EMPLOYING BAYESIAN VECTOR AUTOREGRESSION METHOD AS AN ALTERNATIVE TECHNIQUE FOR FORECASTING TAX REVENUE IN ZAMBIA

By

RUSSELL CHILESHE

A dissertation submitted to the University of Zambia in partial fulfillment of the requirements for the degree of Master of Science in Statistics

THE UNIVERSITY OF ZAMBIA
LUSAKA

2025

Declaration

I, **Russell Chileshe**, hereby declare that this dissertation represents my own work (except where otherwise indicated) and that it has not previously been submitted for a degree, diploma, or other qualification at this or another university.

Signature: .... Date: ....

Certificate of Approval

This dissertation of Russell Chileshe has been approved as fulfilling the requirements or partially fulfilling the requirements for the award of the Master of Science in Statistics degree by the University of Zambia.

....	....	....
Examiner I	Signature	Date
....	....	....
Examiner II	Signature	Date
....	....	....
Examiner III	Signature	Date
....	....	....
Chairperson Board of Examiners	Signature	Date
....	....	....
Supervisor	Signature	Date

Abstract

Tax revenue forecasting is of critical importance for a government in ensuring adequacy and stability in tax and expenditure policies, as it also contributes to the budget and strategic planning of a country. Henceforth, several tax types need to be projected for the specific fiscal year using models that are statistically sound and with a smaller margin of error. This study explored the Bayesian Vector Auto-Regression (BVAR) method as an alternative technique for forecasting tax revenue in Zambia. The study forecasted Corporate Income Tax (CIT), Personal Income Tax (PIT),Value-Added Tax (VAT) and Total tax revenue (TTR) using Bayesian vector autoregression (BVAR) models with quarterly data from 2010Q1 to 2023Q4 and the results were compared with Autoregressive Moving Averages (ARIMA) and Error, Trend, Seasonal (ETS). Based on the RMSE, the results of BVAR model with the Normal-Wishart prior was the best model for measuring the accuracy forecasting of the CIT, PIT, VAT and TTR. In most cases, ETS is the second best after BVAR and was superior to ARIMA. The results suggest that the BVAR model is best suited to be used to forecast tax revenues in Zambia with the ETS model as an alternative. The study suggests that the BVAR forecasting methods may also be extended to other smaller taxes to investigate whether they will fit these taxes accurately as it does for major taxes.

Key Words: ARIMA,BVAR, ETS, RMSE, Personal income taxes, Corporate Income tax, Economic growth, Gross fixed capital formation, Inflation rate, Stock market performance, value added taxes, Zambia.

Dedication

This dissertation is dedicated to my beloved family and friends, particularly my mother, Ms. Lydia Chibwe for her unwavering support towards my academic journey. She has been there for me and encouraged me throughout my academic life. I would not have been able to carry out this work without her support.

Acknowledgment

First of all, I would like to thank my employers, Zambia Airports Corporation Limited, for offering me a study leave to pursue this program. Furthermore, my sincere gratitude also goes to my supervisor, Dr. John Musonda, for his unstoppable guidance, advice, and admonition throughout the period of this study. Despite his numerous responsibilities, he was always available and ready to offer parental and professional advice, which culminated in the successful completion of this study. I thank you for your sacrifice and support that came in handy when I needed you. May the creator bless you without measure.

Furthermore, I also would like to acknowledge all my lecturers who taught me during the coursework of the master's program; Dr. K. Muzundu, Dr. M. Kikonko, Dr. V. Nawa and Mrs. S. Jain, for the knowledge they imparted to me that has been collectively useful in carrying out this dissertation. I also would like to thank the Department of Mathematics, Statistics and Actuarial Science who, through the Eastern African Universities Program (EAUMP) project, granted me the partial financial support.

I also feel very indebted to a lot of people, too many to mention, who contributed in one way or another towards making the writing of this dissertation a success. The Lord Almighty thanks me for the knowledge, ability, and strength given to me to be able to produce this piece of work.

Lusaka, February, 2025
Russell Chileshe

Contents

Declaration iii

Approval iv

Abstract iv

Acknowledgments vi

List of Symbols x

1 Introduction 1

1.1 Background 2
1.2 Statement of the Problem 3
1.3 Aim of the Study 3
1.4 Research Objectives 3
1.5 Research Questions 3
1.6 Significance of the Study 3
1.7 Scope of the study 4
1.8 Limitation of the Study 4
1.9 Organization of the Dissertation 4

2 Literature Review 6

2.1 Introduction 6
2.2 Taxes in Zambia 6
2.3 Types of taxes in Zambia 7
2.4 Personal Income Tax 8
2.5 Corporate Income Tax 9
2.6 Value Added Tax 9
2.7 Custom and excise duty 10
2.8 Tax evasion 10
2.9 The Tax Forecasting process in Zambia 11
2.10 Micro-Simulation 12
2.11 Macro-Econometric tax revenue models 13
2.12 Tax buoyancy approach 13
2.13 Box-Jenkins techniques 14
2.14 Revenue trending and extrapolation 14
2.15 Professional judgement 15

2.16 The Tax Forecasting process in other countries . . . 15
2.17 Studies comparing the performance of BVAR and other Methodologies . . . 16
2.18 Studies based on ARIMA Models . . . 19
2.19 Studies comparing ETS AND ARIMA Models . . . 19
2.20 Chapter Summary . . . 20

3 Methodology **21**
3.1 Introduction . . . 21
3.2 Bayesian vector autoregression . . . 21
3.3 Litterman/Minnesota prior . . . 24
3.4 Normal-Wishart prior . . . 26
3.5 Sims-Zha priors . . . 26
3.6 Selecting priors . . . 28
3.7 Error, Trend, Seasonal Methods . . . 29
3.8 Autoregressive integrated moving average . . . 30
3.9 Identification of the AR and MA process . . . 32
3.10 Ljung-Box Test . . . 33
3.11 Stationarity and unit root tests . . . 34
3.12 Dickey-Fuller test . . . 34
3.13 Phillips-Perron test . . . 35
3.14 Model selection . . . 36
3.15 Measures of accuracy for forecasting . . . 37
3.16 Data and statistical software . . . 38
3.17 Chapter Summary . . . 38

4 Presentation, Interpretation and Discussion of Findings **40**
4.1 Introduction . . . 40
4.2 Description of the tax variables . . . 40
4.3 Description of the Economic Variable . . . 42
4.4 Unit Root Tests . . . 43
4.5 ARIMA Model for CIT . . . 44
4.6 Diagnostic Tests of the ARIMA Model for CIT . . . 46
4.7 Forecasted ARIMA Model for CIT . . . 47
4.8 Error, Trend, Seasonal Models for CIT . . . 47
4.9 Forecasted ETS model for CIT . . . 48
4.10 Bayesian Vector Autoregressive Model for CIT . . . 49
4.11 Forecasted BVAR model for CIT . . . 49
4.12 Discussion of Corporate Income Tax Results . . . 50
4.13 ARIMA Model for PIT . . . 51
4.14 Diagnostic Tests of the ARIMA Model for PIT . . . 53

4.15 Forecasted ARIMA Model for PIT 55
4.16 Error, Trend, Seasonal Models for PIT 55
4.17 Forecasted ETS model for PIT 56
4.18 Bayesian Vector Autoregressive Model for PIT 56
4.19 Forecasted BVAR model for PIT 57
4.20 Discussion of Personal Income Tax Results 58
4.21 ARIMA Model for VAT 59
4.22 Diagnostic Tests of the ARIMA Model for VAT 60
4.23 Forecasted ARIMA Model for VAT 62
4.24 Error, Trend, Seasonal Models for VAT 62
4.25 Forecasted ETS model for VAT 63
4.26 Bayesian Vector Autoregressive Model for VAT 63
4.27 Forecasted BVAR model for VAT 64
4.28 Discussion of Value added Tax Results 65
4.29 ARIMA Model for TTR 66
4.30 Diagnostic Tests of the ARIMA Model for TTR 67
4.31 Forecasted ARIMA Model for TTR 69
4.32 Error, Trend, Seasonal Models for TTR 69
4.33 Forecasted ETS model for TTR 70
4.34 Bayesian Vector Autoregressive Model for TTR 70
4.35 Forecasted BVAR model for TTR 71
4.36 Discussion of Total Tax Revenue Results 72

5 Conclusion and Recommendations 74
5.1 Conclusion 74
5.2 Recommendations 74
5.3 Areas of Future Research 75

References 76

APPENDIX 78

A Unit Root Tests 79

B BVAR, ARIMA and ETS Models for CIT 87

C BVAR, ARIMA and ETS Models for PIT 95

D BVAR, ARIMA and ETS Models for VAT 103

E BVAR, ARIMA and ETS Models for TTR 110

List of Symbols

The following is a list of some symbols used in this dissertation and their meanings.

α	exponential smoothing parameter
β	exponential smoothing parameter
ω	degrees of freedom of the inverse Wishart
Σ	variance-covariance matrix
ε_t	white noise
ρ_τ	autocorrelation function of a AR process of order p
Q^*	Ljung-Box Q-statistic test
$P(\varphi)$	prior distribution of the possible values
$\pi(A_0)$	marginal distribution of A_0
$P\left(\varphi/Y\right)$	posterior distribution of φ , given the observed data Y
$\emptyset(\mu_0, T_0)$	normal density distribution with mean μ_0 and covariance T_0
$L\left(Y/\varphi, \Sigma\right)$	likelihood function of the Bayesian estimation
$P\left(\varphi, \Sigma/y\right)$	marginal posterior distributions conditional on the data

Chapter 1

Introduction

The Zambian government has the role of ensuring the safety and welfare of its citizens. However, for it to realize this role, it must levy taxes. Tax revenue is the crucial source of income for the government, and thus it is necessary to estimate its expenditure before it can be budgeted for its income to meet its commitments. The important issues in the design of comprehensive fiscal policy have been the precision of budget forecasts, mainly tax revenue forecasts (Nandi et al.[28]). Accurate tax revenue forecasting is essential in meeting the expenditure for the purpose of budgeting. Over the years, an important issue in the design of fiscal policy has been the precision of government budget forecasts, particularly those of tax revenues (Auerbach[1]).

In Zambia, the Minister of Finance and National Planning (MOFNP) as the head of the National Treasury is the one who sets goals for tax revenue. The responsibility of collecting revenues is assigned to the Zambia Revenue Authority (ZRA[39]), which is under the Ministry of Finance and National Planning. ZRA presents its revenue projections to a joint committee comprised of the National Treasury, Bank of Zambia, and the Parliamentary Committee known as the Planning and Budgeting Committee (PBC) in the Zambian Parliament is responsible for reviewing and analyzing the country's budget and revenue performance. It is within this technical committee that national estimates are developed. The national targets are then approved by the MOFNP, where, after the effect of any proposed tax policy changes, the target is imputed.

Tax revenue forecasting plays a vital role in a nation budgeting process. Tax revenue is a significant focus area for the government to develop their tax system to improve revenue collections, enhance fairness and efficiency of taxes, and to boost economic growth and exports. Governments need funds to finance their budget expenditures (Jenkins et al.[21]), and tax revenues are the main source of government income. For example, if government expenditure is more than its revenues, this will lead to a deficit, and the shortfall will be funded by borrowing or increasing taxes, and both may have a negative effect on the performance of the Zambian economy.

The contribution of tax revenue to the fiscus signifies the need for ZRA to consider using the combined methods to improve the accuracy of the forecast. Numerous techniques or methodologies are used to estimate several types of tax revenue collection. The famous methodologies used by ZRA are based on growth trends, averages, and contributions of the previous period, as well as expert knowledge, which plays a vital role. In a least developed country such as Zambia, ZRA should consider various factors when forecasting tax revenue. Tax revenue forecasts are developed with regard to economic theories, employed techniques, and, most importantly, the assumptions taken into consideration. These assumptions are derived from economic variables such as national income growth, inflation rate, interest rates, employment rate, exchange rate, and net export (Jenkins et al.[21]).

1.1 Background

Accurate forecasting of tax revenue is crucial for effective fiscal planning, budget formulation, and economic policymaking in Zambia. According to ZRA[39], it relies on forecasting models to estimate future tax revenues, which in turn influence government expenditure decisions. Accurate forecasts help avoid fiscal imbalances and ensure that public resources are efficiently allocated (Granger & Newbold [13]; Hamilton [15]). However, the traditional forecasting methods currently employed by ZRA, such as Autoregressive Integrated Moving Average (ARIMA) and Error, Trend, and Seasonal (ETS) models, have limitations in capturing the complexities and structural changes in economic data, leading to inaccuracies in forecasts (Box & Jenkins [4]; Hyndman & Athanasopoulos [19]; Hyndman [18]). These methods are often inadequate when faced with issues such as structural breakdowns, seasonality, and external shocks that are common in developing economies (Smith & Wallis [33]).

In recent years, Bayesian Vector Auto-Regression (BVAR) has emerged as a promising alternative for improving forecast accuracy by incorporating prior information and managing overparameterization more effectively (Koop [23]; Litterman [25]). BVAR models have been shown to outperform traditional models in various macroeconomic applications, especially in the presence of high-dimensional data and complex interrelationships between variables (Caraiani [6]; Fullerton [11]). Studies such as those by Gupta & Kabundi [14] and Krol [24] have demonstrated the effectiveness of BVAR models in capturing dynamic relationships in the forecasting of tax revenues and improving predictive performance.

Despite the success of BVAR models in other contexts, their potential to forecast tax revenues in Zambia has not yet been fully explored. Zambia's economic landscape, characterized by volatility and structural changes, presents an ideal test ground for evaluating the comparative performance of the BVAR, ARIMA, and ETS models. This study seeks to fill this gap by conducting a thorough analysis of these methods to recommend the most reliable approach to improve tax revenue forecasts, thus contributing to more effective fiscal management and economic stability in Zambia (Granger & Newbold [13]; Hyndman [18]).

1.2 Statement of the Problem

The ZRA faces significant challenges in producing accurate tax revenue forecasts. Inaccurate forecasts can lead to improper fiscal planning, budget deficits, and economic instability. Traditional methods like ARIMA and ETS are currently employed, but may not always yield the most reliable results due to their inherent limitations. The BVAR method, with its ability to incorporate prior knowledge and manage complex relationships between multiple variables, presents a potential improvement. However, its effectiveness in the context of Zambia's tax revenue forecasting needs to be thoroughly evaluated and compared with the existing methods to determine whether it can provide more accurate and reliable forecasts.

1.3 Aim of the Study

The aim of this study is to evaluate the effectiveness of the Bayesian Vector Auto-Regression (BVAR) method as an alternative technique to forecast tax revenue in Zambia, comparing its performance with the ARIMA and ETS methods currently used to recommend the most accurate model to improve tax revenue projections.

1.4 Research Objectives

(i) To Identify / select the most suitable BVAR, ARIMA, and ETS models to forecast tax revenue in Zambia based on the characteristics of the data.

(ii) To Estimate the parameters of the identified BVAR, ARIMA, and ETS models.

(iii) To compare the forecast performance of the identified models using historical tax revenue data from Zambia.

1.5 Research Questions

(i) Which BVAR, ARIMA, and ETS models are the most suitable to forecast tax revenue in Zambia based on the characteristics of the data?

(ii) What are the estimated parameters of the identified BVAR, ARIMA, and ETS models?

(iii) How does the forecast performance of the identified BVAR, ARIMA, and ETS models compare when applied to historical tax revenue data from Zambia?

1.6 Significance of the Study

Accurate tax revenue forecasting is essential for effective fiscal planning and policymaking. By comparing the performance of BVAR with ARIMA and ETS methods, this study seeks to identify the best approach to improving tax revenue forecasts in Zambia. The findings will provide valuable information for the Zambia Revenue Authority (ZRA) and other stakeholders involved in fiscal policy and budget planning, ultimately contributing to better fiscal management and economic stability in Zambia.

1.7 Scope of the study

The purpose of the study is to compare the performance of three models in the forecasting of tax revenues, namely BVAR, ARIMA, and ETS, which are used to forecast the main tax types in Zambia, such as personal income tax (PIT), corporate income tax (CIT), value added tax (VAT), and total tax revenue (TTR). The last mentioned is also incorporated, since the other three tax types contributed approximately 80% to TTR. The study used quarterly data from 2010 to 2023. Tax analysis is multidimensional and is affected by numerous factors, such as economic growth, interest rate, consumption, taxpayer compliance, and behavior. These factors make it impractical for tax forecasting to rely on only one method.

1.8 Limitation of the Study

The main limitation of the study is the inaccessibility of quarterly data for the period prior to 2010. The study is limited to the period between 2010 and 2023. This is the period in which the quarterly data are available. However, the study aims to address the research questions and objectives and ensure that the study depicts the use of the BVAR method as an alternative technique to forecast tax revenue in Zambia.

1.9 Organization of the Dissertation

There are five chapters in this dissertation. The introduction and background information for the research are covered in Chapter 1. It contains the dissertation's goals, objectives, and research questions, in addition to introducing the problem statement. The chapter will also give the significance of the study and finally the scope in which the research will be conducted. Chapter 2 will highlight the overview of the characteristics and development of tax revenue in Zambia, as well as the trend of economic growth in Zambia. Furthermore, the chapter will deal with the relevant literature in terms of the methodology chosen in the study. It outlines the tax forecasting process in Zambia and the techniques used. It also explores the literature on ETS approaches, followed by ARIMA models, and the BVAR technique. The review focuses on the approaches that have been adopted by previous researchers and the limitations of their methods, as well as presenting a discussion of the results of previous studies. The third chapter discusses in detail the methods and procedures used in selecting the best models for PIT, CIT, VAT, and TTR. The measures of accuracy and model selection techniques are also discussed, as well as tests for stationarity of the time-series data. The fourth chapter deals with the analysis, interpretation, and discussion of the results of the three selected approaches to forecasting tax revenue in this study. The fifth chapter will discuss the solutions to the research questions highlighted in Chapter 1. It will present the conclusion of the investigation and provide recommendations based on the findings of the investigation.

Chapter 2

Literature Review

2.1 Introduction

In Zambia, ZRA under the management of the Minister of Finance and National Planning deals with the issues of tax and tax legislation, ZRA is the revenue authority given the mandate to collect and manage all taxes, duties, and levies. Other functions of ZRA, in addition to collecting all revenue that is accrued, ZRA is given the power to ensure maximum compliance with tax and customs legislation. Also provide a customs service that will maximize revenue collection, protect our borders, and facilitate trade (ZRA annual report[39]).

2.2 Taxes in Zambia

The primary purpose of tax collection in Zambia is to collect and provide funds to the government to meet expenditure obligations. A tax can be used as a tool to restructure revenue in a country's economy to help reduce inequality or can be used as a guideline mechanism to encourage or discourage specific actions to improve community welfare. The wealth of any nation is measured by its performance in infrastructure provision through its construction of the manufacturing industry. The construction of a manufacturing industry is large, unstable, and requires huge capital expenditures (ZRA annual report[39]). In the general context, the objective of taxation is comparable throughout the world. However, the resolution to a good tax administration system is most of the time specific to a particular nation. The policy of the country impacts how the taxes and tax administration systems are structured. The strengths of the country are articulated in its ability to make revenue collection equitable, fair, and beneficial to most people in a country.

Furthermore, Wallschutzky [36] argued that tax evaders believed that the government did not use taxpayer funds wisely and that low-income earners bore a disproportionate share of the tax burden. If most taxpayers anticipate that many others will follow suit, then most will avoid paying taxes. In contrast, if most taxpayers believe that few others will evade, then most will pay taxes. Individual and corporate taxpayers engage in these behaviors, although businesses are more prone than individuals to evade taxes. In contrast to a higher value for tax revenues needed for public goods and services, which has a favorable influence on tax compliance, excessive uncertainty and the sunk costs of tax enforcement have a negative effect on tax compliance, according to a study by Bako [2] on tax enforcement, compliance, and morale. Additionally, he discovered that people who use third-party reporting typically avoid taxes at a lower rate than people who self-declare.

In a least developed country such as Zambia, there is a high appetite for government spending due to the many social and economic areas that require attention. The challenges faced by citizens are common in the least developed countries. Common difficulties in such economies include poor health facilities that cannot satisfy the population, poor school systems, dilapidated infrastructure, and a poor road network required to move essential goods and services effectively. According to Bako [2], it was found that in developing countries, road development constitutes a major component of the construction industry. This means that a substantial segment of the national budget for infrastructure development is devoted to road development projects. For this reason certain tax models can be shared and can work effectively in countries with similar economic positions without a lot of customization. In least developed countries, there is a high dependency on tax collections for the economy to thrive. ZRA is constantly seeking to ensure they are collecting the correct amount of taxes by identifying, fixing revenue leakages, and increasing compliance levels of the tax revenue collections.

2.3 Types of taxes in Zambia

The Zambian law states that any individual or company that earns income is required to contribute to the collection of national tax revenue. Zambian taxpayers are engaged in many income-generating activities. For instance, there are others that are high earning, some are medium while others are low income earning. However, it would be unfair to collect the same amount of tax from all taxpayers due to the many disparities in the business that they conduct and the income generated.

The tax revenue collected from a narrow base often involves very high tax rates. This usually happens when it is difficult to trace some taxpayer normally in the informal sector. According to Wallschutzky [36], it is said that in Argentina, more than 80 percent of income goes undeclared. In Ivory Coast, a country where there is reliance on formal sector tax, employment in the private formal sector accounts for only 1.4 percent of the populace. Repetitions to increase tax collection from the formal sector in Ivory Coast have failed, as there is extensive tax evasion, with an average effective tax rate of approximately 48 percent. To make the tax collection process equitable for businesses and individuals earning income, Zambia has a number of different types of tax, and these entities will fall into one or more tax types depending on the types of income-generating activity they are engaged in and the amount of income they generate. The critical attribute of the income distribution problem is the low levels of rural and high urban income and that the actual incomes are influenced by the rural-urban terms of trade.

There are numerous types of tax that are applicable in the country, including rates and other policies. These are normally reviewed throughout each year and may be revised in the annual national budget. Some of the policies may be designed to encourage certain industries, while some policy changes may be meant to give some sort of relief to taxpayers under certain conditions. For example, if the government wanted to discourage a certain popular product that may be destructive to the health of citizens such as cigarettes, it would introduce a new tax type specifically for it or increase the rates already existing that affect that product. According to Doan et al. [10] in their research on the determinants of tax evasion in imported vehicles, it was noted that by presenting certain taxes called'sin tax' to cigarettes, it would be expected that their usage would decrease to encourage a healthier way of life and thus enhance the conducive environment for growth.

Furthermore, if the government wanted to promote a cleaner and greener environment, they would decide to raise taxes on the production or importation of plastic bags. The other comparable example of this was during the COVID 19 pandemic, where numerous businesses suffered due to closures of borders, airports, shops, bars, and restaurants. In such circumstances, the government may decide to provide tax relief in the most affected industries in order to stimulate and enhance production and recovery. However, while other industries suffered during the pandemic, for others an opportunity presented itself and they were able to increase their earnings. For example, the telecommunication industry, which received a very high demand for their goods services due to many companies allowing their employees to work remotely. According to a study by Gupta et al. [14] comparing pre-COVID services with COVID services, it was found that telecommunication services are the only industry that experiences an increase in liquidity. Due to restrictions and social distancing, companies encouraged their employees to work away from their offices while their main offices remained closed. In a situation like this, the government may decide to raise the tax rates or introduce a new tax type for entities in that industry so that the collections can cushion the negative effects of the pandemic in the economy.

2.4 Personal Income Tax

The personal income tax often referred to as the pay-as-you-earn (PAYE) system was established in 1966, as a type of withholding tax from employment income (ZRA [39]). The Zambia Revenue Authority defines Pay as You Earn as the process of withholding taxes from employees' total income according to their wages. Earnings from employment can include salaries and wages, overtime and bonuses, gratuities and allowances, cash benefits, and commissions [16]. With this system, the employer has the authority to determine the amount of tax that each employee must pay, deduct the tax from their wages, and send the money that has been withheld to ZRA. The amount of tax that the employer deducts from any pay depends on the total gross pay of the employee and the current applicable tax rates in the period. The PAYE rates are subject to change according to the annual national budget. According to the annual ZRA report [39], 2020 saw excess collections of ZMW 4,135.1 million, or 13.5 percent, above the ZMW 30,628.3 million yearly objectives reported by the Direct Taxes Division.

2.5 Corporate Income Tax

The corporate Income Tax Act, found in Chapter 323 of the Zambian Laws, governs income taxation in Zambia. Corporate Income Tax is a type of tax that is levied on the earnings of partnerships, limited corporations, self-employed persons, and employees' emoluments. Income tax must be paid on profits by all individuals who generate profits. Income tax applies to entities that have a turnover greater than ZMW 800, 000.00 (ZRA [39]). This indicates that an organization cannot register for turnover tax and income tax simultaneously. The Income Tax Act allows for the deduction of certain expenditure items when determining taxable business profits that decrease the total liability of the taxpayer. Deductions may be made for expenses incurred wholly and solely for the benefit of the business, revenue rather than capital expenditures, losses carried forward from the same source under specific circumstances, donations to recognized public benefit organizations, and other circumstances stipulated by Zambian law. In 2020, the income tax had an audit assessment value of ZMW 1,724.55 million. The income tax category of taxpayers is required to submit in their provisional return before 31^{st} March of the charge year with an estimated liability. They are also required to make provisional payments towards that liability and file the final return accompanied by financial statements the following year before 21^{st} June.

2.6 Value Added Tax

The value added tax is a consumption tax on commodities that are imposed in the production stage, was first introduced in France in 1954 and is known as the value-added tax (VAT). Every country using VAT sets its own unique rules on the administration of this tax, but the basic principle remains the same. VAT is a consumption-based tax that is imposed throughout the supply chain at every stage when value is added to goods or services, according to the Zambia Revenue Authority. Since VAT depends on consumption, the main ways to legally avoid it are to consume only zero-rated or exempt supply or to avoid consuming any goods or services rated standard. According to Byrne [5] , more than 130 nations impose VAT, which typically generates 20% or more of the total tax income. It is widely used in sub-Saharan Africa and other regions, where it has served as the cornerstone of tax reform in numerous developing nations. By many measures, the most important change to tax administration and policy in recent decades has been the introduction of the VA. Typically, companies are required to register for VAT, which entitles them to collect tax on sales and allows them to recover tax on inputs if a certain threshold is reached. Value-added tax is paid in Zambia by the last party in the supply chain who is not registered for VAT. Individuals who are registered for VAT must pay the Zambia Revenue Authority the output VAT that exceeds their input VAT, as well as claim back any input VAT they incurred over the course of their business through the return. Standard-rated goods are subject to a standard VAT rate of 16%; zero-rated goods are subject to a VAT rate of 0%; and exempt goods are not subject to any VAT. In 2020, the value added tax had an audit assessment value of ZMW 1,674.84 million. Declarations and payments for value added tax are due every 18$^{\text{th}}$ of each month.

2.7 Custom and excise duty

Excise taxes are defined as taxes on production although they are typically classified as taxes on consumption. Excise duty is a levy levied at any point along the production or distribution process on specific items or products, regardless of whether they are produced domestically or imported. Calculated on the basis of the weight, strength, or quantity of the products, as well as their value. Every manufacturer of an excisable product in Zambia is required to be registered and licensed with the ZRA except oil marketing companies that are licensed by the Energy Regulation Board, mobile network operators and Internet service providers licensed by the Zambia Information and Communications Authority and importers and distributors of cigarettes. Therefore, any entity manufacturing, distilling, mixing, or brewing all types of spirits, wines, any transportable beverage with a volume percentage of greater than two percent alcohol, opaque beer, cigarettes, and other tobacco products, electrical energy, cosmetics, plastic carrier bags, fuel oils and gases, hydrocarbon oils such as gasoline, diesel need to be registered for excise duty. The license is annual and the taxpayer is expected to apply for renewal before 31st September of every year. Declarations and payments for value added tax are due every 15th of every month. In 2020, the excise duty had an audit assessment value of ZMW 179.47 million (ZRA [39].

2.8 Tax evasion

Tax evasion implies the payment of a suppressed amount or failure to pay any amount from one's taxable income that is required. Furthermore, evasion occurs when a taxpayer deliberately chooses to underreport their income or submit a disproportionately large refund claim. According to the International Monetary Fund (IMF), in 1990, only 60% of the countries around the world had a tax ratio of 10%, and by 2013, that percentage had increased to 75% of all countries. The low compliance rates are intimately related to the act of tax evasion. Wallschutzky [36] revealed that by studying two categories of people who abscond from the mainstream public with respect to the answers obtained from tax evaders, it was discovered that one of the main causes of tax evasion is the belief that taxes are excessive and do not reflect the value for the taxpayer.

Furthermore, Wallschutzky [36] argued that tax evaders believed that the government did not use taxpayer funds wisely and that low-income earners bore a disproportionate share of the tax burden. If most taxpayers anticipate that many others will follow suit, then most will avoid paying taxes. In contrast, if most taxpayers believe that few others will evade, then most will pay taxes. Both individual and corporate taxpayers engage in these behaviors, although businesses are more prone than individuals to evade taxes. In contrast to a higher value for tax revenues needed for public goods and services, which has a favorable influence on tax compliance, excessive uncertainty and the sunk costs of tax enforcements have a negative effect on tax compliance, according to a study by Bako [2] on tax enforcement, compliance, and morale. Additionally, he discovered that people who use third-party reporting typically avoid taxes at a lower rate than people who self-declare.

Tax evasion is committed in different ways with the purpose of avoiding detection by the tax administration. Although some schemes are highly complex and may include many valid procedures both inside the system and outside, others tend to be relatively simple and clear. One such method involves a few separate activities and procedures+ that, taken separately, do not constitute fraud but, when combined, demonstrate that the taxpayer's goal was evasion. The Trends Report published by the Financial Intelligence Center in 2021 showed that in 2021 there were 17 reports of tax evasion cases amounting to ZMW 722 million, entailing an increase in value from ZMW 717 million in the year 2020 and 144 million in the prior year.

2.9 The Tax Forecasting process in Zambia

The ZRA has the mandate to collect tax revenue and is under the Ministry of Finance and National Planning (MOFNP), which is the institution responsible for fiscal forecasting. ZRA uses trend analysis and macroeconomic indicators, taking into account the levels and growth rates of gross domestic product (GDP), to develop an understanding of the revenue situation. It also considers actual collection trends and payment patterns, growth in the tax register, compliance actions, and legislative changes. Statistical modeling is also used to some extent to support revenue estimates.

The revenue forecast methodologies of ZRA are not well documented. Henceforth, the ZRA revenue forecast methodologies mentioned and described in the following were taken from ZRA internal documents of the Revenue Planning, Analysis & Reporting division's Quarterly Bulletin [15] and some research papers produced internally within ZRA. ZRA employs multiple methods to arrive at a point estimate for a specific fiscal year. The main four techniques used are recommended in the revised Code of Good Practice on Fiscal Transparency, as published by the IMF. These four main approaches are the effective rate technique, the elasticity technique, the model-based technique, and the trend and extrapolation techniques. The mentioned approaches may be disaggregated to roughly six techniques currently used at ZRA, these are: Micro-Simulation Model; Macroeconometric Tax Revenue Models; Tax Buoyancy/Elasticity Approach; Box-Jenkins techniques; Revenue Trending & Extrapolation and Professional Judgement. These methods are explained in the following subsections.

2.10 Micro-Simulation

The microsimulation (MS) models involve highly disaggregated data which mainly comprise comprehensive data at the company or individual level to determine tax liability instead of actual tax collection. The MS model uses individuals or companies that file income tax returns and non-filers, which are individuals whose taxable income is below the threshold. MS models are good in their ability to simulate alternative policy proposals and to determine who would benefit from specific policy change or bear the burden of the tax change. This approach was adopted from the Van Heerden and Schoeman [35] paper and was also referenced by Makananisa [26]. Van Heerden and Schoeman [35] utilized data from the publication of Tax Statistics as a proxy to determine the ratio for allowances to be applied to each individual income group. An average allowance ratio τ_{AAR} is developed from taxable income t_i and gross income y_i per taxable income in equation 2.1 below;

$$\tau_{AAR} = \frac{y_i - t_i}{y_i} \tag{2.1}$$

An average allowance ratio in equation 2.1 is then applied to everyone or company gross income group in the equation 2.2

$$A_i = y_i * \tau_{AAR} \tag{2.2}$$

The taxable income is defined as a gross income minus allowable deduction and is given by equation 2.3.

$$t_i = y_i - A_i \tag{2.3}$$

Given the existing tax codes which can be changed for policy simulation purposes tax liability can be determined by using the following equation 2.4,

$$PIT_i = f(t_i, \tau_i) \tag{2.4}$$

in the case of companies, the equation 2.5 may be used:

$$CIT_i = f(t_i, \tau_i) \tag{2.5}$$

Van Heerden and Schoeman [28] stated that this procedure is a static method, and behavioural changes are not taken into consideration.

2.11 Macro-Econometric tax revenue models

ZRA also employs macroeconometric methods, and these models can be categorized under econometric models which are based on the relationship of the dependent variable to several explanatory variables with residual considerations. Macro-based models specify the proxy for various taxes to determine the potential revenue collection for each tax type and are based on the past performance of tax collections and economic growth. The literature specifies numerous proxies for different tax types; CIT may be modeled with company profit where proxy is gross operating surplus. The PIT may be modeled with compensation to employees and employment as explanatory variables, while the value-added tax may take investment and gross domestic expenditure (GDE). Total tax revenue (TTR) can be modeled with gross domestic product (GDP) as an independent variable. These models may be single-equation regression models or multi-variable equation models and may be represented as shown in Equation 2.6,

$$logTTR_i = \beta_0 + \beta_1 log(\phi_1 * \gamma_1) + ... + \beta_i log(\phi_i * \gamma_i) + \epsilon_i \tag{2.6}$$

Here TTR_i represents tax collections at time i ;β_0 is the intercept; ϕ_i is the tax base at time i ; γ_i denotes statutory tax rate at time i and ϵ_i is the disturbance term at time i. Furthermore, macroeconometric models with vector auto-regression models that use the lagged value of the dependent variable as an explanatory variable are generally used, and the vector term implies that two or more variables are involved.

2.12 Tax buoyancy approach

The response of tax revenues to changes in GDP is measured by tax elasticity and tax buoyancy (B), and these concepts aid to describe the overall structure of a tax system and serve as valuable analytical tools for designing tax policy (Jenkins et al. [21]). Tax buoyancy is a measure of the total response of tax revenues to changes in national income and takes into consideration increases in income and discretionary changes introduced by tax body in the system. Discretionary changes can be changes in the tax rate, tax bases, or tax policy. Tax buoyancy can be represented as follows:

$$B = \frac{\Delta T_i}{\Delta Y_i} * \frac{Y_i}{T_i} \tag{2.7}$$

Where B represents Buoyancy of tax revenue to income;T_i is the tax revenue; ΔT_i denotes Change in tax revenue; Y_i is the Income and ΔY_i is the Change in income and represent specific tax type and proxy respectively. If the calculation excludes the effect of changes in tax rates and tax bases and considers only the effects due to changes in income levels, whether changes were made in the tax structure during that period, then the tax elasticity occurs. Tax elasticity is a vital factor for forecasting purposes and its coefficient gives an indication to policymakers of whether tax revenues will rise at the same pace as the income. The tax elasticity is given by the equation,

were made in the tax structure during that period then the tax elasticity occurs. Tax elasticity is a vital factor for forecasting purposes and its coefficient gives an indication to policymakers of whether tax revenues will rise at the same pace as the income. Tax elasticity is given by equation,

$$E_i = \frac{\%\Delta T_i}{\%\Delta Y_i} \tag{2.8}$$

where, E_i is the elasticity of tax revenue to GDP; ΔT_i change in tax revenue; and ΔY_i change in income or GDP. i represent specific tax type and proxy, respectively. An elastic tax system is a highly desirable system, as it provides the government with a sustained fiscal resource base to finance its outlays (Jenkins et al. [21]).

