-value
(Intercept)	-5.1720	3.7159	-1.3919	0.172
d(ODA_diff)	0.0347	1.9992	0.0174	0.986
d(L(ODA_diff, 1))	2.5457	2.3170	1.0987	0.279
d(GCF_diff)	1.7084	0.9825	1.7388	0.090
d(L(GCF_diff, 1))	-4.4179	1.0410	-4.2440	<0.001
d(Infl)	-0.6198	0.6557	-0.9452	0.351
d(GDP_pc_diff)	0.7381	0.1217	6.0667	<0.001
ect	-1.4408	0.1318	-10.9354	<0.001
Summary Statistics				
Residual Standard Error	24.76			
Adjusted R-squared	0.8082			
F-statistic	27.48			
p-value	8.029e-13			

4.3.2 Model Diagnostics and Stability

A series of diagnostic tests were carried out to assess model robustness. The Breusch-Godfrey test indicated no evidence of autocorrelation (p = 0.33), while the Breusch-Pagan test found no heteroskedasticity (p = 0.42). The Jarque-Bera test showed residuals were approximately normally distributed (p = 0.26), and the RESET test did not suggest functional form misspecification (p = 0.62).

Furthermore, stability tests were conducted using CUSUM and CUSUMSQ plots. Both plots remained within the 5% confidence bands throughout the sample period, confirming that the model is structurally stable over time.

Although the Augmented Dickey-Fuller (ADF) test on residuals did not reject the null of a unit root, this result does not undermine the evidence of cointegration, which the more robust bounds testing procedure has already established.

Table 4.4: ARDL Model Diagnostic Tests

Test	Statistic	p-value
Breusch-Godfrey	2.2358	0.3270
Breusch-Pagan	13.3391	0.4220
Jarque-Bera	2.6658	0.2637
Chow Test	1.1631	0.8657
Ramsey RESET	0.2574	0.6156
ADF Test	-2.8993	0.2171

Table 4.5: ECM Model Diagnostic Tests

Test	Statistic	p-value
Breusch-Godfrey	1.7185	0.1899
Breusch-Pagan	6.6484	0.4664
Jarque-Bera	2.6658	0.2637

4.3.3 Causality Analysis

Granger causality tests were employed to assess temporal precedence between variables, particularly examining whether foreign aid (ODA) Granger-causes household consumption (Hh_C). The analysis tested the null hypothesis that ODA does not Granger-cause Hh_C against the alternative that it does, using a lag order of 2. A p-value exceeding the 0.05 threshold indicated a failure to reject the null, suggesting no statistically significant evidence that foreign aid Granger-causes household consumption. This indicates that changes in aid flows do not consistently precede changes in consumption levels within the tested lag structure. However, this does not preclude the existence of other forms of causal influence, such as long-term structural effects or indirect relationships, that may not be captured by Granger causality tests alone.

4.4 Machine Learning Models and Performance Evaluation

To complement the traditional ARDL framework, this study evaluated the predictive performance of three Machine Learning (ML) models, LASSO regression, Random Forest, and XGBoost, in forecasting household consumption, which is used as a proxy for poverty levels. These models were trained using the same feature set as the ARDL model, including lagged variables. They were assessed on their ability to generalise to out-of-sample data using standard predictive performance metrics.

All models were trained on 80% of the data (chronologically split to preserve time structure) and evaluated on the remaining 20%. The metrics used were Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2).

4.4.1 ARDL Model

As the traditional benchmark model, the ARDL was evaluated not only for coefficient inference and long-run relationships but also for predictive accuracy. The final model, based on AIC-driven selection, included two lags for each independent variable.

On the test data, the ARDL model achieved moderate explanatory power, with an MAE of 19.68, RMSE of 24.70, and an R^2 of 0.47 (Table 4.6). However, its predictive accuracy was notably inferior to that of both LASSO regression and the Random Forest ensemble. These results suggest that while the ARDL framework remains valuable for identifying long-term relationships, it may be less suited for short-term forecasting, where machine learning techniques demonstrate superior performance.

4.4.2 LASSO Regression

LASSO regression was implemented with 5-fold cross-validation to optimise the regularisation parameter (lambda.min) to address potential overfitting in high-dimensional settings. This approach performs simultaneous variable selection and coefficient shrinkage, favouring parsimonious models with stronger generalizability.

The final LASSO model retained only the most predictive features, assigning non-zero coefficients to GDP per capita growth (0.56), the first lag of ODA (-1.71), and the second lag of trade openness (0.65), while inflation contributed a marginal negative effect. This selective sparsity enhanced interpretability while reducing overfitting risks.

As shown in Table 4.6, the model achieved superior predictive performance with an MAE of 8.98, RMSE of 11.83, and R^2 of 0.78—outperforming all benchmarked approaches in prediction error minimization and explanatory power. These results position LASSO as the optimal choice for poverty measurement forecasting in this study.

Table 4.6: Model Performance Comparison

Model	MAE	RMSE	R^2
ARDL	68.8080	77.9118	0.0008
LASSO	8.7748	11.9078	0.7751
Random Forest	15.8602	19.1742	0.6653
XGBoost	25.5340	28.9391	0.0302

4.4.3 Random Forest

The Random Forest model was trained using 5-fold cross-validation, with hyperparameter tuning identifying an optimal *mtry* value of 6. Variable importance, assessed through both mean squared error (MSE) increase and node purity, highlighted GDP per capita as the most influential predictor, followed by inflation and several lagged terms (including GDP and trade variables). This ensemble approach effectively captured nonlinear relationships in the data, offering improved flexibility over linear models.

As presented in Table 4.6, the Random Forest achieved competitive performance with an MAE of 15.86, RMSE of 19.17, and R^2 of 0.67. While it outperformed the ARDL benchmark, its predictive accuracy remained slightly inferior to the LASSO model, suggesting a trade-off between interpretability (favoured by LASSO's sparse coefficients) and the ability to model complex interactions.

