Employing Bayesian Vector Auto-Regression (BVAR) Method as an Alternative Technique for Forecasting Tax Revenue in South Africa.

by

Mojalefa Aubrey Molapo

Submitted in accordance with the requirements for the degree of

MASTER OF SCIENCE

in the subject

STATISTICS

at the

UNIVERSITY OF SOUTH AFRICA

Supervisor: Professor John Olaomi

February 2017

DECLARATION

I declare that the dissertation titled "***Employing Bayesian Vector Auto-Regression (BVAR) method as an alternative technique for forecasting tax revenue in South Africa.***" is my own work. The dissertation has not been submitted for any degree or examination at any other university or institution, and all the sources I have used or quoted in this dissertation have been acknowledged by complete references.

Name: Mojalefa A. Molapo

Date of Declaration: 23 February 2017

Signed:________________________

DEDICATION

I would like to dedicate this dissertation to my wife, Tshifularo Molapo; for her an unwavering support she displayed during the course of my studies. To our son Tumelo Molapo; you gave me strength for always asking me, “Daddy you doing school work?” Let me also dedicate this work to our new born twins, Oabetswe Molapo and Keabetswe Molapo, they brought immense joy to the family.

ACKNOWLEDGEMENTS

When writing this dissertation, I was encouraged by the lyrics of the song, **"***He Brought Me This Far"*

"There's been some mountains that I had to climb, There's been some days when the sun just would not shine, and there's been some valley's that I had to go through, And sometimes I didn't know what to do... God said he's going to take me all the way".

Let me first start by thanking the mighty GOD for guiding me throughout this dissertation; it was not an easy journey.

I would like to express my sincere thankfulness to a number of people of which without their contributions and support this work would not have been accomplished.

Special thanks to my dissertation supervisor Professor John Olaomi for his overwhelming support and invaluable guidance. I deeply appreciate your contributions and valuable comments.

Let me pass my sincere gratitude to my former colleague at SARS, Dr Sambulo Malumisa, for your encouragement and advice throughout my study. I appreciate your contributions and willingness to go extra miles to assist me.

I also wish to extend my gratitude to South African Revenue Services (SARS) for their Tax Statistics publication to be available for public; it helps a lot for the purpose of tax data.

How can I forget my editor, Leanne Brown from UNISA for editing my work, thank you, I am highly indebted to you.

This work would not have been possible without the support and encouragements of my wife, Tshifularo Molapo, thank you my love for being there for me.

I would also like to thank my friends for pushing me to complete this work. Last but not the least; I would like to thank my family for the endless support and my mother Rina Molapo for supporting me throughout writing this dissertation.

ABSTRACT

Tax revenue forecasts are important for tax authority as it contributes to the budget and strategic planning of the country. For this reason various tax types need to be projected for the specific fiscal year using models which are statistically sound and with a smaller margin of error.

The aim of this paper is to forecast South Africa's major tax revenues, i.e. Corporate Income Tax (CIT), Personal Income Tax (PIT) and Value-Added Tax (VAT) using Bayesian vector autoregression $(BVAR)$ models with quarterly data from 1998Q1 to 2012Q1 and compare the results with time series models, i.e. Autoregressive Moving Averages $(ARIMA)$ and state space exponential smoothing models, Error, Trend, Seasonal (ETS). Also the total tax revenue (TTR) is forecasted. The out-of-sample data is for the period 2012Q2 TO 2015Q1 and the forecasting accuracy of $(ARIMA)$ and (ETS) models are compared with $(BVAR)$.

Based on RMSE, the results confirm the accuracy of $(BVAR)$ models for forecasting major tax revenues. On the other hand the (ETS) model appears to be accurate for TTR. In most cases (ETS) is the second best to $(BVAR)$ and was superior to $(ARIMA)$. The results suggest that $(BVAR)$ models may be used to forecasts tax revenues in South Africa together with (ETS) as alternatives models.

Key words: Corporate Income Tax; Personal Income Tax; Total Tax Revenue; Forecasting; Bayesian Vector Autoregressive $(BVAR)$,; Error, trend, seasonal (ETS),; Autoregressive moving Average ($(ARMA)$, ;Root mean squared error $(RMSE)$. ; Autocorrelation.

TABLE OF CONTENTS

LIST OF ABBREVIATION AND ACRONYMS

(A, A)	---	Additive Holt-Winters method
(A, M)	---	Multiplicative Holt-Winters method
(A, N)	---	Holt's linear method
(A_D, M)	---	Holt-Winters damped method
(A_D, N)	---	Additive damped trend method
(M, N)	---	Exponential trend method
(M_D, N)	---	Multiplicative damped trend method
(N, N)	---	Simple exponential smoothing
ACF	---	Autocorrelation function
ADF	---	Augmented Dickey-Fuller test
AIC	---	Akaike information criterion
AIC_C	---	Akaike information criterion corrected
AR	---	Autoregressive
ARCH	---	Autoregressive conditional heteroskedasticity
ARIMA	---	Autoregressive integrated moving average
BIC	---	Bayesian information criterion
BVAR	---	Bayesian vector autoregressive
ECM	---	Error correction model
ETS	---	Error trend seasonality
GARCH	---	Generalized autoregressive conditional heteroskedasticity
MAPE	---	Mean absolute percentage error
MPE	---	Mean percentage error
PACF	---	Partial autocorrelation function
RMSE	---	Root mean square error
U	---	Theil inequality coefficient
VAR	---	Vector autoregressive
CIT	---	Corporate Income Tax
LCIT	---	Log of Corporate Income Tax
DLCIT	---	Differenced Log of Corporate Income Tax
PIT	---	Personal Income Tax

LPIT	---	Log of Personal Income Tax
DLPIT	---	Differenced Log of Personal Income Tax
VATP	---	Value-Added Tax Payments
LVATP	---	Log of Value-Added Tax
DLVATP	---	Differenced Log of Value-Added Tax Payments
TTR	---	Total Tax Revenue
LTTR	---	Log of Total Tax Revenue
DLTTR	---	Differenced Log of Total Tax Revenue
PP	---	Phillip Perron

CHAPTER 1

INTRODUCTION AND BACKGROUND

1.1 INTRODUCTION

One of the responsibilities of the government is to ensure the safety and wellbeing of its citizens. However, in order for it to achieve this role, it has to levy taxes. Tax revenue is government's key source of income and the government needs to estimate its expenditure before it can budget for its income to meet its obligations. One of the key issues in the design of sound fiscal policy has been the accuracy of budget forecasts, particularly tax revenue forecasts (Nandi et al, 2014). Accurate revenue forecasting is crucial to meet expenditure for the purpose of budgeting. Over the years, a key issue in the design of fiscal policy rules has been the accuracy of government budget forecasts, particularly those of tax revenues (Auerbach, 1999).

Like most countries, in South Africa, the Minister of Finance as the head of the National Treasury is the one to set targets for tax revenue. The responsibility of revenue collection is mandated to the South African Revenue Service (SARS), which falls under the Ministry of Finance. SARS presents its revenue projections to a joint committee comprised of the National Treasury, South African Reserve Bank and SARS itself, known as the Revenue Analysis Working Committee (RAWC). It is within this technical committee that national estimates are developed. The national targets are then approved by the Minister of Finance, where after the effect of any proposed tax policy changes will be imputed on the target.

Tax revenue forecasting plays a critical role in government's budgeting process. It has become an important focus area for government to develop their tax system, improve revenue collections, enhance fairness and efficiency of taxes, and also to support economic growth and exports. Governments need funds to finance their budget expenditures (Jenkins, 2000) and tax revenues are the main source of government income. If government spending is more than its revenues, the deficit will results and the shortfall will be financed by borrowing or increasing taxes, and both may have a negative impact on the economic growth over time.

The contribution of tax revenue to the fiscus highlights the need for tax authorities to consider using combined methods to improve forecast accuracy. There are various techniques or methodologies used for estimating various tax types' revenue collection. The common methodologies used by tax authorities are based on growth trends, averages and contributions of the previous period, as well as expert knowledge, which plays an important role. In developing revenue forecasts there are various factors that should be taken into account by revenue authorities. Tax revenue forecasts are developed in relation to economic theories, employed techniques and most importantly assumptions undertaking. These assumptions are derived from economic variables such as growth in the national income, rate of inflation, interest rates, employment and including the international environment (Jenkins, 2000).

1.1.1 The importance of revenue forecasting

The performance of tax revenue is ultimately dependent on the performance of the economy, which means that when the economy performs below expectation, the amount of tax available for collection is obviously reduced, (SARS annual report, 2004). Government has to monitor economic performance throughout the fiscus and project revenues.

Revenue forecasts are made by national governments in the course of budget preparation (Golosov & King, 2002). "*Because of the magnitude of the fiscal problems facing many states, forecasting has assumed a more central role in the policy making process; as a result, revenue forecasts are closely examined and accuracy is essential for planning purposes" (Fullerton, 1989).*Developing a fiscally sound budget there is a need to generate reliable forecasts with good precision and use forecasts as a benchmark for how much money is needed by government to be able to provide services to its citizens.

Accurate revenue forecasts are widely regarded as a key element for the design and execution of sound fiscal policies (Danninger, 2005). Most governments use revenue forecasts to set targets, and in turn, targets are used as a performance measure. "*At the micro level, realistic revenue forecasts become effective standards of measurement against which the actual performance of collecting agencies is assessed*" (Gamboa, 2002). Setting a target combines various processes, ranging from the development of revenue forecasting models to experts' knowledge on various tax types, and consultations with other external stakeholders.

Revenue forecasting is also important for planning purposes. It helps to determine available resources and develops budget expenditure. Reliable and accurate expenditure forecasting is very important, due to its huge impact on government's budgeting process and debt management. Over-estimated projections may lead governments to introduce cuts in services, in order to balance the budget.

1.1.2 An overview of the South African tax system

In South Africa, the national treasury under the leadership of Minister of finance deals with the issues of tax and tax legislation, SARS is the revenue authority given the mandate to collect and managed all taxes, duties and levies. Other functions of SARS in addition to collect all revenue that is due, in terms of the SARS act (No.34 of 1997); SARS is given the power to ensure maximum compliance with tax and customs legislation. Also provide a customs services that will maximise revenue collection, protect our borders and facilitate trade (SARS annual report, 2015).

1.1.2.1 The tax system in South Africa

Prior to 1 January 2001, South African taxpayers were taxed on the basis of source taxation. Tax system in South Africa is residence-based, meaning residents are qualified to get certain exclusions. The residents are taxed on their income and capital gains acquired domestically and globally regardless of where their income was earned. The taxpayers who are not the residents of South Africa are taxed on their income gained from a source in South Africa. Taxes from outside the borders of South Africa are credited against tax payable on foreign income. The introduction of income tax in South Africa can be traced back in 1914 with the Income Tax Act No 28. Through the years the income tax act has undergone numerous changes, and the act currently adopted is the Income Tax Act No 58 of 1962. The Act outline provisions for four different types of income tax, namely normal tax, donations tax, secondary tax on companies, and withholding tax.

The income tax system in South Africa is progressive and is based on the principle that the wealthy people should contribute a greater share of tax to the state than the poor, i.e. the more a person earns, the higher the percentage of tax he/she pays. Income tax is levied on an individual's taxable income for the year of assessment, which is determined from the total assessable income less allowable deductions to arrive at taxable income.

Taxpayers have the obligation to submit their tax returns by a specific date each year. SARS publishes this date and encourages people to file their returns through filing season campaigns, in order to meet the deadline. Tax returns must be filed by taxpayers whose taxable income is above the tax threshold (Proposed by Minister of Finance). Individuals claiming relief from tax in terms of a double taxation agreement are also required to submit tax returns for this purpose. Without the submission and assessment of a tax return, a short-term business traveler is not guaranteed such relief.

To avoid double taxation, South Africa has a broad double taxation treaties agreement with its global trading partners. Tax treaties may cover income taxes, inheritance taxes, value-added taxes, or any other taxes. Above and beyond bilateral treaties, there are also multilateral agreements.

1.1.2.2 The tax structure of South Africa

The tax structure of South African is largely dominated by three major tax types, i.e. Personal Income Tax (PIT), followed by Corporate Income Tax (CIT) and then Value-Added Tax (VAT).These three tax types are the major contributors to the total tax collection in South Africa, contributing approximately 80%, and the remaining tax types collectively contribute 20% to the total tax collection. All taxes that are paid or levied on income, capital gains and expenditure in South Africa obtained from SARS website are listed below:

- Air Passenger Tax (APT),
- Capital Gains Tax (CGT),
- Corporate Income Tax (CIT),
- Diamond Export Levy,
- Dividends Tax,
- Donations Tax,
- Estate Duty,

- Excise Duties and Levies,
- Income Tax (IT),
- Mineral and Petroleum Resource Royalty (MPRR),
- Pay As You Earn (PAYE),
- Provisional Tax,
- Retirement Funds Tax,
- Secondary Tax on Companies (STC) replaced by dividend tax,
- Securities Transfer Tax (STT),
- Skills Development Levy (SDL),
- Stamp Duty,
- Transfer Duty,
- Turnover Tax,
- Un-certificated Securities Tax,
- Unemployment Insurance Fund (UIF) and
- Value Added Tax (VAT).

For the purpose of this study, only the major tax types are discussed in detail below, namely income tax (PAYE and CIT) and value-added tax.