2.13 Box-Jenkins techniques

Box-Jenkins methodology is also known as ARIMA techniques, and these methods explain tax revenue as a function of past values of itself taking into account autoregressive (AR) part and random error terms known as moving average (MA) component. These models are discussed in detail in the Methodology.

2.14 Revenue trending and extrapolation

These methods also incorporate constant trend growth techniques, which uses information on revenue collections received year-to-date (YTD) in the present fiscal year and compares it with the collections made in the previous year, same period. The forecast will be based on the assumption that the growth rate does not vary during the fiscal year and will remain approximately the same. The methodology forecasts revenue collections for the current fiscal year and will account for administrative changes at the beginning of the fiscal year. The formula is given by equation;

$$\hat{Y}_t = \left[\frac{YTD_t}{YTD_{t-1}}\right] * Y_{t-1} \tag{2.9}$$

where, $\hat{Y}_t$ denotes the forecast of the current fiscal year; YTD_t denotes the revenue collections to date of the current year; YTD_{t-1} represents the revenue collections to date of the previous fiscal year and Y_{t-1} is the previous revenue collections in full year.

It is obvious that this method has some disadvantages, as it uses the YTD growth rate at any point to project the full current. Makananisa [26] pointed out the limitation of the model that it assumes a constant growth rate throughout the fiscal year, which is not the case as the economy is driven by various factors that will have a negative or positive impact on revenue collections, therefore the revenue collections growth rate will fluctuate.

2.15 Professional judgement

Professional Judgment (PJ) is the other technique used in ZRA. This method involves the expert's assessment of revenue collections from various internal units or departments and government department other than ZRA. The professional judgment forecast considers the estimated results from all forecasting techniques, but also takes into consideration information that is transmitted to cash flow, administrative changes, and other special factors that cannot be automatically incorporated into the other modeling approaches (Boonzaaier[3]). Boonzaaier [3] stated that historically, the PJ method has often been shown to add significant value to the overall forecasting process, especially in the case of corporate income tax, whose collection dynamics is difficult to capture within a linear modeling framework.

Makananisa [26] highlighted the disadvantage of PJ methods, since professional judgement consists of forecasts of different models and is based on different scenarios, there could be some drawbacks when the expected scenario does not hold or the outcome is not as anticipated.

2.16 The Tax Forecasting process in other countries

The Department of Finance is mandated to forecast tax revenue in Ireland using macroeconomic variables as proxies supplied by the Economic Forecasting Unit of the Department of Finance and, where appropriate, certain elasticity factors. The forecasts are made three times a year, the first is in May/June for the Budget Strategy Memorandum (BSM). These forecasts are said to be for the information of the government only and are therefore not publicized. The second round of forecasts is made in September / October for the Pre Budget-Outlook (PBO), and the third round of forecasts is made in November / December for the Budget (Hannon et al. [16]). The tax forecasting methodology in Ireland is generally given by the equation,

$$\emptyset_{t+1} = \left(\emptyset_t - T_t\right)\left(1 + \left(\theta_{t+1}^{GE}\right)\right) + V_{t+1} + W_{t+1} + X_{t+1} \tag{2.10}$$

Where $\emptyset_{t+1}$ is the one year forecast ahead for a specific tax type, $\emptyset_t$ is the current year estimate for that tax type, T_t are once-off items that affect the outcome in the current year, θ_{t+1}^{G} is the estimated growth rate of the relevant macroeconomic driver that have an impact to the specific tax type year ahead, E denotes the elasticity of tax to its proxy, V_{t+1} are once-off items affecting the outcome in the next year, W_{t+1} is the estimated static outcome from any policy changes that impact receipts for a specific tax in the year ahead, and X_{t+1}

is a discretionary factor impose by the Department of Finance. Proxies used for various tax types are nominal personal consumption for VAT, gross operating surplus for CIT, and non-agricultural employment and wages for PAYE.

Hannon et al. [16] argued that some previous work that looked at Irish revenue forecasts in an international context found that the Irish official forecast performance was on the weaker end of the spectrum. Makananisa [26] employed time series methodologies (exponential smoothing and ARIMA) to forecast the main tax types such as personal income tax, corporate income tax, value-added tax, and total tax revenue in South Africa for three years in the future. The researcher used monthly data from January 1995 to March 2010. The results of the Makananisa study suggested that the SARIMA and Holt-Winters models perform well in modeling and forecasting PIT and VAT. Holt-Winters model was found to perform better than the SARIMA model to forecast CIT and TTR. The study concluded that the chosen models are expected to perform better when projecting future values in stable economic conditions, with the assumption that there will be no shocks in the economy. This study also recommended the use of the selected methods when forecasting tax revenue. The researcher further alluded to the fact that if there is no change in collection approaches, the selected techniques will be accurate with limited bias in forecasting tax revenues. The error encountered will be minimal, and fewer model revisions will be performed.

2.17 Studies comparing the performance of BVAR and other Methodologies

The approach has been used extensively to forecast economic variables and its usage in the field of taxation has not been widely explored. There are several studies that used this technique, such as that conducted by Litterman [25], who used BVAR to forecast economic variables such as real GDP, unemployment, and inflation. In using techniques, Litterman [25] was avoiding the problem of over-parameterization and suggested putting weaker restrictions on the coefficients rather than placing zero. Litterman assumption was a normal prior distribution with a mean of zero and small standard deviation, while the mean on a variable's first own lag is one with a larger standard deviation. Furthermore, Theil's mixed estimation approach as described by Doan [10] was used to estimate coefficients. The standard prior has three distinct characteristics; for instance, the prior probabilities on deterministic variables such as seasonal dummy variables are flat or noninformative. The other characteristic is that the prior distribution is an independent normal. Lastly, the mean of the distribution is zero except for the first lag of the dependent variable of the equation, which is equal to one.

Ramos [30] constructed a BVAR for the leading car market in Portugal to project the share of the car market. The author showed the usability of the VAR and BVAR methodologies as a marketing tool that satisfies two requirements, such as the prediction of market share and the provision of information on the changing competitive conditions of the market. Ramos incorporated five marketing variables into the model. The prior selection was based on the accuracy of the out-of-sample forecasts, which was compared with the accuracy of forecasts from an unrestricted VAR model and benchmark forecasts generated from ARIMA techniques. It was concluded that BVAR is the best forecasting tool relative to univariate ARIMA and VAR models, due to its use of few degrees of freedom.

Furthermore, Ramos [30] postulated that BVAR provides significant information to the people responsible for marketing, using impulse response functions and variance decompositions. The researcher indicated some disadvantages of BVAR, that models are highly condensed, and interpretations of structure based on the signs and sizes of estimated parameters should always be avoided. The researcher highlighted certain limitations of BVAR, first, the models are much reduced forms, and impulse response analysis should be used to test the hypotheses about effects. Second, the forecast accuracy depends on the specifications of the previous. If the prior is not specified correctly, an alternative model, such as an unrestricted VAR or ARIMA model, may be used, which may perform well. Third, using the prior that is selected based on some objective function like Theil's statistics for out-of-sample forecasts may not be best beyond the period for which it was chosen. The model functions best in a stable condition with sufficient data available.

There is scarce literature on the adoption of the BVAR approach as a forecasting tool for tax revenue. One study that used the BVAR technique to forecast state tax revenue was that conducted by Krol [24], who applied the models to Californian tax revenue. Krol stated that Bayesian vector autoregressions generally outperform standard vector autoregressions and simple univariate models in forecasting macroeconomic variables. Krol's study sought to determine whether BVAR would also outperform other models when forecasting state revenue. In most cases, Krol's results show that BVAR models have the smallest root mean squared error compared to the other models examined and recommended that tax revenue forecasters consider using Bayesian vector auto-regressions when producing revenue forecasts.

Although there is scarce literature on the application of the BVAR approach to tax forecasting, there are several studies that employ BVAR to forecast economic variables. In the study conducted by Caraiani [6], the BVAR framework was used to forecast the dynamics of the output of the Romanian economy. The various versions of the BVARs were estimated and compared in terms of forecasting statistics with the OLS and the unrestricted VAR, as well as with the naïve forecast. The best BVAR model in terms of forecast accuracy was selected to forecast the dynamics of quarterly GDP for five quarters, ending in quarter four of 2010. The findings confirmed that the Bayesian approach outperforms standard models. The best BVAR model was used to forecast quarterly GDP in the short term. The results indicated that recovery would be slow and that the output gap would continue to be negative for a few quarters, even after the economy started to grow. The study suggested that other more complex models that incorporate an extension to the open economy or the development of models to analyze monetary and fiscal policy may be used.

In developing the priors, Caraiani [6] follows Litterman's stylized facts of macroeconomic time series that most macroeconomic time series are characterized by a trend, the most recent lags matter the most and the own lags of a variable influence a variable much more than the lags of other variables. With these stylized facts, a prior distribution was derived, which is a random walk.

Furthermore, Yao [37] employed Bayesian VAR methods, as proposed by Litterman, to estimate and forecast several North Dakota macroeconomic variables, including employment, income, and tax receipts. The performance of the out-of-sample BVAR methods was evaluated and compared with vector autoregression models. Data from the first quarter of 1998 to the third quarter of 2005 were used as a holdout sample. In his study, the superiority of the BVAR over the VAR was also confirmed, and the results indicated that properly incorporating prior information into the BVAR provides accurate and responsive forecasts.

Yao [37] adopted the prior of Litterman [25] assuming that a reasonable approximation of the behavior of an economic variable is a random walk around an unknown. The prior of Yao reflects the belief that, first, the coefficients have a prior mean of zeros except the first lag of the dependent variable, which has a mean equal to one. Second, the parameters are uncorrelated, meaning the more past, the smaller the standard deviation of the parameters. Third, the prior standard deviation of the dependent variable should be larger, which implies that the parameters for other variables in the equation are believed to be more tightly centered around zero.

Another study using BVAR was that conducted by Carriero and Mumtaz [7]. Their study investigated the performance of BVARs with constant and drifting coefficients in forecasting key fiscal variables, government revenues, expenditures, and interest payments on outstanding debt. The authors used data from Germany, France, the UK, and the US to show that BVARs perform better than autoregressive forecasts. Carriero and Mumtaz [7] investigated the possibility that the techniques used by various previous studies are too small in scale, possibly with the over-parameterization problem, and do not have time variation coefficients and volatilities. The author estimated various specifications, of which allow summarisation of the information contained in a large data set effectively, avoiding the over-parameterization problem, and allow for time variation in coefficients and volatilities. The finding was that, firstly, once over- parameterization is corrected, the use of extra explanatory variables is important in forecasting fiscal variables, and multivariate models perform better than univariate specifications in forecasting; secondly, the large system implementation and the time variation play a very important role in forecasting.

2.18 Studies based on ARIMA Models

Zakai [38] modeled Pakistan's GDP using a set of ARIMA based on the Box- Jenkins technique. The best suited ARIMA model was, among others, ARIMA (1,1,0), and the forecast values for the next few years were generated by applying the selected model that provided the best fit for the data. Sample forecasting was performed for the period 1953 to 2009, and the visual presentation of the forecast values revealed good behavior.

In a study that focuses on developing a mathematical model to estimate and forecast the income tax revenue of the Philippines for the period 2014-2020, Urrutia et al. [34] considered five explanatory variables, namely the growth rate of the real gross domestic product, the employment population, the unemployment rate, the annual prices of domestic crude oil and the inflation rate. The study examined annual data from 1980 to 2013 for each variable, collected from the National Statistical Coordination Board, Department of Labor and Employment, inflation data.com, and World Bank. In forecasting income tax revenue, the ARIMA model was developed, and the best-fitted model that was obtained was ARIMA (0,1,0) this is a random walk model, a special type of ARIMA model. The paired T-Test was used to test the forecast performance of the model and showed that there was no big difference between the predicted and actual values, indicating that the models are the best at predicting the income tax revenue of the Philippines.

2.19 Studies comparing ETS AND ARIMA Models

Skarbovik [32] employed an AR process, an ARIMA process, and an exponential smoothing state space (ETS) model to find an appropriate fit. Henceforth, to improve the accuracy of the single-best model forecast, the forecasts from the three models were combined. The main objective of the Skarbovik [32] study was to project residential house prices in Norway using data from April 2013 to March 2014. The analysis found that the prediction of the model (ETS) was the most precise compared to the other models, and the conclusion was based on the out-of-sample root mean square error (RMSE) and the mean absolute scaled error (MASE).

In modeling and forecasting fish catches, Bako [2] developed the state space approach (ETS). The author used two methods of time series analysis to predict the catch of fish of three commercial fish species found in Malaysian waters. The Box-Jenkins method was used, together with the ETS state-space exponential method. The models were used to model and forecast monthly catches of the three fish species for two years based on collected data spanning from 2007 to 2011. The ETS models were found to perform better for two species and the SARIMA model performed better for one species based on the root mean square error (RMSE) and mean absolute error (MAE). The conclusion of the study was that both models are suitable for projecting monthly fishery dynamics.

2.20 Chapter Summary

The literature review for most studies conducted related to time series forecasting is limited to tax revenue forecasting. Tax authorities around the world mostly depend on the elasticity approach and judgement forecasting, and not much of the time-series techniques is utilized. Moreover, it is evident from the literature review that there is limited use of techniques in tax forecasting, only Krol [10] employed the technique. The importance of the technique is based on the choice of prior probabilities. Most of the literature follows Litterman [25] priors and Sims and Zha [31]. In this study, the three priors considered are the Litterman/Minnesota prior, which assumes a random walk process, the Normal-Wishart prior, which is a conjugate prior normal data, and the Sims-Zha prior, which show how the dummy variables are used to produce the priors for structural models.

Chapter 3

Methodology

3.1 Introduction

The forecasting methodologies are categorized according to two broad approaches, namely time series forecasting and econometric forecasting. Time-series forecasting predicts the variable values from previous observations of that variable, while econometric forecasting is based on models that relate the endogenous variable to several exogenous variables with residual considerations.

The tools and techniques that most developed countries use to forecast various tax revenues consist of macro-based models (Chun-Yan Kuo [8]). These models specify the proxies for tax types, to determine the potential revenue collection for each tax type. The methods are based on the past performance of tax collections and economic growth. In generating the revenue forecasts, discretionary changes should be considered by adjusting for them to consider only the revenue collection associated with economic performance. In this study, discretionary effects such as revenue initiatives and legislative changes are not adjusted due to the absence of distinction between revenue collection related purely to economic performance and collection related to budget policies.

The study will use both time series and econometric approaches to compare their performance and the accuracy of the forecasts using data from the first quarter of 2010 to the first quarter of 2024. The last 12 observations out of a total of 60 are reserved for checking the accuracy of the forecasting methods. The descriptions of the selected techniques used for forecasting in this study are clarified in the sections that follow;

3.2 Bayesian vector autoregression

The BVAR model is a vector autoregression model using the Bayes Theorem based on the prior and posterior distribution. Henceforth, BVAR is simply a VAR model with priors introduced to control coefficients of the variables. The basis of Bayesian statistics is the Bayes Theorem, which states as follows: Suppose we observe a random variable Y and wish to make inferences about random variable φ, where φ is drawn from some distribution $P(\varphi)$, then from the definition of conditional probability;

$$P\left(\varphi/Y\right) = \frac{P\left(\varphi, Y\right)}{P\left(Y\right)} \tag{3.1}$$

The equation 3.1 can be expressed as the joint probability by conditioning on φ , which gives us;

$$P(Y,\varphi) = P\left(Y/\varphi\right) P(\varphi) \tag{3.2}$$

Substituting equation (3.2) in (3.1) gives us Bayes' theorem:

$$P\left(\varphi/Y\right) = \frac{P\left(Y/\varphi\right) P(\varphi)}{P(Y)} \tag{3.3}$$

For n possible outcomes, equation 3.3 may be written as follows;

$$P\left(\varphi_j/Y\right) = \frac{P\left(Y/\varphi_j\right) P\left(\varphi_j\right)}{P(Y)} = \frac{P\left(Y/\varphi_j\right) P\left(\varphi_j\right)}{\sum_{j=1}^{n} P\left(Y/\varphi_j\right) P\left(\varphi_j\right)} \tag{3.4}$$

$P(\varphi)$ is the prior distribution of the possible values, and $P\left(\varphi/Y\right)$is the posterior distribution of φ , given the observed data Y.

Contrary to the point estimators such as means and variances used by classical statistics, Bayesian statistics involves generating the posterior distribution of the unknown parameters, given both the data and some prior density for these parameters. Henceforth, Bayesian statistics provides a much more complete picture of the uncertainty in the estimation of the unknown parameters, especially after the confounding effects of nuisance parameters are removed.

The vector autoregression (VAR) models are widely used to model economic time-series. The main challenge experienced with these models is the issue of handling many parameters, as stated in most of the literature. To overcome this difficulty, Litterman [25] employed a Bayesian VAR approach to solve the over-fitting problem. He suggested that over-fitting may be avoided without imposing an exact zero restriction on the coefficients. Litterman recommended using a Bayesian strategy to estimate the VAR, equation by equation, where a priori, the lags have decreasing relevance.

Doan et al. [10] used VAR models to impose less arbitrary restrictions than traditional econometric models. According to Ciccarelli and Rebucci [9], the VAR model can be represented as:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_{t-\rho} {+} CZ_t {+} \varepsilon_t \qquad where\ t{=}\ 1{,}2{,}{\ldots}{,}T \tag{3.5}$$

where y_I denotes a vector of endogenous variables $(nx1)$ representing a vector of error terms ε_t that is independently, identically and normally distributed with the variance-covariance matrix Σ, thus $\varepsilon_t \sim N(0, \Sigma)$; φ_t for t = 1,2,..., T in the form of the matrix (nxn) and C as the matrix (nxd), and Z_t is a vector $(nx1)$ of exogenous variables. Equation 3.4 may yield imprecisely estimated relations that fit the data well, due to the large number of variables included, this problem is known as over-fitting.

The Bayesian estimation principle can be derived from the equation rewritten in component form as;

$$y_t = X_t\varphi + \varepsilon_i \tag{3.6}$$

where, $X_t = \left(I_n \bigotimes G_{t-1}\right)$ is the matrix (nxn)., $G_{t-1} = \left(y_{t-1}, y_{t-2}, \ldots, y_{t-p}z_t\right)$ is $(kx1)$ and $\varphi = vec\left(\varphi_1, \varphi_{2,} \ldots, \varphi_\rho\right)$ is $(nkx1)$. The unknown parameters of the model are φ and Σ. The likelihood function of the Bayesian estimation of (3.5) given the probability density function of the data conditional on the parameters of the model is given by

$$L\left(Y/\varphi, \Sigma\right) \propto |\Sigma|^{-\frac{T}{2}}\, exp\left[-\frac{1}{2}\sum_{t=1}^{T}\left(y_t - X_t\varphi\right)'\ \Sigma^{-1}(y_t - X_t\varphi)\right] \tag{3.7}$$

And a joint prior distribution on the parameters, $P\left(\varphi, \Sigma/y\right)$ the joint posterior distribution of the parameters conditional on the data is obtained through the Bayes theorem stated in equation (3.3),

$$P\left(\varphi, \Sigma/y\right) = \frac{P\left(\varphi, \Sigma\right) L\left(Y/\varphi, \Sigma\right)}{P\left(y\right)} \propto P\left(\varphi, \Sigma\right) L\left(Y/\varphi, \Sigma\right) \tag{3.8}$$

The joint probability density function of the observations and the parameters, $P\left(\varphi, \Sigma, y\right)$ can be expressed;

$$P\left(\varphi, \Sigma, y\right) = L\left(Y/\varphi, \Sigma\right) P\left(\varphi, \Sigma\right) = P\left(\varphi, \Sigma/y\right) P\left(y\right) \tag{3.9}$$

where $\propto$ denotes 'proportional to'. Given $P\left(\varphi, \Sigma/y\right)$, the marginal posterior distributions conditional on the data, $P\left(\varphi/y\right) andP\left(\Sigma/y\right)$ can then be obtained by integrating out φ and Σ from $P\left(\varphi, \Sigma/y\right)$, respectively. Finally, the location and dispersion of $P\left(\varphi/y\right)$ and $P\left(\Sigma/y\right)$ can be easily analyzed to produce point estimates of the parameters of interest and precision measures.

BVAR model is VAR with priors introduced to the control coefficients of the variables. The VAR(k) model is;

$$y_t = \beta_0 + \sum_{j=1}^{K} A_j y_{t-1} + \varepsilon_t \tag{3.10}$$

ε_t is assumed to be identically and normally distributed with the variance-covariance matrix Σ, where $y_t, t = (1, 2, \ldots, T)$ is a $(nx1)$ vector of observations on n time series variables, β_0 is a $(nx1)$ vector of intercepts, and A_j is a (nxn) coefficients matrix.

Let y be a vector $(Tx1)$, which stacks the observation of T on each dependent variable in columns next to each other.

Denote the following as follows;

$$x_t = (1, y_{t-1}, ..., y_{t-k}), X = \begin{vmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_T \end{vmatrix} \; and \; A = \begin{vmatrix} \beta_0 \\ A_1 \\ A_2 \\ \vdots \\ A_k \end{vmatrix} \tag{3.11}$$

and $\beta = Vec(A))$, which is a $(pmX1)$ vector. Following the above definitions, the VAR model can be represented as

$$Y = XA + E \; where \; E \sim N(0, \Sigma) \tag{3.12}$$

Different choice of priors may be used with the VAR models, three priors were used in this study and the first was the Minnesota prior (Litterman[25]). This prior assumes that Σ is known. The three priors used in this dissertation are defined in the following sections.

3.3 Litterman/Minnesota prior

This was proposed by Litterman [25] at the University of Minnesota and the Federal Reserve Bank of Minneapolis. Henceforth, the name Litterman prior is sometimes referred to as the Minnesota prior. Litterman prior assumes that each variable follows a random-walk process with possible drift. If we want to estimate the $(kx1)$ vector φ_i with the parameters of the equation i^{th} of (3.8) when the variance $\left(\delta_{ii}^2\right)$ of the error term is known, the Litterman prior assumes that the prior of φ is defined as follows;

$$P(\varphi_i) \sim N(\ \hat{\varphi}_i, \hat{\delta}_i) \tag{3.13}$$

Here, $\hat{\varphi}_i$ and $\hat{\delta}_i$ are the prior mean and variance covariance matrix, respectively. Σ is assumed to be a fixed and restricted diagonal matrix with its elements calculated from the estimation of a univariate autoregression model of order AR(p). The observation of i^{th} equation 3.5 can be written as;

$$Y_i = X\varphi_i + \varepsilon_i \qquad where\ i = 1, 2, ..., n \tag{3.14}$$

Where Y_i and ε_i are vectors $(Tx1)$, X substituted X_t in equation (3.5). Litterman [25] assumes that λ is a degenerate random variable with the following structure for the diagonal elements of $\hat{\delta}_i$;

$$var(\ \hat{\varphi}_i) = \begin{cases} \left[\frac{\lambda_1}{l^{\lambda_3}}\right]^2 & for\ (i = j) \\ \left[\frac{\lambda_1 \lambda_2 \sigma_i}{l^{\lambda_3} \sigma_j}\right]^2 & for\ (i \neq j) \end{cases} \tag{3.15}$$

The three scalars, λ_1,λ_2, and λ_3 were chosen to simplify the elements of $\hat{\delta}_i$. Here, λ_1 controls the overall tightness, and λ_2 controls the tightness of the lags of the cross-variable in the equation. λ_3 captures the decay of the lag in the prior variance with i=1,2,...,ρ representing variable's lags.

By changing the hyper-parameter scalar values, one may lead to tightening or loosening of the prior. The choice of the values for the scalars depends on the empirical trials of playing around with different values. Given this choice of prior, the posterior for φ is computed as follows;

$$\hat{\varphi}_i = \hat{\delta}_i(\hat{\delta}_i^{-1}\varphi_i + \hat{\delta}_{ii}^{-1}X'Y_i) \qquad where \qquad \hat{\delta}_i = \left[\widetilde{\delta}_i^{-1} + \delta_{ii}^{-1}X'X\right]^{-1} \tag{3.16}$$

Henceforth, $P(\varphi_i/Y) \sim N(\hat{\beta}_i, \hat{\delta}_i)$, given that $\hat{\delta}_i$,$\hat{\beta}_i$ and δ_{ii}^{-1} is known, $\hat{\beta}_i$ is taken as a point estimate.

3.4 Normal-Wishart prior

The conjugate prior for normal data is the Normal-Wishart if the assumption of a fixed and diagonal variance-covariance matrix of residuals is relaxed, and the conjugate prior for normal data is defined as follows;

$$P(\varphi/\Sigma) \sim N(\hat{\varphi}, \Sigma \bigotimes \hat{\delta}) \tag{3.17}$$

$$P(\Sigma) \sim iW(\Sigma, \omega) \tag{3.18}$$

The prior distribution of φ will be normal with the prior mean $E(\varphi) = \varphi$ and the prior variance $VAR(\varphi) = (\omega - n - 1)^{-1}\hat{\Sigma} \bigotimes \hat{\delta}$, where ω are the degrees of freedom of the inverse Wishart satisfying $\omega > n + 1$

By applying the Bayes rule, the posterior is given as follows;

$$P(\varphi/\Sigma, Y) \sim N(\hat{\varphi}, \Sigma \bigotimes \hat{\delta}) \quad where \quad \hat{\varphi} = \left[\widetilde{\varphi}^{-1} + X'X\right]^{-1} \quad and \ \hat{\beta} = \hat{\varphi}(\widetilde{\varphi}^{-1}\widetilde{\beta} + X'X\widetilde{\beta}) \tag{3.19}$$

$$\hat{\Sigma} = \ \hat{\beta}'X'X \ \hat{\beta} + \ \hat{\beta}'\hat{\varphi}^{-1}\hat{\beta} + \widetilde{\Sigma} + (Y - X\hat{\beta})'(Y - X\hat{\beta}) - \ \hat{\beta}'(\widetilde{\varphi}^{-1} + X'X)\hat{\beta} \tag{3.20}$$

3.5 Sims-Zha priors

Sims and Zha [31] elaborated the estimated structural BVAR model. To set the Sims-Zha priors for the structural parameters, the structural VAR model is suggested;

$$y_t A_0 = \alpha_0 + \sum_{k=1}^{p} A_k y_{t-k} + \varepsilon_t \tag{3.21}$$

where $\varepsilon_t \sim N(0, I)$ and $\Sigma = (A_0^{-1})' A_0^{-1}$. The Bayesian prior is developed for the unrestricted VAR and then will be mapped to the restricted prior parameter. Defining A_t to be a matrix of coefficients in the lagged variable, the equation can be expressed as follows;

$$YA_0 - XA_t = E \tag{3.22}$$

where Y is (HXm), A_0 is (mXm), X is $(H+(mp+1))$,A_t is ((mp+1)Xm) and E is (HXm) Sims-Zha prior to A_0 and A_t is given by;

$$\pi(A_0)\pi(A_t\ /A_0) = \pi(A_0)\emptyset(\beta_0, T_0) \tag{3.23}$$

Where $\pi(A_0)$ denotes a marginal distribution of A_0 and $\emptyset(\mu_0, T_0)$ represents a normal density distribution with mean $\mu_0 = A_t - \mu A_0$ and covariance $T_0 = T(A_0)$.

Henceforth, the conditional likelihood is expressed as;

$$L(Y/A) \propto |A_0|^H\, exp\left(-\frac{1}{2}trace(ZA)'(ZA)\right) \qquad where\ Z = (Y - X)\ and\ A = (A_0A_t)' \tag{3.24}$$

The posterior density is derived by combining equations (3.22) and (3.23) as follows;

$$\pi(\alpha) \propto \pi_0(\alpha_0)\, |A_0|^H\, |T_0|^{-\frac{1}{2}}\, exp\left[-\frac{1}{2}\alpha_0{}'(I \bigotimes Y'Y)\alpha_0 - 2\alpha_t{}'(I \bigotimes X'X)\alpha_t + \beta_t' T_0^{-1}\beta_0\right] \tag{3.25}$$

From the above equation α denotes elements of vector A, this posterior is non-standard, the conditional posterior distribution of A_t/A_0 that is derived as follows;

$$\pi(\alpha_t/\alpha_0) = \emptyset(\hat{\beta}, (I \bigotimes X'Y + T_0^{-1})^{-1}) \quad where\ \hat{\beta} = (I \bigotimes X'X + T_0^{-1})^{-1}((I \bigotimes X'Y)\alpha_0 + \beta_t' T_0^{-1}\beta_0 \tag{3.26}$$

The elements of H_0 for i,j=1,2,...,k and l=1,2,...,p is written as;

$$H_{(ol,ij)} = \left[\frac{\lambda_1\lambda_2}{\delta_j l^{\lambda_3}}\right]^2 \tag{3.27}$$

Where δ_j^2 denotes the j^{th} diagonal element of Σ for the i^{th} lag of the series i in equation j The three hyperparameters λ_1,λ_2 and λ_3 represent the general beliefs about the VAR. λ_1 is overall tightness of beliefs in, λ_2 denotes the standard deviation in A_t and λ_3 is a lag decay. In the case where prior information is considered as dummy variables, Sims and Zha suggest Y^d and X^d as additional dummy variables.

$$Y^d = \begin{vmatrix} Y_1^d \\ Y_2^d \end{vmatrix} \ and \ X^d = \begin{vmatrix} X_1^d \\ X_2^d \end{vmatrix} \tag{3.28}$$

which care for the unit roots of Y_1^d and X_1^d and the trends of Y_2^d and X_2^d, therefore the model is expressed as follows;

$$\begin{vmatrix} Y^d \\ Y \end{vmatrix} A_0 - \begin{vmatrix} X^d \\ X \end{vmatrix} A_t = E \tag{3.29}$$

3.6 Selecting priors

The essential properties of Bayesian statistics are the construction of the parameters based on prior information that the modeler believes. Selecting the prior distribution is the most significant part of BVAR modeling. The prior distribution and sample data are required to get the posterior distribution. The selection of the prior distribution depends on the experience, previous knowledge of the forecaster, and the structure of the available information. Giannone et al. [12] recommended the selection of priors using the marginal data density, for example, the likelihood function integrated over the model parameters, which depends only on the hyperparameters that characterize the relative weight of the prior model and the information in the data.

The BVAR models are important because it depends on the fact that the model parameters are random variables. The rationale is to represent the prior information for all unknown quantities through a prior distribution and to combine them with the objective information derived from observations to obtain the posterior distribution. The posterior distributions are commonly derived by the application of the Bayes theorem. The dissertation considered the prior distributions that are based on the normal distribution and commonly used in the literature, such as the Litterman or Minnesota prior, the normal Wishart prior, and the Sims-Zha normal Wishart prior.

3.7 Error, Trend, Seasonal Methods

The ETS is explained by exponential smoothing techniques in general, as this is the basis on which the ETS methods were developed. The exponential smoothing method is a famous forecasting technique and was developed by Holt [17] and extended by Winters. The concept of exponential smoothing methods is to produce forecasts using weighted averages of past observations, with the weights decaying exponentially as the observations get older. This implies that the more recent the observation, the higher the associated weight.

There are many categories of exponential smoothing methods, the simplest of which is Simple Exponential Smoothing (SES). SES is appropriate for generating forecasts of data with no trend or seasonal pattern. Holt [17] improved the SES to allow data forecasting with a trend named after him, Holt's linear trend methods. It comprises a forecasting equation and two smoothing equations for the level and the trend, as given by the equations;

$$\hat{y}_t = \alpha y_t + (1-\alpha)(\hat{y}_{t-1} + r_{t-1}); \qquad where\ ,\ 0 < \alpha < 1 \tag{3.30}$$

$$r_t = \beta(\hat{y}_t - \hat{y}_{t-1}) + (1-\beta)r_{t-1}; \qquad where\ ,\ 0 < \beta < 1 \tag{3.31}$$

$$\hat{y}_{t+1} = \hat{y}_t + lr_t \tag{3.32}$$

Where $\hat{y}_t$ is the level estimate of the series at time t and r_t represents the slope estimate of the series at time t. The smoothness of the series is determined by two parameters, α and β, these parameters must be between 0 and 1.

The Holt method was extended by Winters to capture seasonality and is called the Holt-Winters seasonal method. This method has a forecast equation and three smoothing equations, one for the level (L_t), one for the trend (b_t), and one for the seasonality component (s_t), with associated smoothing parameters α, β, and γ, respectively. The three smoothing equations for multiplicative seasonality are given as follows (Makridakis and Wheelwright[13]);

$$L_t = \alpha \frac{y_t}{S_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \;\; where \; , \; 0 < \alpha < 1 \tag{3.33}$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \;\; where \; , \; 0 < \beta < 1 \tag{3.34}$$

$$s_t = \gamma \frac{y_t}{L_t} + (1-\gamma)s_{t-s} \;\; where \; , \; 0 < \gamma < 1 \tag{3.35}$$

$$\hat{y}_{t+m} = \; (L_t + b_t m)s_{t-s+m} \tag{3.36}$$

where s represents the duration of seasonality, s_t denotes the seasonal component and is $\hat{y}_{t+m}$ the forecast for m periods in the future.

3.8 Autoregressive integrated moving average

The ARIMA model is used for the purpose of comparing the performance of BVAR, which is aimed at describing the autocorrelation in the data. The ARIMA model is generally used for time series data with trends and auto-regression. The ARIMA is a differencing process that makes the ARMA process stationary.

The seasonal ARIMA, denoted SARIMA, is a generalization and extension of the regular ARIMA. It is used for time series where a pattern repeats itself seasonally over time. ARIMA models (Box et al. [4]) take historical data into account and decompose these data into an autoregressive process (AR), an integrated process (I) and a moving average (MA) process of the forecast errors. The processes are identified by standard notation; for example, the AR order is represented by p and the MA by q, and a combination of AR and MA representations is known as ARIMA (p, d, p).

The autoregression model forecasts the variable using a linear combination of the past values of the variable, for instance, the regression of the variable against itself. The autoregression model of the order p is written as;

$$y_t = c + \theta_1 y_{t-1} + \theta_2 y_{t-2} + ... + \theta_{t-p} y_{t-p} + \varepsilon_t \tag{3.37}$$

where c is a constant and ε_t is a white noise.

In addition to the AR(p) model above, there is also the moving average model, which forecasts the dependent variable using a linear combination of white noise error terms. Generally, the model of order q can be expressed as follows;

$$y_t = c + \varepsilon_t + \emptyset_1 \varepsilon_{t-1} + \emptyset_2 \varepsilon_{t-2} + ... + \emptyset_q \varepsilon_{t-q} \tag{3.38}$$

where c is a constant and ε_t is a white noise.