4.4.4 XGBoost

The XGBoost algorithm was implemented with hyperparameter tuning and early stopping and validated through 5-fold cross-validation. While XGBoost typically excels in large-scale datasets, its performance here proved suboptimal, likely due to the limited sample size and high dimensionality introduced by lagged terms.

As shown in 4.6, the model achieved an MAE of 25.53 and RMSE of 28.94, with a negligible R^2 of 0.03. This poor explanatory power suggests overfitting to training noise rather than capturing generalizable patterns. The results align with theoretical expectations: without sufficient data to constrain its complexity, XGBoost's gradient-boosted trees struggled to outperform simpler regularized (LASSO) or ensemble (Random Forest) alternatives in this context.

Clearly, LASSO regression outperforms all other models, including the traditional ARDL benchmark, in predicting household consumption. Its performance suggests that a smaller number of well-chosen lagged features can more effectively forecast short-term fluctuations in poverty proxies, particularly in settings with limited data.

4.5 Feature Importance Analysis

To uncover which variables most strongly influenced household consumption predictions (used here as a proxy for poverty), feature importance was examined using both permutation importance and SHAP values across the LASSO, Random Forest, and XGBoost models.

Across all models, GDP per capita growth (GDP_pc_diff) consistently emerged as the most influential feature. LASSO additionally identified L_ODA_diff_1 and L_Trade_diff_2 as important, while Random Forest emphasised L_FDI_diff_1, L_Trade_diff_1, and GCF_diff. XGBoost similarly ranked ODA_diff and LGCF_diff_2 highly. These results, detailed in Tables C.1, C.2 and C.3 and Figures C.1,C.3 and C.3 in the Appendix, indicate strong alignment between economic growth and household welfare outcomes.

SHAP value plots (Appendix C.1 to C.6) further confirmed the directional influence of these features. In particular, GDP_pc_diff showed a consistently strong positive contribution to predicted household consumption. Other variables—such as aid flows, trade openness, and investment—had more nuanced effects, varying by observation.

Overall, the results reinforce the centrality of economic growth while highlighting the lagged effects of aid and investment in shaping consumption trends over time.

4.6 Simulation Results

This section presents the results of the simulation study described in Chapter 3. The objective of the simulation was to evaluate the predictive performance and robustness of machine learning models, LASSO regression, Random Forest, and XGBoost, under systematically varied data conditions. These included changes in sample size, noise level, and the functional form of the data-generating process (DGP). Each condition was designed to reflect realistic challenges encountered in macroeconomic modelling, such as small samples, high variance, and potential nonlinearity in predictor-response relationships.

The simulation was conducted on synthetic datasets for which the true DGP was known, allowing performance to be evaluated using out-of-sample predictive metrics: Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and the coefficient of determination (R^2). The ARDL model was excluded from this section, as it was fitted on the full data and is not suited for predictive validation in simulated settings where generalisation is the focus.

Table 4.7, Table 4.8, and Table 4.9 present the results for each simulation scenario.

Table 4.7: Simulation Results: Varying Sample Sizes

Model	**MAE**	**RMSE**	**R²**
n = 30			
LASSO	1.615	1.690	0.031
Random Forest	1.332	1.576	0.192
XGBoost	1.760	1.990	0.025
n = 50			
LASSO	0.832	1.151	0.508
Random Forest	1.043	1.294	0.242
XGBoost	1.391	1.717	0.140
n = 200			
LASSO	1.096	1.299	0.474
Random Forest	1.317	1.572	0.222
XGBoost	1.310	1.583	0.253
n = 10,000			
LASSO	0.938	1.174	0.986
Random Forest	1.055	1.325	0.982
XGBoost	1.101	1.393	0.980

Interpretation

Across all simulation scenarios, LASSO generally outperformed other models in low-noise, linear settings, confirming its suitability for sparse, structured data. However, under nonlinear DGPs or high-noise conditions, tree-based methods—especially XGBoost—offered improved robustness and more stable predictive accuracy.

The impact of sample size was also evident: larger datasets led to improved R^2 and lower RMSE across all models, consistent with the theoretical expectation that model performance improves with more information. In particular, all models showed near-perfect fit when trained on 10,000 observations.

These findings reinforce the value of combining econometric and machine learning methods in applied development contexts, as each model type excels under different data regimes.

Table 4.8: Simulation Results: Varying Noise Levels

Model	MAE	RMSE	R^2
Noise SD = 0.5			
LASSO	0.545	0.727	0.746
Random Forest	0.686	0.833	0.581
XGBoost	0.709	0.924	0.577
Noise SD = 1.0			
LASSO	1.096	1.299	0.474
Random Forest	1.317	1.572	0.222
XGBoost	1.310	1.583	0.253
Noise SD = 2.0			
LASSO	1.382	1.754	0.469
Random Forest	1.648	2.022	0.269
XGBoost	1.777	2.195	0.159
Noise SD = 4.0			
LASSO	3.083	3.855	0.100
Random Forest	3.042	3.811	0.098
XGBoost	3.119	3.843	0.123

Table 4.9: Simulation Results: Linearity vs. Nonlinearity

Model	MAE	RMSE	R^2
Linear DGP			
LASSO	1.096	1.299	0.474
Random Forest	1.317	1.572	0.222
XGBoost	1.310	1.583	0.253
Nonlinear DGP			
LASSO	0.826	1.005	0.772
Random Forest	1.018	1.267	0.634
XGBoost	0.977	1.215	0.657

Chapter 5: Conclusion and Recommendation

5.1 Conclusion

This chapter synthesizes the key insights from our comparative analysis of econometric and machine learning approaches to modeling household consumption in Kenya. Building on our empirical results, we examine three critical dimensions: (1) the relative strengths of ARDL and ML methods in capturing consumption dynamics, (2) the policy implications of our identified determinants, and (3) methodological lessons for development economics research.