1.1.2.3 Brief overview of Value-Added Tax

VAT was introduced to South Africa in 1991, in order to replace General Sales Tax (GST). The VAT system which was implemented in South Africa is a “destination-based, consumption-type” VAT, i.e. the tax is levied on all goods and services sold in the domestic economy. While imposing a tax at a destination principle, imports are taxed in the same way as domestically produced goods, and exports are not subject to tax (Jenkins and Kuo, 1995). The VAT is calculated on the value of credit invoices. In South Africa, the standard VAT rate is 14% of the value of goods sold. In addition to this standard rate, several goods, such as basic foodstuffs, are zero-rated and certain services, such as financial services, transport and education, are exempt from VAT. Zero-rated means that VAT is levied at 0% rate on the selling price of goods and no output tax will be payable, and the VAT vendor can claim all credit (refunds) for VAT paid on inputs. Exempt goods also mean that no VAT is levied on the selling price of the goods and services, however, the VAT vendor is not eligible to claim a refunds on the VAT paid for inputs.

1.1.2.4 Brief overview of Personal Income Tax

Personal income tax (PIT) is the largest source of tax revenue in South Africa and one of government's main sources of income (Tax Statistics publication, 2015). For the past five years, PIT collections have been contributing 34.4% on average to total tax collection. The tax is levied on the taxable income of individuals and trusts, calculated by following a specific framework (gross income less exemptions and allowable deductions). The majority of individuals receive their income as salaries or wages, pension or annuity payments and investment income (interest and dividends) (Tax Statistics publication, 2015). In addition, taxable capital gains are also included as part of taxable income. Business income is taxable as personal income for sole proprietors and partners.

1.1.2.5 Brief overview of Corporate Income Tax

CIT is the third largest tax in South Africa, contributing 19.8% on average for the past five years to total tax revenue. CIT, like PIT, is an income tax, and follows the same principle, but is levied on the taxable income of companies and close corporations. The CGT is a tax levied on capital gains at a specific rate and is incorporated in CIT taxable income. "*All companies are part of the provisional tax system, which requires taxpayers to provide for their final tax liability by paying two amounts, amounting to at least 80% of the final tax liability during the applicable year of assessment (or the lesser of 90% of actual taxable income and the basic amount, if taxable income does not exceed R1 million), and a third voluntary "top-up" payment after the end of the tax year*"(Tax Statistics publication, 2015). The consequence of non-adherence to the payment system is incurring penalties and interest. The law stipulates that the first provisional tax payment must be paid within six months of the beginning of the year of assessment. The second payment must be paid not later than the last business day of the year of assessment.

1.1.2.6 Tax collection link to the national budget process

The national budget is a document that outlines the government income and spending plans. The income is mainly derived from taxes, and government expenses are on goods and services. The budget also has a particular significance, as it outlines the tax proposal for the specific tax period and sets the target for revenue collection for the current and following financial years. For example, the main tax proposals for the fiscal year 2013/14 as stated in the Budget Speech (2013) include the following:

- Personal income tax relief of R7 billion
- Reforms to the tax treatment of contributions to retirement savings
- An employment tax incentive targeted to support young workers and those employed in special economic zones
- Tax relief for small businesses
- Requiring foreign businesses selling digital goods in South Africa to register as VAT vendors
- Increases in fuel and excise taxes
- Alignment of the proposed carbon tax, energy-efficiency savings tax incentive and the electricity levy.

Tax proposals are different for each fiscal year, as it depends on what government wants to achieve or target for a specific period.

1.2 PROBLEM STATEMENT

Tax authorities are faced with the challenges of producing accurate tax revenue forecasts. Tax revenue is an important part of government's budget process and helps Ministries of Finance to set targets for the fiscus. To achieve accurate forecasts, governments should employ various statistical techniques for forecasting purposes, and compare model performances based on their statistical power. A statistical model with a smaller statistical margin of error and sound statistical tests is required to forecast revenue collection.

1.3 RESEARCH OBJECTIVE AND HYPOTHESIS

The objective of the study is to compare the performance of the Bayesian Vector Autoregression (BVAR) approach with the Autoregressive Integrated Moving Average (ARIMA) and Error, Trend, Seasonal (ETS) exponential smoothing methods, in order to recommend the best model to be used to project tax revenue in South Africa. The accuracy of the statistical model will encourage revenue authorities to explore various statistical methodologies as an alternative approach to existing methods. Furthermore, the aim of the study is to fit a suitable model to the quarterly selected tax revenue data, and compute forecasts for the next 12 quarters. The hypothesis is that the BVAR model performs better than ETS and ARIMA in projecting tax revenue.

1.4 METHODOLOGY

The study will examine the performance of the three forecasting models – Bayesian Vector Autoregression, Autoregressive Moving Average (this may include non-seasonal or seasonal, depending on tax type behaviour) and Error, Trend, Seasonal exponential smoothing methods. The latter is an improved version of the exponential smoothing methods developed by Brown (1959). Hyndman et al (2002) develop state space models for smoothing methods of which prediction intervals, maximum likelihood estimation and Akanke's Information Criterion may be calculated. The models generate automatic forecasts with less human interaction, and are known as state space models.

The ARIMA model was popularised by Box and Jenkins, and is a commonly used technique in forecasting. Recently, the technique was among those examined in a study to identify an appropriate model for forecasting tax revenue in Bangladesh. The ARIMA model has two processes, namely autoregressive (AR) and moving average (MA), and the differenced component to make the data stationary. It is difficult to determine the order of the processes, and autocorrelations and partial autocorrelations are important tools to determine the order of AR and MA processes.

BVAR is widely used to forecast economic variables, but there is very little research on its usage in forecasting tax revenue. A recent study that use this approach was the study of Krol (2010), in which the author compared the performance of BVAR and VAR, and found that BVAR performs better based on RMSE.

1.5 DATA COLLECTION AND SOFTWARE

The data used in the study consists of monthly total collections of CIT, VAT, PIT and the total tax collection of all tax types in South Africa. PIT is the largest contributor to tax revenue. For the purpose of this study PIT includes assessed tax, provisional tax and PAYE collected by employers on behalf of employees less refunds. PIT has a direct relationship with employment and recently benefited from above-inflation wage settlements, bonuses paid out, retrenchment packages, and once-off PAYE collections from the vesting of shares. Therefore, the PIT data series is not adjusted for these factors, which may result in large model errors.

CIT revenue comprises of assessed and provisional payments paid by corporates minus the refunds and is levied on profit made by companies. CIT is affected by economic performance - poor economic conditions have an impact on it.

The VAT collection data series used in this study is only domestic excluding import VAT. Generally, VAT is affected by various factors, such as consumer spending, caused by high consumer debt, modest employment, and low growth in disposable income.

The tax revenue data is sourced from an annual SARS tax statistics publication and annual reports. This is a joint publication of SARS and the National Treasury. The monthly tax collection data was converted to quarterly data, in order to match its counterpart economic data.

The quarterly economic data is used in the study specifically for the BVAR technique, which requires dependent variables (revenue) and independent or explanatory variables (economic). The economic data is obtained from two websites, namely Statistics South Africa (STASSA) and the South African Reserve Bank's (SARB) online statistics tool.

The software used to generate the results was Econometric Views (Eviews), which is the software currently used by SARS for model development and to generate revenue forecasts. The other software programmes that will complement Eviews, if need be, are SPSS, R and SAS.

1.6 MOTIVATION FOR THE STUDY

There is a need for tax authorities to use various methodologies to set revenue targets accurately. One of the ways of determining the best approach is to compare the precision of different techniques.

According to the SARS annual report for 2014/15 published in 2015, revenue estimates for the next three years, the medium term, are set or adjusted on three occasions during the financial year: for 2014/15, estimates were announced in the February 2014 Budget (generally referred to as the Printed Estimate), in October 2014 in the Medium Term Budget Policy Statement (MTBPS), and in the February 2015 Budget (the Revised Estimate).

This study may help in reducing the frequency of revising estimates due to large errors, as it has been observed that some of the errors of the printed estimates are more than 5%, and for the revised estimates, the errors are less than 5%, even close to zero.

1.7 SCOPE AND LIMITATIONS OF THE STUDY

The aim of the study is to compare the performance of three models, namely BVAR, ARIMA and ETS, which are used to forecast the main tax types in South Africa, i.e. PIT, CIT, VAT and TTR. The last mentioned is also included, as the other three tax types contributed approximately 80% to TTR. The best technique is recommended to complement the existing methodologies used in the South African Revenue Service, but this will in no way replace the current approaches. Tax analysis is complex and is impacted by various factors, such as economic growth, interest rate, consumption, taxpayer compliance and behaviour. These factors make it impractical for tax forecasting to depend on only one approach.

1.8 OUTLINE OF THE STUDY

The study is divided into five chapters. The first chapter provides an introduction, which highlights the background to the study in terms of the importance of revenue forecasting. It also provides an overview of the South African tax system and its structure, and describes the tax types which will be modelled in the study. Chapter one also contains the problem statement, objectives of the study, methodology, motivation for the study, scope and limitations, and organisation of the study.

The Chapter two deals with relevant literatures in terms of the methodology chosen in the study. It outlines the process of tax forecasting in South Africa 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 from 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 for the three selected approaches to forecasting the tax types in this study. The fifth chapter of the study draws conclusions and future studies based on the analysis.

CHAPTER 2

LITERATURE REVIEW

2.1 INTRODUCTION

There is a variety of literature on forecasting in general, especially in the field of economics, unlike the field of taxation, which is poorly covered in this regard. There is a small body of economics literature that examines state tax revenue forecasting (Krol, 2010), and several studies have been conducted in various parts of the world on tax revenue forecasting. The sections that follow will provide an overview of the methodologies and findings of some of these studies.

2.2 THE TAX FORECASTING PROCESS IN SOUTH AFRICA

In South Africa, the South African Revenue Services (SARS) is given the responsibility of collecting tax revenue, and it falls under the National Treasury, which is the institution responsible for fiscal forecasting. SARS, like any other tax authority in the world, uses trend analysis and macro-economic indicators, taking the levels and growth rates of Gross Domestic Product (GDP) into account, in order 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 (SARS intranet). Statistical modelling is also used to some extent to support revenue estimates.

SARS presents its revenue projections to a joint committee comprised of the National Treasury, South African Reserve Bank and SARS itself, known as the Revenue Analysis Working Committee (RAWC). It is within this technical committee that national estimates are developed. The consensus national targets are then approved by the Minister of Finance, after which the effect of any proposed tax policy changes will be imputed on the target (SARS, 2012).

The estimates are announced by the Minister of Finance during the budget speech in February of each year. The estimates in a particular fiscal year are revised twice that year: in October, during the Medium Term Budget Policy Statement (MTBPS), and in February, when the final Revised Budget is presented. The preparation for October's MTBPS starts in September and gives the authorities the opportunity to amend the original budget estimates if the forecasted economic conditions have changed significantly.

Like any other tax authority in the world, SARS follows both bottom-up and top-down approaches to generate the forecasts. There are numerous methods employed by SARS for the purpose of developing fiscal revenue forecasts to inform the targets.

Revenue Forecast methodologies of SARS are not well documented. SARS revenue forecast methodologies mentioned and described below were taken from SARS internal documents of Revenue Planning, Analysis & Reporting division's Quarterly Bulletin (2012) and some research papers produced internally within SARS. SARS use multiple methods to arrive at a point estimates 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 SARS, these are: Micro-Simulation Model; Macro-econometric Tax Revenue Models; Tax Buoyancy/Elasticity Approach; Box-Jenkins techniques; Revenue Trending & Extrapolation and Professional Judgement. These methods are explained in the following subsections.

2.2.1 Micro-Simulation

The micro-simulation (MS) models require a highly disaggregated data which mainly consisting of comprehensive data at the company or individual level to determine tax liability instead of actual tax collection. The MS model uses individuals or companies who file income tax returns and non-filers, which are individuals whose taxable incomes are below the threshold. MS models are good in its ability to simulate alternative policy proposals and also determining who would benefit from specific policy change or bear the burden of the tax change. This approach was adopted from Van Heerden and Schoeman (2010) paper and was also referenced by Makananisa (2015). Van Heerden and Schoeman (2010) utilized data from Tax Statistics publication as a proxy to determine ratio for allowances to be applied to each individual income group. An average allowance ratio (τ_{allow}) is developed from taxable income (t_i) and gross income (y_i) per taxable income in equation (2.1)

$$\tau_{allow} = \frac{y_i - t_i}{y_i} \quad (2.1)$$

The ratio per income group from equation 2.1 is then applied to each individual or company gross income group in the equation 2.2.

$$allow_i = y_i * \tau_{allow} \quad (2.2)$$

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

$$t_i = y_i - allow_i \quad (2.3)$$

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

$$PIT_i = f(t_i, \tau_{structure})$$

or in the case of companies the equation 2.4 may be used:

$$CIT_i = f(t_i, \tau_{structure}) \quad (2.4)$$

Van Heerden & Schoeman (2010) stated that this procedure is a static method and behavioural changes are not discounted for.

2.2.2 Macro-Econometric tax revenue models

SARS also uses macro-econometric methods. These models may be classified under econometric models which are based on relating the dependent variable to a number of independent variables (sometimes called explanatory variables) with residuals considerations. Macro-based models specify the proxies for various taxes in order to determine the potential revenue collection for each tax types and are based on the past performance of tax collections and economic growth. Literature specifies various proxies for different tax types; CIT may be modelled with company profit where proxy is gross operating surplus. The PIT may be modelled with compensation to employees and employment as explanatory variables, while value-added tax may take Investment and gross domestic expenditure (GDE). The total tax revenue (TTR) may be modelled with gross domestic product (GDP) as an independent variable. These models may be single-equation regression models or multi-equation models and may be represented by the general equation,

$$\log(T_i) = \beta_0 + \beta_i \log(\phi_i.\gamma_i) + ... + \beta_n \log(\phi_n.\gamma_n) + \varepsilon_i$$

Where T_i represents tax collections at time i ; β_0 is the intercept; ϕ_i is the tax base at time i ; γ denotes statutory tax rate at time i and ε_i is the error term at time i .