The general form of the combined processes ARIMA(p,q) is shown below;

$$y_t = \theta_1 y_{t-1} + \theta_2 y_{t-2} + ... + \theta_p y_{t-p} + \varepsilon_t - \emptyset_1 \varepsilon_{t-1} - \emptyset_2 \varepsilon_{t-2} - ... - \emptyset_q \varepsilon_{t-q} \tag{3.39}$$

Time series have a seasonal component; this implies that a time series is seasonal if there is a periodic variation after a certain time interval. The seasonal model, commonly known as the aforementioned, is a generalization and extension of the ordinary model to accommodate seasonality in the data. This seasonal component of the model is denoted by capital letters, SARIMA(p,d,q)(P,D,Q) where the first bracket indicates the nonseasonal parameters and the last bracket indicates the seasonal factor parameters for the order of the autoregressive, integration and moving average parts of the model SARIMA(p,d,q)(P,D,Q). The general form is given as:

$$\left(1 - \emptyset_1 B - \emptyset_2 B^2 - ... - \emptyset_\rho B^\rho\right) \left(1 - \beta B^2 - \beta_2 B^{2s} - ... - \beta_\rho B^{\rho s}\right) (1 - B)^d \left(1 - B^s\right)^D y_t \tag{3.40}$$

=c+(1-$\psi_1 B - \psi_2 B^2 - ... - \psi_q B^q)(1 - \theta B^s - \theta_2 B^{2s} - ... - \theta_Q B^{Qs})\varepsilon_t$

Where;

$\left(1 - \emptyset_1 B - \emptyset_2 B^2 - ... - \emptyset_\rho B^\rho\right)$ *is the nonseasonal autoregression part of order P*

$\left(1 - \beta B^2 - \beta_2 B^{2s} - ... - \beta_\rho B^{\rho s}\right)$ *is the seasonal autoregression part of order P*

$\left(1 - \psi_1 B - \psi_2 B^2 - ... - \psi_q B^q\right)$ *is the nonseasonal moving average part of order q*

$\left(1 - \theta B^s - \theta_2 B^{2s} - ... - \theta_Q B^{Qs}\right)$ *is the seasonal moving average part of order q*

$(1 - B)^d$ *is the seasonal differencing of order*

$(1 - B^s)^D$ *is the seasonal differencing of order*

s is the period of the seasonal patterns

Due to the seasonality of the corporate income tax and the fact that it is volatile, seasonal ARIMA will be applied to generate the forecast.

3.9 Identification of the AR and MA process

The first step in developing model using the Box-Jenkins approach is to determine the stationarity of the series and investigate if there is any significant seasonality that needs to be modeled. Once the issue of stationarity and seasonality has been solved, the next step is to identify the order (p and q) of the autoregressive and moving average terms. It is difficult to tell what the order of AR and MA should be from a time plot and what values of p and q are appropriate for the data. In this regard, the main tools for identification are the autocorrelation function (ACF), the partial autocorrelation function (PACF) and the resulting correlograms, which are plots of ACF and PACFs against lags. The ACF plot shows the autocorrelations that measure the relationship between y_t and y_{t-k} for different values of k. Observations may be correlated to each other, for example, if y_t and y_{t-1} are correlated, then y_{t-1} and y_{t-2} must also be correlated; hence y_t and y_{t-2} may be correlated, because they are both linked to y_{t-1} . To solve this problem, the solution is to use PACF, which measures the relationship between y_t and y_{t-k} by eliminating the effects of other time delays, 1,2,3,...,$k-1$. This means that the first partial autocorrelation is equal to the first autocorrelation.

Theoretically, the autocorrelation function (ρ_τ) of a pure AR process of order p follows a homogeneous difference equation constructed from the AR operator below;

$$\alpha(L) = 1 + \alpha_1 L + ... + \alpha_P L^P \tag{3.41}$$

The autocorrelation function (ρ_τ) is given by

$$\rho_\tau = -(\alpha_1 \rho_{t-1} + \alpha_2 \rho_{t-2} + ... + \alpha_\rho \rho_{t-\rho}) \ for \ all \ \rho \tag{3.42}$$

This equation will generate a sequence of a mixture of damped exponential and sinusoidal functions. The sequence of a sinusoidal will indicate the presence of complex roots in the operator $\alpha(L)$. The partial autocorrelation function (π_τ) clearly identifies a pure AR process. The theoretical π_τ of the AR (p) process has $\pi = 0$ for all $\tau > p$. The elements of the partial autocorrelation function of the sample are expected to be close to zero for lags greater than p, corresponding to estimates of parameters equal to zero. The significance of the partial autocorrelation of the values is checked by the p^{th} order process, of which the standard deviations for all lags greater than p are approximated by $\frac{1}{\sqrt{N}}$. The bounds of $\pm\frac{1.96}{\sqrt{N}}$ are also plotted on the graph of the partial autocorrelation function. For an AR(1) process, the sample autocorrelation function is decreasing exponentially, and, though, higher-order AR processes are often a mixture of exponentially decreasing and damped sinusoidal components.

The theoretical autocorrelation function of a pure moving average process of order q has $\rho_\tau = 0$ for all $\tau > p$. The corresponding partial autocorrelation function (π_τ) is progressively decaying towards zero. To decide whether the corresponding sample autocorrelation function r_τ is zero, we need some standard error for the sample estimates of these quantities.

For an MA(q) process with a sample size of N, the standard deviation of r_τ is given by

$$\frac{1}{\sqrt{N}}\left[1 + 2(r_1^2 + r_2^2 + ... + r_q^2)\right]^{\frac{1}{2}} \qquad for\ all \quad \tau > p \tag{3.43}$$

The autocorrelation function of the MA(q) process becomes zero at lag q+1 and greater, so it is essential to examine the autocorrelation function of the sample to see where it becomes zero. This is done by placing the 95% confidence interval for the sample autocorrelation function on the sample autocorrelation plot. The sample autocorrelation function is given by the limits of $\pm\frac{1.96}{\sqrt{N}}$ which are the approximate 95% confidence intervals for the autocorrelations of a series of white noise.

3.10 Ljung-Box Test

The Ljung-Box test is used mainly in portmanteau tests for ARIMA models and was developed by Ljung and Box. The purpose of this Q-test statistic is to determine whether the set of autocorrelation coefficients is different from zero. The test is done on the residuals of an estimated ARIMA model, instead of the original series. The Q-statistic test is given by Makridakis et al [27] as;

$$Q^* = n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{(n-k)} \tag{3.44}$$

Where, n is the sample size, k represents the number of autocorrelation lags included in the statistic and r_k^2 denotes the autocorrelation of the squared sample in the lag k. The Ljung-

Box test the hypothesis that the residuals of the ARIMA model are without autocorrelation, and in case that the residuals are white noise, the Q- test statistic is asymptotically Chi-square distributed. Care should not be taken to accept a model based solely on portmanteau tests (Makridakis et al. [27]).

3.11 Stationarity and unit root tests

Time series mostly appear to be non-stationary, and there are many unit root tests used to detect if the series is non-stationary. Non-stationarity can be detected by visual examining of the time series graph and by looking at the series correlogram or using unit roots statistical tests. A series can be transformed by differencing once or more times to become stationary. The order of the differencing is the number of times the series needs to be differenced to become a stationary series. A series that is different once is represented by $I(1)$ and $I(0)$ and is a stationary time series of order zero. For this study, two tests for stationarity will be discussed, that is, the Dickey-Fuller test and the Phillips-Perron test.

Consider y_t time series in the form $y_t = \alpha + \beta y_{t-1} + \mu_t$ where $u_t = \rho\mu_{t-1} + \varepsilon_t$,the unit root tests are based on testing the null hypothesis $H_0 : \rho = 1$ against the alternative hypothesis $H_1 : \rho < 1$. The characteristic polynomial has a root equal to unity under the null hypothesis, hence the name unit root tests.

3.12 Dickey-Fuller test

The Dicky-Fuller test is mostly used tests to detect unit roots, and as the name implies, it was discovered by David Dickey and Wayne Fuller in 1979. The test follows the AR (1) process;

$$y_t = \rho\mu_{t-1} + \varepsilon_t \tag{3.45}$$

Here, ε_t is an independent and identical distribution of random variables.

The Dickey fuller Test Hypothesis are as follows;

$\mathrm{H}_0 : \rho = 1$ *against the* $H_1 : \rho < 1$

where y_t is non-stationary under the null hypothesis and is stationary under the alternative hypothesis. Standard t statistics do not follow the t distribution due to the nonstationarity of y_t under the null hypothesis. To test the null hypothesis, the following test statistics equation may be used:

$$DF = \frac{\rho}{se(\rho)} \tag{3.46}$$

The DF test in Equation (3.46) follows the assumption that the error terms are independent and identically distributed, without a drift in the model.

An extension of the DF-test is the augmented Dickey–Fuller test (ADF), which eliminates all autocorrelations in the time series. The procedure for the ADF test is like the Dickey–Fuller test procedure, the only difference is the model where it is applied. The model to which ADF is applied to is shown as below;

$$\Delta y_t = \alpha + \beta t + \phi y_{t-1} + \sigma_1 \Delta y_{t-1} + ... + \sigma_{p-1} \Delta y_{t-p+1} + \varepsilon_t \tag{3.47}$$

where, α denotes a constant, β is the coefficient on a time trend, and represents the lag order of the autoregressive process. Putting the constraints $\alpha = 0$ and $\beta = 0$,this resembles a model with a random walk, and using the constraint resembles a model of a random walk with a drift. The ADF test is performed under the hypothesis;

$\mathrm{H}_0 : \emptyset = 0$ $\quad$ *against the* $\quad$ $H_1 : \emptyset < 0$

The test statistic is computed as;

$$\mathrm{DF}_\tau = \frac{\hat{\emptyset}}{se(\hat{\emptyset})} \tag{3.48}$$

If the test statistic DF_τ is less than the critical value, then the null hypothesis of $\emptyset = 0$ is rejected and no unit root is present.

3.13 Phillips-Perron test

The Phillips-Perron (PP) test is an alternative technique for correcting for serial correlation in unit root testing and was developed in 1988 by Phillips and Perron. The PP test uses the standard DF or ADF test, but modifies the t-ratio to prevent serial correlation to affect the asymptotic distribution of the test statistic. The difference between the PP and ADF tests is in terms of how these tests deal with the issue of serial correlation and heteroskedasticity in the error terms. The test model for the PP test is given as;

$$\Delta y_t = \beta' D_t + \pi y_{t-1} + \mu_t \tag{3.49}$$

where μ_t denotes $I(0)$ which may be heteroskedastic. The PP tests correct for serial correlation and heteroskedasticity in the error terms of the μ_t test model, by directly modifying the Dickey-Fuller test statistics $t_{\pi=0}$ and $T_{\hat{\pi}}$. The test statistics denoted by Z_t and Z_π are given as:

$$Z_t = \left(\frac{\hat{\delta}^2}{\hat{\lambda}^2}\right)^{\frac{1}{2}} * t_{\pi=0} - \frac{1}{2}\left(\frac{\hat{\lambda}^2 - \hat{\delta}^2}{\hat{\lambda}^2}\right)\left(\frac{T * se(\hat{\pi})}{\hat{\delta}^2}\right) \tag{3.50}$$

$$Z_\pi = T * \hat{\pi} - \frac{1}{2}\frac{T^2 * se(\hat{\pi})}{\hat{\delta}^2}(\hat{\lambda}^2 - \hat{\delta}^2) \tag{3.51}$$

The estimates of the variance parameters are as follows;

$$\hat{\lambda}^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E\left(\mu_t^2\right) \tag{3.52}$$

$$\hat{\lambda}^2 = \lim_{T\to\infty} \sum_{t=1}^{T} E(T^{-1} S_T^2) \;\; where \; S_T = \sum_{t=1}^{T} \mu_t \tag{3.53}$$

where $\hat{\delta}^2$ is the sample variance of the least squares residual. $\hat{\mu}_t$ is a consistent estimate of δ^2,and the Newey-West long-run variance estimate of μ_t using $\hat{\mu}_t$ is a consistent estimate of λ^2.Under the null hypothesis that $\pi = 0$, the Philip Pellon test statistic Z_t and Z_π have the same asymptotic distributions as the ADF t statistics and normalized bias statistics. The advantage of PP tests over ADF tests is that PP tests are robust to general forms of heteroskedasticity in the error term μ_t, and the user does not have to specify a lag length for the test regression.

3.14 Model selection

A common approach used for model identification is based on information criteria. The well-known information criterion is the Akaike information criterion (AIC), which is a tool to assess the fit of statistical models for a given set of data. The AIC value of the model is represented by the following equation;

$$AIC = -2ln(L) + 2k \tag{3.54}$$

where L is the numeric value of the maximum likelihood of the model, the number of estimated parameters is denoted by k, and n represents the sample size.

Given a set of competing models for specific data, the preferred model is the one with the smallest AIC value. As an alternative to AIC, if the number of observations is few, AIC_c (corrected AIC) is normally used. The AIC_c is used when the sample size $\frac{n}{k} > 40$, as recommended by Burnham and Anderson [5]. The corrected AIC is represented by formula (3.55).

$$\text{AIC}_C = -2ln(L) + \frac{2kn}{n-k-1} = AIC + \frac{2k(k+1)}{n-k-1} \tag{3.55}$$

The AIC_c depends on the statistical model with the assumption that the model is univariate, linear, and has normally distributed residuals. There are other measures that may be used for the model section, but in this study, the AIC is used to select the best model.

3.15 Measures of accuracy for forecasting

The evaluation of the accuracy of forecasts frequently involves using measures of forecast accuracy that assess the performance of the models in terms of handling the data for tax types over the entire sample period. These measures are independent of the scale of the data and are mean percentage error (MPE) and mean absolute percentage error (MAPE), as shown in the following formulae;

$$MPE = \frac{1}{n}\sum_{t=1}^{n} |y_t - \hat{y}_t| \tag{3.56}$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left|\frac{y_t - \hat{y}_t}{y_t}\right| \tag{3.57}$$

where $\hat{y}_t$ is the forecast value in period t; y_t is the actual value in period t and n is the size of the sample. Other accuracy measures which are scale dependent are also commonly used, such as mean standard error (MSE), mean absolute error (MAE), and root mean square error (RMSE). The formulae for these measures of accuracy are given as;

$$MSE = \frac{1}{n} \sum_{t=1}^{n} \varepsilon_t^2 \tag{3.58}$$

$$MAE = \frac{1}{n} \sum_{t=1}^{n} |\varepsilon_t| \tag{3.59}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^{n} (y_t - \hat{y}_t)^2} \tag{3.60}$$

3.16 Data and statistical software

The data for tax revenue collections was obtained from MOFNP and the economic data was sourced from the Zambia Statistics Agency (Zamstats). The software used to generate the results was R- Studio and Econometric Views (E Views), which are currently used at ZRA and BOZ for model development and to generate revenue forecasts. Hyndman and Kandahar [20] describe two automatic forecasting algorithms appropriate to seasonal and non-seasonal data, which have been used in the R forecast package. The first algorithms are for state-space innovations that underlie exponential smoothing methods, and the second is a stepwise algorithm to forecast with ARIMA models.

3.17 Chapter Summary

This chapter discussed the techniques used in this study, the BVAR model, the ETS model, and the ARIMA model. These techniques were appropriate for capturing the data generation process of past observations of TTR, CIT, VAT, and PIT. The techniques will also be used to generate revenue collection forecasts for these three types of tax.

The prior selection for BVAR was also discussed, and the choice of the prior distribution is the most significant part of the modeling. In this dissertation, prior distributions that are considered were commonly used in literature and are the Litterman/Minnesota prior, The Normal-Wishart prior and The Sims-Zha normal- Wishart prior. These priors are based on the normal distribution. To determine the presence of unit roots in error terms, the ADF and PP tests were used. The best model is selected according to the Akaike Information

Criteria for competing models within the same approach. The best performing methods will be selected on the basis of measures of accuracy within different approaches. Revenue forecasting is not limited to the techniques discussed in this study, nevertheless, as other methods exist which can be employed to generate forecasts.

Chapter 4

Presentation, Interpretation and Discussion of Findings

4.1 Introduction

Based on theoretical and empirical discussion in Chapter 3, Employing Bayesian Vector Auto-Regression (BVAR) Method as an Alternative Technique for Forecasting Tax Revenue in Zambia. The focus of this chapter is on the application of the methodologies discussed earlier, analysis and interpretation of the results of the BVAR, ARIMA, and ETS models employed to generate forecasts of the PIT, VAT, CIT, and TTR.

4.2 Description of the tax variables

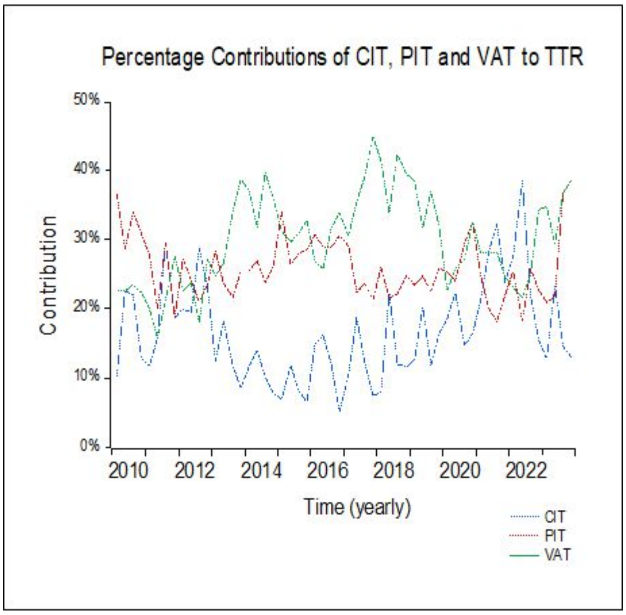


Figure 4.1: Percentage Contribution of CIT PIT and VAT to TTR

The CIT revenue collection has been growing rapidly between 2014 and 2022. In 2012 the CIT collection contracted due to an economic recession that started late in 2008. In 2014 the CIT collection slowly recovered with an improved rate of 12%. The share of CIT collection in total revenue has improved drastically from 10% in 2010 to 19. 5% in 2023, as shown in Figure 4.1. The contribution of CIT to GDP rose from 10% in 2010 to the highest contribution to GDP of 31% in 2022.

The second largest source of tax revenue over the years has been PIT although its share in total revenue has declined from 36% in 2010 to 31% in 2023. In 2010 to 2017, the contribution was hovering around 28.0% on average. In 2014, the PIT share in total tax revenue was surpassed by the Value-Added Share Record 45% in 2018. The percentage of PIT to GDP has decreased from 22% in 2018 to 18% in 2021.

VAT has the highest contribution to tax revenue. However, it has been increasing over the years. In 2010 the VAT share in total tax revenue was 23% reaching its highest peak of 44% in 2018. The average contribution from 2010 to 2023 was 32%.

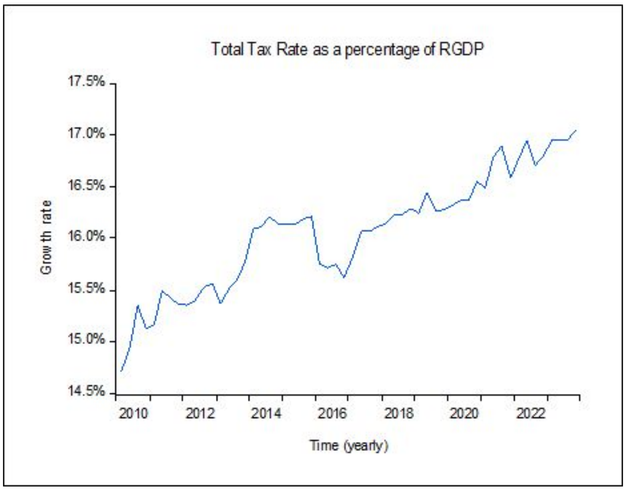


Figure 4.2: Total Tax rate as a percentage of RGDP

The TTR has been growing tremendously, from a rate of 14.6% in 2010 to 17. 1% in 2023. This is mainly driven by PIT, VAT and CIT, as these three tax types contribute approximately 80%.The TTR collection contracted due to an economic recession that started late in 2008. The total tax-to-GDP ratio has improved significantly from 2010 to 2023.

4.3 Description of the Economic Variable

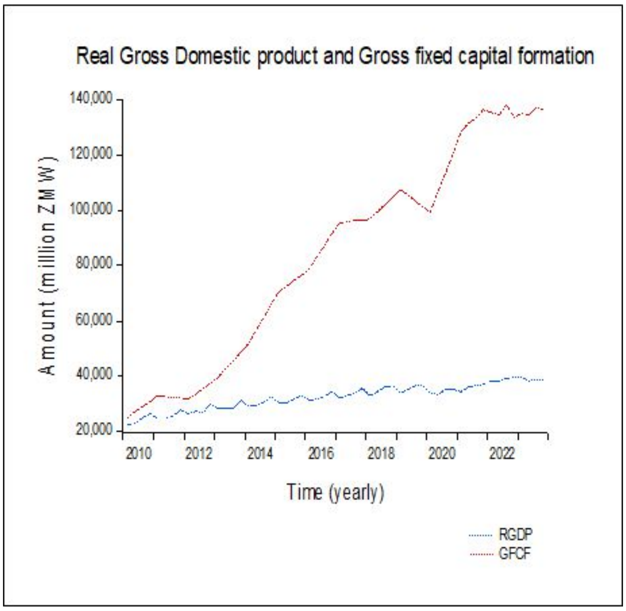


Figure 4.3: Real Gross Domestic product and Gross fixed Capital Formation

The economy of Zambia fell into a deep recession due to the adverse impact of the COVID–19 pandemic. Real GDP contracted by an estimated 4.9% in 2020, after growing by 4.0% in 2018 and 1. 9% in 2019. The economy of Zambia fell into a deep recession due to the adverse impact of the COVID–19 pandemic. Real GDP contracted by an estimated 4.9% in 2020, after growing by 4.0% in 2018 and 1.9% in 2019. The contraction in output is the result of an unprecedented deterioration in all key sectors of the economy.

GFCF as shown above indicates a steady rise from 2010 through the years to 2012, before rising sharply in 2013 to peak initially in 2021. The figure above shows a continuous upward increase in GFCF, with the highest recorded post-Covid 19 pandemic in 2023. Increased rates of GFCF reduce the cost of borrowing; thus, the reduced interest rates lead to a shift of resources from consumables to Stock Market Instruments so that savings and investments are positively affected. This causes an increase in the demand for stock market instruments, which tends to increase the volume of trading, and thus the value of traded stocks. Market capital may therefore increase as the demand for shares increases because of the substitution process.

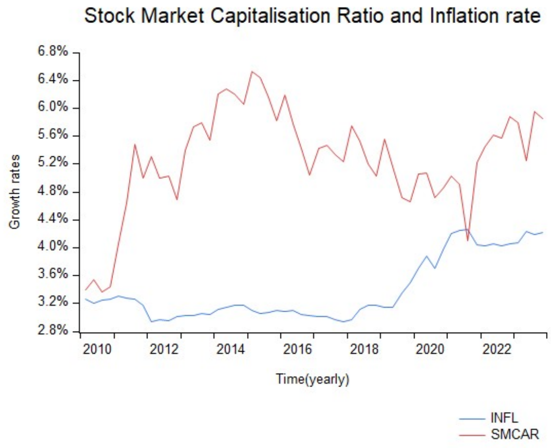


Figure 4.4: Stock Market Capitalisation Ratio and Inflation Rate

Inflation has been increasing, mainly driven by the pass-through effects of the depreciation of the kwacha and elevated food and transport prices. Following the outbreak of COVID-19, the quarterly inflation rate increased from 3.5% in 2010 to 6.6% in 2015 and is projected to remain above the target range of 6% to 8% in 2025. The external position also worsened in 2020 due to dwindling reserves and continuous increases due to copper price and output fluctuations, rising public debt payments, and elevated non-oil imports.

The size of the market is measured by the capitalization ratio of the stock market. The ratio is the value of domestic equities traded on the exchange divided by GDP. Figure 3 shows a continuous decline in the stock market capitalization ratio. A sharp decline was recorded during the Covid 19 pandemic between 2020 and 2021. The trend shows a sharp increase in the market capitalization ratio between 2019 and 2020.The ratio continuously increased in the years that followed after 2022.

4.4 Unit Root Tests

In time series analysis, the variables must be tested for stationarity. Therefore, this study used the conventional Augmented Dicky-Fuller (ADF) test and the Phillips-Perron (PP) test to establish the order of integration of the variables used. The objective was to ensure that the variables employed were stationary to avoid spurious results.

Variables	ADF Level	ADF 2 diff Level	PP Level	PP 2 diff Level
LCIT	0.8165	0.0000	0.9672	0.0000
LPIT	0.9918	0.0000	0.7903	0.0000
LVAT	0.4243	0.0000	0.4468	0.0001
LTTR	0.4099	0.0000	0.4211	0.0000

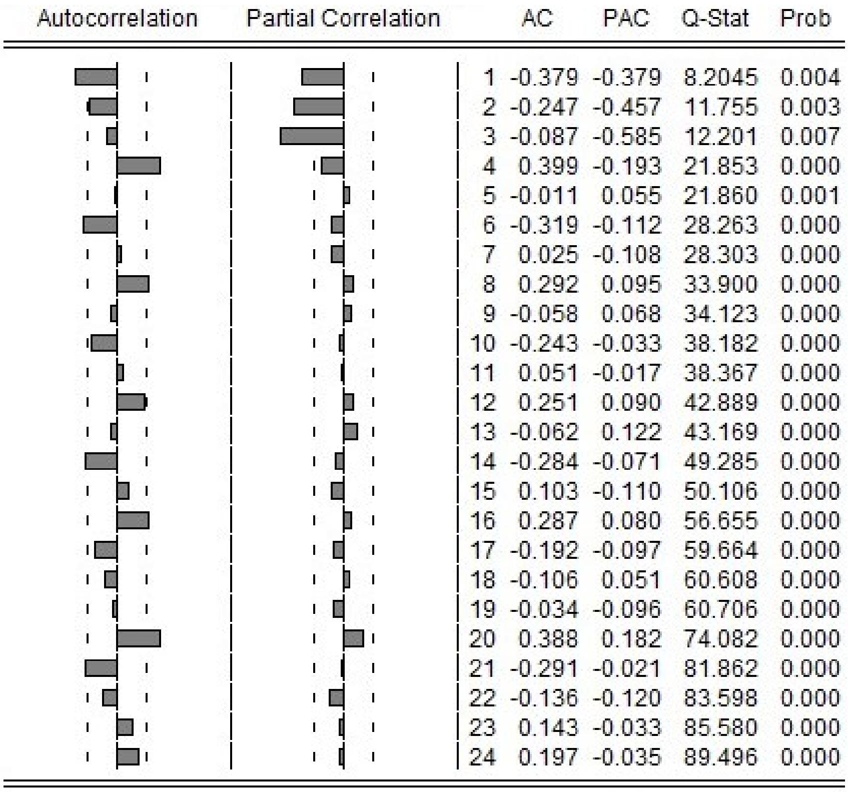
Table 4.1: Unit Root Tests Results

The corporate income tax, personal income tax, Value added tax and Total tax rate were found to be level non-stationary and thus were differenced twice to make them stationary. Table 4.1 gives a summary of the results while (Appendix A) shows the actual output.

4.5 ARIMA Model for CIT

After determining the correct order of differencing required to make the CIT series stationary, we now find an appropriate ARMA form to model the stationary CIT series. The Box-Jenkins method is the traditional and most used methodology which includes examining plots of the sample autocorrelation and partial autocorrelation from a correlogram. In addition to the Box-Jenkins methodology, several other methods of identification have been suggested in the literature.

The stationary data correlogram for DLCIT is shown in Figure 4.6. The correlogram is used to determine the parameters (p, q) of ARIMA. An AR(p) process has a PACF that lengthens at lag p while an MA(q) model has an ACF that lengthens at lag q.



Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.379	-0.379	8.2045	0.004
		2	-0.247	-0.457	11.755	0.003
		3	-0.087	-0.585	12.201	0.007
		4	0.399	-0.193	21.853	0.000
		5	-0.011	0.055	21.860	0.001
		6	-0.319	-0.112	28.263	0.000
		7	0.025	-0.108	28.303	0.000
		8	0.292	0.095	33.900	0.000
		9	-0.058	0.068	34.123	0.000
		10	-0.243	-0.033	38.182	0.000
		11	0.051	-0.017	38.367	0.000
		12	0.251	0.090	42.889	0.000
		13	-0.062	0.122	43.169	0.000
		14	-0.284	-0.071	49.285	0.000
		15	0.103	-0.110	50.106	0.000
		16	0.287	0.080	56.655	0.000
		17	-0.192	-0.097	59.664	0.000
		18	-0.106	0.051	60.608	0.000
		19	-0.034	-0.096	60.706	0.000
		20	0.388	0.182	74.082	0.000
		21	-0.291	-0.021	81.862	0.000
		22	-0.136	-0.120	83.598	0.000
		23	0.143	-0.033	85.580	0.000
		24	0.197	-0.035	89.496	0.000

Figure 4.5: Correlogram of CIT

Looking at Figure 4.5, it can be seen that the ACF is cut off at lag 4 (q=4) and the PACF is cut off at lag 3 (p=3). After determining the range of models $\{ARMA(p,q) : 0 \leqslant p \leqslant 3, 0 \leqslant q \leqslant 4\}$, the best model was selected based on AIC, HQ, and SIC. After identifying the parameters, automatic ARIMA forecasting was performed, and the best top three models are shown in Table 4.2.

ARIMA models	AIC*	SIC	HQ
ARIMA(1,2,1)	1.3363	1.4823	1.3928
ARIMA(2,2,1)	1.1406	1.2866	1.1971
ARIMA(3,2,1)	1.3637	1.5097	1.4201

Table 4.2: Best three ARIMA models based on AIC, SIC and HQ

The models are selected on the basis of AIC; the model with the smallest AIC is the best model. The appropriate model selected is ARIMA(2,2,1) with AIC of 1.1406. The competing models are ARIMA(1,2,1) with AIC of 1.3363 and ARIMA(3,2,1) with AIC of 1.3637.

Table 4.2 indicates that the best model is ARMA(2,2,1) based on AIC of 1.1406. The model output may be seen in Appendix B. The model is stationary at second difference(d=2).The identified model is given as follows:

$$D(LCIT) - D(LCIT)_{t-2} = 0.0386 - 0.5142D(LCIT)_{t-2} + \varepsilon_t - 0.4563\varepsilon_{t-1} \tag{4.1}$$

The ARIMA model for CIT is chosen based on meeting the conditions that are in line with the robustness of the model. The chosen ARIMA model is stable because the inverse roots of the characteristic polynomials are not outside the unit circle, as indicated in Figure 4.6.

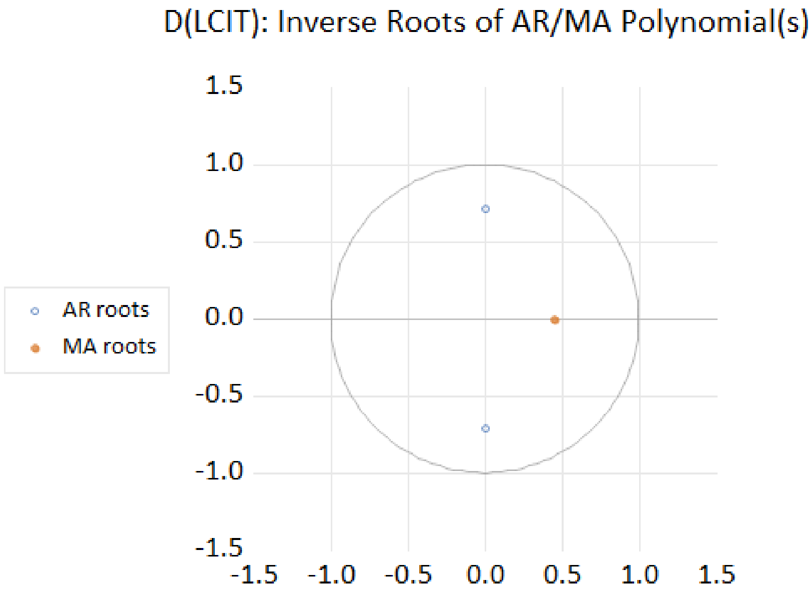


Figure 4.6: Inverse Roots of AR/MA polynomials

4.6 Diagnostic Tests of the ARIMA Model for CIT

Diagnostic tests help to check that the estimated model is statistically sound and acceptable. This is based on some statistical tests, which are performed to check whether the residuals of the models are not autocorrelated and are normally distributed. The Ljung-Box Q statistics show that the values for all 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected, and therefore there is no autocorrelation detected in the residuals series, as shown. The normality test also confirms that the residual for the ARMA (2,2,1) model follows a normal distribution, since the p-value = 0.512 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected. Therefore, the residuals follow a normal distribution as shown in Figure 4.8 below.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.040	-0.040	0.0933	
		2	0.188	0.187	2.1941	
		3	-0.089	-0.078	2.6722	0.102
		4	0.154	0.119	4.1224	0.127
		5	0.011	0.049	4.1306	0.248
		6	-0.129	-0.194	5.2002	0.267
		7	-0.011	-0.001	5.2075	0.391
		8	0.046	0.098	5.3464	0.500
		9	-0.067	-0.115	5.6496	0.581
		10	-0.098	-0.087	6.3156	0.612
		11	-0.018	0.048	6.3396	0.706
		12	0.091	0.068	6.9420	0.731
		13	-0.032	-0.042	7.0177	0.798
		14	-0.123	-0.111	8.1748	0.771
		15	0.023	0.039	8.2177	0.829
		16	0.079	0.073	8.7142	0.849
		17	-0.175	-0.235	11.236	0.736
		18	-0.080	-0.045	11.774	0.759
		19	-0.256	-0.205	17.482	0.422
		20	0.077	-0.031	18.015	0.455
		21	-0.321	-0.223	27.520	0.093
		22	-0.094	-0.120	28.354	0.101
		23	0.007	0.111	28.359	0.130
		24	0.038	-0.047	28.505	0.160

Figure 4.7: Ljung Box Q-Statistic Test

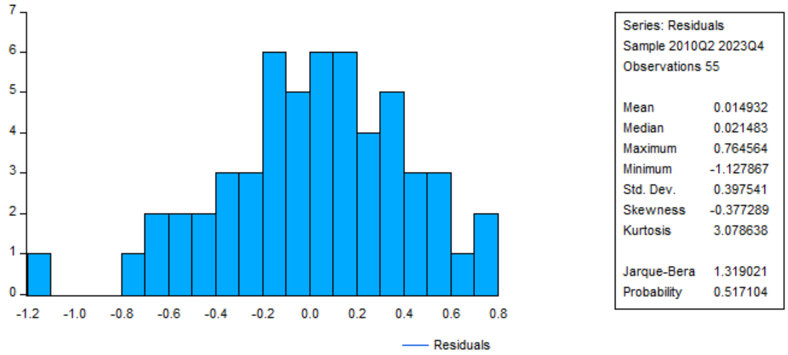


Figure 4.8: Jarque Bera test

4.7 Forecasted ARIMA Model for CIT

Figure 4.9 shows the diagrammatic representation of the quarterly actual corporate income tax rate and the predicted rates. The measures of forecast accuracy are as follows, the RMSE is 0.3833, MAE is 0.3051, MAPE is 2.1487 and Theil statistics are 0.0135.