Our findings demonstrate how traditional time-series analysis and modern predictive modeling can complement each other, ARDL providing structural validation of long-run relationships while ML techniques offer superior short-term forecasting accuracy. This synthesis bridges our empirical results with their broader implications for both economic research and poverty monitoring.

Below are a summary of Key results:

5.1.1 Model Performance

The ARDL model, while grounded in econometric theory and designed to handle lagged dependencies and cointegration, performed poorly in out-of-sample prediction with an R^2 of only 0.0008. This suggests that its linear structure and rigid assumptions may have constrained its ability to capture the full complexity of the data-generating process.

Machine learning models demonstrated significantly superior predictive performance, particularly LASSO ($R^2 \approx 0.78$) and Random Forest ($R^2 \approx 0.66$). These models automatically selected relevant variables and adapted them to nonlinear interactions, which are likely present in economic behaviour over time.

5.1.2 Feature Relevance

Across models, GDP per capita growth (GDP_pc_diff) consistently emerged as the most statistically significant predictor of household consumption. Lagged values of ODA (official development assistance), trade, FDI, and gross capital formation (GCF) were also frequently highlighted by permutation importance scores and SHAP value visualisations. These findings confirm the importance of growth-linked and external-sector indicators in shaping consumption behaviour.

Interestingly, variables such as inflation and population (Box-Cox transformed) appeared less consistently, and the population was ultimately dropped due to unstable stationarity behaviour. However, its effects are partially captured in per capita metrics.

5.1.3 Econometric Perspective: Short-Run vs Long-Run Effects

The ARDL and Error Correction Model (ECM) framework provides critical structural insights that machine learning approaches could not replicate. While ML models achieved superior predictive accuracy, they cannot distinguish short-term fluctuations from long-term equilibrium relationships, a defining strength of the ARDL-ECM methodology. Bounds testing confirmed cointegration between household consumption and key macroeconomic variables, indicating a stable long-run equilibrium despite short-term volatility.

The ECM results underscored three key findings. First, short-run dynamics revealed statistically significant impacts on household consumption from gross capital formation, GDP per capita growth, and lagged ODA. Second, the negative and highly significant error correction term (ECT) indicated a partial annual adjustment toward equilibrium, validating the system's stability. Third, the ARDL long-run coefficients reaffirmed GDP per capita as the most influential persistent determinant of consumption.

Ultimately, while ML models excelled in forecasting, the ARDL-ECM framework delivered causal interpretation and dynamic decomposition, disentangling transient shocks from enduring economic relationships, an analytical frontier beyond pure machine learning techniques.

5.2 Recommendations

This study demonstrates the complementary value of classical econometric and machine learning approaches in economic modelling. The ARDL framework remains indispensable for structural analysis and cointegration testing, while ML models (particularly regularized approaches like LASSO) provide superior predictive accuracy and flexibility in capturing complex relationships.

The choice between methodologies should be context-dependent: ARDL offers theoretical validation when data is limited, but the structure is known, while ML excels when forecasting accuracy is paramount. Future research directions could develop (1) hybrid ARDL-ML frameworks to bridge explanatory and predictive goals, (2) Bayesian or bootstrapped ML methods for improved uncertainty quantification in small samples, and (3) multivariate extensions to enhance generalizability across economic contexts.

Ultimately, this work establishes that household consumption, as a key poverty indicator, can be modelled with high accuracy when rigorous model selection, appropriate lag structures, and the complementary strengths of econometric and machine learning approaches are combined.

5.3 Study Limitations

One major challenge was the limited data availability, as macroeconomic indicators are typically reported annually. With only 47 observations, the data volume was small by machine learning standards, which often require large sample sizes to exploit their capacity for pattern discovery fully.

This limitation likely affected model stability and generalizability, particularly for deeper models such as XGBoost, which underperformed relative to LASSO and Random Forest in this context.

Moreover, the small sample size constrained the application of more sophisticated time series techniques and made it difficult to perform rigorous time-series cross-validation. While care was taken to preserve temporal ordering and avoid leakage, the model evaluation strategy remained conservative.

While machine learning models demonstrated superior predictive performance, the ARDL framework provided indispensable structural insights through cointegration analysis, bounds testing, and long-run equilibrium modelling, capabilities inherently absent in pure ML approaches. This dichotomy suggests significant potential for developing hybrid methodologies that combine the strengths of both paradigms.

A synergistic approach could leverage ML's predictive advantages (automated variable selection, nonlinear pattern recognition) while preserving ARDL's theoretical rigour (error correction mechanisms, long-run causality testing). For instance, ML techniques could first identify relevant predictors and functional forms, followed by ARDL specification to validate economic relationships and test for cointegration. Such integration would bridge the explanatory-predictive divide, offering a more comprehensive analytical framework for economic time series analysis.

Bibliography

Akobeng, E. (2020). Harnessing foreign aid for the poor: role of institutional democracy. *Journal of Economic Studies*, 47(7):1689–1710.

Amin, S. (1972). Underdevelopment and dependence in black africa—origins and contemporary forms. *The Journal of Modern African Studies*, 10(4):503–524.

Anetor, F. O., Esho, E., and Verhoef, G. (2020). The impact of foreign direct investment, foreign aid and trade on poverty reduction: Evidence from Sub-Saharan African countries. *Cogent Economics amp; Finance*, 8(1):1737347.

Asongu, S. A., Nwachukwu, J., and Biekpe, N. (2019). Foreign aid, terrorism and growth: Conditional evidence from quantile regression. *Annals of public and cooperative economics*, 90(3):457–486.

Baltagi, B. H. and Baltagi, B. H. (2008). *Econometric analysis of panel data*, volume 4. Springer.