Other macro-econometric models used are vector autoregression models which uses lagged value of the dependent variable as an explanatory variables and the vector term indicates that two or more variables are involved.

2.2.3 Tax buoyancy/elasticity approach

The response of tax revenues to changes in the GDP is measured by tax elasticity and tax buoyancy and these concepts help to explain the overall structure of a tax system and serve as valuable analytical tools for designing tax policy (Jenkins et al. 2000). Tax buoyancy is a measure of the total response of tax revenues to changes in national income and takes into considerations increases in income and discretionary changes introduced by tax body in the system. Discretionary changes may be changes in tax rate, tax bases or tax policy. Tax buoyancy can be represented as follows:

$$B = \frac{\Delta T_i}{\Delta Y_k} * \frac{Y_k}{T_i}$$

Where B represents Buoyancy of tax revenue to income; T_i is the tax revenue; ΔT_i denotes Change in tax revenue; Y_k is the Income and ΔY_k is the Change in income. i and k represent specific tax type and proxy respectively. If the calculation excludes the impact of changes in tax rates and tax bases and takes into account only the effects due to changes in income levels, whether or not changes were made in the tax structure during that time period then the tax elasticity occurs. Tax elasticity is an important factor for forecasting purposes and its coefficient gives an indication to policy-makers of whether tax revenues will rise at the same pace as the income. Tax elasticity is given by equation,

$$E_i = \frac{\%\Delta T_i}{\%\Delta Y_k}$$

where, $\boldsymbol{E_i}$ is the elasticity of tax revenue to income or GDP; $\boldsymbol{\Delta T_i}$ change in tax revenue; and $\boldsymbol{\Delta Y_k}$ change in income or GDP. $\boldsymbol{i}$ and $\boldsymbol{k}$ 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 for financing its outlays (Jenkins et al. 2000).

2.2.4 Box-Jenkins techniques

Box-Jenkins methodology are 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 chapter 3 section 3.2.3.

2.2.5 Revenue trending and extrapolation

These methods also encompass 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 base on the assumption that the growth rate does not vary during the fiscal year 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

$$\hat{Y}_t = \left(\frac{YTD_t}{YTD_{t-1}}\right) * Y_{t-1}$$

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

It is clearly this method have some disadvantages as it uses YTD growth rate at any point to project full current. Makananisa (2015) pointed the limitation of the model that it assumes the constant growth rate throughout the fiscal year which is not the case as economy is driven by various factors which will have a negative or positive impact on revenue collections, therefore revenue collections growth rate will fluctuate.

2.2.6 Professional judgement (PJ)

Professional Judgement is the other techniques used in SARS. This method entails the expert's assessment of revenue collections from various internal units or departments and government department other than SARS. The forecast from professional judgement considers the estimated outcomes from all the aforementioned forecasting techniques, but also takes into consideration information that relates to cash flow, administrative changes and other special factors that cannot be automatically incorporated into the other modelling approaches (Boonzaaier 2012). Boonzaaier stated that historically, PJ method has often proved 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 modelling framework.

Makananisa (2015) highlighted the disadvantage of PJ methods, that since professional judgement comprises of forecasts of different models and is based on different scenarios, there could be some drawbacks when the expected scenario does not hold or outcome is not as anticipated.

2.3 THE TAX FORECASTING PROCESS IN OTHER COUNTRIES

2.3.1 Revenue Forecasting Process in Australia

In preparing its fiscal forecasts, the Fiscal Advisory Council (FAC) in Australia relies on international best practices requiring forecasters to exercise prudence, adopt transparent and realistic assumptions, and to use recognised forecasting methods (The Australian Fiscal Advisory Council, 2014). The Australian fiscal forecasts are released two times a year by FAC, in the spring and in the fall, with its first forecast traced back to fall 2014. Each forecast covers the current and following year, and the forecasts are developed using bottom-up approach whereby government revenue and expenditure flows are gathered and disaggregated and projected item by item. For instance, the withholding of taxes on wages and pensions is broken down into employee income tax and pension income tax based on wage tax statistics. This breakdown is relevant for the quality of the forecasts, as the growth rates of wage income and pension income may differ.

The best techniques are selected based on three criteria, i.e. un-biasedness, precision and simplicity. Un-biasedness means that, in the past, the forecasting error should have been around zero on average. In other words, the forecasts have tended to be neither too optimistic nor too pessimistic, because underestimated and overestimated results have offset each other on average. Precision is evident when the mean squared errors have been small, which shows that the development of flows has been captured well over time. The simplicity of the methods implies that in the case of doubt, the simpler of two methods should be chosen.

Australia applies the rule that forecasting methods are based on the assumption that all revenue and expenditure categories K_t change over time following the equation

$$K_t = (1 + g_t^K)K_{t-1} + \phi_t \tag{2.5}$$

Where g_t^K is the rate at which the flows captured in category K are forecasted to grow from $t-1$ to t and ϕ_t denotes discretionary changes considered in year t. The revenue forecasts are generated based on the four methodologies explained below (special cases of equation 2.5):

- ***Adjusted Trend Projections*:** This is the method of projecting the trend with adjusted discretionary measures. The adjusted growth rates for the years leading the forecast period $h < t$ are calculated by using the equation

$$g_h^X = \frac{X_t - \phi_h}{X_{h-1}} - 1 \tag{2.6}$$

  The result of the equation 2.6 is influenced by two assumptions: firstly, the functional relationship defined in the equation assumes that the discretionary part will continue to grow at the adjusted growth rate. Secondly, the adjustment is by definition based on an estimate of the actual budgetary impact of the discretionary measure, which is typically difficult to establish, as the counterfactual outcome (i.e. what would K_t have been in the absence of the discretionary measures) is unobservable. The choice of the period for calculating the adjusted trend is dependent on structural breaks in the respective revenue and expenditure categories.

- ***Elasticity-Based Projections*****:** This technique uses elasticity to project the changes in revenue and expenditure categories that are due to changes in underlying macroeconomic, fiscal, structural and socio-demographic measures. The projection rate for the revenue or expenditure category $K(g_t^K)$ is calculated according to the growth rate of the underlying variable X against the last year (g_t^X) and the elasticity of K with regard to $X(\varepsilon^{K,X})$. The equation is given by:

$$g_t^k = g_t^X . \varepsilon^{K,Y}$$

  The choice of the underlying variables is a qualitative decision based on the convergence of economic factors. These variables will be highly correlated with the budget category and are limited to the variables for which forecasts exist or may be derived through forecasting methods. Revenue forecasts are only based on macroeconomic variables, whereas expenditure forecasts are also derived from other variables, especially socio-economic variables.

- ***Carry-Forward Projections*****:** This approach carries forward the previous year's figures based on a symmetric random walk assumption, i.e. the best guess estimate of future fiscal flows of a given category is the past year's level. This method is used for small and erratic budget categories for which the sign of the previous trend is not robust for the choice of the forecast period.

- ***Ad Hoc Projections Reflecting Expert Judgment:*** The aim of this method is to choose an ad hoc measure of K_t that is informed by research and expert judgement. The method is used for income and expenditure categories dominated by discretionary measures which are not directly linked to macroeconomic developments and do not follow a stable trend.

2.3.2 Revenue Forecasting Process in Ireland

In Ireland, tax revenues are forecast by the Department of Finance 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 done three times a year, the first of which 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 publicised. The second round of the forecasts is performed in September/October for the Pre Budget Outlook (PBO), and the third round of forecasts is done in November/December for the Budget. The tax forecasting methodology in Ireland is generally given by the equation

$$\theta_{t+1} = (\theta_t - T_t)(1 + (B_{t+1}^G E)) + T_{t+1} + M_{t+1} + J_{t+1}$$

Where θ_{t+1} is the one year forecast ahead for a specific tax type, θ_t is the current year estimate for that tax type, T_t are once-off items that affect the outcome in the current year, B_{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, T_{t+1} are once-off items affecting the outcome in the next year, M_{t+1} is the estimated static outcome from any policy changes that impact receipts for a specific tax in the year ahead, and J_{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 (2015) stated that some previous work that looked at Irish revenue forecasts in an international context found that the Irish official forecasting performance was on the weaker end of the spectrum.

2.3.3 Revenue Forecasting Process in New Zealand

Like Australia, New Zealand Treasury produces economic and fiscal forecasts two times in a year. The first forecast is prepared for annual budget and are named the Budget Economic and Fiscal Update (BEFU) during May or June. The second forecast is published in December, a week or two weeks before Christmas and called the Half-Year Economic and Fiscal Update (HYEFU) (Keene & Thompson, 2007). The forecasts are developed by the macroeconomic forecasting team; they take into account development in economic data, the projections done by other forecasters and also deliberate the state of the economy with business people in

New Zealand. The forecasts period covers present year and four years' period in the future. The tax forecasts are prepared simultaneously with the economic forecasts in June of every year.

The models used in New Zealand are based on spreadsheet and follows the procedure outlined below in phases as stated in (Keene & Thompson, 2007):

Phase1: Determine the nominal tax revenue for the last available year which is the base year.

Phase 2: Adjust the nominal tax revenue for the base year by removing any known anomalies to establish the true underlying tax position for that year.

Phase 3: Apply the forecast growth rates of relevant macroeconomic variable(s) to forecast tax for 1 to 5 years ahead, applying elasticities if required.

Phase 4: Adjust the tax forecast for anomalies such as tax policy changes, expected shifts in payment dates or taxpayer behaviour, and include any judgemental forecasting adjustments that may be deemed appropriate.

A more detailed description of some of the major tax types may be found in "An analysis of tax revenue forecast errors in New Zealand (Keene & Thompson, 2007).

2.4 STUDIES SPECIFIC TO SOUTH AFRICA TAX REVENUE FORECAST

Boonzaaier (2012) working paper was the first to attempt formalizing revenue forecasting practices that are currently employed by SARS. But his paper was an open-ended discussion document which over time is expected to evolve as improved techniques becomes available or further research are conducted. Boonzaaier (2012) was mainly to inform members of the Revenue Analysis, Planning and Reporting division regarding the revenue forecasting process. The researcher outlines the revenue forecasting process and practices around the world. He also mentioned and explains the techniques used for revenue forecasting; among the methods stated are Single-Equation Regression, Vector Autoregression, Micro-simulation, Box-Jenkins Methodology (ARIMA), Constant Trend Growth Methodology and Professional Judgment.

Makananisa (2015) apply time series methodologies (exponential smoothing and ARIMA) to forecast major tax types (Personal Income Tax, Corporate Income Tax and Value-Added Tax) and total tax revenue in South Africa for three years ahead. The researcher used monthly data from January 1995 to March 2010. The results of Makananisa study suggested that SARIMA and Holt-Winters models perform well in modelling and forecasting PIT and VAT. Holt-Winters model was found to perform better than the SARIMA model for forecasting CIT and TTR. The study concluded that the chosen models are expected to perform better when projecting the future values in stable economic conditions, with the assumption that there will be no shocks in the economy. This study further 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 done.

2.5 STUDIES COMPARING THE PERFORMANCE OF BVAR AND OTHER METHODOLOGIES

The *BVAR* approach has been largely used to predict or forecast economic variables, and its usage in the field of taxation has not been widely explored. There are several studies that employed this technique, such as that conducted by Litterman (1986), who used the *BVAR* to forecast economic variables (Real GDP, unemployment and Inflation).

In using *BVAR* techniques Litterman (1986) was avoiding problem of over-parameterization and suggested putting weaker restrictions on the coefficients rather than placing zero. Litterman assumption was 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 (2007) was used to estimate coefficients. The standard prior has three distinct characteristics, i.e., the prior probabilities on deterministic variables such as seasonal dummy variables are flat or non-informative. The other characteristic is that the prior distribution is 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. Follows these characteristics Krol (2010) specify the standard deviation of the prior probabilities as was specified by Doan (2007) indicated by formula

$$S(i,j,1)=\{[\gamma g(1)f(i,j)]s_i\}/s_j, \qquad f(i,i)=g(1)=1.0$$

where s_i represents the standard error of a regression of variable i on lags of itself. The $[\gamma g(1)f(i,j)]$ represents the tightness of the prior distribution on coefficient i in equation j for lag l. γ is the overall tightness of the standard deviation of the prior distribution. The smaller value γ takes results in a smaller $S(i,j,1)$ and a tighter standard deviation for the prior.

In order to project the car market share, Ramos (1996) constructed a *BVAR* for the leader car market in Portugal. The author showed the usability of *VAR* and *BVAR* methodologies as a tool for marketing that satisfies two requirements, i.e. market share predicting and providing information about the competitive changing conditions of the marketplace. Ramos incorporated five marketing variables in the model. The selection of prior 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.

Ramos (1996) further showed that *BVAR* provides important information for the people who are responsible for marketing, by utilising impulse response functions and decompositions of variance. 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 forecasts accuracy is depended on the specifications of the prior. If the prior is not specified correctly, an alternative model like 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 environment with sufficient data available.

There is sparse literature on the adoption of the $BVAR$ approach as a forecasting tool for tax revenue. One study which employed the $BVAR$ technique to forecast state tax revenue was that conducted by Krol (2010), who applied the models to Californian tax revenue. Krol stated that Bayesian vector auto-regressions generally outperform standard vector auto-regressions and simple univariate models in forecasting macroeconomic variables. Krol's study sought to determine whether or not $BVAR$ would also outperform other models when forecasting state revenue. In most cases, Krol's results show that the $BVAR$ models have the smallest root mean squared error compared to the other models examined, and recommended that tax revenue forecasters should consider using Bayesian vector auto-regressions when producing revenue forecasts.