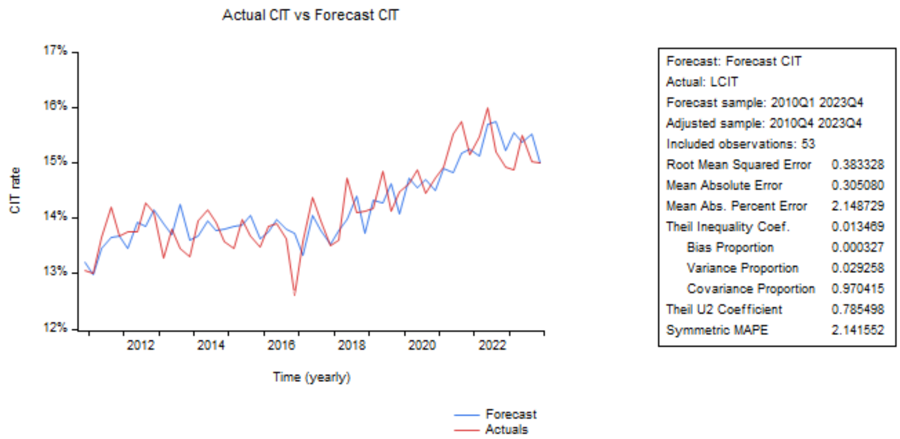


Figure 4.9: Actual CIT vs Forecast CIT

4.8 Error, Trend, Seasonal Models for CIT

The selected model based on AIC has a level smoothing parameter estimate $\alpha = 0.49$, the trend parameter is zero, indicating that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.31$. The output of the best model is shown in Appendix B. The best four ETS model based on AIC is shown in Table 4.3.

ETS Models	AIC*	SIC	HQ
ETS(A,A,A)	123.7779	139.9808	130.0598
ETS(A,M,A)	123.9811	140.1839	130.2629
ETS(A,M$_D$, A)	123.3255	141.5536	130.3925
ETS(A,A$_D$, A)	123.3225	141.5507	130.3896

Table 4.3: Best Four ETS models based on AIC, SIC and HQ

The three appropriate models that are competing as shown in Table 4.3 are ETS(A,M,A) with AIC of 123.9811; ETS(A,A, A) with AIC of 123.7779; ETS(A,M_D,A) with AIC of 123.3255 and ETS(A, A_D,A) with AIC of 123.3225. The models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The best selected model is ETS (A, A_D, A) with an AIC of 123.3225.

4.9 Forecasted ETS model for CIT

The best ETS model (A, A_D, A) is used to generate the CIT forecasts shown on the graph, as shown in Figure 4.10. Observing the graph, it can be seen that the forecast of the CIT series is closer to the actual series except in quarter three of 2016 and 2021.

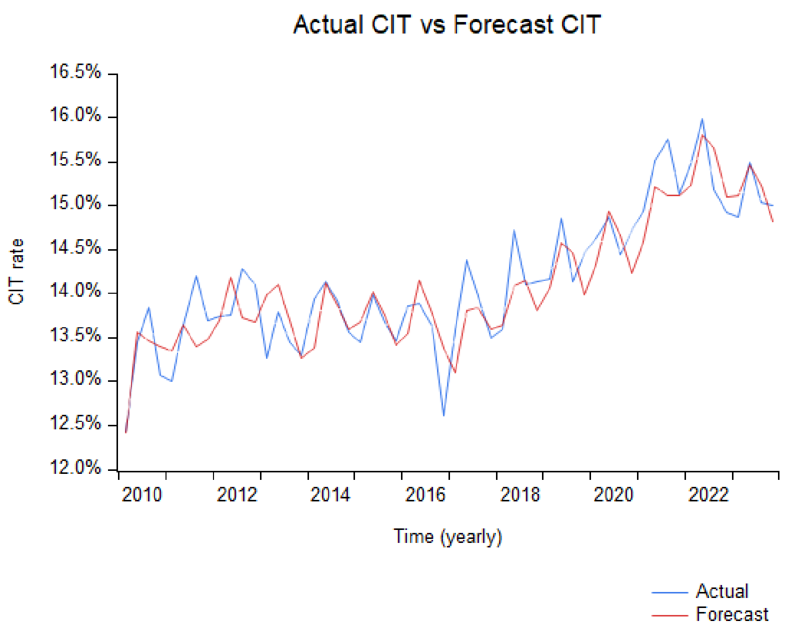


Figure 4.10: Actual CIT vs Forecast CIT

The closeness of the forecast series to the actual series suggests that the selected model has better prediction power and is appropriate for the CIT forecast. This is confirmed by the calculated precision measures, the RMSE with a value of 0.3422.

4.10 Bayesian Vector Autoregressive Model for CIT

The table 4.4 below shows the results of BVAR models for CIT using Minnesota prior, Normal-Wishart prior, and Sims-Zha prior with regard to the RMSE (the full models are shown, refer to Appendix section B).

BVAR Model	**Minnesota prior**	**Normal Whishart prior**	**Simz-Zha Prior**
RMSE	0.3571	0.3162	0.3566

Table 4.4: BVAR models based on RMSE

The best BVAR model for CIT was selected by comparing the RMSE of the out sampling forecasts accuracy, and the smallest is that of BVAR Normal Whishart prior with RMSE of 0.3162, compared to BVAR Minnesota prior and Sims-Zha prior with RMSE of 0.3571 and 0.3566 respectively.

4.11 Forecasted BVAR model for CIT

The best model used was $BVAR_{NWP}$ to generate the quarterly CIT forecasts, as represented in Figure 4.12.

Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LCIT	56	0.316173	0.243874	1.715331	0.011152
LRGDP	56	0.021486	0.016847	0.162127	0.001036
LGFCF	56	0.032608	0.024314	0.217745	0.001453
INFL	56	0.041268	0.028669	8.290509	0.055879
SMCAR	56	0.000881	0.000701	12.86047	0.083171

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure 4.11: Forecast Evaluation of $BVAR_{NWP}$

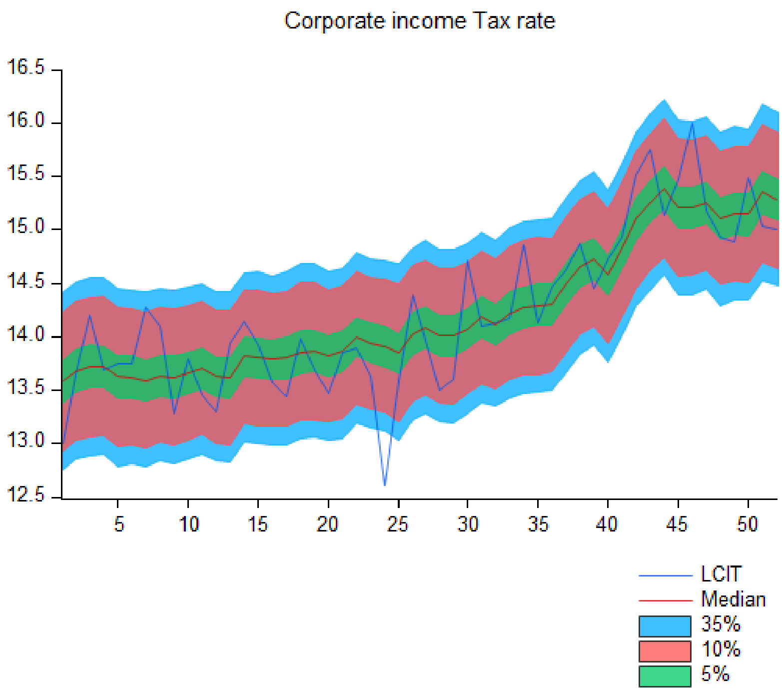


Figure 4.12: Actual CIT vs Forecast CIT

Figure 4.12, indicates that, when we compare the predicted CIT with the actual CIT. The predicted CIT appears to perform well in predicting the results up to the time period when the forecast value is within the 95% confidence interval. After that , they perform poorly when the forecast value is outside the predicted range mentioned above, since the credible intervals 95% for the output gap do not contain the observed CIT for the part of the time period.

4.12 Discussion of Corporate Income Tax Results

The appropriate ARIMA models for the CIT models were developed based on AIC, SIC, and HQ. First, the three appropriate ARIMA models were selected based on AIC. The ARIMA model (2,2,1) was found to have the lowest AIC of 1.1406 compared to the other two ARIMA models. In terms of prediction accuracy, ARIMA (2,2,1) was found to have a minimum RMSE of 0.3833. The results conclude that ARIMA(2,2,1) is the best model to fit the CIT data better compared to the other ARIMA.

The best ETS model for CIT was a specification of additive, multiplicative-dampened error, additive (A, A_D, A) with a level smoothing parameter estimate $\alpha = 0.49$, trend parameter is zero, indicating that trend components do not change from its starting value) and seasonal parameter $\gamma = 0.31$. The AIC of the (A, A_D, A) model was 123.3225 smaller than the other two competing models. The forecast evaluation was performed for the out-of-sample period starting from the second quarter of 2016 and the first quarter of 2021 and (A, A_D, A) was found to have a minimum RMSE of 0.3422 These results suggest that ETS (A, A_D, A) performs better than the other two ETS models for CIT.

The best BVAR model for CIT with three different priors was based on RMSE. The BVAR of the normal Whishart prior was the best model with RMSE of 0.3162. The results suggest that $BVAR_{NWP}$ performed better than Bayesian vector autoregression with normal Wishart prior and Bayesian vector autoregression with Sims and Zha priors.

The best final model was selected on the basis of RMSE. In conclusion $BVAR_{NWP}$ was superior to the selected ARIMA (2,2,1) and ETS (A, A_D, A) in handling the CIT data set. Therefore, it is the best method that can be used to forecast corporate income tax. Table 4.5 shows the RMSE of the three approaches.

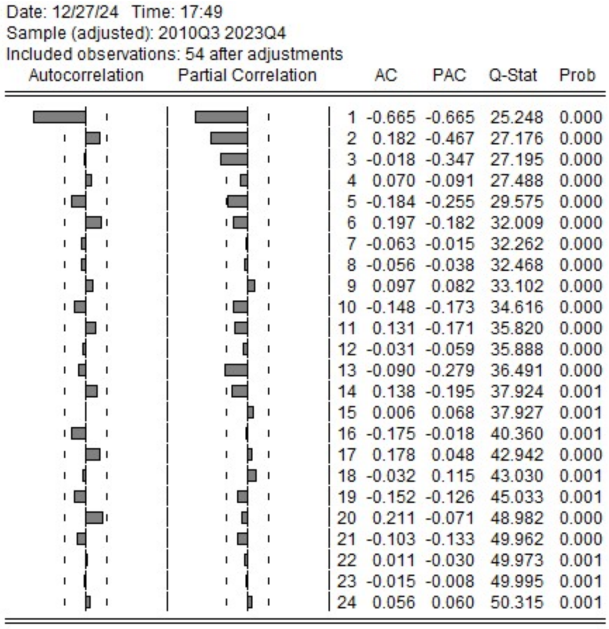
Models for CIT	**RMSE**
$BVAR_{NWP}$	0.3162
ETS(A, A_D, A)	0.3422
ARIMA(2,2,1)	0.3833

Table 4.5: RMSE for the best model of CIT

4.13 ARIMA Model for PIT

After finding the correct order of differencing required to make the PIT series stationary, we now find an appropriate ARMA form to model the stationary PIT series. The Box-Jenkins method shows plots of the sample autocorrelation and partial autocorrelation on a correlogram.

The stationary data correlogram for DLPIT is shown in Figure 4.13. The correlogram is used to determine the parameters (p, q) of ARIMA. An AR(p) process has a PACF that lengthens at lag p while an MA(q) model has an ACF that lengthens at lag q.



Date: 12/27/24 Time: 17:49
Sample (adjusted): 2010Q3 2023Q4
Included observations: 54 after adjustments

Lag	AC	PAC	Q-Stat	Prob
1	-0.665	-0.665	25.248	0.000
2	0.182	-0.467	27.176	0.000
3	-0.018	-0.347	27.195	0.000
4	0.070	-0.091	27.488	0.000
5	-0.184	-0.255	29.575	0.000
6	0.197	-0.182	32.009	0.000
7	-0.063	-0.015	32.262	0.000
8	-0.056	-0.038	32.468	0.000
9	0.097	0.082	33.102	0.000
10	-0.148	-0.173	34.616	0.000
11	0.131	-0.171	35.820	0.000
12	-0.031	-0.059	35.888	0.000
13	-0.090	-0.279	36.491	0.000
14	0.138	-0.195	37.924	0.001
15	0.006	0.068	37.927	0.001
16	-0.175	-0.018	40.360	0.001
17	0.178	0.048	42.942	0.000
18	-0.032	0.115	43.030	0.001
19	-0.152	-0.126	45.033	0.001
20	0.211	-0.071	48.982	0.000
21	-0.103	-0.133	49.962	0.000
22	0.011	-0.030	49.973	0.001
23	-0.015	-0.008	49.995	0.001
24	0.056	0.060	50.315	0.001

Figure 4.13: Correlogram of PIT

Examining Figure 4.13, it can be observed that ACF is cut off at lag 1 (q = 1) and PACF is cut off at lag 5 (p=5). After determining the range of models $\{ARMA(p, q) : 0 \leqslant p \leqslant 5, 0 \leqslant q \leqslant 1\}$, the best model was selected based on AIC,HQ, and SIC. After identifying the parameters, automatic ARIMA forecasting was performed, and the best top three models are shown in Table 4.6.

ARIMA models	**AIC***	**SIC**	**HQ**
ARIMA(1,2,1)	-0.2871	-0.1411	-0.2306
ARIMA(2,2,1)	-0.2857	-0.1397	-0.2292
ARIMA(3,2,1)	-0.2862	-0.1402	-0.2297
ARIMA(5,2,1)	-0.3003	-0.1543	-0.2438

Table 4.6: Best Four ARIMA models based on AIC, SIC and HQ

The models are selected on the basis of AIC; the model with the smallest AIC is the best model. The appropriate model selected is ARIMA(5,2,1) with AIC of -0.3003. The competing models are ARIMA(1,2,1) with AIC of -0.2871 ,ARIMA(3,2,1) with AIC of -0.2862, and ARIMA (2,2,1) with AIC of -0.2857

Table 4.6 indicates that the best model is ARMA(5,2,1) based on AIC of -0.3003. The model output can be seen in Appendix B. The model is stationary at second difference(d=2).The identified model is given as follows:

$$D(LPIT) - D(LPIT)_{t-2} = 0.0399 + \varepsilon_t - 0.3623\varepsilon_{t-1} \tag{4.2}$$

The ARIMA model for PIT is chosen based on conditions that are in line with model robustness. The chosen ARIMA model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle, as indicated in Figure 4.14.

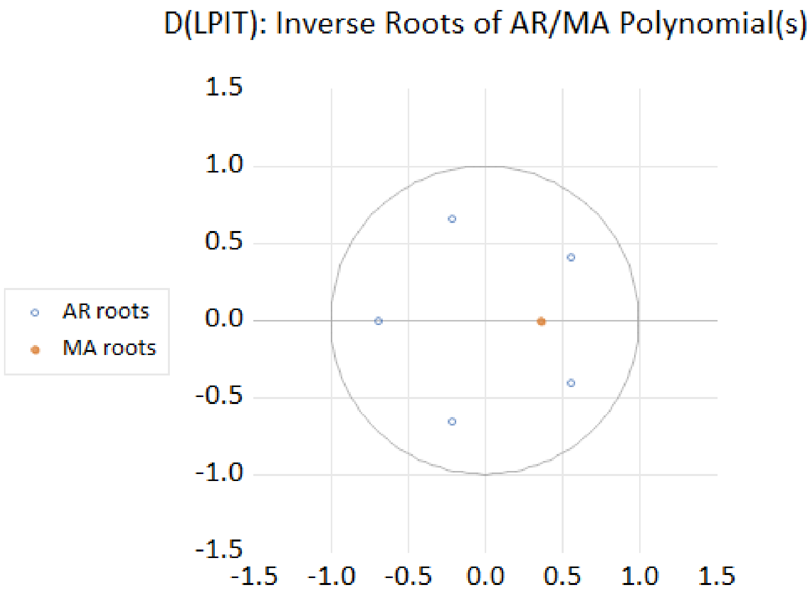


Figure 4.14: Inverse Roots of AR/MA polynomials

4.14 Diagnostic Tests of the ARIMA Model for PIT

The Ljung-Box Q statistics show that the values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected; therefore, no autocorrelation is detected in the residual series. The normality test also confirms that the residual for the ARMA (5,2,1) model follows a normal distribution, since the p-value = 0.2660 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected. Therefore, the residuals follow a normal distribution as shown in Figure 4.16 below.

Sample (adjusted): 2010Q2 2023Q4
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.028	-0.028	0.0470	
		2	0.061	0.060	0.2669	
		3	0.026	0.029	0.3071	0.579
		4	-0.036	-0.039	0.3882	0.824
		5	-0.022	-0.027	0.4173	0.937
		6	0.092	0.096	0.9627	0.915
		7	-0.072	-0.063	1.3013	0.935
		8	-0.187	-0.206	3.6244	0.727
		9	-0.098	-0.113	4.2782	0.747
		10	-0.214	-0.197	7.4650	0.487
		11	0.018	0.017	7.4868	0.587
		12	0.012	0.016	7.4971	0.678
		13	-0.036	-0.034	7.5922	0.749
		14	0.140	0.165	9.0816	0.696
		15	0.038	0.058	9.1972	0.758
		16	-0.120	-0.155	10.347	0.736
		17	0.101	0.002	11.180	0.740
		18	-0.006	-0.087	11.184	0.798
		19	-0.050	-0.099	11.404	0.835
		20	0.199	0.152	14.963	0.664
		21	0.008	0.052	14.969	0.725
		22	-0.022	0.069	15.016	0.775
		23	-0.043	-0.033	15.200	0.813
		24	-0.072	-0.075	15.731	0.829

Figure 4.15: Ljung Box Q-Statistic Test

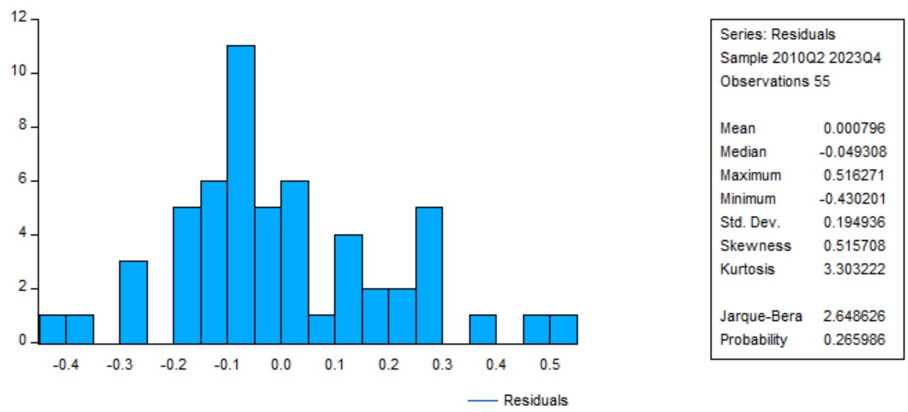


Figure 4.16: Jarque Bera test

4.15 Forecasted ARIMA Model for PIT

Figure 4.17 shows the quarterly actual corporate income tax rate and its forecast rates. The forecast accuracy measures are as follows: RMSE is 0.1853, MAE is 0.1383, MAPE is 0.9372 and Theil statistics are 0.0063.

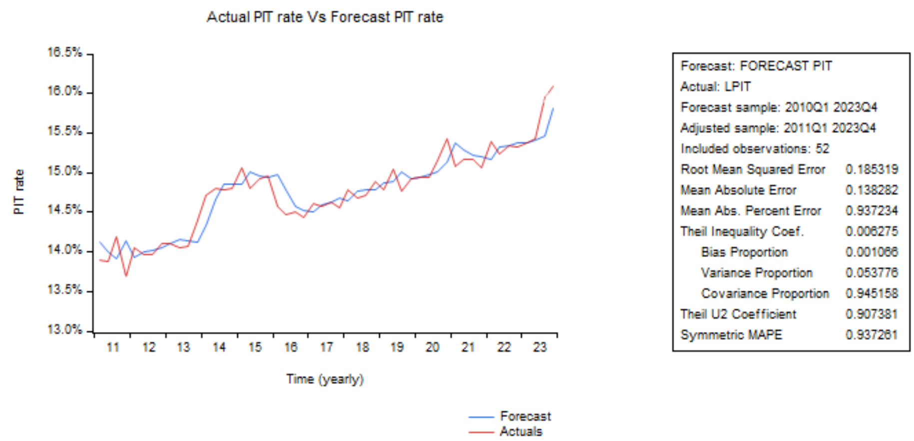


Figure 4.17: Actual PIT vs Forecast PIT

4.16 Error, Trend, Seasonal Models for PIT

The appropriate model based on AIC has a level smoothing parameter estimate $\alpha = 0.65$, the trend parameter is zero, indicating that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0$. The output of the best model is shown in Appendix B. The best four ETS model based on AIC is shown in Table 4.7.

ETS Models	AIC*	SIC	HQ
ETS(A,A,A)	54.5607	70.7635	60.8425
ETS(A,M,A)	54.5771	70.7799	60.8589
ETS(A,M_D, A)	56.5771	74.8053	63.6442
ETS(A,A_D, A)	56.5607	74.7889	63.6277

Table 4.7: Best Four ETS models based on AIC, SIC and HQ

The three appropriate models that are competing as shown in Table 4.7 are ETS(A,M,A) with AIC of 54.5771; ETS(A, A_D,A) with AIC of 56.5607; ETS(A,M_D,A) with AIC of 56.5771 and ETS(A,A, A) with AIC of 54.5607. The models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The best selected model is ETS(A,A,A) with AIC of 54.5607.

4.17 Forecasted ETS model for PIT

The best ETS model (A, M, A) is used to generate the PIT forecasts plotted on the graph as shown in Figure 4.18. Observing the graph, it shows that the forecast of the PIT series is closer to the actual series except in the second quarter of 2014 and the fourth quarter of 2021.

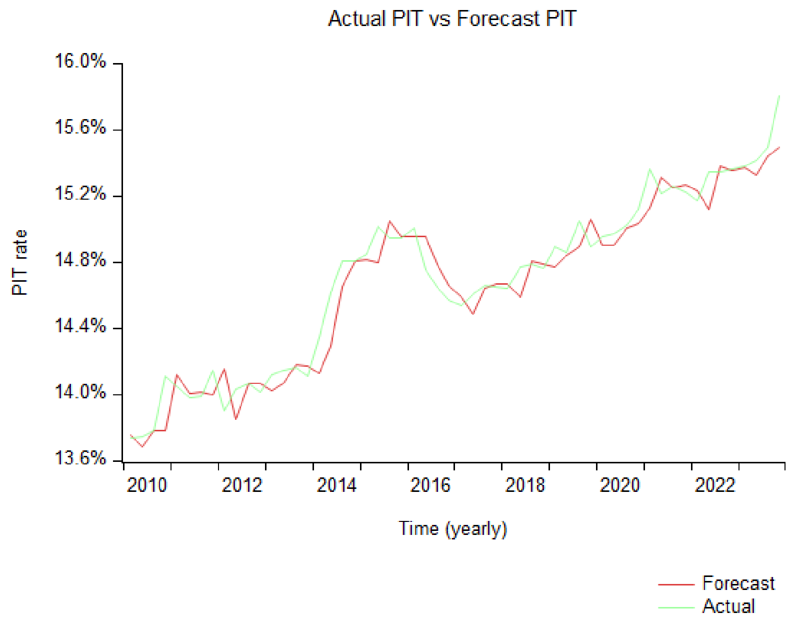


Figure 4.18: Actual PIT vs Forecast PIT

The closeness of the predicted series to the actual series suggests that the selected model has better prediction power and is appropriate for the PIT forecast. This is confirmed by the calculated precision measures, the RMSE with a value of 0.1886.

4.18 Bayesian Vector Autoregressive Model for PIT

The table 4.8 below indicts the results of appropriate BVAR models for PIT using Minnesota prior,Normal-Wishart prior and Sims-Zha prior with regard to the RMSE (the full models are shown refer to Appendix section C)

BVAR Model	Minnesota prioe	Normal Whishart prior	Simz-Zha Prior
RMSE	0.1812	0.1638	0.1810

Table 4.8: BVAR models based on RMSE

The best BVAR model for PIT was selected by comparing the RMSE of the out sampling forecasts accuracy, and the smallest is that of BVAR Normal Whishart prior with RMSE of

0.1638, compared to BVAR Minnesota prior and Sims-Zha prior with RMSE of 0.1812 and 0.1810 respectively.

4.19 Forecasted BVAR model for PIT

The best model used was $BVAR_{NWP}$ to generate the quarterly forecasts of PIT as represented in Figure 4.20.

Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LPIT	56	0.163804	0.121556	0.822864	0.005570
LRGDP	56	0.022587	0.017634	0.169639	0.001089
INFL	56	0.041716	0.030130	9.048963	0.056501
LGFCF	56	0.032532	0.025120	0.224920	0.001450
SMCAR	56	0.000745	0.000559	10.36945	0.070197

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure 4.19: Forecast Evaluation of $BVAR_{NWP}$

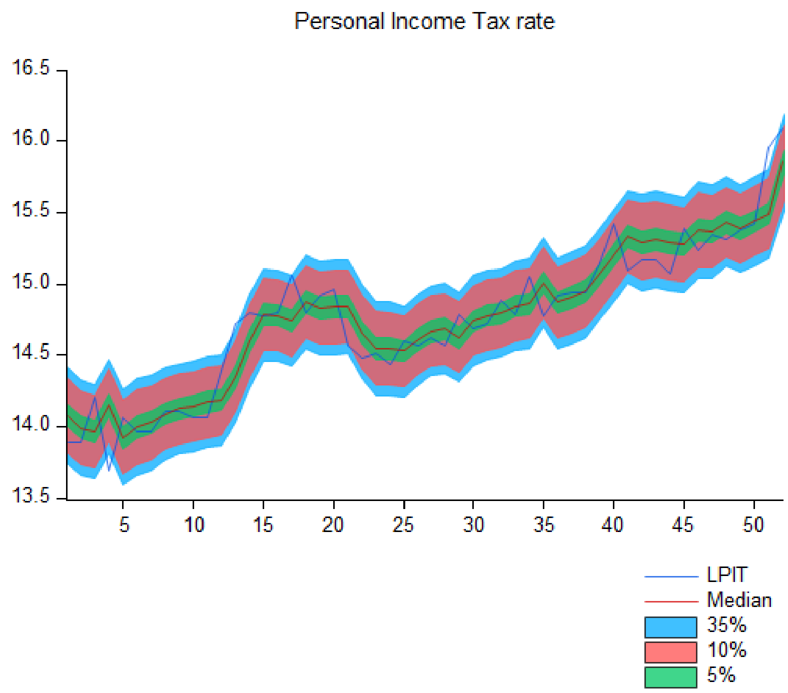


Figure 4.20: Actual PIT vs Forecast PIT

Figure 4.20, indicates that, when we compare the predicted PIT with the actual PIT. The predicted PIT appears to perform well in predicting the results up to the time period when the forecast value is within the 95% confidence interval. After that , they perform poorly when the forecast value is outside the predicted range mentioned above, since the credible intervals 95% for the output gap do not contain the observed PIT for the part of the time period.

4.20 Discussion of Personal Income Tax Results

The appropriate ARIMA model for PIT models was developed on the basis of AIC, SIC, and HQ. First, the four appropriate ARIMA models were selected based on AIC. The ARIMA (5,2,1) model was found to have the lowest AIC of -0.3003 compared to the other three ARIMA models. In terms of forecast accuracy, ARIMA (5,2,1) was found to have a minimum RMSE of 0.1853. The results conclude that ARIMA(5,2,1) is the best model to fit PIT data better compared to the other ARIMA.

The best ETS model for PIT was a specification of additive, additive and additive ETS (A, A, A) with level smoothing parameter estimate $\alpha = 0.65$, trend parameter is zero indicating that the trend components do not change from its starting value) and seasonal parameter $\gamma = 0.0$. The AIC of the ETS(A,A, A) model was 54.5607 smaller than the other three competing models. The forecast evaluation was performed for an out-of-sample period starting from the second quarter of 2014 and the fourth quarter of 2021, and ETS (A,A,A) was found to have a minimum RMSE of 0.1886 These results suggest that ETS(A,A,A) performs better than the other three ETS models for PIT.

The best BVAR model for PIT with three different priors was based on RMSE. The BVAR of the normal Whishart prior was the best model with a RMSE of 0.1638. The results suggest that $BVAR_{NWP}$ performed better than Bayesian vector autoregression with normal Wishart prior and Bayesian vector autoregression with Sims and Zha priors.

The best final model was selected on the basis of RMSE. In conclusion $BVAR_{NWP}$ was superior to the selected ARIMA (5,2,1) and ETS (A, A, A) in handling the PIT data set, henceforth, it is the best method that can be used to forecast personal income tax. Table 4.9 shows the RMSE of the three approaches.

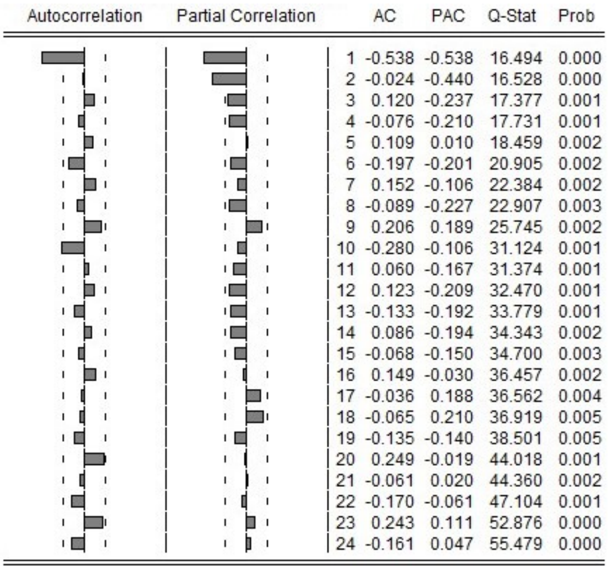
Models for PIT	**RMSE**
$BVAR_{NWP}$	0.1638
ARIMA(5,2,1)	0.1853
ETS(A,A,A)	0.1886

Table 4.9: RMSE for the best model of PIT

4.21 ARIMA Model for VAT

After determining the correct order of differencing needed to make the VAT series stationary, we now find an appropriate ARMA form to model the stationary VAT series. The Box-Jenkins method shows plots of the sample autocorrelation and partial autocorrelation on a correlogram.

The stationary data correlogram FOR DLVAT is shown in Figure 4.21. The correlogram is used to determine the parameters (p, q) of ARIMA. An AR(p) process has a PACF that lengthens at lag p while an MA(q) model has an ACF that lengthens at lag q.



Lag	AC	PAC	Q-Stat	Prob
1	-0.538	-0.538	16.494	0.000
2	-0.024	-0.440	16.528	0.000
3	0.120	-0.237	17.377	0.001
4	-0.076	-0.210	17.731	0.001
5	0.109	0.010	18.459	0.002
6	-0.197	-0.201	20.905	0.002
7	0.152	-0.106	22.384	0.002
8	-0.089	-0.227	22.907	0.003
9	0.206	0.189	25.745	0.002
10	-0.280	-0.106	31.124	0.001
11	0.060	-0.167	31.374	0.001
12	0.123	-0.209	32.470	0.001
13	-0.133	-0.192	33.779	0.001
14	0.086	-0.194	34.343	0.002
15	-0.068	-0.150	34.700	0.003
16	0.149	-0.030	36.457	0.002
17	-0.036	0.188	36.562	0.004
18	-0.065	0.210	36.919	0.005
19	-0.135	-0.140	38.501	0.005
20	0.249	-0.019	44.018	0.001
21	-0.061	0.020	44.360	0.002
22	-0.170	-0.061	47.104	0.001
23	0.243	0.111	52.876	0.000
24	-0.161	0.047	55.479	0.000

Figure 4.21: Correlogram of VAT

Figure 4.21 shows that ACF terminates at lag 10 (q = 10) and PACF terminates at lag 2 (p=2). After determining the range of models $\{ARMA(p, q) : 0 \leqslant p \leqslant 2, 0 \leqslant q \leqslant 10\}$, the best model was selected based on AIC,HQ, and SIC. After identifying the parameters, automatic ARIMA forecasting was performed, and the best top three models are shown in Table 4.10.

ARIMA models	AIC*	SIC	HQ
ARIMA(1,2,10)	-0.0909	0.0550	-0.0345
ARIMA(2,2,10)	-0.0644	0.08156	-0.0079

Table 4.10: Best two ARIMA models based on AIC, SIC and HQ

The appropriate models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The appropriate model selected is ARIMA(1,2,10) with AIC of -0.0909.The competing models ARIMA(2,2,10) with AIC of -0.0644.

Table 4.10 shows that the best model is ARMA(1,2,10) based on AIC of -0.0909. The model output can be seen in Appendix B. The model is stationary at second difference(d=2).The identified model is given as follows:

$$D(LVAT) - D(LVAT)_{t-2} = 0.0499 - 0.3358D(LVAT)_{t-10} + \varepsilon_t \tag{4.3}$$

The ARIMA model for VAT is chosen on the basis of conditions that are in line with the robustness of the model. The chosen ARIMA model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle, as indicated in Figure 4.22.

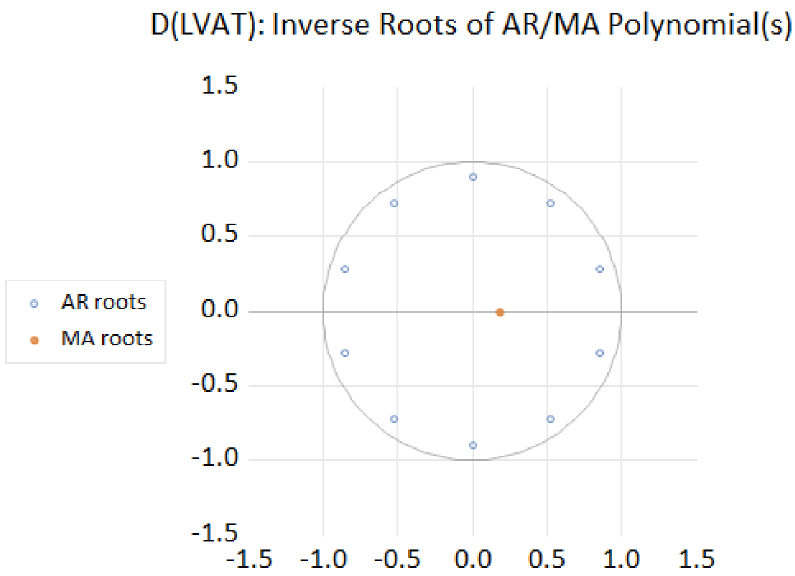


Figure 4.22: Inverse Roots of AR/MA polynomials

4.22 Diagnostic Tests of the ARIMA Model for VAT

The Ljung-Box Q statistics show that the values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected; therefore, no autocorrelation is detected in the residual series. The normality test also confirms that the residual for the ARMA (1,2,10) model follows a normal distribution, since the p-value = 0.7080 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected. Therefore, the residuals follow a normal distribution as shown in Figure 4.24 below.