Banda, L. G. (2020). Limitations of the use of modernization theory in formulating and implementing development policies in africa–the case of tanzania and malawi. *Journal of Development Economics, Forthcoming*.

Baran, P. (1973). 1957. the political economy of growth. *New York: Monthly Review*.

Bienefeld, M. (1988). Dependency theory and the political economy of africa's crisis. *Review of African Political Economy*, 15(43):68–87.

Boone, P. (1996). Politics and the effectiveness of foreign aid. *European economic review*, 40(2):289–329.

Borter, D. K. (2017). The aid effectiveness architecture in africa: An analysis of aid structures in kenya's agriculture sector. *International Journal of African Development*, 4(2):4.

Bräutigam, D. A. and Knack, S. (2004). Foreign aid, institutions, and governance in sub-saharan africa. *Economic development and cultural change*, 52(2):255–285.

Chenery, H. B. (1967). Foreign assistance and economic development. In *Capital movements and economic development*, pages 268–292. Springer.

Chiamogu, A. P., Nweke, E., and Chukwudi, M. C. (2012). Making poverty a history in africa: The role of the g8. *International Journal of Sustainable Development & World Policy*, 1(1):29–59.

Dos, S. T. et al. (1970). The structure of dependence. *American economic review*, 60(2):231.

Durokifa, A. A. and Ijeoma, E. C. (2018). Neo-colonialism and Millennium Development Goals (MDGs) in Africa: A blend of an old wine in a new bottle. *African Journal of Science, Technology, Innovation and Development*, 10(3):355–366.

Edward, S. P. and Karamuriro, H. T. (2020). Effects of foreign aid inflow on economic growth of uganda: autoregressive distributed lag modelling (ardl). *International Journal of Sustainability Management and Information Technologies*, 6(2):36–49.

Farah, A. M., Önder, M., and Ayhan, E. (2018). How foreign aid affect developing countries: The case of ethiopia. *Avrasya Etüdleri*, 53(1):7–38.

Gitaru, K. (2015). Impact of foreign aid on economic growth.

Gunder, F. A. (1989). The development of underdevelopment. *Monthly Review*, 41(2):37–51.

Hamilton, J. D. (2020). *Time series analysis*. Princeton university press.

Haruna, M. A., Hassan, S. B., and Ahmad, H. S. (2023). How responsive is the poverty to the foreign direct investment inflows in nigeria? evidence from linear and non-linear ardl. *International Journal of Social Economics*, 50(1):73–96.

Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.

Iwegbu, O. and Dauda, R. O. (2022). Effectiveness of foreign aid in poverty reduction in africa: The role of fiscal policy. *CBN Journal of Applied Statistics*, 13(1):55–92.

Kabonga, I. (2016). Dependency theory and donor aid: a critical analysis. *Africanus: Journal of Development Studies*, 46(2):29–39.

Khainga, D. and Oduor, J. (2009). Effectiveness of foreign aid on poverty reduction in kenya. Technical Report Working Paper No. 34, Global Development Network, Nairobi, Kenya.

Kohnert, D. (2025). Impact of Trump 2.0 on Sub-Saharan Africa. Working Paper 123754, Munich Personal RePEc Archive (MPRA). MPRA Paper No. 123754.

Lopes, C. (2024). *The Self-deception Trap: Exploring the Economic Dimensions of Charity Dependency Within Africa-Europe Relations*. Springer Nature.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Mahembe, E. and Odhiambo, N. M. (2019). Foreign aid, poverty and economic growth in developing countries: A dynamic panel data causality analysis. *Cogent Economics & Finance*, 7(1):1626321.

Mahembe, E. and Odhiambo, N. M. (2021). Does foreign aid reduce poverty? a dynamic panel data analysis for sub-saharan african countries. *The Journal of Economic Inequality*, 19:875–893.

M'amanja, D. and Morrissey, O. (2006). Foreign aid, investment and economic growth in kenya: a time series approach. Technical report, CREDIT research paper.

Maruta, A. A., Banerjee, R., and Cavoli, T. (2020). Foreign aid, institutional quality and economic growth: Evidence from the developing world. *Economic Modelling*, 89:444–463.

Mohameda, M. R. and Mzeea, S. S. (2017). Foreign aid and human development: A quantile regression approach. *International Journal of Economics, Management and Accounting*, 25(1):27–41.

Moyo, D. (2010). *Dead Aid: Why Aid Is Not Working and How There Is a Better Way for Africa*. Farrar, Straus and Giroux, New York.

Narayan, P. K. (2005). The saving and investment nexus for china: evidence from cointegration tests. *Applied economics*, 37(17):1979–1990.

Ndikumana, L. and Pickbourn, L. (2017). The impact of foreign aid allocation on access to social services in sub-saharan africa: The case of water and sanitation. *World Development*, 90:104–114.

Nkoro, E., Uko, A. K., et al. (2016). Autoregressive distributed lag (ardl) cointegration technique: application and interpretation. *Journal of Statistical and Econometric methods*, 5(4):63–91.

Ogbodo, J. C. and Attamah, N. (2019). Analysis of the impact of foreign aid on poverty reduction in nigeria, 1981-2017. *ESUT JOURNAL OF SOCIAL SCIENCES*, 4(1).

Parsons, T. (1960). *Structure and process in modern societies*. Free Press of Glencoe.

Pesaran, M. and Smith, R. (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics*, 68(1):79–113.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3):289–326.

Phillips, P. C. and Hansen, B. E. (1990). Statistical inference in instrumental variables regression with i (1) processes. *The review of economic studies*, 57(1):99–125.

Prebisch, R. (1962). The economic development of latin america and its principal problems. Report E/CN.12/89/Rev.1, United Nations.

Ravallion, M. and Chen, S. (2017). Welfare-consistent global poverty measures. Technical report, National bureau of economic research.