Although there is sparse literature on the application of the $BVAR$ approach to tax forecasting, there are numerous studies on its usage for forecasting economic variables. In the study conducted by Caraiani (2010), the $BVAR$ framework was used to forecast the dynamics of output for the Romanian economy. The several versions of $BVARs$ 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 forecasting 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 for forecasting the quarterly GDP in the short run. The results indicated that the 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 other more complex models may be used that incorporate an extension to the open economy or the development of models to analyse monetary and fiscal policy.

In developing the priors Caraiani (2010) follows Litterman's stylised facts of time series from macroeconomics that, most of the 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 stylised facts a prior distribution was derived which is a random walk.

In another study, Yao (2011) employed Bayesian *VAR* methods, as proposed by Litterman (1986), to estimate and forecast several North Dakota macroeconomic variables, including employment, income and tax receipts. The out-of-sample performance of the *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 hold-out sample. In his study, the superiority of the *BVAR* to the *VAR* was also confirmed, and the results indicated that properly incorporating prior information into the BVAR provides accurate and responsive forecasts.

Yao (2011) adopted Litterman (1986)'s prior, by assuming a reasonable approximation of the behaviour of an economic variable is a random walk around an unknown. Yao's prior reflects the belief that, first, the coefficients are having prior mean of zeroes 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 the parameters for other variables in the equation is believed to centre more tightly around zero.

The *BVAR* model was again used in Romania to provide an analysis of the transmission mechanism of the monetary policy in a study conducted by Spulbăr et al (2012). The *BVAR* model was developed for the Romanian economy in order to identify the major shock in Romania economy over the last 10 years and to provide information concerning the evolution of the economic response to these shocks. The authors estimated the *BVAR* using the technique used by Sims and Zha (1998), as well as the KoKo Minnesota/Litterman (2010) which highlighted the core factors that have an impact to the Romanian economy for the last ten years period. The variables that related to the development of industrial production were added, these are exchange rate, inflation, real estate prices, monetary aggregate M2[1], and the interest rate. One of the conclusions of the study was that the exchange rate is a

[1]Includes savings deposits, money market mutual funds and other time deposits, which are less liquid and not as suitable as exchange mediums but can be quickly converted into cash or checking deposits.

vital mechanism that considerably impacts the real economy variables. In addition, the other factors that contributed to the increase of real estate prices, is the monetary aggregate M2, together with the appreciation of the national currency. The positive aspect attributed from this study is associated to the absence of output puzzle and price puzzle, with the channel of the interest rate increasing and becoming more consistent over the past few years.

Another study which used $BVAR$ was that conducted by Carriero and Mumtaz (2012). 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, UK and US to show that $BVARs$ perform better than the autoregressive forecasts.

Carriero and Mumtaz (2012) investigated the possibility that the $VARs$ employed by various past studies are too small in scale, possibly with over- parameterisation problem, and do not have time variation coefficients and volatilities. The author estimated various specifications of $BVARs$ which allow summarisation of the information contained in a large data set effectively, avoiding the over-parameterisation problem, and allow for time variation in coefficients and volatilities. The finding was that, firstly, once over- parameterisation is corrected, the use of extra explanatory variables is important in forecasting fiscal variables, and multivariate models performs better than univariate specifications in forecasting; secondly, the large system implementation and the time variation play a very important role in forecasting.

2.6 STUDIES BASED ON ARIMA MODELS

The study conducted by Nazmi and Leuthold (1985) developed a time series model for predicting state income tax receipts using the Hannan-Quinn criteria. The authors determined the linear and log linear versions of the $ARIMA(1,0,0)$ model and used a Box-Cox transformation to select linear version of time series model. When compared with the forecasts from an econometric model, the forecasts obtained from the linear time series model were better suited to exploring the percentage root mean square error criterion. This was because the econometric model uses more information than the time series approach. The econometric model relies on personal income for both the in-sample and out-of-sample periods, relative to the time series model which employed tax receipts data. The time series model consistently outperforms the econometric model in forecasting state tax receipts based on the percentage root mean square error test. The study concluded that time series analysis a best technique for projecting state tax receipts.

In their study, Meylar et al (1998) outlined the steps which required to be performed in order to use *ARIMA* time series models for projecting Irish inflation. The Box Jenkins techniques and the objective penalty function methods were considered as ways to identify *ARIMA*. *ARIMA* Modelling procedure was followed, which includes data collection and examination; determining the order of integration; model identification; diagnostic checking; and forecast performance evaluation. Issues in *ARIMA* forecasting were demonstrated in relation to the harmonised index of consumer prices (HICP) and its major sub-components. A range of models was retained based on the robustness of the approach, which performs optimally in the model identification and diagnostic checking stages, for use in forecast performance evaluation phase. The authors commented that *ARIMA* models are not performing good with more volatile data series, and are using historical data and not good at forecasting turning points. In the long run better-specified multivariate models usually are superior to *ARIMA* models. To improve forecasting performance, *ARIMA* model was fit to a 'noiseless' version of the HICP series. The finding was that a preliminary analysis of a developed 'noiseless' series indicates that the optimal *ARIMA* model does indeed superior than the *ARIMA* model fitted to the noisy series, but there is little variation.

Contreras et al (2000) used the *ARIMA* method to forecast the next-day electricity prices of Spain and California electricity markets respectively.

ARIMA techniques were employed to analyse time series data and previously have been mainly used for load forecasting, due to their accuracy and mathematical soundness. The hours needed to project future prices by Spanish model were five hours while the Californian model requires two hours. These differences may be associated with a different bidding structures and ownership. "Average errors in the Spanish market are around 10%, with and without explanatory variables, and around

5% in the stable period of the Californian market (around 11% considering the three weeks, and without explanatory variables)". In Spain, dependent variables are only required in months with a strong correlation between available hydro production and price. In any other month, the effect is cancelled out. The errors are reasonable for both markets, considering the complex nature of price series data and the previously reported results in the literature, especially from Artificial Neural Networks.

Legeida and Sologoub (2003) tested different methodologies for forecasting VAT revenues in Ukraine. They employed the effective rate approach and econometric method. The finding was that there was a stable empirical long-run relationship between VAT revenues and the VAT base. The econometric model was not developed due to econometric problems which arose as a result of incorporating economic variables into the model, such as multicollinearity, endogeneity, etc. The authors acknowledged that the effective rate approach is the most burdensome of all methods to estimate VAT revenues, and requires a huge amount of statistical information. They also developed a suitable $ARIMA$ model for predicting VAT revenue in the short-run, and the forecast was consistent with government projections for the budget. The study recommended that all methodologies should be applied simultaneously. All estimates should be compared and combined to come up with a reasonable forecast number, in order to account for the merits and shortcomings of each of the methods. The econometric methodologies do not account for discretionary government policies, which could influence the revenue forecast.

Lu (2009) study attempted to develop a time series to generate forecasts for gross domestic product (GDP) in China up to the first quarter of 2009, using data from 1962 to 2008. The study compared $ARIMA$ with other models, and the $ARIMA(4,1,0)$ model was selected as the best model. To test for the presence of a break point the Chow test was used. The test revealed that there was evidence of a data break point between the fourth quarter of 1977 and the first quarter of 1978. The GDP was modelled using $ARIMA$ models based on the Box-Jenkins method. The selected $ARIMA$ model was used to generate an out-of-sample forecasting for the 1st quarter of 2009 GDP value. The study concluded that $ARIMA(4,1,0)$ is a suitable model to forecast GDP, and may be applicable for forecasting purposes. The forecast outcome for the fourth quarter of 2009 was acceptable.

Koirala (2011) also used the $ARIMA$ technique as one of the tools to forecast government revenues. The level data of monthly revenue data series for the period 1997 to 2012 was used to project government revenues. The five methods were used, the Winter and Seasonal $ARIMA$ methods were found to be superior in forecasting the monthly revenue series of the Nepal government. However, the $SARIMA$ method was found to be performing better than Winter method based on MPE and MAPE criterion. The study found that the results of forecasted revenues may vary depending on the more sophisticated methods of forecasting employed which capture cyclical components of the revenue data series.

The study conducted by Mehmood and Ahmad (2012) aimed to forecast Pakistan's exports to SAARC for the years ahead using an $ARIMA$ model. The authors found $ARIMA(1,1,4)$ to be the precise model amongst other $ARIMA$ models for predicting Pakistan's exports to SAARC. The study concluded that exports from Pakistan to the SAARC region would increase over the forecasted period.

Dadzie (2013) used $ARIMA$ models to forecast the domestic and import VAT of Ghana using data from 1999 to 2009. The author followed the Box- Jenkins technique and found that $ARIMA(2,1,2)$ was the best fit to forecast domestic VAT revenue in Ghana, while $ARIMA(2,1,1)$ was the appropriate model to forecast import VAT revenue.

Zakai (2014) modelled Pakistan's GDP using a set of $ARIMA$ based on the Box-Jenkins technique. The best-suited $ARIMA$ model, amongst others was $ARIMA(1,1,0)$, and forecast values for the next few years were generated by applying the selected model which provided the best fit for the data. Sample forecasting was done for the period 1953 to 2009, and the visual presentation of forecast values revealed good behaviour.

In a study which 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 (2015) considered five explanatory variables, namely real gross domestic product growth rate, employment population, unemployment rate, annual domestic crude oil prices, and inflation rate. The study examined annual data from 1980 to 2013 for each variable, which was collected from the National Statistical Coordination Board, Department of Labour and Employment, inflationdata.com and World Bank. In forecasting income tax revenue, $ARIMA$ 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 forecasting performance of the model, and it showed that there was no big difference between predicted and actual values, signifying that the models is the best in predicting the income tax revenue of the Philippines.

In addition, the above study used multiple linear regression, and the authors identified the factors affecting the income tax revenue of the Philippines. Based on the results that were obtained, there are three significant factors that can actually predict income tax revenue, namely employment population, annual domestic crude oil prices, and inflation rate. Correspondingly, the Granger Causality Test was used to verify the causal relationship between factors affecting income tax revenue, and it was found that a uni-directional Granger causal relationship existed between income tax revenue and domestic crude oil prices and real gross domestic product.

2.7 STUDIES COMPARING ETS AND ARIMA MODELS

"Exponential smoothing (ETS) has received significant attention in recent years due to the invention of its state space formulation" (Yang et al, 2015). Guizzi et al (2015) analysed and forecasted temperature, pressure and humidity using four years of time series data. They compared three methods, namely the $ARIMA$ model, the Holt-Winters additive seasonal model, and the (ETS) model described in Hyndman et al. (2008). The study found that the $ARIMA$ model is the best temperature forecast method.

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

Huselius and Walled (2014) also compared the performance of univariate time series methods to predict the Swedish inflation rate. Exponential smoothing and $ARIMA$ models, both regular and from an underlying state space model, were fitted, and the forecasts were compared with those of the National Institute of Economic Research (NIER). The results showed that a state space $MA(9)$ performed best in relation to NIER, and had lower specification errors. In cases of a varying pattern, an original $ARIMA(1,0,11)$ model with and without seasonality of 12 often performed well, but at too high a level. The finding was also that in times of stagnation, the ETS models performed well, by capturing the accurate level. The conclusion of this study was that different univariate models can perform well in different economic cycles, but multivariate state space models would probably be better for longer periods.

In modelling and forecasting fish catches, Bako (2014) developed the state space approach (ETS). The author used two methods of time series analysis to predict the fish catch of three commercial fish species found in Malaysian waters. The Box-Jenkins method, together with the ETS state space exponential method, was used. 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 best suggested models for various species of fish were $SARIMA(1,1,1,)(0,0,1)_{12}$, $SARIMA(1,1,4,)(0,0,1)_{12}$, $SARIMA(2,1,1,)(0,0,1)_{12}$ and $ETS(M,A,M)$, $ETS(M,N,M)$, and $ETS(M,N,M)$. It was found that the ETS models performed 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 appropriate in projecting monthly fishery dynamics.

2.8 CONCLUSION

It is clear from the literature review that most studies conducted relating to time series forecasting are limited for tax revenue forecasting. Tax authorities around the world mostly rely on elasticity approach and judgement forecasting not much of time series techniques is used.

And also evident from the literature review there is a limited use of $BVAR$ techniques in tax forecasting, only Krol (2010) employed the technique. The importance of $BVAR$ technique is based on the choice of prior probabilities. Most literature follows Litterman (1986) priors and Sims and Zha 1984). In this study the three priors considered are Litterman/Minnesota prior, which assumes random walk process, Normal-Wishart prior which is a conjugate prior normal data and Sims-Zha prior which show how the dummy variables are used to produce the priors for structural VAR models. The priors are discussed in chapter 3.

Also from literature review there is no evidence of forecasting tax revenue using ETS techniques, this paper will be the first to explore the use of ETS in the taxation environment.

CHAPTER 3

METHODOLOGY

3.1 INTRODUCTION

Generally, forecasting methodologies can be classified 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 dependent variable to a number of independent variables (explanatory variables) with residuals considerations.

The tools which most developed countries use to forecast various tax revenues consist of macro-based models (Chun-Yan Kuo, 2000). These models specify the proxies for tax types, in order 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 taken into account by adjusting for them to consider only the revenue collection associated with economic performance. In this study, the discretionary effects (revenue initiatives and legislative changes) are not adjusted due to a lack of distinction between revenue collection related purely to economic performance and collection linked to budget policies.