Sample (adjusted): 2010Q2 2023Q4
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.000	0.000	2.E-09	
		2	-0.039	-0.039	0.0887	
		3	0.097	0.097	0.6544	0.419
		4	-0.074	-0.076	0.9886	0.610
		5	-0.025	-0.017	1.0293	0.794
		6	-0.184	-0.203	3.2055	0.524
		7	-0.016	0.000	3.2218	0.666
		8	-0.152	-0.182	4.7550	0.576
		9	-0.033	0.007	4.8274	0.681
		10	0.014	-0.043	4.8411	0.774
		11	-0.078	-0.057	5.2793	0.809
		12	0.108	0.046	6.1370	0.804
		13	0.048	0.031	6.3102	0.852
		14	0.177	0.152	8.6992	0.728
		15	0.150	0.132	10.460	0.656
		16	0.130	0.157	11.812	0.621
		17	-0.008	-0.037	11.817	0.693
		18	-0.219	-0.193	15.876	0.462
		19	-0.170	-0.243	18.385	0.365
		20	0.098	0.188	19.252	0.376
		21	-0.122	-0.048	20.631	0.358
		22	-0.121	0.015	22.017	0.340
		23	0.073	0.047	22.538	0.369
		24	-0.075	-0.094	23.102	0.396

Figure 4.23: Ljung Box Q-Statistic Test

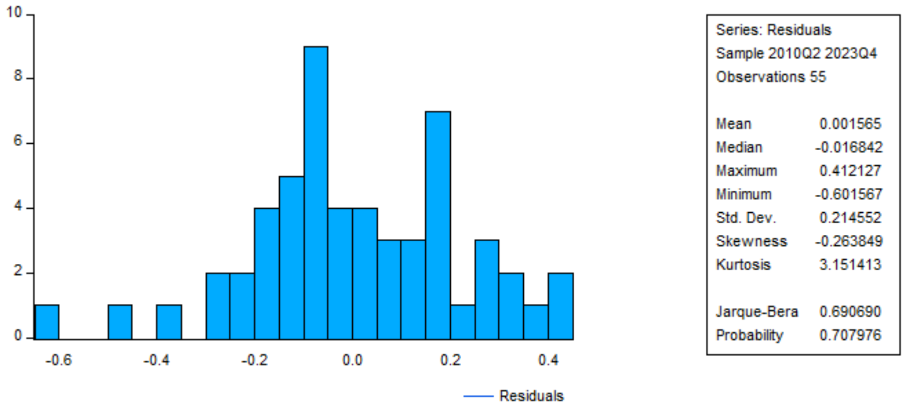


Figure 4.24: Jarque Bera test

4.23 Forecasted ARIMA Model for VAT

Figure 4.25 shows the quarterly actual value added tax rate and its forecast rates. The forecast accuracy measures are as follows: RMSE is 0.2150, MAE is 0.1701, MAPE is 1.1325 and Theil statistics are 0.0071.

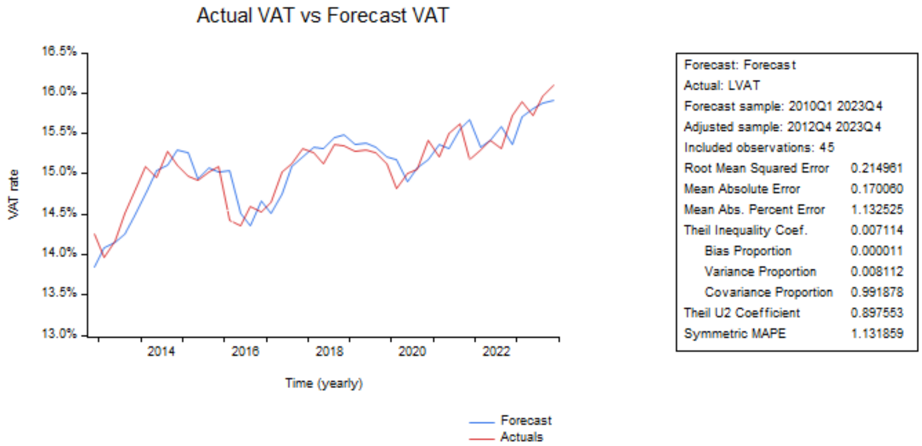


Figure 4.25: Actual VAT vs Forecast VAT

4.24 Error, Trend, Seasonal Models for VAT

The appropriate model based on AIC has a level smoothing parameter estimate $\alpha = 0.83$, the trend parameter is zero, indicating that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0$. The output of the best model is shown in Appendix B. The best four ETS model based on AIC is shown in Table 4.11.

ETS Models	**AIC***	**SIC**	**HQ**
ETS(A,A,A)	64.1071	80.3099	70.3889
ETS(A,M,A)	64.3907	80.5935	70.6725
ETS(A,M_D, A)	65.9336	84.1617	73.0006
ETS(A,A_D, A)	65.7424	83.9706	72.8094

Table 4.11: Best Four ETS models based on AIC, SIC and HQ

The three appropriate models that are competing as shown in Table 4.11 are ETS(A,M,A) with AIC of 64.3907; ETS(A, A_D,A) with AIC of 65.7424; ETS(A,M_D,A) with AIC of 65.9336 and ETS(A,A, A) with AIC of 64.1071. The models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The best selected model is ETS(A,A,A) with AIC of 64.1071.

4.25 Forecasted ETS model for VAT

The best ETS model (A, M, A) is used to generate the PIT forecasts plotted on the graph as shown in Figure 4.26. Observing the graph, it shows that the PIT series forecast is closer to the actual series, except in the third quarter of 2013 and the second quarter of 2016.

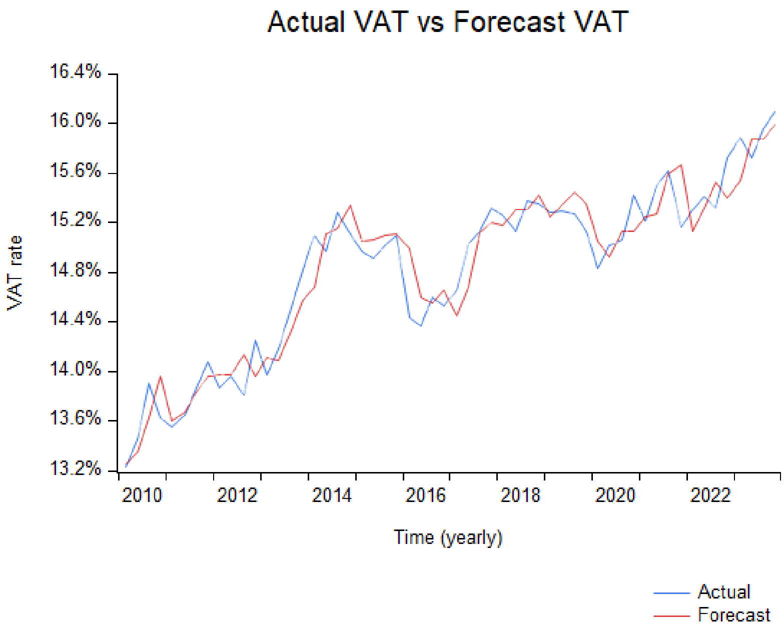


Figure 4.26: Actual VAT vs Forecast VAT

The closeness of the forecast series to the actual series suggests that the selected model has better prediction power and is appropriate for the forecast of VAT. This is confirmed by the calculated precision measures, the RMSE with a value of 0.2053.

4.26 Bayesian Vector Autoregressive Model for VAT

Table 4.12 below indicts the results of appropriate BVAR models for VAT using Minnesota prior, Normal-Wishart prior, and Sims-Zha prior with regard to the RMSE (the full models are shown; refer to Appendix Section D).

BVAR Model	**Minnesota prioe**	**Normal Whishart prior**	**Simz-Zha Prior**
RMSE	0.2150	0.1665	0.2151

Table 4.12: BVAR models based on RMSE

The best BVAR model for VAT was selected by comparing the RMSE of the out sampling forecasts accuracy, and the smallest is that of BVAR Normal Whishart prior with RMSE of

0.1665, compared to BVAR Minnesota prior and Sims-Zha prior with RMSE of 0.2150 and 0.2151 respectively.

4.27 Forecasted BVAR model for VAT

The best model used was $BVAR_{NWP}$ to generate the quarterly VAT forecasts as represented by Figure 4.28.

Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LVAT	56	0.166471	0.128442	0.865180	0.005606
LGFCF	56	0.032774	0.024760	0.221470	0.001461
LRGDP	56	0.024672	0.019587	0.188350	0.001189
INFL	56	0.043633	0.028865	8.578180	0.059127
SMCAR	56	0.000763	0.000628	11.65416	0.072250

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure 4.27: Forecast Evaluation of $BVAR_{NWP}$

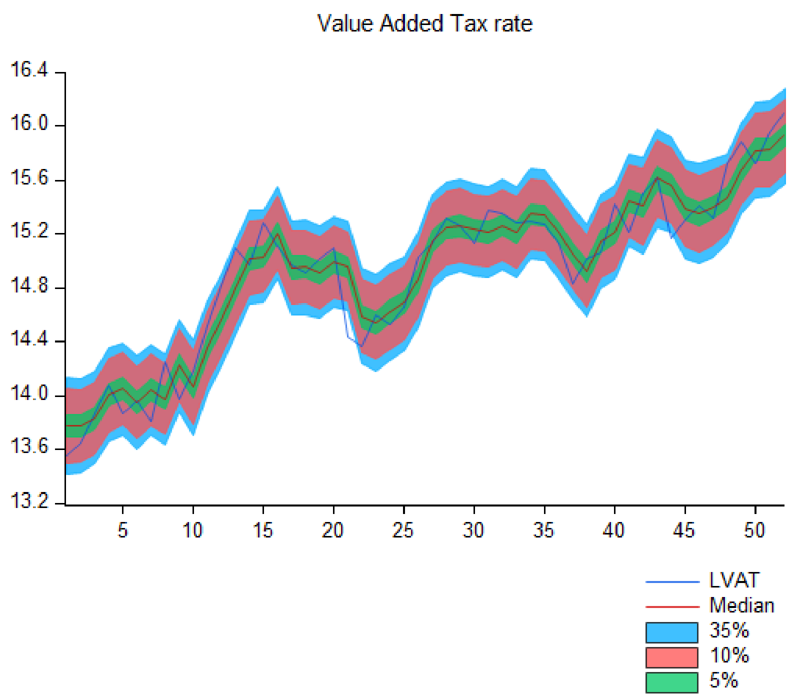


Figure 4.28: Actual VAT vs Forecast VAT

Figure 4.28, indicates that, when we compare the predicted VAT with the actual VAT. The predicted VAT appears to perform well in predicting the results up to the time period when the forecast value is within the 95% confidence interval. After that , they perform poorly when the forecast value is outside the predicted range mentioned above, since the credible intervals 95% for the output gap do not contain the VAT observed for the part of the time period.

4.28 Discussion of Value added Tax Results

The appropriate ARIMA model for VAT models was developed based on AIC, SIC and HQ. First, the two appropriate ARIMA models were selected based on AIC. The ARIMA (1,2,10) model was found to have the lowest AIC of -0.0909 compared to the ARIMA (2,2,10) model. In terms of forecast accuracy, ARIMA (1,2,10) was found to have a minimum RMSE of 0.2150. The results conclude that ARIMA(1,2,10) is the best model to fit VAT data better compared to the other ARIMA.

The best ETS model for VAT was a specification of Additive, Additive, and Additive ETS (A, A, A) with a level smoothing parameter estimate $\alpha = 0.83$, trend parameter is zero indicating that trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.0$. The AIC of the ETS(A,A, A) model was 64.1071 smaller than the other three competing models. The forecast evaluation was performed for the out-of-sample period starting from the third quarter of 2013 and the second quarter of 2016 and it was found that ETS (A,A,A) had a minimum RMSE of 0.2053 These results suggest that ETS(A,A,A) performs better than the other three ETS models for VAT.

The best BVAR model for VAT with three different priors was based on RMSE. The BVAR with normal Whishart prior was the best model with RMSE of 0.1665. The results suggest that $BVAR_{NWP}$ performed better than Bayesian vector autoregression with normal Wishart prior and Bayesian vector autoregression with Sims and Zha priors.

The best final model was selected on the basis of RMSE. In conclusion $BVAR_{NWP}$ was superior to the selected ARIMA (1,2,10) and ETS (1,A, A) in handling the VAT data set, henceforth, it is the best method that can be used to forecast personal income tax. Table 4.9 shows the RMSE of the three approaches.

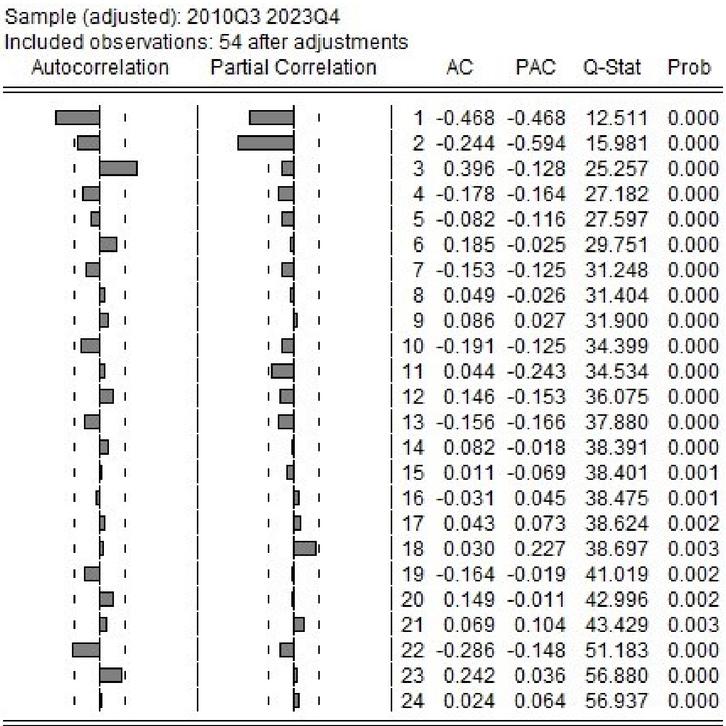
Models for VAT	**RMSE**
$BVAR_{NWP}$	0.1665
ETS(A,A,A)	0.2053
ARIMA(1,2,10)	0.2150

Table 4.13: RMSE for the best model of VAT

4.29 ARIMA Model for TTR

After finding the correct order of differencing needed to make the TTR series stationary, we now find an appropriate ARMA form to model the stationary TTR series. The Box-Jenkins method shows plots of the sample autocorrelation and partial autocorrelation on a correlogram.

The stationary data correlogram FOR DLTTR is shown in Figure 4.29. The correlogram is used to determine the parameters (p, q) of ARIMA. An AR(p) process has a PACF that lengthens at lag p while an MA(q) model has an ACF that lengthens at lag q.



Sample (adjusted): 2010Q3 2023Q4
Included observations: 54 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.468	-0.468	12.511	0.000
		2	-0.244	-0.594	15.981	0.000
		3	0.396	-0.128	25.257	0.000
		4	-0.178	-0.164	27.182	0.000
		5	-0.082	-0.116	27.597	0.000
		6	0.185	-0.025	29.751	0.000
		7	-0.153	-0.125	31.248	0.000
		8	0.049	-0.026	31.404	0.000
		9	0.086	0.027	31.900	0.000
		10	-0.191	-0.125	34.399	0.000
		11	0.044	-0.243	34.534	0.000
		12	0.146	-0.153	36.075	0.000
		13	-0.156	-0.166	37.880	0.000
		14	0.082	-0.018	38.391	0.000
		15	0.011	-0.069	38.401	0.001
		16	-0.031	0.045	38.475	0.001
		17	0.043	0.073	38.624	0.002
		18	0.030	0.227	38.697	0.003
		19	-0.164	-0.019	41.019	0.002
		20	0.149	-0.011	42.996	0.002
		21	0.069	0.104	43.429	0.003
		22	-0.286	-0.148	51.183	0.000
		23	0.242	0.036	56.880	0.000
		24	0.024	0.064	56.937	0.000

Figure 4.29: Correlogram of TTR

Figure 4.29 indicates that ACF terminates at lag 22 (q = 22) and PACF terminates at lag 2 (p=2). After determining the range of models $\{ARMA(p,q) : 0 \leqslant p \leqslant 2, 0 \leqslant q \leqslant 22\}$, the best model was selected based on AIC,HQ, and SIC. After identifying the parameters, automatic ARIMA forecasting was performed, and the best top three models are shown in Table 4.14.

ARIMA models	**AIC***	**SIC**	**HQ**
ARIMA(1,2,1)	-0.8275	-0.6815	-0.7710
ARIMA(2,2,3)	-0.8673	-0.7212	-0.8108
ARIMA(1,2,22)	-0.8929	-0.7469	-0.8365

Table 4.14: Best two ARIMA models based on AIC, SIC and HQ

The appropriate models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The appropriate model selected is ARIMA(1,2,22) with AIC of -0.8929.The competing models are ARIMA(1,2,1) with AIC of -0.8275 and ARIMA(2,2,3) with AIC of -0.8673.

Table 4.14 shows that the best model is ARMA (1,2,22) based on an AIC of -0.8929. The model output can be seen in Appendix B. The model is stationary at second difference(d=2).The identified model is given as follows:

$$D(LTTR) - D(LTTR)_{t-2} = 0.0385 - 0.4179D(LTTR)_{t-22} + \varepsilon_t \tag{4.4}$$

The ARIMA model for TTR is chosen based on conditions that are in line with model robustness. The chosen ARIMA model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle, as indicated in Figure 4.30.

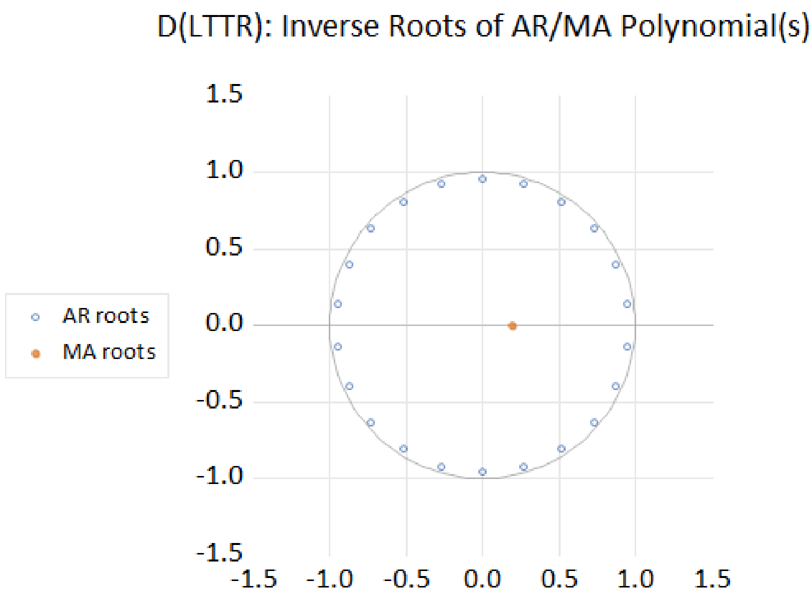


Figure 4.30: Inverse Roots of AR/MA polynomials

4.30 Diagnostic Tests of the ARIMA Model for TTR

The Ljung-Box Q statistics show that the values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected; therefore, no autocorrelation is detected in the residual series. The normality test also confirms that the residual for the ARMA (1,2,22) model follows a normal distribution, since the p-value = 0.0.2535 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected.

Therefore, the residuals follow a normal distribution as shown in Figure 4.32 below.

Sample (adjusted): 2010Q2 2023Q4
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.042	0.042	0.1037	
		2	-0.202	-0.204	2.5217	
		3	0.140	0.166	3.7073	0.054
		4	-0.104	-0.179	4.3729	0.112
		5	-0.105	-0.015	5.0649	0.167
		6	0.029	-0.050	5.1188	0.275
		7	-0.123	-0.119	6.1056	0.296
		8	0.003	0.030	6.1062	0.411
		9	-0.097	-0.200	6.7429	0.456
		10	-0.194	-0.143	9.3567	0.313
		11	-0.029	-0.130	9.4156	0.400
		12	0.100	0.051	10.140	0.428
		13	0.072	0.044	10.527	0.484
		14	0.152	0.128	12.304	0.422
		15	0.068	0.026	12.667	0.474
		16	0.069	0.104	13.049	0.523
		17	0.023	-0.008	13.092	0.595
		18	-0.089	-0.056	13.759	0.617
		19	-0.112	-0.138	14.851	0.606
		20	-0.052	-0.113	15.093	0.656
		21	0.011	0.031	15.104	0.716
		22	0.055	0.093	15.395	0.753
		23	0.020	0.140	15.436	0.800
		24	-0.071	-0.013	15.948	0.818

Figure 4.31: Ljung Box Q-Statistic Test

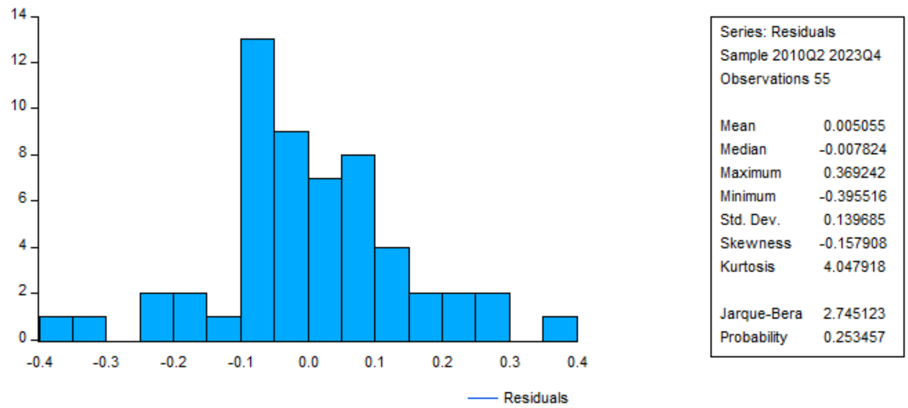


Figure 4.32: Jarque Bera test

4.31 Forecasted ARIMA Model for TTR

Figure 4.25 shows the quarterly actual value added tax rate and its forecast rates. The forecast accuracy measures are as follows: RMSE is 0.1359, MAE is 0.1006, MAPE is 0.6156 and Theil statistics are 0.0041.

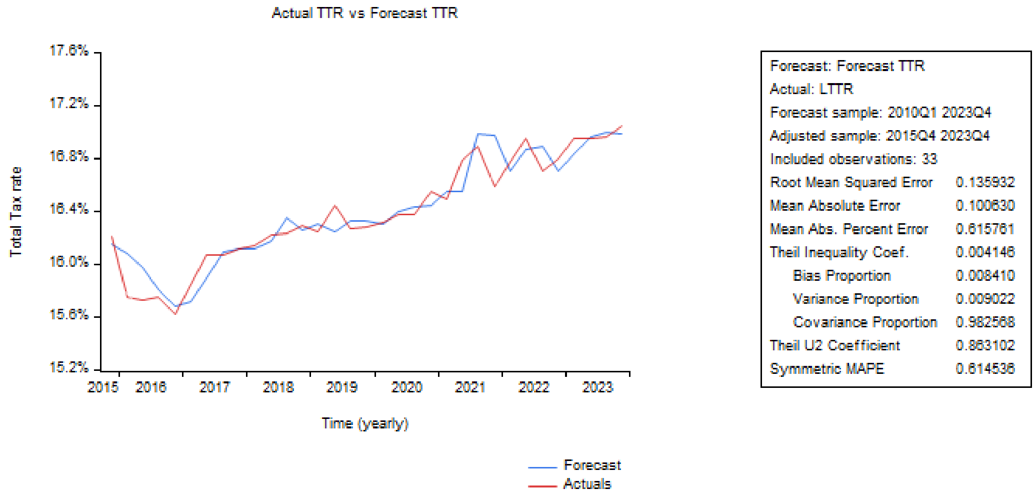


Figure 4.33: Actual TTR vs Forecast TTR

4.32 Error, Trend, Seasonal Models for TTR

The appropriate model based on AIC has a level smoothing parameter estimate $\alpha = 0.82$, the trend parameter is zero, indicating that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0$. The output of the best model is shown in Appendix B. The best four ETS model based on AIC is shown in Table 4.15.

ETS Models	AIC*	SIC	HQ
ETS(A,A,A)	24.8911	41.0939	31.1729
ETS(A,M,A)	25.1562	41.359	31.4380
ETS(A,M$_D, A$)	26.4693	44.6976	33.5364
ETS(A,A$_D, A$)	26.2758	44.5040	33.3428

Table 4.15: Best Four ETS models based on AIC, SIC and HQ

The three appropriate models that are competing as shown in Table 4.11 are ETS(A,M,A) with AIC of 25.1562; ETS(A,A_D,A) with AIC of 26.2758; ETS(A,M_D,A) with AIC of 26.4693 and ETS(A,A, A) with AIC of 24.8911. The models are selected on the basis of the AIC; the model with the smallest AIC is the best model. The best selected model is ETS (A, A, A) with an AIC of 24.8911.

4.33 Forecasted ETS model for TTR

The best ETS model (A, M, A) is used to generate the PIT forecasts plotted on the graph as shown in Figure 4.34. Observing the graph, it shows that the PIT series forecast is closer to the actual series, except in the first quarter of 2013 and the third quarter of 2016.

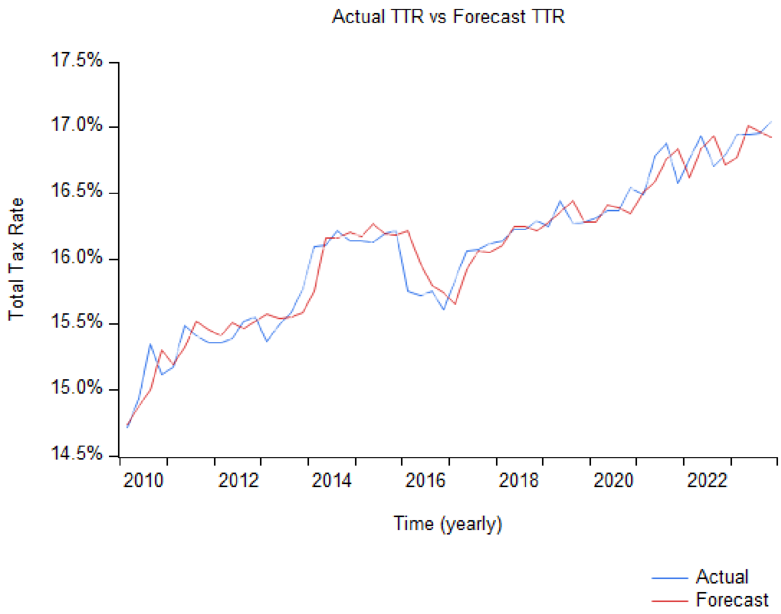


Figure 4.34: Actual TTR vs Forecast TTR

The closeness of the forecast series to the actual series suggests that the selected model has better prediction power and is appropriate for the forecast of VAT. This is confirmed by the calculated precision measures, the RMSE with a value of 0.1447.

4.34 Bayesian Vector Autoregressive Model for TTR

Table 4.16 below indicts the results of appropriate BVAR models for TTR using Minnesota prior,Normal-Wishart prior and Sims-Zha prior with regard to the RMSE (the full models are shown; refer to Appendix Section E)

BVAR Model	Minnesota prior	Normal Whishart prior	Simz-Zha Prior
RMSE	0.1355	0.1091	0.1349

Table 4.16: BVAR models based on RMSE

The best BVAR model for TTR was selected by comparing the RMSE of the out sampling forecasts accuracy, and the smallest is that of BVAR Normal Whishart prior with RMSE of 0.1091, compared to BVAR Minnesota prior and Sims-Zha prior with RMSE of 0.1355 and 0.1349 respectively.

4.35 Forecasted BVAR model for TTR

The best model used was $BVAR_{NWP}$ to generate the quarterly forecast of TTR as represented in Figure 4.35.

Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LTTR	56	0.109137	0.080410	0.499575	0.003396
LGFCF	56	0.032224	0.024188	0.217114	0.001436
LRGDP	56	0.024487	0.019882	0.191250	0.001180
INFL	56	0.043489	0.028875	8.492806	0.058930
SMCAR	56	0.000770	0.000636	11.91733	0.072726

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure 4.35: Forecast Evaluation of $BVAR_{NWP}$

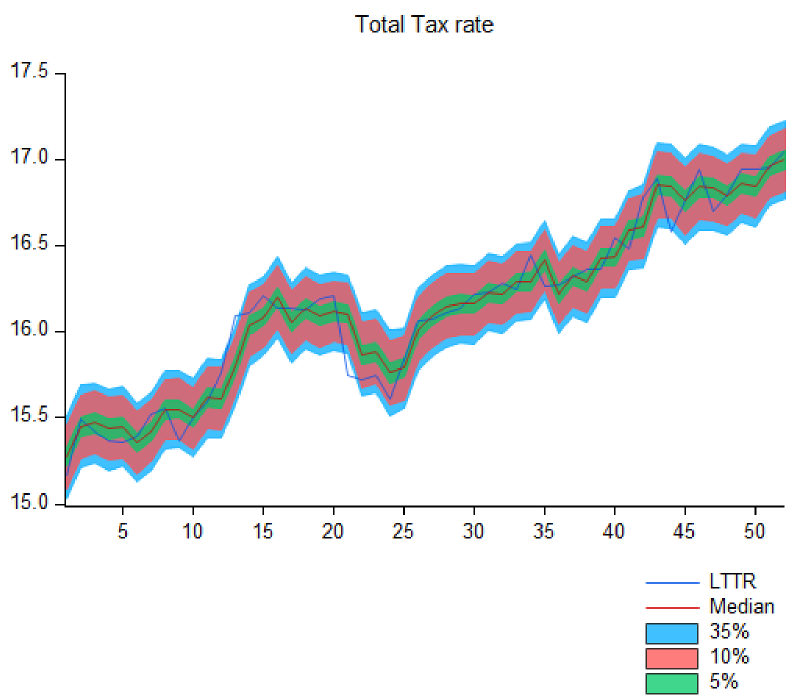


Figure 4.36: Actual TTR vs Forecast TTR

Figure 4.33, indicates that, when we compare the predicted TTR with the actual TTR. The predicted TTR appears to perform well in predicting the results up to the time period when the forecast value is within the 95% confidence interval. After that , they perform poorly when the forecast value is outside the predicted range mentioned above, since the credible intervals 95% for the output gap do not contain the TTR observed for the part of the time period.

4.36 Discussion of Total Tax Revenue Results

The appropriate ARIMA model for the TTR models was developed based on AIC, SIC, and HQ. First, the three appropriate ARIMA models were selected based on AIC. The ARIMA (1,2,22) model was found to have the lowest AIC of -0.8929 compared to the ARIMA (1,2,1) model and ARIMA(2,2,3). In terms of forecast accuracy, ARIMA (1,2,22) was found to have a minimum RMSE of 0.1359. The results conclude that ARIMA(1,2,22) is the best model to fit TTR data better compared to the other ARIMA.

The best ETS model for VAT was a specification of Additive, Additive, and Additive ETS (A, A, A) with a level smoothing parameter estimate $\alpha = 0.82$, trend parameter is zero indicating that trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.0$. The AIC of the ETS(A,A, A) model was 24.8911 smaller than the other three competing models. The forecast evaluation was performed for the out-of-sample period starting from the first quarter of 2013 and the third quarter of 2016 and it was found that ETS (A,A,A) had a minimum RMSE of 0.1447. These results suggest that ETS(A,A,A) performs better than the other three ETS models for TTR.

The best BVAR model for VAT with three different priors was based on RMSE. The BVAR with normal Whishart prior was the best model with RMSE of 0.1091. The results suggest that $BVAR_{NWP}$ performed better than Bayesian vector autoregression with normal Wishart prior and Bayesian vector autoregression with Sims and Zha priors.

The best final model was selected on the basis of RMSE. In conclusion $BVAR_{NWP}$ was superior to the selected ARIMA (1,2,22) and ETS (A, A, A) in handling the TTR data set. Henceforth, it is the best method that can be used to forecast personal income tax. Table 4.17 shows the RMSE of the three approaches.

Models for TTR	**RMSE**
$BVAR_{NWP}$	0.1091
ARIMA(1,2,22)	0.1359
ETS(A,A,A)	0.1447

Table 4.17: RMSE for the best model of TTR

These results compare favorably with the results of various studies on the same subject such as Ramos [30] who concluded that BVAR is the best forecasting tool relative to univariate ARIMA and VAR models, due to its use of few degrees of freedom. The findings coincide with the study conducted by Caraiani [6], who confirmed that the Bayesian approach outperforms standard models. The results indicated that recovery would be slow and that the output gap would continue to be negative for a few quarters, even after the economy started to grow. The study suggested that other more complex models that incorporate an extension to the open economy or the development of models to analyze monetary and fiscal policy may be used.

Furthermore, the study conducted by Yao [37] coincides with the finding of this study, confirms the superiority of BVAR over VAR, and the results indicated that the proper incorporation of prior information into BVAR provides accurate and responsive forecasts. Finally, the study conducted by Carriero and Mumtaz [7] is in line with the findings of this study; the finding was that, once over- parameterization is corrected, the use of extra explanatory variables is important in forecasting fiscal variables, and multivariate models perform better than univariate specifications in forecasting. The large implementation of the system and the variation in time play a very important role in forecasting.

Chapter 5

Conclusion and Recommendations

5.1 Conclusion

This study explored the Bayesian Vector Auto-Regression Method as an Alternative Technique for Forecasting Tax Revenue in Zambia from 2010Q1 to 2023q4. The study predicted revenue collection for Corporate Income Tax, Personal Income Tax, Value-Added Tax, and Total Tax Revenue by comparing the three outcome methods employed. The three methods are Autoregressive Moving Averages (ARIMA), Error, Trend, and Seasonal Models (ETS) and Bayesian Vector Autoregression (BVAR). ARIMA and ETS models use historical data for specific taxes, while BVAR uses selected variables such as Real Gross Domestic Product (RGDP), Gross fixed capital formation (GFCF), inflation rate (INFL) and Stock Market Capitalization ratio (SMCAR) as response variables for the tax rate. The variables are chosen on the basis of economic theory and literature.

The Akaike Information Criterion (AIC) was used to compare models derived from the ETS and ARIMA methods; however, the BVAR method used the root mean squared error (RMSE). The root mean squared error was used to compare models from ARIMA, ETS and BVAR. BVAR models are estimated using three priors, the Minnesota prior, the Normal-Wishart prior, and the Sims-Zha prior. The forecast performances of the appropriate three methods are evaluated based on the minimum RMSE.

The results revealed the accuracy of Bayesian Vector Autoregression for predicting tax data. BVAR using Normal-Wishart prior performs better than ARIMA and ETS in all taxes under consideration including total tax revenue. The total tax revenue was best fitted by the BVAR model with Normal-Wishart prior. The BVAR model outperforms the ETS models and ARIMA/SARIMA in forecasting total tax revenue. Henceforth, the BVAR models can be improved by selecting more appropriate variables that explain various taxes and by including more explanatory variables.

5.2 Recommendations

The study revealed that the ETS models in most cases are second best to BVAR and in the forecasting of tax collection. The economic variables which influence tax revenue should be explored in forecasting revenue with BVAR techniques. Furthermore, we advise policymakers to incorporate the BVAR method among the existing methods used to forecast tax revenues in Zambia, since it produces more robust and precise estimates.

5.3 Areas of Future Research

Future research should focus on the suitability of ETS methods to forecast tax revenue as an alternative approach to existing tax models. ETS methods are not used adequately in tax revenue forecasting. The methodology mainly used is simple exponential smoothing.Furthermore, studies on BVAR forecasting methods may also be extended to other smaller taxes to investigate whether they will fit these taxes accurately, as it does for major taxes.

References

[1] Auerbach, A. J. (1999). *On the Performance and Use of Government Revenue Forecasts.* National Tax Journal, **54**(4) 765–782.

[2] Bako, H. Y. (2014). *Forecasting pelagic fish in Malaysia using ETS state space approach.* University Tun Hussein on Malaysia.

[3] Boonzaaier, W. (2012). *Revenue forecasting practices: current international developments and the case of south Africa. South African Revenue Service.* Revenue Analysis, Planning and Reporting.

[4] Box, G. E. P., Jenkins, G. M. (1976). *Time Series Analysis: Forecasting and Control.* Holden-Day.