Rostow, W. W. (1990). *The stages of economic growth: A non-communist manifesto*. Cambridge university press.

SULE, A. (2019). Reassessing the efficacy of foreign aid and grants in poverty reduction in nigeria. *Asian Economic and Financial Review*, 9(4):450–460.

Tang, K.-B., Bundhoo, D., et al. (2017). Foreign aid and economic growth in developing countries: Evidence from sub-saharan africa. *Theoretical Economics Letters*, 7(05):1473.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Ugwuanyi, U. B., Ezeaku, H. C., and Ibe, I. G. (2017). The impact of official aid on poverty reduction: Empirical evidence from nigeria (1981-2014) using the ardl and bound test approach. *European Journal of Sustainable Development*, 6(2):111–111.

Wallerstein, I. (1974). *The Modern World-System I: Capitalist Agriculture and the Origins of the European World-Economy in the Sixteenth Century*. Academic Press, New York, 1st edition. With a New Prologue in later editions.

Appendix A: Similarity Report

Statistical and Machine Learning Approaches to Assessing Foreign Aid Effectiveness in Kenya An ARDL Framework.pdf

by Noel Jepkosgei Rutto

Submission date: 02-Apr-2025 08:00AM (UTC+0300)

Submission ID: 2632700040

File name:
36057_Noel_Jepkosgei_Rutto_Statistical_and_Machine_Learning_Approaches_to_Assessing_Foreign_Aid_Effectiveness_in_Ke_226980732.pdf (568.14K)

Word count: 12461

Character count: 70276

Statistical and Machine Learning Approaches to Assessing Foreign Aid Effectiveness in Kenya An ARDL Framework.pdf

ORIGINALITY REPORT

3%	4%	3%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	su-plus.strathmore.edu Internet Source	2%
2	Submitted to Strathmore University Student Paper	1%
3	www.pressacademia.org Internet Source	1%

Exclude quotes	Off	Exclude matches	< 1%
Exclude bibliography	Off		

Appendix B: Ethical Clearance Confirmation

Strathmore UNIVERSITY

1st April 2025

Mrs Rutto Noel,
noel.rutto@strathmore.edu

Dear Mrs Rutto,

RE: Statistical and Machine Learning Approaches to Assessing Foreign Aid Effectiveness in Kenya: An ARDL Framework

This is to inform you that SU-ISERC has reviewed and **approved** your above **SU-masters** proposal. Your application reference number is **SU-ISERC2809/25.** The approval period is from **1st April 2025 to 31st March 2026.**

This approval is subject to compliance with the following requirements:

i. Only approved documents including (informed consents, study instruments, MTA) will be used.
ii. All changes including (amendments, deviations, and violations) are submitted for review and approval by SU-ISERC.
iii. Death and life-threatening problems and serious adverse events or unexpected adverse events whether related or unrelated to the study must be reported to SU-ISERC within 72 hours of notification.
iv. Any changes anticipated or otherwise that may increase the risks or affected safety or welfare of study participants and others or affect the integrity of the research must be reported to SU-ISERC within 72 hours.
v. Clearance for the export of biological specimens must be obtained from relevant institutions.
vi. Submission of a request for renewal of approval at least 60 days prior to the expiry of the approval period. Attach a comprehensive progress report to support the renewal.
vii. Submission of an executive summary report within 90 days of completion of the study to SU-ISERC.

Before commencing your study, you will be expected to obtain a research license from National Commission for Science, Technology, and Innovation (NACOSTI) https://research-portal.nacosti.go.ke/ and obtain other clearances needed.

Yours sincerely,

Mr Ambrose Rachier,
Chairperson; SU-ISERC

Appendix C: Simulation and Model Code

The following listing reproduces the R Markdown file used for simulation and modelling.

This R Markdown file contains the simulation code used to evaluate the performance of econometric and machine learning models in predicting poverty-related outcomes. It reflects the modelling strategies applied in the main thesis analysis, with adjustments for a synthetic data context.

> **Note:** The full data cleaning and empirical analysis code applied to the actual WDI dataset is available upon request.

Load Required Libraries

The following packages are required to run the simulations and analyses in this file.

```
library(ARDL)
library(caret)
library(glmnet)
library(randomForest)
library(xgboost)
library(dplyr)
library(ggplot2)
library(knitr)
library(tidyr)
```

---

1. Simulated Data Generation

This section defines the function used to generate synthetic datasets for simulation. The datasets include controlled linear and nonlinear relationships, with configurable sample size and noise levels.

```
simulate_data <- function(n, noise_sd = 1, nonlinear = FALSE, seed = 42) {
  set.seed(seed)
  time <- 1:n
  ODA <- cumsum(rnorm(n, 0, 0.2)) + 2
  FDI <- cumsum(rnorm(n, 0, 0.3)) + 1.5
  GDPpc <- cumsum(rnorm(n, 0, 0.5)) + 5
  Inflation <- rnorm(n, mean = 10, sd = 3)
  GCF <- cumsum(rnorm(n, 0, 0.2)) + 3
  Trade <- cumsum(rnorm(n, 0, 0.25)) + 4
  HhC <- numeric(n)
  HhC[1:2] <- 0

  for (t in 3:n) {
    nonlinear_term <- if (nonlinear) 0.1 * (GDPpc[t - 1]^2) else 0
    HhC[t] <- 0.4 * HhC[t - 1] - 0.2 * HhC[t - 2] +
      0.6 * ODA[t - 1] +
      0.3 * FDI[t - 2] +
      0.5 * GDPpc[t - 1] -
      0.4 * Inflation[t - 1] +
```

```
      0.2 * GCF[t - 2] +
      0.1 * Trade[t - 1] +
      nonlinear_term +
      rnorm(1, mean = 0, sd = noise_sd)
  }

  data.frame(ODA, FDI, GDPpc, Inflation, GCF, Trade, HhC)
}
```