In our study, both time-series and econometric approaches are used to compare their performance and the accuracy of forecasts using data from the first quarter of 1998 to the first quarter of 2015. The last 12 observations out of a total of 69 are reserved for checking the forecasting methods' accuracy. Explanations of the selected techniques used for forecasting in this study are explained in the sections that follow.

3.2 SELECTED METHODOLOGIES

Enders (2003) states that "the task facing the modern time-series econometrician is to develop reasonably simple models capable of forecasting, interpreting, and testing hypotheses concerning economic data". This statement highlights the importance of selecting the best model for generating forecasts.

Three methodologies are employed in respect of ensuring precision and accuracy in handling data for PIT, CIT, VAT and TTR. Various measures of accuracy will be used to select the best models, and will be discussed in this chapter.

When modelling time-series data, the main objective is to develop models that are as close as possible to the true, but unknown, data generating process (Skarbovik, 2013).

A time-series may be represented by the following additive and multiplicative equations with four components:

$$y_t = T_t + S_t + C_t + I_t \qquad (additive)$$

$$y_t = T_t * S_t * C_t * I_t \qquad (multiplicative)$$

where, y_t is the time-series (here in this dissertation it can be any tax or economic variables).

T_t is the trend component, which represents the long-term behaviour, increase or decrease of the series;

S_t is the seasonal component, which represents variations that recur during a specific period of the year to the same extent, regular, as well as relatively short-term, repetitive, up-and-down fluctuations of the series;

C_t is the cyclical component, which represents the regular periodic movements, and potential up-and-down swings of the series; and the I_t or irregular component, which represents uncontrollable shock caused by unexpected events.

3.2.1 Bayesian vector autoregression

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

$$P(\phi / X) = \frac{P(X, \phi)}{P(X)} \qquad (3.1)$$

We may express equation 3.1 as the joint *probability by conditioning on* ϕ*,* which gives us*:*

$$P(X, \phi) = P(X / \phi) P(\phi) \qquad (3.2)$$

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

$$P(\phi / X) = \frac{P(X/\phi)/P(\phi)}{P(X)} \qquad (3.3)$$

With *n* possible outcomes $(\phi_1, ..., \phi_n)$, equation 3.3 may be written as follows:

$$P(\phi_j / X) = \frac{P(X / \phi_j) P(\phi_j)}{P(X)} = \frac{P(X / \phi_j) P(\phi_j}{\sum_{i=1}^{n} P(\phi_j) P(X / \phi_i)}$$

$P(\phi)$ is the prior distribution of the possible ϕ values, and $P(\phi / X)$ is the posterior distribution of ϕ, given the observed data X .

As opposed to the point estimators (means, variances) used by classical statistics, Bayesian statistics is concerned with generating the posterior distribution of the unknown parameters, given both the data and some prior density for these parameters. As such, 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.

Vector autoregression (VAR) models are broadly used to model economic time-series. The main difficulty experienced with these models is the issue of handling a large number of parameters, as stated in most literature. To overcome this difficulty, a Bayesian VAR approach was employed by Litterman (1980) to solve the over-fitting problem. He suggested that over-fitting may be avoided without imposing an exact zero restriction on the coefficients. Litterman (1986) recommended using a Bayesian strategy to estimate the VAR, equation by equation, where a prior, the lags have decreasing importance.

Doan et al. (1986) used VAR models to impose less arbitrary restrictions than traditional econometric models. According to Ciccarelli and Rebucci (2003), the VAR model can be represented as:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + KZ_T + \varepsilon_t \qquad t = 1, \ldots, T \qquad (3.4)$$

where y_t denotes a $(n\,\text{x}1)$ vector of endogenous variables with $\boldsymbol{\varepsilon}_t$ represents a $(n\,\text{x}1)$ vector of error terms independently, identically and normally distributed with variance-covariance matrix Σ, i.e. $\varepsilon_t \sim IIN(0, \Sigma)$; ϕ_t $(t = 1, \ldots, p)$ in the form $(\boldsymbol{n}\ \text{x}\ \boldsymbol{n})$ matrix and $\boldsymbol{K}$ as $(\boldsymbol{n}\ \text{x}\,\boldsymbol{d})$ matrix, and $\mathbf{z_t}$ is a $(\boldsymbol{d}\ \text{x}\,\mathbf{1})$ vector of exogenous variables.
The 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 overfitting.

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

$$y_t = X_t \phi + \varepsilon_t \qquad t = 1, \ldots, T \qquad (3.5)$$

where, $X_t = (I_n \otimes G_{t-1})$ is $(n$ x $n)$, $G_{t-1} = (y'_{t-1}, \ldots, y'_{t-p} z'_t)'$ is $(k$ x $1)$ and $\phi = vec(\varphi_1, \varphi_2, \ldots, \varphi_p, D)$ is $(nk\,\text{x}1)$. 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 model's parameters is given by,

$$L(y/\phi, \Sigma), \propto |\Sigma|^{\frac{-T}{2}} \exp\{-\frac{1}{2}\sum_t (y_T - X_t\phi)'\Sigma^{-1}(y_t - X_t\phi)\}$$

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

$$P(\phi, \Sigma / y) = \frac{P(\phi, \Sigma) L(y/\phi, \Sigma)}{P(y)} \propto P(\phi, \Sigma) L(y/\phi, \Sigma)$$

The joint probability density function of the observations and the parameters, $P(\phi, \Sigma, y)$ can be expressed as

$$P(\phi, \Sigma, y) = L(y / \phi, \Sigma) P(\phi, \Sigma) = P(\phi, \Sigma / y) P(y)$$

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

BVAR model is VAR with priors introduced to control coefficients of the variables.

The VAR (k) model is:

$$y_t = \beta_0 + \sum_{i=1}^{k} \boldsymbol{A_j} y_{t-1} + \varepsilon_t$$

ε_t is assumed to be $\boldsymbol{i.i.d \sim N(0, \Sigma)}$, where y_t ($\boldsymbol{t = 1, \ldots, T}$) is a $\boldsymbol{(n\,x\,1)}$ vector of observations on $\boldsymbol{n}$ time series variables, $\boldsymbol{\beta_0}$ is a $\boldsymbol{(n\,x\,1)}$ vector of intercepts and $\boldsymbol{A_i}$ is a $\boldsymbol{(n\,x\,n)}$ coefficients matrix.

By defining $\boldsymbol{y}$ to be a $\boldsymbol{(T\,x\,1)}$ matrix, which stacks the observation of $\boldsymbol{T}$ on each dependent variable in columns next to one another. Denote by

$$x_t = (1, y_{t-1}, \ldots, y_{t-k}) \quad and \quad X = \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_T \end{bmatrix}, \quad A = \begin{bmatrix} \beta_0 \\ A_1 \\ . \\ . \\ . \\ A_k \end{bmatrix},$$

and $\beta = vec(A)$ which is a $(pm\;x\;1)$vector. Following the above definitions, the VAR model can be represented as:

$$\boldsymbol{Y = XA + E} \qquad \text{where } E \sim N(0, \Sigma)$$

Different choice of priors may be used with the VAR models, three prior was used in this study and the first was the Minnesota prior (Litterman, 1986). This prior assumes that Σ is known. The three priors used in this dissertation are defined in the following sections.

3.2.1.1 Litterman/Minnesota prior

This was proposed by researchers (Litterman 1986) at the University of Minnesota and the Federal Reserve Bank of Minneapolis hence the name Litterman prior or the Minnesota prior. Litterman prior assumes that each variable follows a random walk process with a possible drift. If we want to estimate the $(k \times 1)$ vector ϕ_i with the parameters of the i^{th} equation of (3.8) when the variance (σ_{ii}^2) of the error term is known, the Litterman prior assumes that the prior of ϕ is

$$p(\phi_i) \sim N(\bar{\phi}_i, \bar{\sigma}_i)$$

Where $\bar{\phi}_i$ and $\bar{\sigma}_i$ is the prior mean and variance-covariance matrix of ϕ_i respectively. Σ is assumed to be fixed and restricted diagonal matrix with its elements calculated from the estimation of a univariate autoregression model of order AR(p). The observation of the i^{th} equation 3.5 can be written as

$$Y_i = X\phi_i + \varepsilon_i \qquad\qquad i = 1, \dots, n;$$

Where Y_i and ε_i are $(T \times 1)$ vectors, X replaced X_t in equation (3.5).

Litterman (1986) assumes that λ is a degenerate random variable with the following structure for the diagonal elements of $\bar{\sigma}_i$.

$$\operatorname{var}(\bar{\phi}_j) = \begin{cases} \left(\dfrac{\lambda_1}{l^{\lambda_3}}\right)^2 & for \quad (i = j) \\[2ex] \left(\dfrac{\lambda_1 \lambda_2 \sigma_i}{l^{\lambda_3} \sigma_j}\right)^2 & for \quad (i \neq j) \end{cases}$$

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

By changing the hyper-parameter scalar values may lead to tightening or loosening the prior. The choice of the values for the scalars depends on the empirical trials of playing around different values. With this choice of prior, the posterior for ϕ is given by where

$$\tilde{\phi}_i = \tilde{\sigma}_i(\bar{\sigma}_i^{-1}\phi_i + \sigma_{ii}^{-1}X'Y_i)$$

and

$$\tilde{\sigma}_i = (\bar{\sigma}_i^{-1} + \sigma_{ii}^{-1}X'X)^{-1}$$

Therefore $p(\phi_i | Y) = N(\tilde{\beta}_i, \tilde{\sigma}_i)$, if $\bar{\sigma}_i^{-1}$, β_i and σ_{ii}^{-1} is known, $\tilde{\beta}_i$ may be taken as a point estimate.

3.2.1.2 Normal-Wishart prior

When the assumption of a fixed and diagonal variance-covariance matrix of residuals is relaxed, the conjugate prior for normal data is the normal-Wishart,

$$p(\phi | \Sigma) = N(\bar{\phi}, \Sigma \otimes \bar{\sigma})$$

$$p(\Sigma) = iW(\Sigma, \alpha)$$

The prior distribution of ϕ will be normal with prior mean $E(\phi) = \bar{\phi}$ and prior variance $V(\phi) = (\alpha - n - 1)^{-1} \bar{\Sigma} \otimes \bar{\sigma}$, where α is the degrees of freedom of the inverse-Wishart satisfying $\alpha > n + 1$.

From Bayes rule, the posterior is

$$p(\phi | \Sigma, Y) = N(\tilde{\phi}, \Sigma \otimes \bar{\sigma})$$

Where,

$$\tilde{\phi} = (\bar{\phi}^{-1} + X'X)^{-1}$$

and

$$\tilde{B} = \tilde{\phi}(\bar{\phi}^{-1}\bar{B} + X'X\hat{B}_{ols})$$

$$\tilde{\Sigma} = \hat{B}_{ols}'X'X\hat{B}_{ols} + \bar{B}'\bar{\phi}^{-1}\bar{B} + \bar{\Sigma} + (Y - X\hat{B}_{ols})'(Y - X\hat{B}_{ols}) - \tilde{B}'(\bar{\phi}^{-1} + X'X)\tilde{B}$$

3.2.1.3 Sims-Zha priors

Sims and Zha (1998) show the estimated structural BVAR model. To set the Sims-Zha priors for the structural parameters, the structural VAR model is suggested:

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

where $\varepsilon_t \sim N(0, I_j)$ and $\sum = A_0^{-1'} A_0^{-1}$. The Bayesian prior is developed for the unrestricted VAR then will be mapped to the restricted prior parameter. By defining A_τ to be a matrix of the coefficients on the lagged variable the equation may be written in the following form

$$YA_0 - XA_\tau = E \tag{3.6}$$

Where Y is $(H \text{ x } m)$, A_0 is $(m \text{ x } m)$, X is $(H + (mp+1))$, A_τ is $((mp+1) \text{ x } m)$ and E is $(H \text{ x } m)$.

Sims-Zha prior on A_0 and A_τ is given by:

$$\pi(A_0)\pi(A_\tau|A_0) = \pi(A_0)\varphi(\beta_{0,}T_0) \tag{3.7}$$

Where $\pi(A_0)$ denoting a marginal distribution of A_0 and $\varphi(\beta_0, T_0)$ represents a normal density with mean $\beta_0 = A_\tau - \mu(A_0)$ and covariance $T_0 = T(A_0)$.
The conditional likelihood can be expressed as

$$L(Y \mid A) \propto | A_0 |^H \exp(-.5 trace(ZA)^{'}(ZA)$$

where,

$$Z = [Y - X]$$

$$A = [A_0 A_\tau]^{'}$$

The posterior density is derived by combining equation (3.6) and 3.7) as follows:

$$\pi(\alpha) \propto \pi_0(\alpha_0) \mid A_0 \mid^H \mid T_0 \mid^{-1/2} \exp[-0.5(\alpha_0^{'}(I \otimes Y^{'}Y)\alpha_0 - 2\alpha_\tau^{'}(I \otimes X^{'}X)\alpha_\tau + \beta_0^{'}T_0^{-1}\beta_0)]$$

Where α denotes elements of vector A, this posterior is nonstandard, the conditional posterior distribution $A_\tau \mid A_0$ can be derived by:

$$\pi(\alpha_\tau \mid \alpha_0) = \phi(\overline{\beta}, (I \otimes X^{'}X + T_0^{-1})^{-1})$$

where,

$$\overline{\beta} = (I \otimes X^{'}X + T_0^{-1})^{-1}((I \otimes X^{'}Y)\alpha_0 + T_0^{-1}\beta_0)$$

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

$$H_{ol,ij} = \left(\frac{\lambda_0 \lambda_1}{\sigma_j l^{\lambda_3}} \right)^2$$

where, σ_j^2 denotes the j^{th} diagonal element of Σ for the l^{th} lag of the series i in equation j. The three hyper-parameters λ_0 , λ_1 and λ_3 represents the general beliefs about the VAR. λ_0 is overall tightness of beliefs on A_0, λ_1 denotes standard deviation on A_τ and λ_3 is a lag decay. In the case where prior information considered as dummy variables, Sims and Zha suggest Y^d and X^d as extra dummy variables.

$$Y^d = \begin{bmatrix} Y_1^d \\ Y_2^d \end{bmatrix}, \qquad X^d = \begin{bmatrix} X_1^d \\ X_2^d \end{bmatrix}$$

Which take care for unit roots (Y_1^d and X_1^d) and trends (Y_2^d and X_2^d), the model can be written as

$$\begin{bmatrix} Y^d \\ Y \end{bmatrix} A_0 - \begin{bmatrix} X^d \\ X \end{bmatrix} A_\tau = E$$

There are many other priorities which are not discussed and are beyond the scope of this paper.