[5] Byrne W. J.(1977). *The elasticity of the tax system of Zambia.* World Development, **11**(2) 153–162.

[6] Caraiani, P. (2010). *The Performance of Bayesian Vector Autoregressions with Imprecise Prior Information.* Romanian Journal of Economic Forecasting, **34**(4) 48–63.

[7] Carreiro, A., Mumtaz, H. (2012). *The forecasting ability of structural models with financial frictions.* Journal of Money, Credit and Banking, **44**(8) 1461–1481.

[8] Chun-Yan Kuo, G. (2000). *Estimation of tax revenue and tax capacity.* Harvard University.

[9] Ciccarelli, M., Rebucci, A. (2003). *Measuring contagion with a Bayesian time-varying coefficient model.* International Monetary Fund, **44**(8) 1461–1481.

[10] Doan, T., Litterman, R. B., Sims, C. A. (1984). *Forecasting and conditional projection using realistic prior distributions.* International Monetary Fund, **3**(1) 1–100.

[11] Fullerton, T. M. (1989). *A Composite Approach to Forecasting State Government Revenues: Case study of Idaho sales tax.* International journal of forecasting, **5**(3) 373–380.

[12] Giannone, D., Lenza, M., Primiceri, G. (2012). *Prior selection for vector autoregressions.* Review of Economics and Statistics, **94**(2) 436–451.

[13] Granger, C. W. J., Newbold, P. (1977). *Forecasting Economic Time Series.* Academic Press.

[14] Gupta, R., Kabundi, A. (2010). *Bayesian Vector Autoregressions and Forecasting South African Tax Revenue.* South African Journal of Economics, **78**(2) 232–245.

[15] Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press.

[16] Hannon, A., Leahy, E., O'Sullivan, R. (2015). *An analysis of tax forecasting errors in Ireland*. The Economic and Social Review, **46**(3) 377–399.

[17] Holt, C. C. (1957). *Forecasting seasonals and trends by exponentially weighted moving averages*. Carnegie Institute of Technology.

[18] Hyndman, R. J. (2006). *Another Look at Forecasts of Seasonal Time Series*. International Journal of Forecasting, **22**(2) 165–180.

[19] Hyndman, R. J., Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*. OTexts.

[20] Hyndman, R. J., Khandakar, Y. (2008). *Automatic time series forecasting: The forecast package R*. Journal of statistical software, **27**(3) 1–22.

[21] Jenkins, G. P., Kuo, C., Shukla, G. P. (2000) *Tax Analysis and Revenue Forecasting. Issues and Techniques*. Harvard Institute for International Development.

[22] Kaliba C., Muya M., and Mumba K (2009). *Cost escalation and schedule delays in road construction projects in Zambia*. International Journal of Project Management, **27**(5) 522–531.

[23] Koop, G. (2003). *Bayesian Econometrics*.Wiley.

[24] Krol, R. (1996). *Bayesian Vector Autoregressions and State Tax Revenue Forecasting*. Journal of Forecasting, **15**(6) 533–550.

[25] Litterman, R. (1986). *Forecasting with Bayesian Vector Autoregressions: Five Years of Experience*. Journal of Business Economic Statistics, **4**(1) 25–38.

[26] Makananisa, M. P. (2015). *Forecasting annual tax revenue of the South African taxes using time series Holt-Winters and ARIMA/SARIMA Models*. University of South Africa, Pretoria.

[27] Makridakis, S., Wheelwright, S. C., Hyndman, R. J. (1998). *Forecasting methods and applications* 3^{rd}.United States of America: John Wiley Sons, Inc.

[28] Nandi, B. K., Chaudhury, M., Hasan, G. Q. (2014). *Univariate time series forecasting*.International Journal of Business and Economic Development, **2**(1) 1–13

[29] Nhekairo W (2018). *The taxation system in Zambia*.Accessed: Aug. 29, 2024. [Online]. Available: https://www.taxjustice-andpoverty.

[30] Ramos, F. R. (1996). *Forecasting Market Shares Using VAR and BVAR Models: A Comparison of their Forecasting Performance*.University of Porto , Faculty of Economics, Porto.

[31] Sims, C., and Zha, T. (1998). *Bayesian Methods for Dynamic Multivariate Mod-*

*els.*International Economic Review, **39**(1) 949–968

[32] Skarbovik, L. F. (2013) *Forecasting house prices in Noway: A univariate approach.*Tromso University Business School, Norway.

[33] Smith, J. N., Wallis, K. F. (2009). *The Structure of Bayesian VARs.*Journal of Econometrics, **149**(1) 102–113

[34] Urrutia, J. D., Mingo, F. L., Balmaceda, C. N. (2015). *Forecasting income tax revenue of the Philippines using autoregressive integrated moving average modelling: A time series analysis.* .International Multi-Disciplinary Research Journal, **1**(9) 1938–1992

[35] Van Heerden, Y., Schoeman, N. J. (2010). *An empirical dissemination of the personal income tax regime in South Africa using a Microsimulation Tax Model.*University of Pretoria, Department of Economics, Pretoria.

[36] Wallschutzky, I. G. (1984). *Possible causes of tax evasion.*Journal of Economic Psychology, **5**(4) 371–384

[37] Yao, F. (2011). *Forecasting several North Dakota microeconomic variables-A Bayesian vector autoregression approach.*West Virginia University, Economics department, Morgan Town.

[38] Zakai, M. (2014). *A time series modeling on GDP of Pakistan..*Journal of Contemporary Issues in Business Research, **3**(4) 200–210.

[39] ZRA (2023). *Annual report 2023.*J[Online]. Available: https://www.zra.org.zm/wp-content/uploads/2023/05/Annual-Report-2023.pdf. [Accessed 27 August 2024].

Appendix A

Unit Root Tests

Null Hypothesis: LCIT has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-0.779290	0.8165
Test critical values:	1% level	-3.562669	
	5% level	-2.918778	
	10% level	-2.597285	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LCIT)
Method: Least Squares
Date: 12/29/24 Time: 01:33
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LCIT(-1)	-0.063241	0.081152	-0.779290	0.4397
D(LCIT(-1))	-0.497203	0.143409	-3.467024	0.0011
D(LCIT(-2))	-0.592701	0.124349	-4.766437	0.0000
D(LCIT(-3))	-0.412900	0.127261	-3.244517	0.0022
C	0.989233	1.149156	0.860835	0.3937

R-squared	0.433631	Mean dependent var	0.037392
Adjusted R-squared	0.385430	S.D. dependent var	0.488745
S.E. of regression	0.383150	Akaike info criterion	1.010429
Sum squared resid	6.899768	Schwarz criterion	1.198048
Log likelihood	-21.27115	Hannan-Quinn criter.	1.082358
F-statistic	8.996209	Durbin-Watson stat	1.862717
Prob(F-statistic)	0.000018		

Figure A.1

Null Hypothesis: D(LCIT,2) has a unit root
Exogenous: Constant
Lag Length: 3 (Automatic - based on SIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.152190	0.0000
Test critical values:	1% level	-3.568308	
	5% level	-2.921175	
	10% level	-2.598551	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LCIT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:30
Sample (adjusted): 2011Q3 2023Q4
Included observations: 50 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LCIT(-1),2)	-4.885850	0.533845	-9.152190	0.0000
D(LCIT(-1),3)	2.668463	0.418179	6.381152	0.0000
D(LCIT(-2),3)	1.427067	0.272801	5.231163	0.0000
D(LCIT(-3),3)	0.398198	0.132672	3.001361	0.0044
C	-0.007911	0.061146	-0.129384	0.8976

R-squared	0.899362	Mean dependent var	-0.005086
Adjusted R-squared	0.890416	S.D. dependent var	1.304632
S.E. of regression	0.431879	Akaike info criterion	1.253295
Sum squared resid	8.393362	Schwarz criterion	1.444498
Log likelihood	-26.33238	Hannan-Quinn criter.	1.326106
F-statistic	100.5364	Durbin-Watson stat	2.038912
Prob(F-statistic)	0.000000		

Figure A.2

Null Hypothesis: LCIT has a unit root
Exogenous: None
Bandwidth: 15 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		1.523682	0.9672
Test critical values:	1% level	-2.607686	
	5% level	-1.946878	
	10% level	-1.612999	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.253841
HAC corrected variance (Bartlett kernel)	0.044813

Phillips-Perron Test Equation
Dependent Variable: D(LCIT)
Method: Least Squares
Date: 12/29/24 Time: 01:36
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LCIT(-1)	0.002573	0.004841	0.531573	0.5972

R-squared	-0.003454	Mean dependent var	0.046926
Adjusted R-squared	-0.003454	S.D. dependent var	0.507595
S.E. of regression	0.508470	Akaike info criterion	1.503195
Sum squared resid	13.96128	Schwarz criterion	1.539692
Log likelihood	-40.33786	Hannan-Quinn criter.	1.517309
Durbin-Watson stat	2.286503		

Figure A.3

Null Hypothesis: D(LCIT,2) has a unit root
Exogenous: None
Bandwidth: 12 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-24.11111	0.0000
Test critical values:	1% level	-2.609324	
	5% level	-1.947119	
	10% level	-1.612867	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.506636
HAC corrected variance (Bartlett kernel)	0.053121

Phillips-Perron Test Equation
Dependent Variable: D(LCIT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:38
Sample (adjusted): 2010Q4 2023Q4
Included observations: 53 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LCIT(-1),2)	-1.380756	0.127743	-10.80884	0.0000

R-squared	0.691923	Mean dependent var	0.020214
Adjusted R-squared	0.691923	S.D. dependent var	1.294658
S.E. of regression	0.718595	Akaike info criterion	2.195651
Sum squared resid	26.85173	Schwarz criterion	2.232827
Log likelihood	-57.18476	Hannan-Quinn criter.	2.209947
Durbin-Watson stat	2.341652		

Figure A.4

Null Hypothesis: LPIT has a unit root
Exogenous: None
Lag Length: 1 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		2.153667	0.9918
Test critical values:	1% level	-2.608490	
	5% level	-1.946996	
	10% level	-1.612934	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LPIT)
Method: Least Squares
Date: 12/29/24 Time: 01:43
Sample (adjusted): 2010Q3 2023Q4
Included observations: 54 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LPIT(-1)	0.004085	0.001897	2.153667	0.0359
D(LPIT(-1))	-0.375071	0.129123	-2.904757	0.0054

R-squared	0.137900	Mean dependent var	0.044680
Adjusted R-squared	0.121321	S.D. dependent var	0.213977
S.E. of regression	0.200578	Akaike info criterion	-0.338894
Sum squared resid	2.092039	Schwarz criterion	-0.265228
Log likelihood	11.15014	Hannan-Quinn criter.	-0.310484
Durbin-Watson stat	1.859204		

Figure A.5

Null Hypothesis: D(LPIT,2) has a unit root
Exogenous: None
Lag Length: 1 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.79784	0.0000
Test critical values:	1% level	-2.610192	
	5% level	-1.947248	
	10% level	-1.612797	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LPIT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:44
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LPIT(-1),2)	-2.372628	0.219732	-10.79784	0.0000
D(LPIT(-1),3)	0.438249	0.119070	3.680585	0.0006

R-squared	0.877523	Mean dependent var	0.010013
Adjusted R-squared	0.875073	S.D. dependent var	0.615163
S.E. of regression	0.217429	Akaike info criterion	-0.176187
Sum squared resid	2.363768	Schwarz criterion	-0.101139
Log likelihood	6.580852	Hannan-Quinn criter.	-0.147415
Durbin-Watson stat	2.360835		

Figure A.6

Null Hypothesis: LPIT has a unit root
Exogenous: Constant
Bandwidth: 0 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-0.870284	0.7903
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.043586
HAC corrected variance (Bartlett kernel)	0.043586

Phillips-Perron Test Equation
Dependent Variable: D(LPIT)
Method: Least Squares
Date: 12/29/24 Time: 01:48
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LPIT(-1)	-0.048036	0.055196	-0.870284	0.3881
C	0.748022	0.810144	0.923319	0.3600

R-squared	0.014089	Mean dependent var	0.043407
Adjusted R-squared	-0.004513	S.D. dependent var	0.212197
S.E. of regression	0.212675	Akaike info criterion	-0.222416
Sum squared resid	2.397228	Schwarz criterion	-0.149422
Log likelihood	8.116433	Hannan-Quinn criter.	-0.194188
F-statistic	0.757395	Durbin-Watson stat	2.645413
Prob(F-statistic)	0.388070		

Figure A.7

Null Hypothesis: D(LPIT,2) has a unit root
Exogenous: Constant
Bandwidth: 2 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-22.92427	0.0001
Test critical values:	1% level	-3.560019	
	5% level	-2.917650	
	10% level	-2.596689	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.061598
HAC corrected variance (Bartlett kernel)	0.028288

Phillips-Perron Test Equation
Dependent Variable: D(LPIT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:49
Sample (adjusted): 2010Q4 2023Q4
Included observations: 53 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LPIT(-1),2)	-1.680814	0.099277	-16.93062	0.0000
C	-0.001301	0.034769	-0.037405	0.9703

R-squared	0.848954	Mean dependent var	-0.018964
Adjusted R-squared	0.845993	S.D. dependent var	0.644710
S.E. of regression	0.253009	Akaike info criterion	0.126218
Sum squared resid	3.264679	Schwarz criterion	0.200569
Log likelihood	-1.344784	Hannan-Quinn criter.	0.154810
F-statistic	286.6459	Durbin-Watson stat	2.434962
Prob(F-statistic)	0.000000		

Figure A.8

Null Hypothesis: LVAT has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.702663	0.4243
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LVAT)
Method: Least Squares
Date: 12/29/24 Time: 01:54
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LVAT(-1)	-0.077325	0.045414	-1.702663	0.0945
C	1.197156	0.673217	1.778262	0.0811

R-squared	0.051862	Mean dependent var	0.052086
Adjusted R-squared	0.033973	S.D. dependent var	0.231580
S.E. of regression	0.227612	Akaike info criterion	-0.086662
Sum squared resid	2.745784	Schwarz criterion	-0.013668
Log likelihood	4.383210	Hannan-Quinn criter.	-0.058435
F-statistic	2.899061	Durbin-Watson stat	2.230956
Prob(F-statistic)	0.094490		

Figure A.9

Null Hypothesis: D(LVAT,2) has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-10.08830	0.0000
Test critical values:	1% level	-3.562669	
	5% level	-2.918778	
	10% level	-2.597285	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LVAT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:55
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LVAT(-1),2)	-2.185043	0.216592	-10.08830	0.0000
D(LVAT(-1),3)	0.434516	0.123919	3.506449	0.0010
C	0.001747	0.036677	0.047632	0.9622

R-squared	0.823404	Mean dependent var	0.012084
Adjusted R-squared	0.816196	S.D. dependent var	0.616530
S.E. of regression	0.264321	Akaike info criterion	0.232658
Sum squared resid	3.423421	Schwarz criterion	0.345230
Log likelihood	-3.049116	Hannan-Quinn criter.	0.275816
F-statistic	114.2345	Durbin-Watson stat	2.266841
Prob(F-statistic)	0.000000		

Figure A.10

Null Hypothesis: LVAT has a unit root
Exogenous: Constant
Bandwidth: 1 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.657660	0.4468
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.049923
HAC corrected variance (Bartlett kernel)	0.043856

Phillips-Perron Test Equation
Dependent Variable: D(LVAT)
Method: Least Squares
Date: 12/29/24 Time: 01:57
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LVAT(-1)	-0.077325	0.045414	-1.702663	0.0945
C	1.197156	0.673217	1.778262	0.0811

R-squared	0.051862	Mean dependent var	0.052086
Adjusted R-squared	0.033973	S.D. dependent var	0.231580
S.E. of regression	0.227612	Akaike info criterion	-0.086662
Sum squared resid	2.745784	Schwarz criterion	-0.013668
Log likelihood	4.383210	Hannan-Quinn criter.	-0.058435
F-statistic	2.899061	Durbin-Watson stat	2.230956
Prob(F-statistic)	0.094490		

Figure A.11

Null Hypothesis: D(LVAT,2) has a unit root
Exogenous: Constant
Bandwidth: 17 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-27.94241	0.0001
Test critical values:	1% level	-3.560019	
	5% level	-2.917650	
	10% level	-2.596689	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.087780
HAC corrected variance (Bartlett kernel)	0.012331

Phillips-Perron Test Equation
Dependent Variable: D(LVAT,3)
Method: Least Squares
Date: 12/29/24 Time: 01:58
Sample (adjusted): 2010Q4 2023Q4
Included observations: 53 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LVAT(-1),2)	-1.538392	0.117457	-13.09745	0.0000
C	-0.005640	0.041487	-0.135948	0.8924

R-squared	0.770831	Mean dependent var	-0.006140
Adjusted R-squared	0.766338	S.D. dependent var	0.624822
S.E. of regression	0.302030	Akaike info criterion	0.480427
Sum squared resid	4.652339	Schwarz criterion	0.554778
Log likelihood	-10.73131	Hannan-Quinn criter.	0.509019
F-statistic	171.5433	Durbin-Watson stat	2.335019
Prob(F-statistic)	0.000000		

Figure A.12

Null Hypothesis: LTTR has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.731734	0.4100
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LTTR)
Method: Least Squares
Date: 12/29/24 Time: 02:00
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LTTR(-1)	-0.065267	0.037689	-1.731734	0.0891
C	1.089312	0.604887	1.800851	0.0774

R-squared	0.053553	Mean dependent var	0.042429
Adjusted R-squared	0.035695	S.D. dependent var	0.157260
S.E. of regression	0.154428	Akaike info criterion	-0.862493
Sum squared resid	1.263943	Schwarz criterion	-0.789499
Log likelihood	25.71855	Hannan-Quinn criter.	-0.834266
F-statistic	2.998904	Durbin-Watson stat	2.194619
Prob(F-statistic)	0.089138		

Figure A.13

Null Hypothesis: D(LTTR,2) has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic - based on AIC, maxlag=10)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-12.82606	0.0000
Test critical values:	1% level	-3.562669	
	5% level	-2.918778	
	10% level	-2.597285	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LTTR,3)
Method: Least Squares
Date: 12/29/24 Time: 02:02
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LTTR(-1),2)	-2.285902	0.178223	-12.82606	0.0000
D(LTTR(-1),3)	0.577431	0.103533	5.577281	0.0000
C	-0.001842	0.021784	-0.084544	0.9330

R-squared	0.848832	Mean dependent var	0.013744
Adjusted R-squared	0.842662	S.D. dependent var	0.395525
S.E. of regression	0.156889	Akaike info criterion	-0.810601
Sum squared resid	1.206087	Schwarz criterion	-0.698029
Log likelihood	24.07563	Hannan-Quinn criter.	-0.767444
F-statistic	137.5710	Durbin-Watson stat	2.178809
Prob(F-statistic)	0.000000		

Figure A.14

Null Hypothesis: LTTR has a unit root
Exogenous: Constant
Bandwidth: 3 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.709160	0.4211
Test critical values:	1% level	-3.555023	
	5% level	-2.915522	
	10% level	-2.595565	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.022981
HAC corrected variance (Bartlett kernel)	0.018553

Phillips-Perron Test Equation
Dependent Variable: D(LTTR)
Method: Least Squares
Date: 12/29/24 Time: 02:03
Sample (adjusted): 2010Q2 2023Q4
Included observations: 55 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LTTR(-1)	-0.065267	0.037689	-1.731734	0.0891
C	1.089312	0.604887	1.800851	0.0774

R-squared	0.053553	Mean dependent var	0.042429
Adjusted R-squared	0.035695	S.D. dependent var	0.157260
S.E. of regression	0.154428	Akaike info criterion	-0.862493
Sum squared resid	1.263943	Schwarz criterion	-0.789499
Log likelihood	25.71855	Hannan-Quinn criter.	-0.834266
F-statistic	2.998904	Durbin-Watson stat	2.194619
Prob(F-statistic)	0.089138		

Figure A.15

Null Hypothesis: D(LTTR,2) has a unit root
Exogenous: Constant
Bandwidth: 2 (Newey-West automatic) using Bartlett kernel

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-15.94119	0.0000
Test critical values:	1% level	-3.560019	
	5% level	-2.917650	
	10% level	-2.596689	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.042963
HAC corrected variance (Bartlett kernel)	0.017004

Phillips-Perron Test Equation
Dependent Variable: D(LTTR,3)
Method: Least Squares
Date: 12/29/24 Time: 02:04
Sample (adjusted): 2010Q4 2023Q4
Included observations: 53 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LTTR(-1),2)	-1.469383	0.122872	-11.95869	0.0000
C	-0.007960	0.029029	-0.274221	0.7850

R-squared	0.737127	Mean dependent var	-0.002007
Adjusted R-squared	0.731973	S.D. dependent var	0.408142
S.E. of regression	0.211301	Akaike info criterion	-0.234064
Sum squared resid	2.277047	Schwarz criterion	-0.159713
Log likelihood	8.202691	Hannan-Quinn criter.	-0.205472
F-statistic	143.0102	Durbin-Watson stat	2.343137
Prob(F-statistic)	0.000000		

Figure A.16

Appendix B

BVAR, ARIMA and ETS Models for CIT

Bayesian VAR Estimates
Date: 12/28/24 Time: 07:24
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Litterman / Minnesota
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L1: 0.1, L2: 0.99, L3: 1, L4: 10
Standard errors in ( )

	LCIT	LRGDP	LGFCF	INFL	SMCAR
LCIT(-1)	0.106485	0.013928	-0.001471	0.012122	-5.13E-05
	(0.08438)	(0.00567)	(0.00516)	(0.00947)	(8.1E-05)
LCIT(-2)	-0.003466	0.001815	-0.000236	0.003143	-7.85E-06
	(0.04777)	(0.00320)	(0.00292)	(0.00535)	(4.6E-05)
LCIT(-3)	0.009180	-0.000842	-3.27E-05	0.002313	6.48E-06
	(0.03257)	(0.00218)	(0.00199)	(0.00365)	(3.1E-05)
LCIT(-4)	0.013235	3.08E-05	3.36E-05	0.001721	4.20E-06
	(0.02464)	(0.00165)	(0.00150)	(0.00276)	(2.4E-05)
LRGDP(-1)	-0.163569	0.542094	0.030483	-0.053214	0.000527
	(0.80351)	(0.05471)	(0.04953)	(0.09080)	(0.00077)
LRGDP(-2)	0.684903	0.139770	0.019079	-0.000740	0.000230
	(0.62440)	(0.04261)	(0.03849)	(0.07056)	(0.00060)
LRGDP(-3)	0.266163	0.086734	0.005871	-0.005759	2.77E-05
	(0.44961)	(0.03070)	(0.02771)	(0.05081)	(0.00043)
LRGDP(-4)	-0.003290	0.069880	0.006023	-0.003992	-0.000110
	(0.35001)	(0.02391)	(0.02158)	(0.03955)	(0.00034)
LGFCF(-1)	-0.171774	0.045478	0.818862	0.074671	0.000278
	(0.80879)	(0.05474)	(0.05024)	(0.09140)	(0.00078)
LGFCF(-2)	0.165317	0.000270	0.106535	0.006282	-7.25E-05
	(0.71060)	(0.04809)	(0.04423)	(0.08030)	(0.00068)
LGFCF(-3)	0.147968	-0.007043	0.017964	-0.004257	-0.000177
	(0.48876)	(0.03308)	(0.03041)	(0.05523)	(0.00047)
LGFCF(-4)	0.128082	-0.003764	0.004870	-0.001420	-0.000196
	(0.36756)	(0.02487)	(0.02286)	(0.04154)	(0.00035)
INFL(-1)	1.365105	-0.029776	0.029517	0.562542	-0.000185
	(0.57498)	(0.03889)	(0.03542)	(0.06535)	(0.00055)
INFL(-2)	0.318697	-0.001769	-0.005582	0.105852	6.00E-05
	(0.40094)	(0.02713)	(0.02471)	(0.04573)	(0.00039)
INFL(-3)	0.162079	-0.001420	-0.005975	0.039802	0.000103
	(0.27893)	(0.01888)	(0.01719)	(0.03182)	(0.00027)
INFL(-4)	0.094662	0.001043	-0.006256	0.022071	7.02E-05
	(0.21230)	(0.01437)	(0.01309)	(0.02422)	(0.00020)
SMCAR(-1)	-6.671043	7.735378	5.096404	-7.982293	0.283388
	(79.3928)	(5.37330)	(4.89418)	(8.97274)	(0.07699)
SMCAR(-2)	-32.77961	1.299595	2.236235	-3.971011	0.059629
	(47.9343)	(3.24409)	(2.95490)	(5.41723)	(0.04660)
SMCAR(-3)	-19.69590	-0.136001	1.553821	-1.674865	0.025633
	(33.0886)	(2.23931)	(2.03972)	(3.73939)	(0.03218)
SMCAR(-4)	-4.878164	-0.917616	0.980496	-0.892672	0.016347
	(25.1723)	(1.70355)	(1.55173)	(2.84475)	(0.02449)
C	1.020702	1.051216	-0.054608	-0.283420	-0.001138
	(4.31213)	(0.29211)	(0.26583)	(0.48732)	(0.00415)
R-squared	0.747503	0.930888	0.995856	0.923536	0.516564
Adj. R-squared	0.584602	0.886300	0.993183	0.874205	0.204670
Sum sq. resids	7.123700	0.057573	0.051208	0.122501	7.57E-06
S.E. equation	0.479371	0.043095	0.040643	0.062862	0.000494
F-statistic	4.588696	20.87740	372.4995	18.72106	1.656216
Mean dependent	14.23101	10.39428	11.27978	0.335577	0.005408
S.D. dependent	0.743772	0.127805	0.492248	0.177238	0.000554
Data marginal log-likelihood		-1861.884			

Figure B.1

Bayesian VAR Estimates
Date: 12/28/24 Time: 07:35
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Normal-Wishart
Hyper-parameters: Mu1: 0, C1: 0.1, C2: 0.1, C3: 6
Standard errors in ( )

	LCIT	LRGDP	LGFCF	INFL	SMCAR
LCIT(-1)	0.295024	0.028445	-0.009043	0.009310	-0.000209
	(0.13578)	(0.02182)	(0.02569)	(0.02690)	(0.01754)
LCIT(-2)	-0.186333	0.018865	0.003678	0.001982	-8.13E-05
	(0.14445)	(0.02321)	(0.02733)	(0.02862)	(0.01866)
LCIT(-3)	0.025909	-0.014324	0.001932	0.016393	4.22E-06
	(0.14370)	(0.02309)	(0.02718)	(0.02847)	(0.01856)
LCIT(-4)	0.314732	-0.012235	0.010625	0.037465	0.000209
	(0.12967)	(0.02084)	(0.02453)	(0.02569)	(0.01675)
LRGDP(-1)	0.008678	0.209689	0.001321	-0.008243	0.000336
	(0.79552)	(0.12784)	(0.15050)	(0.15762)	(0.10274)
LRGDP(-2)	1.095030	0.206707	0.008643	-0.047929	1.82E-05
	(0.77745)	(0.12494)	(0.14708)	(0.15404)	(0.10041)
LRGDP(-3)	-0.021140	0.194986	2.26E-05	-0.080208	0.000394
	(0.80465)	(0.12931)	(0.15222)	(0.15943)	(0.10392)
LRGDP(-4)	-0.431662	0.372131	0.039848	0.010949	-0.000183
	(0.77382)	(0.12436)	(0.14639)	(0.15332)	(0.09994)
LGFCF(-1)	-0.373982	0.057841	0.664170	0.052809	0.000950
	(0.73621)	(0.11831)	(0.13928)	(0.14587)	(0.09508)
LGFCF(-2)	-0.039040	0.026478	0.286241	0.032930	0.000449
	(0.88218)	(0.14177)	(0.16689)	(0.17479)	(0.11393)
LGFCF(-3)	0.181616	-0.036606	0.043640	-0.015216	-0.000161
	(0.87883)	(0.14123)	(0.16626)	(0.17412)	(0.11350)
LGFCF(-4)	0.279087	-0.059948	-0.043607	-0.027975	-0.001166
	(0.72650)	(0.11675)	(0.13744)	(0.14394)	(0.09383)
INFL(-1)	0.744276	-0.099606	0.077927	0.587071	-0.001330
	(0.70883)	(0.11391)	(0.13410)	(0.14044)	(0.09155)
INFL(-2)	0.716796	-0.025812	-0.001032	0.188042	-0.000295
	(0.81983)	(0.13175)	(0.15510)	(0.16244)	(0.10588)
INFL(-3)	0.287732	0.010091	-0.026618	-0.006224	0.000540
	(0.83690)	(0.13449)	(0.15832)	(0.16582)	(0.10809)
INFL(-4)	-0.195545	0.083562	-0.098442	-0.028285	0.001047
	(0.74543)	(0.11979)	(0.14102)	(0.14769)	(0.09627)
SMCAR(-1)	-0.019500	0.001134	0.002714	-0.003336	0.000101
	(1.07374)	(0.17255)	(0.20313)	(0.21274)	(0.13867)
SMCAR(-2)	-0.034968	0.000875	0.003564	-0.003885	9.93E-05
	(1.07373)	(0.17255)	(0.20313)	(0.21274)	(0.13867)
SMCAR(-3)	-0.036020	0.000802	0.004368	-0.003357	9.33E-05
	(1.07372)	(0.17255)	(0.20313)	(0.21274)	(0.13867)
SMCAR(-4)	-0.022618	-0.000827	0.004494	-0.003026	8.57E-05
	(1.07373)	(0.17255)	(0.20313)	(0.21274)	(0.13867)
C	0.117045	0.040967	-0.010774	-0.005745	-0.000226
	(1.06950)	(0.17187)	(0.20233)	(0.21190)	(0.13813)
R-squared	0.801884	0.969334	0.995141	0.940246	0.244878
Adj. R-squared	0.674067	0.949550	0.992005	0.901695	-0.242298
Sum sq. resids	5.589464	0.025546	0.060051	0.095731	1.18E-05
S.E. equation	0.424624	0.028706	0.044013	0.055571	0.000618
F-statistic	6.273689	48.99538	317.4186	24.38956	0.502647
Mean dependent	14.23101	10.39428	11.27978	0.335577	0.005408
S.D. dependent	0.743772	0.127805	0.492248	0.177238	0.000554
Data marginal log-likelihood		189.8867			

Figure B.2

Bayesian VAR Estimates
Date: 12/28/24 Time: 07:52
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Sims-Zha (normal-Wishart)
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L0: 1, L1: 0.1, L3: 1, L4: 10, L5: 10, C3: 6
Standard errors in ( )

	LCIT	LRGDP	LGFCF	INFL	SMCAR
LCIT(-1)	0.105759	0.014135	-0.001496	0.012293	-5.20E-05
	(0.07412)	(0.00892)	(0.00844)	(0.01230)	(7.9E-05)
LCIT(-2)	-0.003613	0.001847	-0.000239	0.003195	-7.98E-06
	(0.04195)	(0.00505)	(0.00477)	(0.00696)	(4.5E-05)
LCIT(-3)	0.009083	-0.000857	-3.33E-05	0.002356	6.59E-06
	(0.02860)	(0.00344)	(0.00326)	(0.00475)	(3.1E-05)
LCIT(-4)	0.013170	3.02E-05	3.35E-05	0.001754	4.28E-06
	(0.02163)	(0.00260)	(0.00246)	(0.00359)	(2.3E-05)
LRGDP(-1)	-0.171692	0.541980	0.030421	-0.053686	0.000529
	(0.70990)	(0.08548)	(0.08080)	(0.11779)	(0.00076)
LRGDP(-2)	0.694201	0.139771	0.019249	-0.000639	0.000232
	(0.55285)	(0.06656)	(0.06292)	(0.09173)	(0.00059)
LRGDP(-3)	0.269970	0.086621	0.005890	-0.005845	2.78E-05
	(0.39834)	(0.04796)	(0.04534)	(0.06610)	(0.00043)
LRGDP(-4)	-0.003679	0.069809	0.006088	-0.004045	-0.000112
	(0.31023)	(0.03735)	(0.03531)	(0.05148)	(0.00033)
LGFCF(-1)	-0.178306	0.045897	0.818655	0.075334	0.000284
	(0.71562)	(0.08616)	(0.08145)	(0.11874)	(0.00077)
LGFCF(-2)	0.166659	0.000153	0.106559	0.006120	-7.26E-05
	(0.62990)	(0.07584)	(0.07169)	(0.10452)	(0.00068)
LGFCF(-3)	0.149477	-0.007239	0.017991	-0.004500	-0.000179
	(0.43312)	(0.05215)	(0.04930)	(0.07187)	(0.00046)
LGFCF(-4)	0.129639	-0.003869	0.004895	-0.001565	-0.000199
	(0.32563)	(0.03921)	(0.03706)	(0.05403)	(0.00035)
INFL(-1)	1.371129	-0.030321	0.030105	0.561844	-0.000187
	(0.50809)	(0.06118)	(0.05783)	(0.08431)	(0.00054)
INFL(-2)	0.319988	-0.001804	-0.005674	0.105690	6.13E-05
	(0.35530)	(0.04278)	(0.04044)	(0.05895)	(0.00038)
INFL(-3)	0.163221	-0.001459	-0.006081	0.039731	0.000105
	(0.24724)	(0.02977)	(0.02814)	(0.04102)	(0.00027)
INFL(-4)	0.095421	0.001051	-0.006376	0.022036	7.15E-05
	(0.18820)	(0.02266)	(0.02142)	(0.03123)	(0.00020)
SMCAR(-1)	-6.329503	7.817939	5.134608	-8.046278	0.283307
	(70.1737)	(8.44919)	(7.98682)	(11.6438)	(0.07523)
SMCAR(-2)	-33.16048	1.308570	2.264317	-4.024584	0.059637
	(42.4713)	(5.11372)	(4.83388)	(7.04720)	(0.04553)
SMCAR(-3)	-19.96699	-0.141874	1.577411	-1.695756	0.025618
	(29.3297)	(3.53141)	(3.33816)	(4.86662)	(0.03144)
SMCAR(-4)	-4.909798	-0.936840	0.996307	-0.903834	0.016332
	(22.3170)	(2.68705)	(2.54001)	(3.70302)	(0.02392)
C	1.010782	1.050953	-0.055370	-0.282547	-0.001150
	(3.79002)	(0.45633)	(0.43136)	(0.62887)	(0.00406)
R-squared	0.748057	0.931070	0.995860	0.923543	0.517052
Adj. R-squared	0.585514	0.886599	0.993189	0.874216	0.205473
Sum sq. resids	7.108073	0.057422	0.051163	0.122490	7.57E-06
S.E. equation	0.478845	0.043039	0.040625	0.062859	0.000494
F-statistic	4.602193	20.93652	372.8296	18.72286	1.659456
Mean dependent	14.23101	10.39428	11.27978	0.335577	0.005408
S.D. dependent	0.743772	0.127805	0.492248	0.177238	0.000554
Data marginal log-likelihood		453.6754			

Figure B.3

Forecast Evaluation
Date: 12/28/24 Time: 07:33
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LCIT	56	0.357075	0.271613	1.919403	0.012595
LRGDP	56	0.031986	0.023373	0.225231	0.001542
LGFCF	56	0.030287	0.023589	0.210709	0.001350
INFL	56	0.046898	0.033468	10.26031	0.063927
SMCAR	56	0.000368	0.000295	5.596629	0.034655

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure B.4

Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LCIT	56	0.316173	0.243874	1.715331	0.011152
LRGDP	56	0.021486	0.016847	0.162127	0.001036
LGFCF	56	0.032608	0.024314	0.217745	0.001453
INFL	56	0.041268	0.028669	8.290509	0.055879
SMCAR	56	0.000881	0.000701	12.86047	0.083171

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure B.5

Forecast Evaluation
Date: 12/28/24 Time: 07:57
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LCIT	56	0.356628	0.271559	1.919069	0.012580
LRGDP	56	0.032070	0.023442	0.225873	0.001546
LGFCF	56	0.030048	0.023397	0.208892	0.001339
INFL	56	0.046746	0.033254	10.18209	0.063695
SMCAR	56	0.000369	0.000296	5.616376	0.034754

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure B.6

Dependent Variable: D(LCIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 09:50
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 31 iterations
Coefficient covariance computed using outer product of gradients
d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.038909	0.016685	2.331984	0.0237
AR(1)	0.239073	0.267684	0.893119	0.3760
MA(1)	-0.814401	0.125034	-6.513418	0.0000
SIGMASQ	0.190144	0.040275	4.721202	0.0000

R-squared	0.248346	Mean dependent var	0.046926
Adjusted R-squared	0.204131	S.D. dependent var	0.507595
S.E. of regression	0.452833	Akaike info criterion	1.336344
Sum squared resid	10.45793	Schwarz criterion	1.482332
Log likelihood	-32.74946	Hannan-Quinn criter.	1.392799
F-statistic	5.616787	Durbin-Watson stat	1.728890
Prob(F-statistic)	0.002094		

Inverted AR Roots	.24
Inverted MA Roots	.81

Figure B.7

Dependent Variable: D(LCIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 09:54
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 35 iterations
Coefficient covariance computed using outer product of gradients
d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.038646	0.022235	1.738066	0.0882
AR(2)	-0.514163	0.136985	-3.753435	0.0004
MA(1)	-0.456386	0.179488	-2.542705	0.0141
SIGMASQ	0.155388	0.033442	4.646524	0.0000

R-squared	0.385739	Mean dependent var	0.046926
Adjusted R-squared	0.349606	S.D. dependent var	0.507595
S.E. of regression	0.409360	Akaike info criterion	1.140613
Sum squared resid	8.546349	Schwarz criterion	1.286601
Log likelihood	-27.36686	Hannan-Quinn criter.	1.197068
F-statistic	10.67553	Durbin-Watson stat	2.011453
Prob(F-statistic)	0.000015		

Inverted AR Roots	-.00+.72i	-.00-.72i
Inverted MA Roots	.46	

Figure B.8

Dependent Variable: D(LCIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 09:52
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 35 iterations
Coefficient covariance computed using outer product of gradients
d.f. adjustment for standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.038428	0.019737	1.946975	0.0571
AR(3)	0.035738	0.207573	0.172170	0.8640
MA(1)	-0.716778	0.117088	-6.121676	0.0000
SIGMASQ	0.195467	0.038827	5.034279	0.0000

R-squared	0.227306	Mean dependent var	0.046926
Adjusted R-squared	0.181853	S.D. dependent var	0.507595
S.E. of regression	0.459127	Akaike info criterion	1.363664
Sum squared resid	10.75066	Schwarz criterion	1.509652
Log likelihood	-33.50076	Hannan-Quinn criter.	1.420119
F-statistic	5.000949	Durbin-Watson stat	1.590626
Prob(F-statistic)	0.004075		

Inverted AR Roots	.33	-.16-.29i	-.16+.29i
Inverted MA Roots	.72		

Figure B.9

ETS Smoothing
Original series: LCIT
Date: 12/28/24 Time: 11:53
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,A,A - Additive Error, Additive Trend, Additive Season (Holt-Winters - additive seasonal), Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.443332
Beta:	0.000000
Gamma:	0.000000

Initial Parameters	
Initial level:	12.93062
Initial trend:	0.040530
Initial state 1:	-0.238399
Initial state 2:	0.081136
Initial state 3:	0.327036
Initial state 4:	-0.169773

Compact Log-likelihood	-53.88898
Log-likelihood	-20.63969
Akaike Information Criterion	123.7780
Schwarz Criterion	139.9808
Hannan-Quinn Criterion	130.0598
Sum of Squared Residuals	6.852451
Root Mean Squared Error	0.349807
Average Mean Squared Error	0.140814

Figure B.10

ETS Smoothing
Original series: LCIT
Date: 12/28/24 Time: 11:56
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,AD,A - Additive Error, Additive-Dampened Trend, Additive Season, Cycle=4
Convergence achieved on boundaries.