2. Model Evaluation (LASSO, RF, XGBoost)

This chunk runs the models used in the thesis for comparison. Evaluation is done using a standard 80/20 split.

```
evaluate_models <- function(df) {
  df_lagged <- df %>%
    mutate(across(c(ODA, FDI, GDPpc, Inflation, GCF, Trade), ~lag(.x, 1),
.names = "{.col}_lag1")) %>%
    na.omit()

  train_index <- createDataPartition(df_lagged$HhC, p = 0.8, list = FALSE)
  train <- df_lagged[train_index, ]
  test <- df_lagged[-train_index, ]

  ardl_rmse <- ardl_mae <- ardl_r2 <- NA
  tryCatch({
    df_ts <- ts(df[, c("HhC", "ODA", "FDI", "GDPpc", "Inflation", "GCF",
"Trade")])
    ardl_fit <- auto_ardl(HhC ~ ODA + FDI + GDPpc + Inflation + GCF + Trade,
data = df_ts, max_order = 2)
    ardl_pred <- predict(ardl_fit$best_model)
    actual <- df_ts[, "HhC"][length(df_ts[, "HhC"]) - length(ardl_pred) +
1:length(df_ts[, "HhC"])]
    ardl_rmse <- sqrt(mean((actual - ardl_pred)^2))
    ardl_mae <- mean(abs(actual - ardl_pred))
    ardl_r2 <- cor(actual, ardl_pred)^2
  }, error = function(e) {})

  x_train <- model.matrix(HhC ~ ., data = train[, grep("_lag1|HhC",
names(train))])[, -1]
  y_train <- train$HhC
  x_test <- model.matrix(HhC ~ ., data = test[, grep("_lag1|HhC",
names(test))])[, -1]
  y_test <- test$HhC

  lasso <- cv.glmnet(x_train, y_train, alpha = 1)
  lasso_pred <- predict(lasso, s = "lambda.min", newx = x_test)
  lasso_rmse <- sqrt(mean((y_test - lasso_pred)^2))
  lasso_mae <- mean(abs(y_test - lasso_pred))
```

```
  lasso_r2 <- cor(y_test, lasso_pred)^2

  rf <- randomForest(x = x_train, y = y_train)
  rf_pred <- predict(rf, x_test)
  rf_rmse <- sqrt(mean((y_test - rf_pred)^2))
  rf_mae <- mean(abs(y_test - rf_pred))
  rf_r2 <- cor(y_test, rf_pred)^2

  dtrain <- xgb.DMatrix(data = x_train, label = y_train)
  dtest <- xgb.DMatrix(data = x_test, label = y_test)
  xgb_fit <- xgboost(data = dtrain, objective = "reg:squarederror", nrounds =
50, verbose = 0)
  xgb_pred <- predict(xgb_fit, dtest)
  xgb_rmse <- sqrt(mean((y_test - xgb_pred)^2))
  xgb_mae <- mean(abs(y_test - xgb_pred))
  xgb_r2 <- cor(y_test, xgb_pred)^2

  data.frame(
    Model = c("ARDL", "LASSO", "Random Forest", "XGBoost"),
    RMSE = c(ardl_rmse, lasso_rmse, rf_rmse, xgb_rmse),
    MAE = c(ardl_mae, lasso_mae, rf_mae, xgb_mae),
    R2 = c(ardl_r2, lasso_r2, rf_r2, xgb_r2)
  )
}
```

Section 1: Varying Sample Sizes

```
sizes <- c(30, 50, 200, 10000)
results_size <- lapply(sizes, function(n) evaluate_models(simulate_data(n =
n)))
names(results_size) <- paste0("n = ", sizes)
df_size <- do.call(rbind, Map(cbind, Dataset = names(results_size),
results_size))

#kable(df_size, digits = 3)
```

Section 2: Varying Noise Levels

```
noises <- c(0.5, 1, 2, 4)
results_noise <- lapply(noises, function(sd) evaluate_models(simulate_data(n
= 200, noise_sd = sd)))
names(results_noise) <- paste0("noise = ", noises)
df_noise <- do.call(rbind, Map(cbind, Dataset = names(results_noise),
results_noise))

#kable(df_noise, digits = 3)
```

Section 3: Linearity vs Nonlinearity

```
results_nl <- list(
  "Linear" = evaluate_models(simulate_data(n = 200, noise_sd = 1, nonlinear =
FALSE)),
  "Nonlinear" = evaluate_models(simulate_data(n = 200, noise_sd = 1,
nonlinear = TRUE))
)
df_nl <- do.call(rbind, Map(cbind, Dataset = names(results_nl), results_nl))

#kable(df_nl, digits = 3)
```

Plot: RMSE by Scenario

```
df_size$Scenario <- "Sample Size"
df_noise$Scenario <- "Noise"
df_nl$Scenario <- "Nonlinearity"

all_perf <- bind_rows(df_size, df_noise, df_nl)

ggplot(all_perf, aes(x = Dataset, y = RMSE, fill = Model)) +
  geom_bar(stat = "identity", position = "dodge") +
  facet_wrap(~Scenario, scales = "free_x") +
  labs(title = "RMSE Comparison by Model and Simulation Scenario",
       x = "Scenario", y = "RMSE") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```

3. ARDL Estimation

While the ARDL model was not used for forecasting in the simulation, I include the structure here to reflect how it was applied in the empirical part of the thesis.

```
# ARDL estimation based on the dynardl package
# library(dynardl)
# df_ts <- ts(sim_data, frequency = 1)
# ardl_fit <- auto_ardl(Y ~ X1 + X2 + X3, data = df_ts, max_order = 2)
# summary(ardl_fit$best_model)
```

4. Stationarity Checks

Unit root testing is a key step in time-series analysis. Below is an example of how I applied the Augmented Dickey-Fuller (ADF) test in the empirical component.