3.2.1.4 Selecting priors

One of the important feature of Bayesian statistics is the construction of the parameters based on prior information which the modeller beliefs. Selecting the prior distribution is the most important part of BVAR modelling. 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 on the structure of the available information. Giannone et al (2012) recommended selection of priors using the marginal data density (i.e. the likelihood function integrated over the model parameters), which only depends on the hyper-parameters that characterize the relative weight of the prior model and the information in the data. The core of the BVAR models lies in the fact that the model parameters are random variables. The idea is to represent the prior information for all the unknown quantities through a prior distribution and combines them with the objective information coming from observations to obtain the posterior distribution (Sevinç and Ergün, 2009). The posterior distributions are commonly derived by the application of Bayes' theorem. In this dissertation, prior distributions which are considered and commonly used in literature are: The Litterman or Minnesota prior, The Normal-Wishart prior and The Sims-Zha normal-Wishart prior. These priors are based on the normal distribution.

3.2.2 Error, Trend, Seasonal Methods (ETS)

In explaining the ETS methods, it is important to start by discussing exponential smoothing techniques in general, as this is the basis on which ETS methods was developed. The exponential smoothing method is a popular technique in the forecasting environment- it was developed by Holt (1957) and extended by Winters (1960). The idea 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 suggests that the more recent the observation, the higher the associated weight will be.

There are several 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 (1957) improved SES to allow for the forecasting of data with a trend named after him, Holt's linear trend methods. It involves a forecasting equation and two smoothing equations for the level and the trend, as given by the equations:

$$
\begin{aligned}
&\tilde{y}_t = \alpha y_t + (1-\alpha)(\tilde{y}_{t-1} + r_{t-1}); && where, 1 > \alpha > 0 \\
&r_t = \beta(\tilde{y}_t - \tilde{y}_{t-1}) + (1-\beta) r_{t-1}; && where, 1 > \beta > 0 \\
&\hat{y} =_{t+1} = \tilde{y}_T + l r_t
\end{aligned}
$$

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

Holt's method was extended by Winters (1960) 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, 1998):

$$L_t = \alpha \frac{y_t}{s_{t-s}} + (1-\alpha)(L_{t-1} + b_{t-1}) \qquad where,\ 1 > \alpha > 0$$

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \qquad where,\ 1 > \beta > 0$$

$$s_t = \gamma \frac{y_t}{L_t} + (1-\gamma)s_{t-s} \qquad where,\ 1 > \gamma > 0$$

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

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

The seasonal methods may be additive or multiplicative, depending on seasonal variations, whether constant or changing in proportion to the level of the series. The development of smoothing methods was the work of Pegels (1969), who classified exponential smoothing methods according to taxonomy. Later, Gardner (1985) and Taylor (2003) extended smoothing methods. The combination of the trend and seasonal components results of fifteen possible exponential smoothing methods as shown in Table 3.1 (the equations for recursive calculations and point forecasts are given in Appendix A, Table 3-A1).

Table 3.1: Classification of exponential smoothing methods

Trend component	**Seasonal component**		
	N	**A**	**M**
	(None)	**(Additive)**	**(Multiplicative)**
N (None)	N,N	N,A	N,M
A (Additive)	A,N	A,A	A,M
A_d (Additive damped)	A_d,N	A_d,A	A_d,M
M (Multiplicative)	M,N	M,A	M,M
M_d (Multiplicative damped)	M_d,N	M_d,A	M_d,M

Source: Hyndman, R. and Athanasopoulos, G. (2012). Forecasting principles and practice

The exponential smoothing methods discussed in Table 3.1 only generate point forecasts. These methods were improved to generate point forecasts as well as forecast intervals, the so-called innovations state space models for exponential smoothing. This was the combined work of Ord et al. (1997), Hyndman et al. (2002) and Hyndman et al. (2005b), which showed that all exponential smoothing methods are optimal forecasts from innovations state space models. There are 30 state space models in total, 15 with additive errors (See Appendix B, Table 3-B1) and 15 with multiplicative errors (See Appendix C, Table 3-C1).

The error correction form of $\boldsymbol{ETS(A,N,N)}$ is derived from the error correction of simple exponential smoothing. The SES error correction equation is given as

$$y_t = \ell_{t-1} + \alpha\varepsilon_t$$

where,

$$\varepsilon_t = y_t - \ell_{t-1} \ and \ y^{t|t-1|} = \ell_{t-1}$$

Therefore, $\varepsilon_t = y_t - y^{t|t-1|}$ is a one-step forecast error, and is written $y_t = \ell_{t-1} + \varepsilon_t$.

The probability distribution of $\boldsymbol{\varepsilon_t}$ needs to be specified in order to make the equation an innovation state space. For a model with additive errors, the assumption is that one-step forecast errorsε_t are normally distributed i.e. $\boldsymbol{\varepsilon_t \sim NID(0,\sigma^2);}$ $\boldsymbol{NID}$ stands for "normally and independently distributed".

Then the equation of the model can be written as

$$y_t = \ell_{t-1} + \varepsilon_t \qquad (3.8)$$

$$\ell = \ell_{t-1} + \alpha\varepsilon_t \qquad (3.9)$$

Equation 3.8 is referred to as the measurement (or observation) and (3.9) as the state (or transition) equation. Equations (3.8) and (3.9), together with the error distribution, form a fully specified innovations state space model underlying simple exponential smoothing. Simple exponential smoothing with multiplicative errors is derived, similar to ETS (A,N,N), by writing the one-step random errors as relative errors:

$$\varepsilon_t = y_t - y^{t|t-1} y^{t|t-1}$$

where, $\varepsilon_t \sim NID(0, \sigma^2$. Substituting $y^{t|t-1} = \ell_{t-1}$ gives $y_t = \ell_{t-1} + \ell_{t-1}\varepsilon_t$ and $\varepsilon_t = y_t - y^{t|t-1} = \ell_{t-1}\varepsilon_t$.

Now, the state space model of multiplicative form is given as

$$y_t \ell_t = \ell_{t-1}(1+\varepsilon_t) = \ell_{t-1}(1+\alpha\varepsilon_t)$$

The two innovations state space models were developed by Hyndman *et al*. (2008b), one corresponding to a model with additive errors and the other to a model with multiplicative errors. “The versatile and fully automatic ETS framework requires neither stationarity nor strict linearity to produce contemporaneous time-series for variable time horizons” (Yusof & Kane, 2012). A complete and detailed explanation of *ETS* models can be checked in Hyndman et al. (2005).

3.2.3 Autoregressive integrated moving average (*ARIMA*)

Another approach used for the purpose of comparing the performance of *BVAR* is the *ARIMA* model, 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 differenced process to make the ARMA process stationary.

The seasonal *ARIMA*, denoted as *SARIMA*, is a generalisation and extension of the regular *ARIMA*. It is used for time-series where a pattern repeats itself seasonally over time (Machiwal and Jha, 2012). $ARIMA$ Models (Box et al., 1994) take historical data into account and decompose this data into an autoregressive process (AR), an integrated (I) process and moving average (MA) process of the forecast errors. The processes are identified by standard notation, i.e. the AR order is represented by p **,** I by d and MA by q and a combination of AR , I 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, i.e. the regression of the variable against itself. The autoregression model of the order *p* is written as:

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

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

Apart from the AR model, there is also the moving average model, which forecasts the dependent variable using a linear combination of white noise error terms. Generally, the MA model of the order *q* can be expressed as

$$y_t = c + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + ... + \phi_q \varepsilon_{t-q}$$

Where ε_t is a white noise. The general form of the combined processes $(ARIMA)$ is:

$$y_t + \beta_1 y_{t-1} + \beta_2 y_{t-2} + ... + \beta_p y_{t-p} + \varepsilon_t - \phi_t \varepsilon_{t-1} - \phi_t \varepsilon_{t-2} - ... - \phi_q \varepsilon_{t-q}$$

Most time-series have a seasonal component, and a time-series is seasonal if there is a periodic variation after a certain time interval. The seasonal $ARIMA$ model (Machiwal and Jha, 2012), commonly known as $SARIMA$, is a generalisation and extension of the ordinary $ARIMA$ model, in order to accommodate seasonality in the data. This seasonal component of the $ARIMA$ model is denoted by capital letters, $SARIMA(p,d,q)(P,D,Q)$, where the first bracket indicates the non-seasonal parameters and the last bracket indicates the seasonal factor parameters for the order of the autoregressive, integration and moving average parts of the model. The general $SARIMA(p,d,q)(P,D,Q)$ is given as:

$$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)(1-\beta B^s-\beta_2 B^{2s}-...-\beta_p B^{ps})(1-B)^d(1-B^s)^D y_t$$
$$=c+(1-\psi_1 B-\psi_2 B^2-...-\psi_q B^q)(1-\theta B^s-\theta_2 B^{2s}-...-\theta_Q B^{Qs})\varepsilon_t$$

where,

$(1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p)$ *is the nonseasonal autoregressive part of order* $p\,(AR(p))$
$(1-\beta B^s-\beta_2 B^{2s}-...-\beta_p B^{ps})$ *is the seasonal autoregressive part of order* $P(AR(P))$
$(1-\psi_1 B-\psi_2 B^2-...-\psi_q B^q)$ *is the nonseasonal moving average part of the order* $q(MA(q))$
$(1-\theta B^s-\theta_2 B^{2s}-...-\theta_Q B^{Qs})$ *is the seasonal moving average part of order* $Q(MA_s(Q))$
$(1-B)^d$ *is the differencing of order* $d\,(I(d))$
$(1-B^s)^D$ *is the seasonal differencing of order* $D(I_s(D))$
s *is the period of the seasonal pattern*

Because corporate income tax is volatile seasonal *ARIMA* will be applied to generate forecast.

3.2.3.1 Identification of the AR and MA process

The first step in developing model using Box-Jenkins approach is to determine the stationarity of the series and investigate if there is any significant seasonality that needs to be modelled. Once the issue of stationarity and seasonality have 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), partial autocorrelation function (PACF), and the resulting correlograms, which are plots of ACF and PACFs against the lags. The ACF plot shows the autocorrelations, which measure the relationship between y_t and y_{t-k} for different values of k. Observations may be correlated to each other, e.g. 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 measure the relationship between y_t and y_{t-k} by eliminating the effects of other time lags, $1,2,3,...,k-1$. 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 $\alpha(L) = 1 + \alpha_1 L + \cdots + \alpha_p L^p$.

The autocorrelation (ρ_τ) is given by

$$\rho_\tau = -(\alpha_1 \rho_{\tau-1} + \cdots + \alpha_p \rho_{\tau-p}) \text{ for all } \rho \geq p.$$

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 (π_τ) function identify a pure *AR* process clearly. The theoretical π_τ of *AR* (*p*) process has $\pi = 0$ for all $\tau > p$. The sample partial autocorrelation function elements are expected to be close to zero for lags greater than p, corresponding to the estimates of parameters that are equal to zero. The partial autocorrelation significance of the values is checked by the pth order process of which 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 $\tau > q$. 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

$$\frac{1}{\sqrt{N}}\left\{1 + 2(r_1^2 + r_2^2 + \cdots + r_q^2)\right\}^{1/2} \quad \text{for} \quad \tau > q.$$

The MA (q) process autocorrelation function becomes zero at lag *q+1* and greater, it is essential to examine the sample autocorrelation function 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 white-noise series.