Parameters	
Alpha:	0.494738
Beta:	0.000000
Gamma:	0.000000
Phi:	0.307993
Initial Parameters	
Initial level:	10.37131
Initial trend:	7.104277
Initial state 1:	-0.244149
Initial state 2:	0.068441
Initial state 3:	0.317748
Initial state 4:	-0.142040
Compact Log-likelihood	-52.66127
Log-likelihood	-19.41198
Akaike Information Criterion	123.3225
Schwarz Criterion	141.5507
Hannan-Quinn Criterion	130.3896
Sum of Squared Residuals	6.558486
Root Mean Squared Error	0.342222

Figure B.11

ETS Smoothing
Original series: LCIT
Date: 12/28/24 Time: 11:59
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,M,A - Additive Error, Multiplicative Trend, Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.443748
Beta:	0.000000
Gamma:	0.000000
Initial Parameters	
Initial level:	12.94160
Initial trend:	1.002780
Initial state 1:	-0.238300
Initial state 2:	0.081244
Initial state 3:	0.327040
Initial state 4:	-0.169985
Compact Log-likelihood	-53.99054
Log-likelihood	-20.74125
Akaike Information Criterion	123.9811
Schwarz Criterion	140.1839
Hannan-Quinn Criterion	130.2629
Sum of Squared Residuals	6.877352
Root Mean Squared Error	0.350442
Average Mean Squared Error	0.141527

Figure B.12

ETS Smoothing
Original series: LCIT
Date: 12/28/24 Time: 12:01
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,MD,A - Additive Error, Multiplicative
-Dampened Trend, Additive Season, Cycle=4
Convergence achieved after 0 iterations

Parameters	
Alpha:	0.494903
Beta:	0.000000
Gamma:	0.000000
Phi:	0.294836

Initial Parameters	
Initial level:	10.50987
Initial trend:	1.829393
Initial state 1:	-0.244118
Initial state 2:	0.068440
Initial state 3:	0.317585
Initial state 4:	-0.141907

Compact Log-likelihood	-52.66273
Log-likelihood	-19.41344
Akaike Information Criterion	123.3255
Schwarz Criterion	141.5536
Hannan-Quinn Criterion	130.3925
Sum of Squared Residuals	6.558827
Root Mean Squared Error	0.342231
Average Mean Squared Error	0.205476

Figure B.13

Appendix C

BVAR, ARIMA and ETS Models for PIT

Bayesian VAR Estimates
Date: 12/28/24 Time: 04:33
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Litterman / Minnesota
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L1: 0.1, L2: 0.99, L3: 1, L4: 10
Standard errors in ( )

	LPIT	LRGDP	LGFCF	INFL	SMCAR
LPIT(-1)	0.167511	0.000169	0.010606	0.033770	0.000200
	(0.08606)	(0.01141)	(0.01039)	(0.01906)	(0.00016)
LPIT(-2)	0.029072	0.001039	0.002518	0.005362	2.95E-05
	(0.04823)	(0.00638)	(0.00581)	(0.01066)	(9.1E-05)
LPIT(-3)	0.007646	0.001430	-2.78E-05	0.002063	3.09E-06
	(0.03278)	(0.00434)	(0.00395)	(0.00724)	(6.2E-05)
LPIT(-4)	0.001520	-0.000108	0.000246	0.000300	1.03E-05
	(0.02475)	(0.00327)	(0.00298)	(0.00547)	(4.7E-05)
LRGDP(-1)	0.480921	0.549000	0.023405	-0.058323	0.000391
	(0.40647)	(0.05463)	(0.04946)	(0.09068)	(0.00077)
LRGDP(-2)	0.061721	0.137957	0.016790	-0.007013	0.000197
	(0.31646)	(0.04263)	(0.03851)	(0.07060)	(0.00060)
LRGDP(-3)	0.096264	0.091690	0.004651	-0.003560	-2.27E-06
	(0.22711)	(0.03062)	(0.02764)	(0.05067)	(0.00043)
LRGDP(-4)	0.062595	0.073091	0.005090	-0.003390	-0.000134
	(0.17711)	(0.02389)	(0.02155)	(0.03951)	(0.00034)
LGFCF(-1)	0.504104	0.037430	0.811715	0.043061	0.000161
	(0.41334)	(0.05522)	(0.05067)	(0.09220)	(0.00079)
LGFCF(-2)	-0.022901	0.003242	0.105706	0.007565	-9.40E-05
	(0.35988)	(0.04808)	(0.04422)	(0.08029)	(0.00068)
LGFCF(-3)	-0.081373	-0.003422	0.018075	0.001134	-0.000181
	(0.24737)	(0.03305)	(0.03038)	(0.05519)	(0.00047)
LGFCF(-4)	-0.031393	-0.000398	0.005378	0.005211	-0.000192
	(0.18598)	(0.02485)	(0.02284)	(0.04149)	(0.00035)
INFL(-1)	0.534270	-0.004949	0.019328	0.571487	-0.000403
	(0.28376)	(0.03790)	(0.03452)	(0.06369)	(0.00054)
INFL(-2)	0.153384	0.003866	-0.007705	0.109267	1.39E-05
	(0.20222)	(0.02702)	(0.02461)	(0.04553)	(0.00038)
INFL(-3)	0.037820	0.001232	-0.006930	0.041895	8.25E-05
	(0.14096)	(0.01883)	(0.01715)	(0.03175)	(0.00027)
INFL(-4)	0.023015	0.002660	-0.006753	0.023666	5.99E-05
	(0.10736)	(0.01434)	(0.01307)	(0.02418)	(0.00020)
SMCAR(-1)	54.48113	7.670075	4.160388	-10.81749	0.266207
	(40.7871)	(5.44838)	(4.96253)	(9.09843)	(0.07807)
SMCAR(-2)	12.79494	1.377606	2.072539	-4.396513	0.056330
	(24.3092)	(3.24774)	(2.95822)	(5.42335)	(0.04665)
SMCAR(-3)	7.747912	-0.328651	1.487909	-2.100231	0.024683
	(16.7644)	(2.23973)	(2.04010)	(3.74012)	(0.03219)
SMCAR(-4)	2.553708	-1.082301	0.950489	-1.180216	0.016073
	(12.7478)	(1.70311)	(1.55132)	(2.84403)	(0.02448)
C	-0.364330	1.058184	-0.061776	-0.307975	-0.001300
	(2.18413)	(0.29208)	(0.26581)	(0.48727)	(0.00415)
R-squared	0.867169	0.919732	0.995880	0.929482	0.529780
Adj. R-squared	0.781471	0.867946	0.993223	0.883987	0.226412
Sum sq. resids	1.847461	0.066867	0.050908	0.112976	7.37E-06
S.E. equation	0.244122	0.046444	0.040524	0.060369	0.000487
F-statistic	10.11894	17.76023	374.7070	20.43021	1.746330
Mean dependent	14.75494	10.39428	11.27978	0.335577	0.005408
S.D. dependent	0.522219	0.127805	0.492248	0.177238	0.000554
Data marginal log-likelihood		-1529.542			

Figure C.1

Bayesian VAR Estimates
Date: 12/28/24 Time: 05:12
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Normal-Wishart
Hyper-parameters: Mu1: 0, C1: 0.1, C2: 0.1, C3: 6
Standard errors in ( )

	LPIT	LRGDP	INFL	LGFCF	SMCAR
LPIT(-1)	0.552549	-0.035808	0.069779	0.044802	0.000994
	(0.14134)	(0.04380)	(0.05299)	(0.04943)	(0.03429)
LPIT(-2)	0.232453	0.016350	0.014201	0.035259	0.000110
	(0.15868)	(0.04918)	(0.05949)	(0.05550)	(0.03850)
LPIT(-3)	0.015118	0.038972	0.005425	-0.022797	-0.000131
	(0.15314)	(0.04746)	(0.05742)	(0.05356)	(0.03716)
LPIT(-4)	-0.103817	-0.034729	-0.047572	0.004057	0.000611
	(0.14650)	(0.04540)	(0.05493)	(0.05124)	(0.03555)
LRGDP(-1)	0.073918	0.236125	-0.032217	-0.010162	-0.000203
	(0.38918)	(0.12061)	(0.14592)	(0.13612)	(0.09443)
LRGDP(-2)	-0.271471	0.131531	-0.017119	0.012062	0.000194
	(0.39251)	(0.12165)	(0.14717)	(0.13728)	(0.09524)
LRGDP(-3)	0.230317	0.192815	-0.010841	-0.007624	0.000361
	(0.39353)	(0.12196)	(0.14755)	(0.13764)	(0.09548)
LRGDP(-4)	0.274629	0.464592	-0.022615	0.011920	-0.001212
	(0.39270)	(0.12171)	(0.14724)	(0.13735)	(0.09528)
INFL(-1)	0.046862	-0.059076	0.605519	0.044079	-0.001993
	(0.36846)	(0.11419)	(0.13815)	(0.12887)	(0.08940)
INFL(-2)	0.419886	-0.000900	0.185626	0.001634	-0.000684
	(0.42993)	(0.13325)	(0.16120)	(0.15037)	(0.10432)
INFL(-3)	-0.097281	0.029958	0.056430	-0.021865	0.000146
	(0.43987)	(0.13633)	(0.16492)	(0.15385)	(0.10673)
INFL(-4)	0.066128	0.079032	0.043125	-0.100421	0.001084
	(0.37781)	(0.11709)	(0.14165)	(0.13214)	(0.09167)
LGFCF(-1)	0.260106	0.032415	-0.011864	0.625891	0.000370
	(0.40006)	(0.12399)	(0.15000)	(0.13992)	(0.09707)
LGFCF(-2)	0.010154	0.024667	0.004832	0.264372	8.17E-06
	(0.46956)	(0.14553)	(0.17606)	(0.16423)	(0.11393)
LGFCF(-3)	-0.301913	-0.015271	-0.017972	0.037562	-0.000398
	(0.46753)	(0.14490)	(0.17529)	(0.16352)	(0.11344)
LGFCF(-4)	0.146170	-0.048208	0.052043	-0.007250	-0.000718
	(0.38150)	(0.11823)	(0.14304)	(0.13343)	(0.09256)
SMCAR(-1)	0.008069	0.001073	-0.004221	0.001394	7.08E-05
	(0.57154)	(0.17713)	(0.21429)	(0.19990)	(0.13867)
SMCAR(-2)	0.011018	0.001349	-0.004350	0.002282	6.46E-05
	(0.57154)	(0.17713)	(0.21429)	(0.19990)	(0.13867)
SMCAR(-3)	0.015414	0.000808	-0.004086	0.003214	5.74E-05
	(0.57154)	(0.17713)	(0.21429)	(0.19990)	(0.13867)
SMCAR(-4)	0.010032	-0.001370	-0.003139	0.003485	5.47E-05
	(0.57154)	(0.17713)	(0.21429)	(0.19990)	(0.13867)
C	-0.107346	0.041854	-0.013483	-0.012137	-0.000244
	(0.56910)	(0.17637)	(0.21337)	(0.19904)	(0.13808)
R-squared	0.891968	0.965818	0.939504	0.995160	0.450426
Adj. R-squared	0.822270	0.943766	0.900475	0.992038	0.095863
Sum sq. resids	1.502545	0.028475	0.096919	0.059810	8.61E-06
S.E. equation	0.220157	0.030307	0.055914	0.043925	0.000527
F-statistic	12.79759	43.79603	24.07158	318.7028	1.270367
Mean dependent	14.75494	10.39428	0.335577	11.27978	0.005408
S.D. dependent	0.522219	0.127805	0.177238	0.492248	0.000554
Data marginal log-likelihood		238.1610			

Figure C.2

Bayesian VAR Estimates
Date: 12/28/24 Time: 05:28
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Sims-Zha (normal-Wishart)
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L0: 1, L1: 0.1, L3: 1, L4: 10, L5: 10, C3: 6
Standard errors in ( )

	LPIT	LRGDP	INFL	LGFCF	SMCAR
LPIT(-1)	0.166776	0.000132	0.034331	0.010722	0.000203
	(0.07673)	(0.01843)	(0.02465)	(0.01691)	(0.00016)
LPIT(-2)	0.028906	0.001052	0.005457	0.002551	2.99E-05
	(0.04299)	(0.01033)	(0.01381)	(0.00947)	(8.8E-05)
LPIT(-3)	0.007573	0.001454	0.002102	-3.53E-05	3.11E-06
	(0.02922)	(0.00702)	(0.00939)	(0.00644)	(6.0E-05)
LPIT(-4)	0.001485	-0.000112	0.000306	0.000248	1.05E-05
	(0.02206)	(0.00530)	(0.00709)	(0.00486)	(4.5E-05)
LRGDP(-1)	0.482462	0.549027	-0.058890	0.023250	0.000391
	(0.36451)	(0.08756)	(0.11712)	(0.08032)	(0.00075)
LRGDP(-2)	0.061097	0.137939	-0.007095	0.016927	0.000198
	(0.28441)	(0.06832)	(0.09138)	(0.06267)	(0.00059)
LRGDP(-3)	0.097112	0.091652	-0.003559	0.004648	-2.82E-06
	(0.20424)	(0.04906)	(0.06562)	(0.04501)	(0.00042)
LRGDP(-4)	0.063261	0.073071	-0.003420	0.005134	-0.000137
	(0.15935)	(0.03828)	(0.05120)	(0.03511)	(0.00033)
INFL(-1)	0.537850	-0.005003	0.570844	0.019725	-0.000410
	(0.25459)	(0.06115)	(0.08180)	(0.05610)	(0.00052)
INFL(-2)	0.154640	0.003935	0.109145	-0.007831	1.44E-05
	(0.18191)	(0.04369)	(0.05845)	(0.04008)	(0.00037)
INFL(-3)	0.037829	0.001246	0.041852	-0.007051	8.41E-05
	(0.12682)	(0.03046)	(0.04075)	(0.02795)	(0.00026)
INFL(-4)	0.023052	0.002702	0.023654	-0.006881	6.11E-05
	(0.09661)	(0.02321)	(0.03104)	(0.02129)	(0.00020)
LGFCF(-1)	0.506918	0.037598	0.042964	0.811462	0.000164
	(0.37119)	(0.08916)	(0.11927)	(0.08179)	(0.00076)
LGFCF(-2)	-0.024268	0.003200	0.007468	0.105722	-9.43E-05
	(0.32382)	(0.07778)	(0.10405)	(0.07136)	(0.00067)
LGFCF(-3)	-0.083124	-0.003521	0.001063	0.018102	-0.000183
	(0.22252)	(0.05345)	(0.07150)	(0.04903)	(0.00046)
LGFCF(-4)	-0.031957	-0.000410	0.005280	0.005406	-0.000194
	(0.16725)	(0.04017)	(0.05374)	(0.03686)	(0.00034)
SMCAR(-1)	55.11342	7.759324	-10.95554	4.177395	0.265859
	(36.5988)	(8.79110)	(11.7594)	(8.06485)	(0.07534)
SMCAR(-2)	12.94964	1.392220	-4.459247	2.096907	0.056279
	(21.8633)	(5.25160)	(7.02477)	(4.81776)	(0.04500)
SMCAR(-3)	7.871619	-0.339522	-2.133237	1.509987	0.024646
	(15.0839)	(3.62318)	(4.84652)	(3.32386)	(0.03105)
SMCAR(-4)	2.593637	-1.106113	-1.199864	0.965560	0.016047
	(11.4721)	(2.75563)	(3.68605)	(2.52798)	(0.02361)
C	-0.371811	1.058177	-0.307431	-0.062482	-0.001312
	(1.94859)	(0.46805)	(0.62609)	(0.42939)	(0.00401)
R-squared	0.867390	0.919790	0.929580	0.995884	0.530397
Adj. R-squared	0.781835	0.868041	0.884148	0.993228	0.227428
Sum sq. resids	1.844382	0.066819	0.112818	0.050867	7.36E-06
S.E. equation	0.243918	0.046427	0.060327	0.040508	0.000487
F-statistic	10.13841	17.77423	20.46091	375.0106	1.750662
Mean dependent	14.75494	10.39428	0.335577	11.27978	0.005408
S.D. dependent	0.522219	0.127805	0.177238	0.492248	0.000554
Data marginal log-likelihood		493.1657			

Figure C.3

Forecast Evaluation
Date: 12/28/24 Time: 05:14
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LPIT	56	0.163804	0.121556	0.822864	0.005570
LRGDP	56	0.022587	0.017634	0.169639	0.001089
INFL	56	0.041716	0.030130	9.048963	0.056501
LGFCF	56	0.032532	0.025120	0.224920	0.001450
SMCAR	56	0.000745	0.000559	10.36945	0.070197

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure C.4

Forecast Evaluation
Date: 12/28/24 Time: 05:32
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LPIT	56	0.181016	0.139110	0.937714	0.006156
LRGDP	56	0.034629	0.025478	0.245469	0.001669
INFL	56	0.044774	0.032814	10.38656	0.061013
LGFCF	56	0.030048	0.023673	0.211346	0.001339
SMCAR	56	0.000364	0.000292	5.571111	0.034250

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure C.5

Forecast Evaluation
Date: 12/28/24 Time: 04:37
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LPIT	56	0.181241	0.139091	0.937515	0.006163
LRGDP	56	0.034470	0.025459	0.245298	0.001661
LGFCF	56	0.030165	0.023767	0.212232	0.001344
INFL	56	0.045023	0.033118	10.48521	0.061381
SMCAR	56	0.000362	0.000291	5.543666	0.034143

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure C.6

Dependent Variable: D(LPIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/27/24 Time: 17:06
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 29 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.043920	0.023313	1.883912	0.0653
AR(3)	0.067617	0.191810	0.352520	0.7259
AR(1)	-0.371794	0.120678	-3.080878	0.0033
SIGMASQ	0.037917	0.006824	5.556176	0.0000

R-squared	0.142318	Mean dependent var	0.043407
Adjusted R-squared	0.091866	S.D. dependent var	0.212197
S.E. of regression	0.202215	Akaike info criterion	-0.286158
Sum squared resid	2.085443	Schwarz criterion	-0.140171
Log likelihood	11.86936	Hannan-Quinn criter.	-0.229704
F-statistic	2.820855	Durbin-Watson stat	2.017365
Prob(F-statistic)	0.048022		

Inverted AR Roots	.31	-.34+.31i	-.34-.31i

Figure C.7

Dependent Variable: D(LPIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/27/24 Time: 17:08
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.039934	0.016708	2.390096	0.0206
AR(5)	-0.157901	0.145431	-1.085748	0.2827
MA(1)	-0.362328	0.129913	-2.788995	0.0074
SIGMASQ	0.037310	0.007600	4.908913	0.0000

R-squared	0.156054	Mean dependent var	0.043407
Adjusted R-squared	0.106410	S.D. dependent var	0.212197
S.E. of regression	0.200589	Akaike info criterion	-0.300276
Sum squared resid	2.052043	Schwarz criterion	-0.154288
Log likelihood	12.25758	Hannan-Quinn criter.	-0.243821
F-statistic	3.143465	Durbin-Watson stat	2.016391
Prob(F-statistic)	0.033016		

Inverted AR Roots	.56-.41i	.56+.41i	-.21+.66i	-.21-.66i
	-.69			
Inverted MA Roots	.36			

Figure C.8

Dependent Variable: D(LPIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/27/24 Time: 17:02
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 13 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.042155	0.019156	2.200640	0.0323
AR(1)	-0.208309	0.574778	-0.362417	0.7185
MA(1)	-0.187335	0.587853	-0.318676	0.7513
SIGMASQ	0.037884	0.006533	5.798863	0.0000

R-squared	0.143078	Mean dependent var	0.043407
Adjusted R-squared	0.092671	S.D. dependent var	0.212197
S.E. of regression	0.202126	Akaike info criterion	-0.287060
Sum squared resid	2.083593	Schwarz criterion	-0.141072
Log likelihood	11.89414	Hannan-Quinn criter.	-0.230605
F-statistic	2.838453	Durbin-Watson stat	1.972406
Prob(F-statistic)	0.047047		

Inverted AR Roots	-.21
Inverted MA Roots	.19

Figure C.9

Dependent Variable: D(LPIT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/27/24 Time: 17:04
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 17 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.041958	0.018780	2.234129	0.0299
AR(2)	0.079336	0.231602	0.342552	0.7333
MA(1)	-0.390188	0.127834	-3.052293	0.0036
SIGMASQ	0.037939	0.006512	5.826259	0.0000

R-squared	0.141824	Mean dependent var	0.043407
Adjusted R-squared	0.091343	S.D. dependent var	0.212197
S.E. of regression	0.202273	Akaike info criterion	-0.285655
Sum squared resid	2.086642	Schwarz criterion	-0.139667
Log likelihood	11.85551	Hannan-Quinn criter.	-0.229200
F-statistic	2.809465	Durbin-Watson stat	1.984329
Prob(F-statistic)	0.048664		

Inverted AR Roots	.28	-.28
Inverted MA Roots	.39	

Figure C.10

ETS Smoothing
Original series: LPIT
Date: 12/27/24 Time: 19:02
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,A,A - Additive Error, Additive Trend, Additive Season (Holt-Winters - additive seasonal), Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.645340
Beta:	0.000000
Gamma:	0.000000

Initial Parameters	
Initial level:	13.69057
Initial trend:	0.041122
Initial state 1:	0.000813
Initial state 2:	0.031330
Initial state 3:	-0.049753
Initial state 4:	0.017610

Compact Log-likelihood	-19.28036
Log-likelihood	13.96893
Akaike Information Criterion	54.56072
Schwarz Criterion	70.76354
Hannan-Quinn Criterion	60.84252
Sum of Squared Residuals	1.990895
Root Mean Squared Error	0.188552
Average Mean Squared Error	0.049878

Figure C.11

ETS Smoothing
Original series: LPIT
Date: 12/27/24 Time: 18:59
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,AD,A - Additive Error, Additive-Dampened Trend, Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.645341
Beta:	0.000000
Gamma:	0.000000
Phi:	1.000000
Initial Parameters	
Initial level:	13.69057
Initial trend:	0.041122
Initial state 1:	0.000812
Initial state 2:	0.031331
Initial state 3:	-0.049752
Initial state 4:	0.017610
Compact Log-likelihood	-19.28036
Log-likelihood	13.96893
Akaike Information Criterion	56.56072
Schwarz Criterion	74.78889
Hannan-Quinn Criterion	63.62774
Sum of Squared Residuals	1.990895
Root Mean Squared Error	0.188552
Average Mean Squared Error	0.049878

Figure C.12

ETS Smoothing
Original series: LPIT
Date: 12/27/24 Time: 18:57
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,M,A - Additive Error, Multiplicative Trend, Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.644010
Beta:	0.000000
Gamma:	0.000000
Initial Parameters	
Initial level:	13.69534
Initial trend:	1.002797
Initial state 1:	0.000773
Initial state 2:	0.031394
Initial state 3:	-0.049701
Initial state 4:	0.017534
Compact Log-likelihood	-19.28857
Log-likelihood	13.96072
Akaike Information Criterion	54.57714
Schwarz Criterion	70.77996
Hannan-Quinn Criterion	60.85894
Sum of Squared Residuals	1.991479
Root Mean Squared Error	0.188579
Average Mean Squared Error	0.049975

Figure C.13

ETS Smoothing
Original series: LPIT
Date: 12/27/24 Time: 19:01
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,MD,A - Additive Error, Multiplicative -Dampened Trend, Additive Season, Cycle=4
Convergence achieved on boundaries.

Parameters	
Alpha:	0.644010
Beta:	0.000000
Gamma:	0.000000
Phi:	1.000000
Initial Parameters	
Initial level:	13.69533
Initial trend:	1.002797
Initial state 1:	0.000773
Initial state 2:	0.031394
Initial state 3:	-0.049702
Initial state 4:	0.017534
Compact Log-likelihood	-19.28857
Log-likelihood	13.96072
Akaike Information Criterion	56.57714
Schwarz Criterion	74.80531
Hannan-Quinn Criterion	63.64416
Sum of Squared Residuals	1.991479
Root Mean Squared Error	0.188579
Average Mean Squared Error	0.049975

Figure C.14

Appendix D

BVAR, ARIMA and ETS Models for VAT

Bayesian VAR Estimates
Date: 12/28/24 Time: 12:23
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Litterman / Minnesota
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L1: 0.1, L2: 0.99, L3: 1, L4: 10
Standard errors in ( )

	LVAT	LGFCF	LRGDP	INFL	SMCAR
LVAT(-1)	0.214944	0.006200	-0.000900	0.007475	0.000182
	(0.08332)	(0.00858)	(0.00942)	(0.01574)	(0.00013)
LVAT(-2)	0.031314	0.002283	-0.002218	0.001285	5.16E-05
	(0.04773)	(0.00491)	(0.00539)	(0.00900)	(7.7E-05)
LVAT(-3)	0.006908	0.000629	-0.000151	9.55E-05	1.61E-05
	(0.03261)	(0.00335)	(0.00368)	(0.00614)	(5.2E-05)
LVAT(-4)	-0.001788	0.000644	0.000623	-0.000390	5.47E-06
	(0.02467)	(0.00253)	(0.00278)	(0.00465)	(4.0E-05)
LGFCF(-1)	0.701388	0.812794	0.040576	0.061407	0.000123
	(0.48473)	(0.05066)	(0.05521)	(0.09219)	(0.00079)
LGFCF(-2)	0.033495	0.104773	0.003738	0.008007	-0.000120
	(0.42227)	(0.04423)	(0.04810)	(0.08032)	(0.00068)
LGFCF(-3)	-0.078189	0.017170	-0.003347	-0.000583	-0.000198
	(0.29013)	(0.03038)	(0.03305)	(0.05519)	(0.00047)
LGFCF(-4)	-0.058117	0.004615	-0.000539	0.002873	-0.000202
	(0.21805)	(0.02283)	(0.02484)	(0.04148)	(0.00035)
LRGDP(-1)	0.219355	0.024146	0.551345	-0.045745	0.000351
	(0.47765)	(0.04956)	(0.05474)	(0.09085)	(0.00077)
LRGDP(-2)	-0.036673	0.018913	0.138939	0.000685	0.000236
	(0.37088)	(0.03848)	(0.04260)	(0.07055)	(0.00060)
LRGDP(-3)	0.122062	0.005765	0.091700	-0.000621	2.36E-05
	(0.26644)	(0.02764)	(0.03062)	(0.05068)	(0.00043)
LRGDP(-4)	0.114910	0.005595	0.073016	-0.001458	-0.000125
	(0.20769)	(0.02155)	(0.02388)	(0.03951)	(0.00034)
INFL(-1)	0.257397	0.024514	-0.003089	0.591767	-0.000332
	(0.32747)	(0.03398)	(0.03730)	(0.06269)	(0.00053)
INFL(-2)	0.060436	-0.006422	0.004209	0.113738	3.66E-05
	(0.23681)	(0.02457)	(0.02698)	(0.04547)	(0.00038)
INFL(-3)	-0.005126	-0.006220	0.001266	0.043952	9.63E-05
	(0.16522)	(0.01714)	(0.01882)	(0.03173)	(0.00027)
INFL(-4)	-0.002166	-0.006310	0.002667	0.024721	6.92E-05
	(0.12590)	(0.01306)	(0.01434)	(0.02418)	(0.00020)
SMCAR(-1)	90.13373	4.285051	8.016059	-8.587676	0.261899
	(47.8369)	(4.96261)	(5.44852)	(9.09832)	(0.07807)
SMCAR(-2)	17.34435	1.973236	1.462071	-4.154896	0.052309
	(28.5956)	(2.96688)	(3.25724)	(5.43920)	(0.04679)
SMCAR(-3)	3.185435	1.448020	-0.277126	-1.944198	0.022960
	(19.6858)	(2.04253)	(2.24239)	(3.74456)	(0.03223)
SMCAR(-4)	-1.034599	0.950968	-1.059040	-1.068569	0.015769
	(14.9531)	(1.55151)	(1.70331)	(2.84437)	(0.02448)
C	-0.595203	-0.042459	1.057130	-0.274553	-0.000832
	(2.56399)	(0.26605)	(0.29235)	(0.48771)	(0.00416)
R-squared	0.866688	0.995840	0.920515	0.923451	0.543229
Adj. R-squared	0.780681	0.993156	0.869234	0.874065	0.248539
Sum sq. resids	2.589683	0.051406	0.066214	0.122637	7.16E-06
S.E. equation	0.289030	0.040722	0.046216	0.062897	0.000480
F-statistic	10.07689	371.0639	17.95055	18.69852	1.843389
Mean dependent	14.93009	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.617169	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		-1581.256			

Figure D.1

Bayesian VAR Estimates
Date: 12/28/24 Time: 12:30
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Normal-Wishart
Hyper-parameters: Mu1: 0, C1: 0.1, C2: 0.1, C3: 6
Standard errors in ( )

	LVAT	LGFCF	LRGDP	INFL	SMCAR
LVAT(-1)	0.673620	0.010022	0.013151	0.008131	0.000593
	(0.13121)	(0.04299)	(0.03827)	(0.04691)	(0.02966)
LVAT(-2)	0.136010	0.021352	-0.014713	0.007917	0.000312
	(0.15543)	(0.05093)	(0.04533)	(0.05557)	(0.03513)
LVAT(-3)	0.085769	-0.012692	-0.017302	0.000658	0.000123
	(0.15530)	(0.05089)	(0.04529)	(0.05553)	(0.03510)
LVAT(-4)	-0.262659	0.026938	0.012785	-0.023374	0.000226
	(0.13293)	(0.04356)	(0.03877)	(0.04753)	(0.03004)
LGFCF(-1)	0.489433	0.644643	0.033548	0.041910	0.000367
	(0.42240)	(0.13840)	(0.12318)	(0.15103)	(0.09547)
LGFCF(-2)	0.179553	0.265571	0.016651	0.016892	3.57E-05
	(0.50358)	(0.16500)	(0.14686)	(0.18005)	(0.11381)
LGFCF(-3)	-0.283020	0.030076	-0.019125	-0.025610	-0.000504
	(0.50169)	(0.16438)	(0.14631)	(0.17938)	(0.11339)
LGFCF(-4)	-0.103399	-0.035694	-0.040419	0.023243	-0.001041
	(0.40658)	(0.13322)	(0.11857)	(0.14537)	(0.09189)
LRGDP(-1)	-0.691643	-0.006155	0.222640	-0.057116	-0.000286
	(0.42835)	(0.14035)	(0.12492)	(0.15316)	(0.09681)
LRGDP(-2)	-0.603352	0.030911	0.154173	-0.006221	0.000464
	(0.44225)	(0.14491)	(0.12897)	(0.15812)	(0.09995)
LRGDP(-3)	0.576674	0.010406	0.210265	0.001528	0.000687
	(0.43724)	(0.14327)	(0.12751)	(0.15633)	(0.09882)
LRGDP(-4)	0.945970	0.007810	0.428476	0.015403	-0.000893
	(0.44394)	(0.14546)	(0.12947)	(0.15873)	(0.10033)
INFL(-1)	-0.090465	0.074692	-0.059510	0.622466	-0.001678
	(0.38968)	(0.12768)	(0.11364)	(0.13933)	(0.08807)
INFL(-2)	0.222018	-0.005452	0.010800	0.188631	-0.000579
	(0.45879)	(0.15033)	(0.13380)	(0.16404)	(0.10369)
INFL(-3)	-0.227542	-0.006446	0.006108	0.050450	0.000660
	(0.46842)	(0.15348)	(0.13661)	(0.16748)	(0.10587)
INFL(-4)	0.288043	-0.085457	0.077560	0.065879	0.001324
	(0.40340)	(0.13218)	(0.11764)	(0.14423)	(0.09117)
SMCAR(-1)	0.016464	0.001462	0.000888	-0.002389	6.22E-05
	(0.61358)	(0.20104)	(0.17894)	(0.21938)	(0.13867)
SMCAR(-2)	0.010301	0.001960	0.001018	-0.003129	5.16E-05
	(0.61358)	(0.20104)	(0.17894)	(0.21938)	(0.13867)
SMCAR(-3)	0.003567	0.002795	0.000466	-0.002802	4.75E-05
	(0.61358)	(0.20104)	(0.17894)	(0.21938)	(0.13867)
SMCAR(-4)	-0.000907	0.003506	-0.001698	-0.001927	5.51E-05
	(0.61357)	(0.20104)	(0.17893)	(0.21938)	(0.13867)
C	-0.100200	-0.007518	0.044420	-0.019340	-6.11E-05
	(0.61109)	(0.20023)	(0.17821)	(0.21849)	(0.13811)
R-squared	0.920748	0.995160	0.958220	0.933513	0.494660
Adj. R-squared	0.869617	0.992038	0.931266	0.890619	0.168633
Sum sq. resids	1.539535	0.059807	0.034804	0.106517	7.92E-06
S.E. equation	0.222851	0.043923	0.033507	0.058618	0.000505
F-statistic	18.00782	318.7223	35.54941	21.76295	1.517239
Mean dependent	14.93009	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.617169	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		228.0785			