```
# ADF test for unit root
# adf_test <- ur.df(sim_data$Y, type = "drift", selectlags = "AIC")
# summary(adf_test)
```

5. Model Diagnostics

This section includes basic diagnostics used to check residual assumptions in the fitted models.

```
# Residual diagnostics on ARDL or linear model
# bgtest(ardl_fit$best_model)                # Serial correlation
# bptest(ardl_fit$best_model)                # Heteroskedasticity
# jarque.bera.test(resid(ardl_fit$best_model))  # Normality
```

Additional Results Tables & Plots

Table C.1: LASSO Permutation Importance

Feature	Importance.05	Importance	Importance.95	Permutation Error
GDP_pc_diff	1.5036001	2.6170549	3.098914	30.95977
L_ODA_diff_1	0.9953382	1.0531857	1.073460	12.45919
Year	1.0000000	1.0000000	1.000000	11.83000
FDI_diff	1.0000000	1.0000000	1.000000	11.83000
ODA_diff	1.0000000	1.0000000	1.000000	11.83000
Trade_diff	1.0000000	1.0000000	1.000000	11.83000
GCF_diff	1.0000000	1.0000000	1.000000	11.83000
Pop_boxcox	1.0000000	1.0000000	1.000000	11.83000
L_Hh_C_diff_1	1.0000000	1.0000000	1.000000	11.83000
L_Hh_C_diff_2	1.0000000	1.0000000	1.000000	11.83000
L_ODA_diff_2	1.0000000	1.0000000	1.000000	11.83000
L_Trade_diff_1	1.0000000	1.0000000	1.000000	11.83000
L_GCF_diff_1	1.0000000	1.0000000	1.000000	11.83000
L_GCF_diff_2	1.0000000	1.0000000	1.000000	11.83000
L_Infl_1	1.0000000	1.0000000	1.000000	11.83000
L_Infl_2	1.0000000	1.0000000	1.000000	11.83000
L_GDP_pc_diff_1	1.0000000	1.0000000	1.000000	11.83000
L_GDP_pc_diff_2	1.0000000	1.0000000	1.000000	11.83000
L_FDI_diff_1	1.0000000	1.0000000	1.000000	11.83000
L_FDI_diff_2	1.0000000	1.0000000	1.000000	11.83000
Infl	0.9993352	0.9999117	1.000291	11.82896
L_Trade_diff_2	0.9317982	0.9375371	1.027801	11.09107

Table C.2: Random Forest Permutation Importance

Feature	**Importance.05**	**Importance**	**Importance.95**	**Permutation Error**
GDP_pc_diff	1.4023476	1.4829541	1.5229815	28.43447
L_ODA_diff_1	1.0029244	1.0254038	1.0305428	19.66130
L_Trade_diff_1	0.9917088	1.0152670	1.0212527	19.46694
L_ODA_diff_2	0.9904278	1.0148104	1.0390431	19.45818
ODA_diff	0.9920374	1.0133647	1.0216572	19.43046
FDI_diff	1.0064417	1.0086455	1.0093756	19.33998
L_Trade_diff_2	0.9970842	1.0066190	1.0361580	19.30112
L_FDI_diff_1	0.9930998	1.0052659	1.0096802	19.27518
L_Infl_1	1.0018362	1.0022368	1.0062345	19.21710
GCF_diff	0.9515188	1.0010304	1.0337589	19.19396
L_Infl_2	0.9973982	1.0001936	1.0054287	19.17792