3.2.3.2 Ljung-Box Test

The Ljung-Box test is a commonly used portmanteau test for $ARIMA$ models, and was developed by Ljung and Box (1978). The purpose of this Q-statistics test is to determine whether the set of autocorrelation coefficients is different from a zero. The test is done to the residuals of an estimated $ARIMA$ model, instead of to the original series. The Q-statistic is given by Makridakis et al. (1998) as:

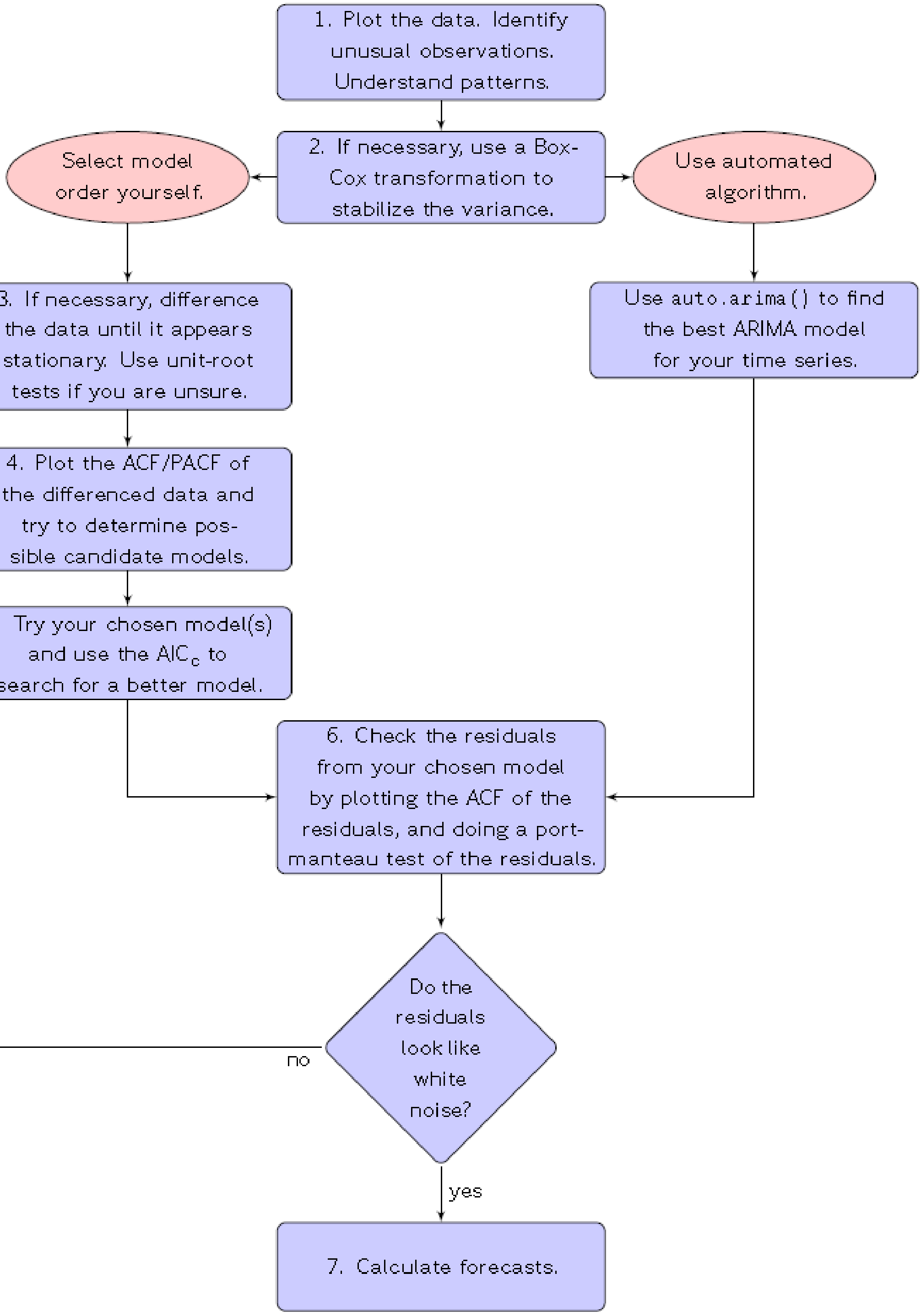
$$Q^* = n(n+2)\sum_{k=1}^{h}\frac{r_k^2}{(n-k)}$$

Where, n is the sample size, k represents the number of autocorrelation lags included in the statistic and r_t^2 denotes the squared sample autocorrelation at lag t.

The Ljung-Box test the hypothesis that the residuals from 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 be taken not to accept a model on the basis of portmanteau tests alone"* (Makridakis et al., 1998).

3.2.3.3 General process for forecasting using an ARIMA model

Figure 3.1 shows the common procedure followed when fitting an ARIMA model.



Source: Hyndman, R.J. and Athanasopoulos, G, (2012). Forecasting: principles and practice

Figure 3.1: The common procedure followed when fitting an ARIMA model

3.3 STATIONARITY AND UNIT ROOT TESTS

Many time-series appear in reality to be not stationary, and there are various unit root tests used to detect if the series non-stationary. Non-stationarity can be detected by visual examining of the time-series graph and by looking at the series correlogram, or by the use of 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 differenced once is represented by $I(1)$ and $I(0)$ is a stationary time-series with the order zero. For the purpose of this study, two tests for stationarity will be discussed, that is, Dickey-Fuller test and Phillips-Perron test.

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

3.3.1 Dickey-Fuller test

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

$$y_t = \rho y_{t-1} + \mu_t$$

where,u_t is an IID series of random variables. The DF test hypotheses are

$$H_0 : \rho = 1, \quad against$$

$$H_1 : \rho < 1$$

y_t is non-stationary under the null hypothesis, and is a stationary under the alternative hypothesis. The standard t-statistics does not follow t-distribution because of the non-stationarity of y_t under the null hypothesis. To test the null hypothesis, the following test statistics equation may be used:

$$DF = \frac{\rho - 1}{s.e(\rho)} \tag{3.10}$$

The DF test in equation (3.10) 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 eliminate all the autocorrelations in the time-series. The procedure for the ADF test is similar to the Dickey–Fuller test procedure, the only difference is the model where is applied. The model where ADF is applied to is shown as,

$$\Delta y_t = \alpha + \beta t + \phi y_{t-1} + \delta_1 \Delta y_{t-1} + ... + \delta_{p-1} \Delta y_{t-p+1} + \varepsilon_t$$

where, α denotes a constant, β is the coefficient on a time trend, and p 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 $\beta = 0$ resembles a modelling of random walk with a drift. The ADF test is performed under the hypothesis

$$H_0 : \phi = 0, \quad against$$

$$H_1 : \phi < 0$$

The test statistic is computed as:

$$DF_\tau = \frac{\hat{\phi}}{SE(\hat{\phi})}$$

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

3.3.2 Phillips-Perron test

The Phillips-Perron (PP) test is an alternative technique for correcting for serial correlation in unit root testing, and was developed by Phillips and Perron in 1988. The PP test uses the standard DF or ADF test, but modifies the t-ratio as 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 deals with the issue of serial correlation and heteroskedasticity in the error terms. The test model for the PP test is given as

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

Where μ_t denotes $I(0)$ which may be heteroskedastic. The PP tests correct for serial correlation and heteroskedasticity in the error terms μ_t of the 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:

$$Z_t = \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2}\right)^{1/2} . t_{\pi=0} - \frac{1}{2}\left(\frac{\lambda^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right) . \left(\frac{T.SE(\hat{\pi})}{\hat{\sigma}^2}\right)$$

$$Z_\pi = T\hat{\pi} - \frac{1}{2}\frac{T^2.SE(\hat{\pi})}{\hat{\sigma}^2}(\hat{\lambda}^2 - \hat{\sigma}^2)$$

The estimates of the variance parameters of

$$\sigma^2 = \lim_{T\to\infty} T^{-1}\sum_{t=1}^{T} E\left[\mu_t^2\right]$$

$$\lambda^2 = \lim_{T\to\infty} \sum_{t=1}^{T} E\left[T^{-1} S_T^2\right]$$

are $\hat{\sigma}^2$ and $\hat{\lambda}^2$. Where $S_T = \sum_{t=1}^{T} \mu_t$. The sample variance of the least squares residual $\hat{\mu}_t$ is a consistent estimate of σ^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 PP Z_t and Z_t statistics have the same asymptotic distributions as the ADF t-statistics and normalised bias statistics. The advantage of the PP tests over the ADF tests is that the 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.

3.4 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 of assessing the statistical models fit for a given set of data. The AIC value of the model is represented by the following equation:

$$AIC = -2In(L) + 2k$$

where L is the numeric value of the maximum likelihood for 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 observation are 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 (2002). The corrected AIC is represented by the formula (3.55) (Hurvich and Tsai, 1989).

$$AIC_C = -2In(L) + \frac{2kn}{n-k-1} = AIC + \frac{2k(k+1)}{n-k-1}$$

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 which may be used for model section, but in this study, the AIC is used to select the best model.

3.5 MEASURES OF ACCURACY FOR FORECASTING

In evaluating the accuracy of forecasts, frequently used measures of forecast accuracy are employed to 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:

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

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

where, $\hat{y}_t$ is the forecasted value in the period t, y_t is the actual value in the 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:

$$MSE = \frac{\sum_{t=1}^{n} e_t^2}{n}$$

$$MAE = \frac{\sum_{t-1}^{n} | e_t |}{n}$$

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

3.6 DATA AND STATISTICAL SOFTWARE

The data for tax revenue collections was sourced from an annual SARS tax statistics publication, which is a joint publication of SARS and the National Treasury. Economic data was sourced from the website of Statistics South Africa (StatsSA) and the South African Reserve Bank's (SARB) online statistics tool.

The software used to generate the results was Econometric Views (Eviews), which is currently used in SARS for model development and to generate revenue forecasts. The other software programs that will complement Eviews, if need be, are SPSS, R and SAS. Hyndman and Kandahar (2008) describe two automatic forecasting algorithms appropriate to seasonal and non-seasonal data, which have been employed in the forecast package for R. The first algorithms are for innovations state space models underlying exponential smoothing methods, and the second is a step-wise algorithm for forecasting with $ARIMA$ models.

3.7 CONCLUSION

This chapter discussed the techniques used in this study, namely Bayesian Vector Autoregression, Error, Trend, Seasonal models (state space models) and Autoregressive Integrated Moving Average (this may be with or without seasonality, and may be differenced depending on the behaviour of the data). These techniques were considered to be appropriate for capturing the data generating process of past observations of Total Tax Revenue (TTR), CIT, VAT and PIT. The techniques will also be used to generate the revenue collection forecasts of these three tax types.

The prior's selection for *BVAR* was also discussed and the choice of the prior distribution is the most important part of *BVAR* modelling. In this dissertation, prior distributions which 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 explore the possibility of the presence of unit roots in the error terms, ADF and PP tests will be used and the results are discussed in chapter 4 of this study. 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 based on measures of accuracy within different approaches.

Revenue forecasting is not limited to the techniques discussed in this study, however, as other methods exist which may be used to generate forecasts, such as ARCH, GARCH, ECM, etc.

CHAPTER 4

DATA ANALYSIS AND RESULTS

4.1 INTRODUCTION

The chapter deals with data analysis and behaviour of the selected tax and economic variables. The statistical methods employed in the analysis of the data were discussed in Chapter 3. The focus of this chapter is on the application of the methodologies discussed earlier, analysis and interpretation of the results of the three methodologies employed to generate forecasts of the three selected tax types.

4.2 BEHAVIOUR OF THE TAX VARIABLES

The CIT revenue collection has been growing rapidly between 1997/98 and 2013/14. In 2009/10 the CIT collection contracted due to economic recession that started late in 2008. In 2010/11 the CIT collection slowly recovered with an improved contraction of 1.5%. CIT collection share to total revenue has improved drastically from 12.9% in 1997/98 to 19.7% in 2012/13 as depicted by Figure 4.1. The CIT contribution to GDP rose from a 3.0% in 1997/98 to the highest ever of 6.9% of GDP in 2008/09. The CIT-GDP ratio declined to 5.3% in 2009/10 due to local and global economic downturn. CIT collection is highly volatile and sensitive to economic condition.

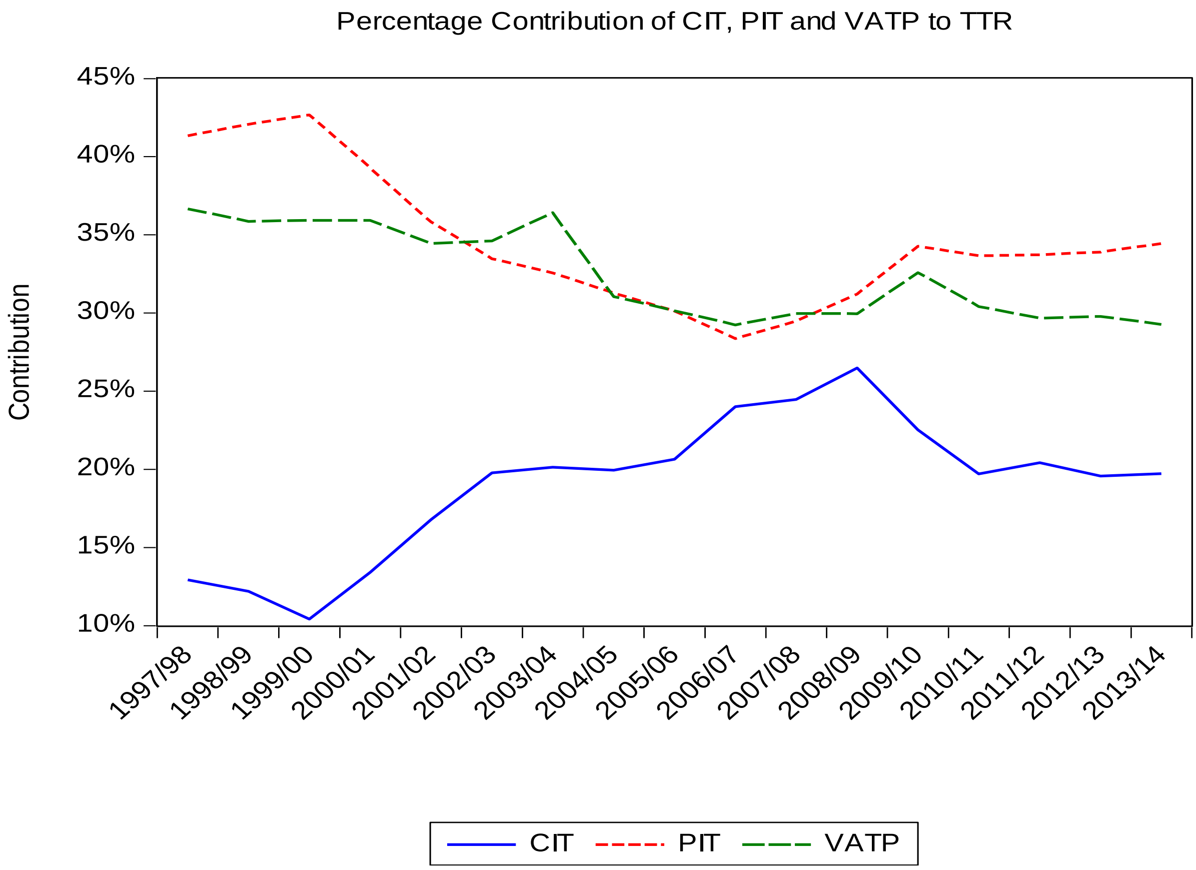


Figure 4.1: Graph showing percentage contribution of CIT, PIT, and VATP to total tax revenue (TTR).