Figure D.2

Bayesian VAR Estimates
Date: 12/28/24 Time: 12:38
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Sims-Zha (normal-Wishart)
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L0: 1, L1: 0.1, L3: 1, L4: 10, L5: inf, C3: 6
Standard errors in ( )

	LVAT	LGFCF	LRGDP	INFL	SMCAR
LVAT(-1)	0.214459	0.006242	-0.000935	0.007629	0.000184
	(0.07580)	(0.01397)	(0.01521)	(0.02075)	(0.00013)
LVAT(-2)	0.031226	0.002310	-0.002264	0.001313	5.24E-05
	(0.04342)	(0.00800)	(0.00871)	(0.01189)	(7.4E-05)
LVAT(-3)	0.006850	0.000634	-0.000157	0.000100	1.64E-05
	(0.02966)	(0.00547)	(0.00595)	(0.00812)	(5.0E-05)
LVAT(-4)	-0.001821	0.000654	0.000633	-0.000395	5.56E-06
	(0.02244)	(0.00414)	(0.00450)	(0.00614)	(3.8E-05)
LGFCF(-1)	0.704812	0.812562	0.040796	0.061678	0.000126
	(0.44419)	(0.08188)	(0.08911)	(0.12162)	(0.00076)
LGFCF(-2)	0.032727	0.104779	0.003706	0.007885	-0.000121
	(0.38774)	(0.07148)	(0.07778)	(0.10616)	(0.00066)
LGFCF(-3)	-0.079958	0.017186	-0.003444	-0.000721	-0.000201
	(0.26633)	(0.04910)	(0.05343)	(0.07292)	(0.00045)
LGFCF(-4)	-0.059176	0.004632	-0.000554	0.002861	-0.000205
	(0.20011)	(0.03689)	(0.04014)	(0.05479)	(0.00034)
LRGDP(-1)	0.219522	0.023990	0.551409	-0.046160	0.000349
	(0.43715)	(0.08059)	(0.08770)	(0.11969)	(0.00074)
LRGDP(-2)	-0.038722	0.019093	0.138930	0.000792	0.000239
	(0.34016)	(0.06271)	(0.06824)	(0.09313)	(0.00058)
LRGDP(-3)	0.123345	0.005786	0.091657	-0.000556	2.38E-05
	(0.24452)	(0.04508)	(0.04905)	(0.06695)	(0.00042)
LRGDP(-4)	0.116584	0.005652	0.072992	-0.001441	-0.000127
	(0.19069)	(0.03515)	(0.03825)	(0.05221)	(0.00032)
INFL(-1)	0.260075	0.024991	-0.003114	0.591489	-0.000337
	(0.29983)	(0.05527)	(0.06015)	(0.08209)	(0.00051)
INFL(-2)	0.060965	-0.006530	0.004281	0.113696	3.76E-05
	(0.21740)	(0.04008)	(0.04361)	(0.05952)	(0.00037)
INFL(-3)	-0.005513	-0.006328	0.001277	0.043945	9.83E-05
	(0.15171)	(0.02797)	(0.03043)	(0.04154)	(0.00026)
INFL(-4)	-0.002403	-0.006429	0.002707	0.024727	7.06E-05
	(0.11562)	(0.02131)	(0.02319)	(0.03166)	(0.00020)
SMCAR(-1)	91.13610	4.307886	8.113926	-8.668997	0.261508
	(43.8031)	(8.07480)	(8.78729)	(11.9931)	(0.07456)
SMCAR(-2)	17.52011	1.995728	1.479282	-4.214243	0.052216
	(26.2462)	(4.83830)	(5.26521)	(7.18611)	(0.04467)
SMCAR(-3)	3.192864	1.469442	-0.286300	-1.973788	0.022906
	(18.0758)	(3.33214)	(3.62615)	(4.94908)	(0.03077)
SMCAR(-4)	-1.084099	0.966326	-1.082157	-1.085290	0.015742
	(13.7326)	(2.53151)	(2.75488)	(3.75994)	(0.02337)
C	-0.601692	-0.043061	1.057061	-0.273785	-0.000838
	(2.33442)	(0.43033)	(0.46830)	(0.63915)	(0.00397)
R-squared	0.866937	0.995843	0.920586	0.923466	0.543959
Adj. R-squared	0.781091	0.993161	0.869351	0.874089	0.249740
Sum sq. resids	2.584843	0.051371	0.066155	0.122614	7.15E-06
S.E. equation	0.288759	0.040708	0.046196	0.062891	0.000480
F-statistic	10.09866	371.3185	17.96798	18.70240	1.848820
Mean dependent	14.93009	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.617169	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		477.6708			

Figure D.3

Forecast Evaluation
Date: 12/28/24 Time: 12:27
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LVAT	56	0.215016	0.165072	1.108741	0.007241
LGFCF	56	0.030302	0.023564	0.210299	0.001351
LRGDP	56	0.034366	0.025383	0.244555	0.001656
INFL	56	0.046924	0.034143	10.74145	0.063999
SMCAR	56	0.000357	0.000286	5.444275	0.033657

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure D.4

Forecast Evaluation
Date: 12/28/24 Time: 12:32
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LVAT	56	0.166471	0.128442	0.865180	0.005606
LGFCF	56	0.032774	0.024760	0.221470	0.001461
LRGDP	56	0.024672	0.019587	0.188350	0.001189
INFL	56	0.043633	0.028865	8.578180	0.059127
SMCAR	56	0.000763	0.000628	11.65416	0.072250

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure D.5

Forecast Evaluation
Date: 12/28/24 Time: 12:39
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LVAT	56	0.215062	0.164952	1.107705	0.007243
LGFCF	56	0.030399	0.023727	0.211789	0.001355
LRGDP	56	0.034113	0.025168	0.242491	0.001644
INFL	56	0.046751	0.033936	10.67857	0.063740
SMCAR	56	0.000359	0.000287	5.468391	0.033778

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure D.6

Dependent Variable: D(LVAT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 14:07
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 13 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.049959	0.019309	2.587374	0.0126
AR(10)	-0.335895	0.157707	-2.129874	0.0380
MA(1)	-0.183199	0.153562	-1.192998	0.2384
SIGMASQ	0.045198	0.008818	5.125826	0.0000

R-squared	0.141606	Mean dependent var	0.052086
Adjusted R-squared	0.091113	S.D. dependent var	0.231580
S.E. of regression	0.220778	Akaike info criterion	-0.090985
Sum squared resid	2.485888	Schwarz criterion	0.055003
Log likelihood	6.502085	Hannan-Quinn criter.	-0.034530
F-statistic	2.804434	Durbin-Watson stat	1.976073
Prob(F-statistic)	0.048950		

Inverted AR Roots	.85+.28i	.85-.28i	.53+.73i	.53-.73i
	.00+.90i	-.00-.90i	-.53-.73i	-.53+.73i
	-.85-.28i	-.85+.28i		
Inverted MA Roots	.18			

Figure D.7

Dependent Variable: D(LVAT)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 14:21
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 45 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.050580	0.023165	2.183502	0.0336
AR(10)	-0.326687	0.181898	-1.795990	0.0784
MA(2)	-0.067816	0.172097	-0.394058	0.6952
SIGMASQ	0.046493	0.008987	5.173171	0.0000

R-squared	0.117003	Mean dependent var	0.052086
Adjusted R-squared	0.065062	S.D. dependent var	0.231580
S.E. of regression	0.223920	Akaike info criterion	-0.064425
Sum squared resid	2.557138	Schwarz criterion	0.081563
Log likelihood	5.771697	Hannan-Quinn criter.	-0.007971
F-statistic	2.252619	Durbin-Watson stat	2.320349
Prob(F-statistic)	0.093395		

Inverted AR Roots	.85+.28i	.85-.28i	.53+.72i	.53-.72i
	.00+.89i	-.00-.89i	-.53-.72i	-.53+.72i
	-.85-.28i	-.85+.28i		
Inverted MA Roots	.26	-.26		

Figure D.8

ETS Smoothing
Original series: LVAT
Date: 12/28/24 Time: 14:33
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,A,A - Additive Error, Additive Trend, Additive Season (Holt-Winters - additive seasonal), Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.831192
Beta:	0.000000
Gamma:	0.000000
Initial Parameters	
Initial level:	13.28981
Initial trend:	0.048492
Initial state 1:	0.076042
Initial state 2:	0.052713
Initial state 3:	-0.051746
Initial state 4:	-0.077008
Compact Log-likelihood	-24.05356
Log-likelihood	9.195725
Akaike Information Criterion	64.10713
Schwarz Criterion	80.30994
Hannan-Quinn Criterion	70.38892
Sum of Squared Residuals	2.360931
Root Mean Squared Error	0.205328
Average Mean Squared Error	0.073275

Figure D.9

ETS Smoothing
Original series: LVAT
Date: 12/28/24 Time: 14:35
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,AD,A - Additive Error, Additive-Dampened Trend, Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.827028
Beta:	0.000000
Gamma:	0.000000
Phi:	0.975817
Initial Parameters	
Initial level:	13.24424
Initial trend:	0.089088
Initial state 1:	0.076446
Initial state 2:	0.052080
Initial state 3:	-0.052272
Initial state 4:	-0.076255
Compact Log-likelihood	-23.87120
Log-likelihood	9.378089
Akaike Information Criterion	65.74240
Schwarz Criterion	83.97057
Hannan-Quinn Criterion	72.80942
Sum of Squared Residuals	2.345604
Root Mean Squared Error	0.204660
Average Mean Squared Error	0.072211

Figure D.10

ETS Smoothing
Original series: LCIT
Date: 12/28/24 Time: 11:59
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,M,A - Additive Error, Multiplicative Trend, Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.443748
Beta:	0.000000
Gamma:	0.000000

Initial Parameters	
Initial level:	12.94160
Initial trend:	1.002780
Initial state 1:	-0.238300
Initial state 2:	0.081244
Initial state 3:	0.327040
Initial state 4:	-0.169985

Compact Log-likelihood	-53.99054
Log-likelihood	-20.74125
Akaike Information Criterion	123.9811
Schwarz Criterion	140.1839
Hannan-Quinn Criterion	130.2629
Sum of Squared Residuals	6.877352
Root Mean Squared Error	0.350442
Average Mean Squared Error	0.141527

Figure D.11

ETS Smoothing
Original series: LVAT
Date: 12/28/24 Time: 14:31
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,MD,A - Additive Error, Multiplicative -Dampened Trend, Additive Season, Cycle=4
Convergence achieved on boundaries.

Parameters	
Alpha:	0.827673
Beta:	0.000000
Gamma:	0.000000
Phi:	0.971110

Initial Parameters	
Initial level:	13.24620
Initial trend:	1.006641
Initial state 1:	0.076545
Initial state 2:	0.052143
Initial state 3:	-0.052327
Initial state 4:	-0.076360

Compact Log-likelihood	-23.96678
Log-likelihood	9.282513
Akaike Information Criterion	65.93355
Schwarz Criterion	84.16172
Hannan-Quinn Criterion	73.00057
Sum of Squared Residuals	2.353624
Root Mean Squared Error	0.205010
Average Mean Squared Error	0.072729

Figure D.12

Appendix E

BVAR, ARIMA and ETS Models for TTR

Bayesian VAR Estimates
Date: 12/28/24 Time: 14:58
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Litterman / Minnesota
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L1: 0.1, L2: 0.99, L3: 1, L4: 10
Standard errors in ( )

	LTTR	LGFCF	LRGDP	INFL	SMCAR
LTTR(-1)	0.241300	0.006701	0.021170	0.039862	0.000263
	(0.08124)	(0.01365)	(0.01498)	(0.02502)	(0.00021)
LTTR(-2)	0.033387	0.001395	0.003182	0.007380	7.53E-05
	(0.04742)	(0.00795)	(0.00872)	(0.01457)	(0.00012)
LTTR(-3)	0.015027	-6.81E-05	0.000179	0.004275	3.60E-05
	(0.03244)	(0.00543)	(0.00597)	(0.00996)	(8.5E-05)
LTTR(-4)	0.000171	0.000913	0.001128	0.002507	1.27E-05
	(0.02457)	(0.00411)	(0.00452)	(0.00754)	(6.4E-05)
LGFCF(-1)	0.403899	0.816990	0.031065	0.051492	0.000189
	(0.29574)	(0.05039)	(0.05490)	(0.09168)	(0.00078)
LGFCF(-2)	0.007945	0.105144	0.000407	0.002971	-0.000127
	(0.25917)	(0.04425)	(0.04811)	(0.08035)	(0.00068)
LGFCF(-3)	-0.051242	0.017147	-0.004510	-0.002823	-0.000208
	(0.17807)	(0.03039)	(0.03306)	(0.05520)	(0.00047)
LGFCF(-4)	-0.043879	0.004391	-0.000624	0.001902	-0.000212
	(0.13382)	(0.02283)	(0.02484)	(0.04149)	(0.00035)
LRGDP(-1)	0.325836	0.023577	0.532908	-0.075468	0.000242
	(0.29735)	(0.05027)	(0.05552)	(0.09216)	(0.00079)
LRGDP(-2)	0.189812	0.018053	0.135215	-0.006618	0.000188
	(0.22777)	(0.03852)	(0.04264)	(0.07061)	(0.00060)
LRGDP(-3)	0.125078	0.004516	0.089426	-0.006364	-2.35E-05
	(0.16364)	(0.02767)	(0.03065)	(0.05073)	(0.00043)
LRGDP(-4)	0.049531	0.005042	0.071266	-0.005128	-0.000148
	(0.12758)	(0.02157)	(0.02391)	(0.03955)	(0.00034)
INFL(-1)	0.488560	0.020908	-0.020419	0.559641	-0.000516
	(0.20904)	(0.03533)	(0.03879)	(0.06518)	(0.00055)
INFL(-2)	0.117263	-0.007522	0.000506	0.106146	-1.55E-05
	(0.14594)	(0.02468)	(0.02709)	(0.04566)	(0.00039)
INFL(-3)	0.044595	-0.006861	-0.000313	0.040404	6.94E-05
	(0.10156)	(0.01717)	(0.01885)	(0.03179)	(0.00027)
INFL(-4)	0.025856	-0.006761	0.001836	0.022702	5.18E-05
	(0.07732)	(0.01307)	(0.01435)	(0.02420)	(0.00020)
SMCAR(-1)	42.77634	4.574232	6.315942	-11.02859	0.260828
	(29.4420)	(4.97784)	(5.46516)	(9.12668)	(0.07832)
SMCAR(-2)	1.863540	2.086640	1.013079	-4.746575	0.053306
	(17.5300)	(2.96433)	(3.25445)	(5.43462)	(0.04675)
SMCAR(-3)	-1.837176	1.530992	-0.423561	-2.076824	0.024454
	(12.0641)	(2.04011)	(2.23974)	(3.74014)	(0.03219)
SMCAR(-4)	-0.812939	0.992238	-1.083927	-1.069859	0.016661
	(9.17173)	(1.55101)	(1.70276)	(2.84344)	(0.02448)
C	0.292345	-0.057370	1.055105	-0.299882	-0.001267
	(1.57110)	(0.26570)	(0.29196)	(0.48707)	(0.00415)
R-squared	0.921225	0.995832	0.920163	0.925601	0.540008
Adj. R-squared	0.870403	0.993143	0.868656	0.877601	0.243240
Sum sq. resids	1.023725	0.051508	0.066508	0.119193	7.21E-06
S.E. equation	0.181723	0.040762	0.046319	0.062008	0.000482
F-statistic	18.12640	370.3248	17.86460	19.28359	1.819626
Mean dependent	16.13707	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.504793	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		-1498.919			

Figure E.1

Bayesian VAR Estimates
Date: 12/28/24 Time: 15:02
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Normal-Wishart
Hyper-parameters: Mu1: 0, C1: 0.1, C2: 0.1, C3: 6
Standard errors in ( )

	LTTR	LGFCF	LRGDP	INFL	SMCAR
LTTR(-1)	0.654372	0.009605	0.012454	0.045547	0.000695
	(0.13087)	(0.06409)	(0.05691)	(0.06964)	(0.04413)
LTTR(-2)	-0.070598	0.021393	0.026067	-0.033687	0.000426
	(0.14532)	(0.07117)	(0.06319)	(0.07733)	(0.04900)
LTTR(-3)	0.298410	-0.025210	-0.024540	0.008959	0.000517
	(0.14349)	(0.07027)	(0.06240)	(0.07636)	(0.04838)
LTTR(-4)	-0.292518	0.050629	-0.007135	0.012540	0.000328
	(0.12129)	(0.05940)	(0.05274)	(0.06455)	(0.04090)
LGFCF(-1)	0.353012	0.659793	0.026715	0.033763	0.000690
	(0.28047)	(0.13735)	(0.12196)	(0.14925)	(0.09457)
LGFCF(-2)	0.089389	0.274884	0.014982	0.003400	0.000202
	(0.33762)	(0.16534)	(0.14682)	(0.17967)	(0.11384)
LGFCF(-3)	-0.123861	0.030138	-0.024974	-0.026335	-0.000453
	(0.33661)	(0.16485)	(0.14638)	(0.17913)	(0.11350)
LGFCF(-4)	-0.151881	-0.042002	-0.032426	0.020265	-0.001256
	(0.27406)	(0.13422)	(0.11918)	(0.14584)	(0.09241)
LRGDP(-1)	-0.249162	0.001423	0.219515	-0.035191	-0.000372
	(0.28418)	(0.13917)	(0.12358)	(0.15123)	(0.09582)
LRGDP(-2)	0.175327	0.026730	0.136465	0.013698	0.000137
	(0.28465)	(0.13940)	(0.12378)	(0.15148)	(0.09598)
LRGDP(-3)	0.364086	-0.018168	0.206776	-0.032684	2.47E-05
	(0.28816)	(0.14112)	(0.12531)	(0.15335)	(0.09716)
LRGDP(-4)	0.157020	-0.006986	0.440622	-0.025407	-0.001356
	(0.28827)	(0.14118)	(0.12536)	(0.15340)	(0.09720)
INFL(-1)	0.150588	0.061531	-0.070048	0.615076	-0.002013
	(0.26534)	(0.12995)	(0.11539)	(0.14120)	(0.08947)
INFL(-2)	0.125154	-0.002315	0.000495	0.195739	-0.000892
	(0.30992)	(0.15178)	(0.13477)	(0.16492)	(0.10450)
INFL(-3)	-0.036265	-0.022968	0.012416	0.040083	0.000129
	(0.31431)	(0.15393)	(0.13668)	(0.16726)	(0.10598)
INFL(-4)	0.227092	-0.110248	0.080573	0.051024	0.000668
	(0.27406)	(0.13422)	(0.11918)	(0.14585)	(0.09241)
SMCAR(-1)	0.007408	0.001544	0.000617	-0.003319	6.04E-05
	(0.41127)	(0.20141)	(0.17884)	(0.21886)	(0.13867)
SMCAR(-2)	0.001873	0.002026	0.000930	-0.004332	5.04E-05
	(0.41127)	(0.20141)	(0.17884)	(0.21886)	(0.13867)
SMCAR(-3)	0.000403	0.002972	0.000226	-0.003648	4.91E-05
	(0.41126)	(0.20141)	(0.17884)	(0.21886)	(0.13867)
SMCAR(-4)	0.001861	0.003588	-0.001882	-0.002700	5.47E-05
	(0.41126)	(0.20141)	(0.17884)	(0.21886)	(0.13867)
C	-0.045541	-0.010695	0.044693	-0.016579	-0.000166
	(0.40944)	(0.20052)	(0.17805)	(0.21789)	(0.13806)
R-squared	0.948645	0.995350	0.958944	0.933775	0.494271
Adj. R-squared	0.915512	0.992350	0.932456	0.891050	0.167995
Sum sq. resids	0.667392	0.057465	0.034202	0.106097	7.92E-06
S.E. equation	0.146727	0.043055	0.033216	0.058502	0.000506
F-statistic	28.63197	331.7745	36.20296	21.85522	1.514885
Mean dependent	16.13707	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.504793	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		259.2476			

Figure E.2

Bayesian VAR Estimates
Date: 12/28/24 Time: 15:06
Sample (adjusted): 2011Q1 2023Q4
Included observations: 52 after adjustments
Prior type: Sims-Zha (normal-Wishart)
Initial residual covariance: Univariate AR
Constant included in covariance calculation
Hyper-parameters: Mu1: 0, L0: 1, L1: 0.1, L3: 1, L4: 10, L5: inf, C3: 6
Standard errors in ( )

	LTTR	LGFCF	LRGDP	INFL	SMCAR
LTTR(-1)	0.240335	0.006728	0.021414	0.040456	0.000266
	(0.07862)	(0.02226)	(0.02397)	(0.03237)	(0.00021)
LTTR(-2)	0.033154	0.001398	0.003214	0.007497	7.65E-05
	(0.04586)	(0.01299)	(0.01399)	(0.01888)	(0.00012)
LTTR(-3)	0.014910	-8.07E-05	0.000171	0.004356	3.66E-05
	(0.03137)	(0.00888)	(0.00957)	(0.01292)	(8.2E-05)
LTTR(-4)	0.000106	0.000924	0.001143	0.002559	1.30E-05
	(0.02376)	(0.00673)	(0.00724)	(0.00978)	(6.2E-05)
LGFCF(-1)	0.405987	0.816807	0.031187	0.051740	0.000194
	(0.28818)	(0.08161)	(0.08788)	(0.11867)	(0.00075)
LGFCF(-2)	0.007335	0.105161	0.000314	0.002743	-0.000128
	(0.25305)	(0.07166)	(0.07717)	(0.10420)	(0.00066)
LGFCF(-3)	-0.052381	0.017167	-0.004624	-0.003011	-0.000210
	(0.17381)	(0.04922)	(0.05300)	(0.07157)	(0.00045)
LGFCF(-4)	-0.044706	0.004410	-0.000631	0.001868	-0.000216
	(0.13058)	(0.03698)	(0.03982)	(0.05377)	(0.00034)
LRGDP(-1)	0.325661	0.023465	0.532733	-0.076339	0.000240
	(0.28933)	(0.08193)	(0.08823)	(0.11914)	(0.00076)
LRGDP(-2)	0.191278	0.018230	0.135162	-0.006626	0.000190
	(0.22213)	(0.06290)	(0.06774)	(0.09147)	(0.00058)
LRGDP(-3)	0.126435	0.004512	0.089366	-0.006430	-2.45E-05
	(0.15968)	(0.04522)	(0.04869)	(0.06575)	(0.00042)
LRGDP(-4)	0.050017	0.005087	0.071226	-0.005198	-0.000151
	(0.12455)	(0.03527)	(0.03798)	(0.05129)	(0.00033)
INFL(-1)	0.491877	0.021382	-0.020750	0.558840	-0.000524
	(0.20350)	(0.05763)	(0.06206)	(0.08380)	(0.00053)
INFL(-2)	0.118008	-0.007638	0.000523	0.105980	-1.55E-05
	(0.14245)	(0.04034)	(0.04344)	(0.05866)	(0.00037)
INFL(-3)	0.044830	-0.006978	-0.000324	0.040340	7.08E-05
	(0.09916)	(0.02808)	(0.03024)	(0.04083)	(0.00026)
INFL(-4)	0.026003	-0.006888	0.001866	0.022678	5.28E-05
	(0.07550)	(0.02138)	(0.02302)	(0.03109)	(0.00020)
SMCAR(-1)	43.38480	4.604369	6.371891	-11.16867	0.260395
	(28.6694)	(8.11865)	(8.74251)	(11.8054)	(0.07501)
SMCAR(-2)	1.885573	2.111756	1.018146	-4.820220	0.053210
	(17.1089)	(4.84494)	(5.21723)	(7.04507)	(0.04476)
SMCAR(-3)	-1.872691	1.554407	-0.436938	-2.108628	0.024409
	(11.7789)	(3.33556)	(3.59188)	(4.85028)	(0.03082)
SMCAR(-4)	-0.828473	1.008517	-1.107703	-1.085573	0.016638
	(8.95655)	(2.53633)	(2.73123)	(3.68811)	(0.02343)
C	0.282805	-0.058030	1.054949	-0.298918	-0.001276
	(1.52100)	(0.43072)	(0.46382)	(0.62631)	(0.00398)
R-squared	0.921348	0.995835	0.920198	0.925635	0.540728
Adj. R-squared	0.870606	0.993148	0.868713	0.877657	0.244424
Sum sq. resids	1.022126	0.051467	0.066479	0.119139	7.20E-06
S.E. equation	0.181581	0.040746	0.046308	0.061993	0.000482
F-statistic	18.15718	370.6230	17.87305	19.29312	1.824907
Mean dependent	16.13707	11.27978	10.39428	0.335577	0.005408
S.D. dependent	0.504793	0.492248	0.127805	0.177238	0.000554
Data marginal log-likelihood		505.5937			

Figure E.3

Forecast Evaluation
Date: 12/28/24 Time: 14:56
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LTTR	56	0.135466	0.098021	0.610535	0.004216
LGFCF	56	0.030325	0.023629	0.210985	0.001352
LRGDP	56	0.034450	0.025190	0.242654	0.001660
INFL	56	0.046210	0.033620	10.42531	0.062990
SMCAR	56	0.000359	0.000285	5.443585	0.033787

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure E.4

Forecast Evaluation
Date: 12/28/24 Time: 12:32
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LVAT	56	0.166471	0.128442	0.865180	0.005606
LGFCF	56	0.032774	0.024760	0.221470	0.001461
LRGDP	56	0.024672	0.019587	0.188350	0.001189
INFL	56	0.043633	0.028865	8.578180	0.059127
SMCAR	56	0.000763	0.000628	11.65416	0.072250

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure E.5

Forecast Evaluation
Date: 12/28/24 Time: 15:07
Sample: 2010Q1 2023Q4
Included observations: 56

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
LTTR	56	0.134896	0.097485	0.607061	0.004198
LGFCF	56	0.030426	0.023771	0.212316	0.001356
LRGDP	56	0.034270	0.025059	0.241401	0.001652
INFL	56	0.045961	0.033411	10.36554	0.062637
SMCAR	56	0.000360	0.000287	5.465436	0.033902

RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Figure E.6

Dependent Variable: D(LTTR)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 15:26
Sample: 2010Q2 2023Q4
Included observations: 55
Failure to improve objective (non-zero gradients) after 13 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.035086	0.004646	7.551616	0.0000
AR(1)	0.736006	0.144446	5.095360	0.0000
MA(1)	-1.000000	1156.923	-0.000864	0.9993
SIGMASQ	0.021249	0.543817	0.039073	0.9690

R-squared	0.124889	Mean dependent var	0.042429
Adjusted R-squared	0.073412	S.D. dependent var	0.157260
S.E. of regression	0.151378	Akaike info criterion	-0.827460
Sum squared resid	1.168676	Schwarz criterion	-0.681472
Log likelihood	26.75515	Hannan-Quinn criter.	-0.771005
F-statistic	2.426113	Durbin-Watson stat	1.938254
Prob(F-statistic)	0.076190		

Inverted AR Roots	.74
Inverted MA Roots	1.00

Figure E.7

Dependent Variable: D(LTTR)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 15:32
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.041640	0.021213	1.962921	0.0551
AR(2)	-0.214859	0.183823	-1.168838	0.2479
MA(3)	0.253126	0.183940	1.376135	0.1748
SIGMASQ	0.021153	0.003383	6.252243	0.0000

R-squared	0.128821	Mean dependent var	0.042429
Adjusted R-squared	0.077575	S.D. dependent var	0.157260
S.E. of regression	0.151037	Akaike info criterion	-0.867280
Sum squared resid	1.163425	Schwarz criterion	-0.721292
Log likelihood	27.85019	Hannan-Quinn criter.	-0.810825
F-statistic	2.513789	Durbin-Watson stat	2.106735
Prob(F-statistic)	0.068750		

Inverted AR Roots	-.00+.46i	-.00-.46i	
Inverted MA Roots	.32+.55i	.32-.55i	-.63

Figure E.8

Dependent Variable: D(LTTR)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 16:45
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 23 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.038492	0.012735	3.022475	0.0039
AR(22)	-0.417919	0.106971	-3.906838	0.0003
MA(1)	-0.197676	0.149641	-1.321001	0.1924
SIGMASQ	0.019183	0.003479	5.513843	0.0000

R-squared	0.209973	Mean dependent var	0.042429
Adjusted R-squared	0.163501	S.D. dependent var	0.157260
S.E. of regression	0.143831	Akaike info criterion	-0.892906
Sum squared resid	1.055050	Schwarz criterion	-0.746919
Log likelihood	28.55493	Hannan-Quinn criter.	-0.836452
F-statistic	4.518247	Durbin-Watson stat	1.885014
Prob(F-statistic)	0.006936		

Inverted AR Roots	.95-.14i	.95+.14i	.87+.40i	.87-.40i
	.73-.63i	.73+.63i	.52+.81i	.52-.81i
	.27+.92i	.27-.92i	.00+.96i	.00-.96i
	-.27-.92i	-.27+.92i	-.52-.81i	-.52+.81i
	-.73-.63i	-.73+.63i	-.87+.40i	-.87-.40i
	-.95+.14i	-.95-.14i		
Inverted MA Roots	.20			

Figure E.9

Dependent Variable: D(LTTR)
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/28/24 Time: 16:45
Sample: 2010Q2 2023Q4
Included observations: 55
Convergence achieved after 23 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.038492	0.012735	3.022475	0.0039
AR(22)	-0.417919	0.106971	-3.906838	0.0003
MA(1)	-0.197676	0.149641	-1.321001	0.1924
SIGMASQ	0.019183	0.003479	5.513843	0.0000

R-squared	0.209973	Mean dependent var	0.042429
Adjusted R-squared	0.163501	S.D. dependent var	0.157260
S.E. of regression	0.143831	Akaike info criterion	-0.892906
Sum squared resid	1.055050	Schwarz criterion	-0.746919
Log likelihood	28.55493	Hannan-Quinn criter.	-0.836452
F-statistic	4.518247	Durbin-Watson stat	1.885014
Prob(F-statistic)	0.006936		

Inverted AR Roots	.95-.14i	.95+.14i	.87+.40i	.87-.40i
	.73-.63i	.73+.63i	.52+.81i	.52-.81i
	.27+.92i	.27-.92i	.00+.96i	.00-.96i
	-.27-.92i	-.27+.92i	-.52-.81i	-.52+.81i
	-.73-.63i	-.73+.63i	-.87+.40i	-.87-.40i
	-.95+.14i	-.95-.14i		
Inverted MA Roots	.20			

Figure E.10

ETS Smoothing
Original series: LTTR
Date: 12/28/24 Time: 17:02
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,A,A - Additive Error, Additive Trend, Additive Season (Holt-Winters - additive seasonal), Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.821389
Beta:	0.000000
Gamma:	0.000000

Initial Parameters	
Initial level:	14.74453
Initial trend:	0.041142
Initial state 1:	-0.019828
Initial state 2:	0.028937
Initial state 3:	0.036372
Initial state 4:	-0.045480

Compact Log-likelihood	-4.445556
Log-likelihood	28.80373
Akaike Information Criterion	24.89111
Schwarz Criterion	41.09392
Hannan-Quinn Criterion	31.17291
Sum of Squared Residuals	1.172068
Root Mean Squared Error	0.144671
Average Mean Squared Error	0.033818

Figure E.11

ETS Smoothing
Original series: LTTR
Date: 12/28/24 Time: 16:59
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,MD,A - Additive Error, Multiplicative -Dampened Trend, Additive Season, Cycle=4
Convergence achieved on boundaries.

Parameters	
Alpha:	0.814519
Beta:	0.000000
Gamma:	0.000000
Phi:	0.969005

Initial Parameters	
Initial level:	14.70225
Initial trend:	1.005409
Initial state 1:	-0.019382
Initial state 2:	0.028408
Initial state 3:	0.035830
Initial state 4:	-0.044855

Compact Log-likelihood	-4.234694
Log-likelihood	29.01460
Akaike Information Criterion	26.46939
Schwarz Criterion	44.69755
Hannan-Quinn Criterion	33.53641
Sum of Squared Residuals	1.163275
Root Mean Squared Error	0.144128
Average Mean Squared Error	0.033435

Figure E.12

ETS Smoothing
Original series: LTTR
Date: 12/28/24 Time: 17:00
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,M,A - Additive Error, Multiplicative Trend,
Additive Season, Cycle=4
Convergence achieved after 1 iteration

Parameters	
Alpha:	0.823810
Beta:	0.000000
Gamma:	0.000000

Initial Parameters	
Initial level:	14.74912
Initial trend:	1.002511
Initial state 1:	-0.019847
Initial state 2:	0.028999
Initial state 3:	0.036412
Initial state 4:	-0.045564

Compact Log-likelihood	-4.578120
Log-likelihood	28.67117
Akaike Information Criterion	25.15624
Schwarz Criterion	41.35905
Hannan-Quinn Criterion	31.43804
Sum of Squared Residuals	1.177630
Root Mean Squared Error	0.145014
Average Mean Squared Error	0.034142

Figure E.13

ETS Smoothing
Original series: LPIT
Date: 12/27/24 Time: 19:01
Sample: 2010Q1 2023Q4
Included observations: 56
Model: A,MD,A - Additive Error, Multiplicative
-Dampened Trend, Additive Season, Cycle=4
Convergence achieved on boundaries.

Parameters	
Alpha:	0.644010
Beta:	0.000000
Gamma:	0.000000
Phi:	1.000000

Initial Parameters	
Initial level:	13.69533
Initial trend:	1.002797
Initial state 1:	0.000773
Initial state 2:	0.031394
Initial state 3:	-0.049702
Initial state 4:	0.017534

Compact Log-likelihood	-19.28857
Log-likelihood	13.96072
Akaike Information Criterion	56.57714
Schwarz Criterion	74.80531
Hannan-Quinn Criterion	63.64416
Sum of Squared Residuals	1.991479
Root Mean Squared Error	0.188579
Average Mean Squared Error	0.049975

Figure E.14