Year	1.0000000	1.0000000	1.0000000	19.17421
Pop_boxcox	1.0000000	1.0000000	1.0000000	19.17421
L_GCF_diff_2	0.9884609	0.9999183	1.0158773	19.17264
Infl	0.9909838	0.9993381	1.0057130	19.16151
L_GDP_pc_diff_2	0.9969014	0.9989797	1.0093362	19.15464
L_Hh_C_diff_2	0.9907799	0.9981186	1.0022584	19.13813
Trade_diff	0.9826539	0.9962401	1.0026114	19.10211
L_Hh_C_diff_1	0.9717752	0.9858193	0.9965483	18.90230
L_GDP_pc_diff_1	0.9659262	0.9837263	1.0049328	18.86217
L_FDI_diff_2	0.9658124	0.9813482	0.9931808	18.81657
L_GCF_diff_1	0.9613796	0.9647231	1.0158949	18.49780

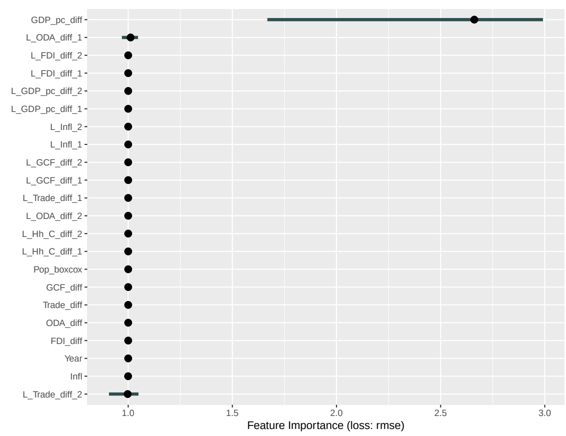

Figure C.1: Permutation importance for the LASSO model.

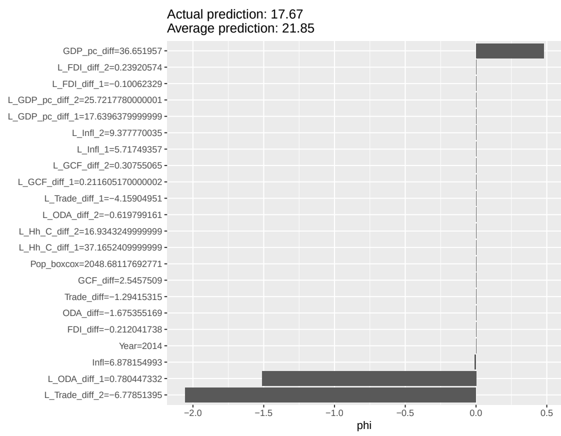

Figure C.2: SHAP values for the LASSO model.

Table C.3: XGBoost Permutation Importance

Feature	Importance.05	Importance	Importance.95	Permutation Error
GDP_pc_diff	1.0409173	1.0690823	1.0735951	30.93825
L_GCF_diff_2	0.9919404	1.0059881	1.0161637	29.11236
ODA_diff	1.0000000	1.0042413	1.0065434	29.06181
L_Trade_diff_1	1.0000000	1.0008297	1.0158488	28.96308
Year	1.0000000	1.0000000	1.0000000	28.93907
Infl	1.0000000	1.0000000	1.0000000	28.93907
FDI_diff	1.0000000	1.0000000	1.0000000	28.93907
GCF_diff	1.0000000	1.0000000	1.0000000	28.93907
Pop_boxcox	1.0000000	1.0000000	1.0000000	28.93907
L_Hh_C_diff_1	1.0000000	1.0000000	1.0000000	28.93907
L_ODA_diff_2	1.0000000	1.0000000	1.0000000	28.93907
L_Trade_diff_2	1.0000000	1.0000000	1.0000000	28.93907
L_Infl_1	0.9844907	1.0000000	1.0078843	28.93907
L_Infl_2	1.0000000	1.0000000	1.0000000	28.93907
L_GDP_pc_diff_2	1.0000000	1.0000000	1.0000000	28.93907
L_GCF_diff_1	0.9875646	0.9984991	1.0020872	28.89563
L_ODA_diff_1	0.9808822	0.9933966	1.0088017	28.74797
Trade_diff	0.9930990	0.9930990	1.0000000	28.73936
L_Hh_C_diff_2	0.9901829	0.9928485	0.9990711	28.73211
L_GDP_pc_diff_1	0.9602820	0.9733833	0.9955618	28.16881
L_FDI_diff_2	0.9671399	0.9671399	1.0000000	27.98813
L_FDI_diff_1	0.9490416	0.9601444	0.9769401	27.78569

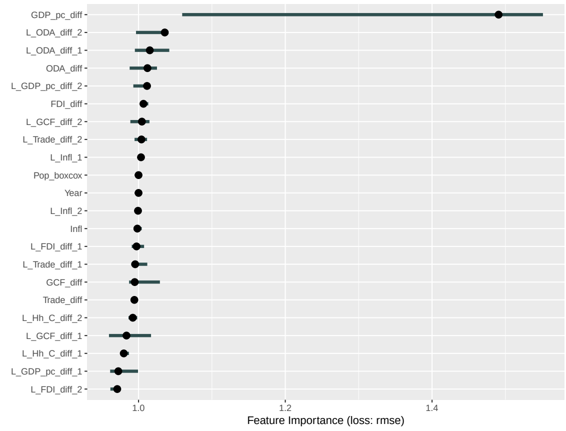

Figure C.3: Permutation importance for the Random Forest model.

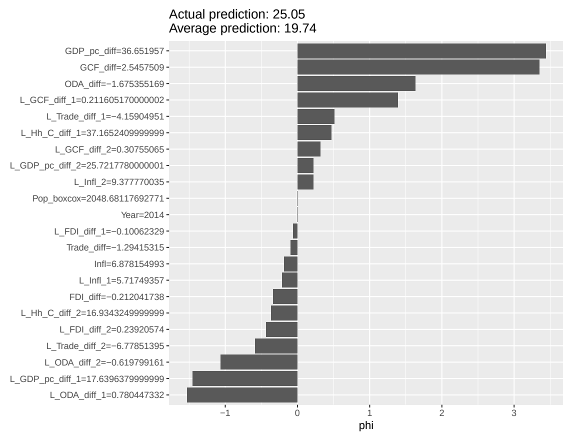

Figure C.4: SHAP values for the Random Forest model.

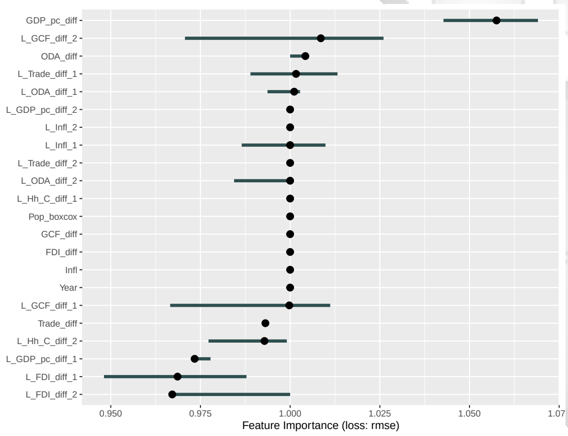

Figure C.5: Permutation importance for the XGBoost model.

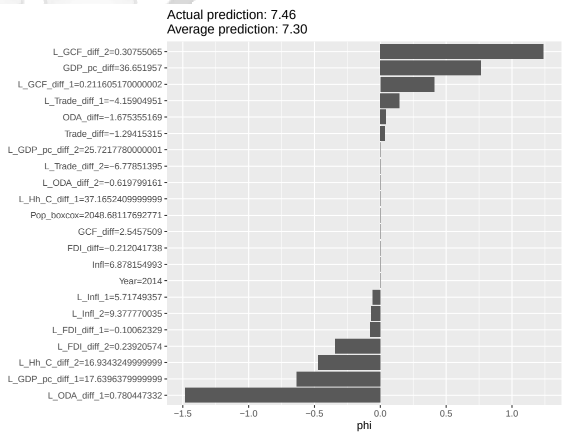

Figure C.6: SHAP values for the XGBoost model.