The largest source of tax revenue over the years has been PIT although its share to total revenue has declined from 41.3% in 1997/98 to 34.4% in 2013/14. In 2009/10 through 2013/14 contribution was hovering around 34.0% on average. In 2002/03 and 2003/04 PIT share to total tax revenue was surpassed by Value-Added share recording 33.5% (VATP: 34.6%) and 32.6% (VATP (36.4%) respectively (See Figure 4.1). A share of PIT to GDP has decline from 9.5% in 1997/98 to 8.6% in2013/14.

VAT payments is the second largest following PIT, its contribution has been decreasing over the years like PIT, in 1997/98 the VATP share to total tax revenue was 36.7% reaching 29.3% in 2013/14 (See Figure 4.1). The average contribution from 2009/10 to 2013/14 was 30.3%. The VATP-GDP ratio has slightly fell below the levels of 1997/98 from 8.4% to 7.3% in 2013/14. The South Africa VATP rate has remained 14% since around 1993/94 to date.

The total tax revenue has been growing tremendously since 1997/98 driven by PIT, VATP and CIT as these three tax types contribute approximately 80%. Total tax to GDP ratio has improved significantly since the born of democracy (See Figure 4.3).

The total tax revenue (TTR) collection has been growing significantly from R165bn in 1994/97 to R900bn in 2013/14. In some years TTR experienced single growths due to poor/weak economic conditions most evident in 2009/10 (-4.2%) following recession in 2008/09.

4.3 BEHAVIOUR OF THE ECONOMIC VARIABLES

All economic variables data was converted to fiscal year (April – March) as tax revenue data is in fiscal years. Gross Domestic Product (GDP) was growing faster before recession (GDP) with a four year average of 12.4%, while after recession the rate of growth was 9.1% on a four year average. The Gross Operating Surplus (GOS) trend follows that of GDP with declining trend with a pre-recession four year average of 12.8% and post-recession four year average of 7.8%. Both GDP and GOS show a downward trend as depicted by Figure 4.2. The pre-recession four-year average of Compensation of Employees (CoE), Gross Domestic Expenditure (GDE), Private Consumption Expenditure (PCE), Employment, Consumer Price Index (CPI) and Exchange Rate are higher than post-recession.

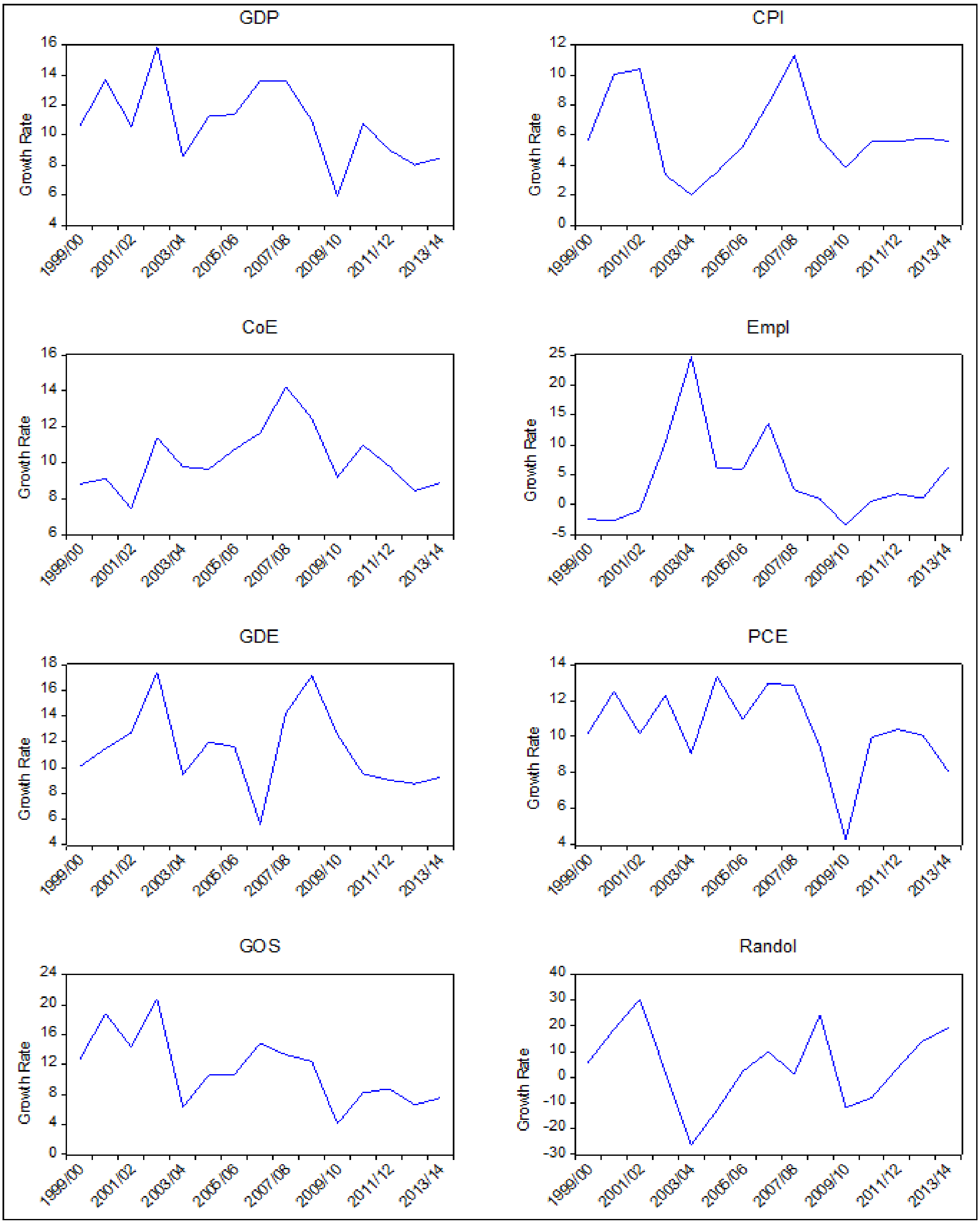
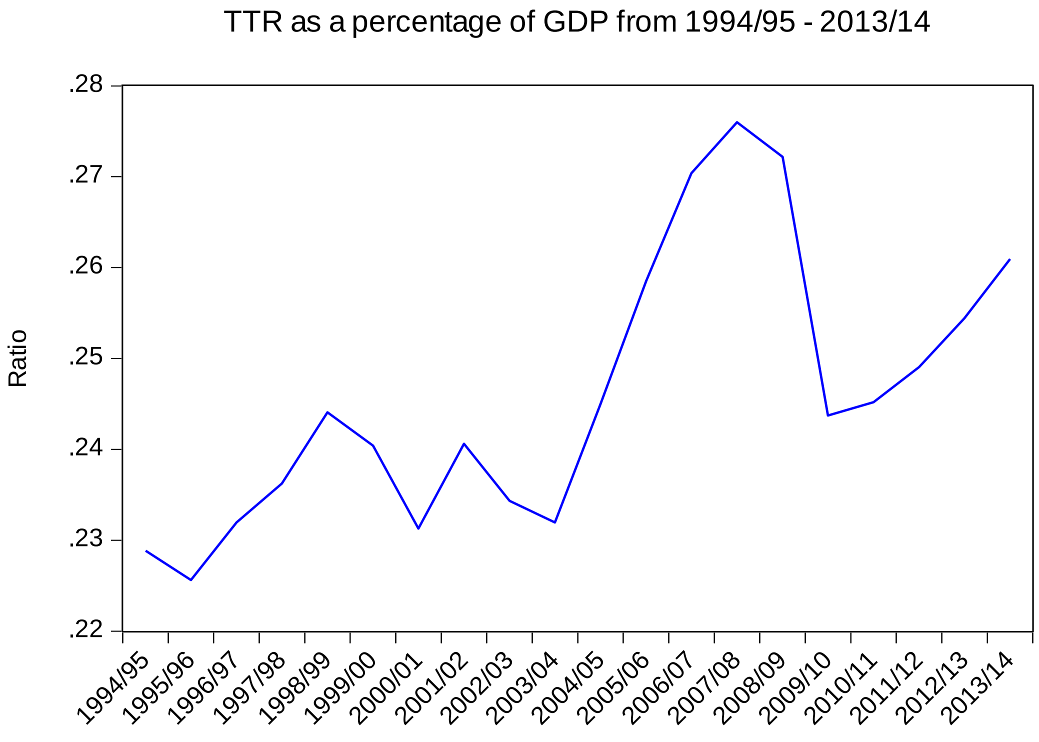


Figure 4.2: The graph showing trends of percentage growth of Economic variables.

CPI averaged 70.1 index points from second quarter of 1998 until first quarter of 2012. Rand/dollar exchange rate averaged R7.4 per dollar from second quarter of 1998 until first quarter of 2012. During this period employment averaged 6.7m. Employment had an upward trend 2003/04 then the trend quickly drops until 2009/10 and slowly recovered through the years up to 2013/14 financial year. Compensation to employee trend was increasing up to 2007/08 financial year, and then decreasing till 2012/13 financial year.

4.3.1 The ratio of total tax revenue collection to gross domestic products in South Africa

The ratio of Tax to GDP is a crucial economic indicator and it is used globally for the purpose of comparative analysis between various countries. Many institutions around the world use this ratio; they include International Monetary Fund (IMF), the World Bank, the Organisation for Economic Co-operation and Development (OECD). Since the fiscal year 1994/95 the South African tax-to-GDP ratio has significantly improved from 22.9% to 26.1% in 2013/14. The ratio was driven by contribution of PIT, VATP and to some extends by CIT. The Figure 4.3 shows the TTR-GDP ratio.

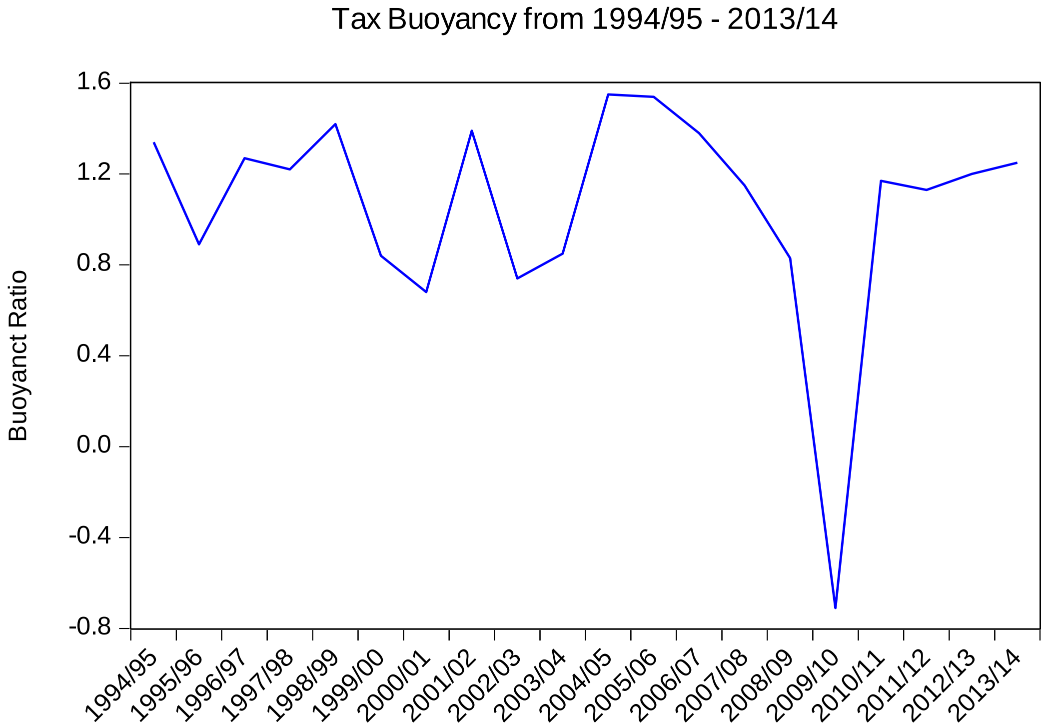


Source: National treasury and SARS, 2014 Tax Statistics.

Figure 4.3: The graph showing Total Tax Revenue (TTR) as a percentage of GDP from 1994/95 – 2013/14.

South Africa total revenue collection has been increasing since 1994/95 till 2013/14 except in 2009/10 due to recession (economic downturn). The reasons cited for increased revenue in the later years were above-inflation wage settlements increment, increases in domestic consumption, increased commodity prices, growth in the value of imports and also improved tax administration and compliance through modernization.

Another instrument for measuring the tax revenues performance relative to changes in economic growth is the buoyancy/elasticity of taxes. South African year-on-year revenue buoyancy has been fluctuating through the years reaching -0.61 in 2009/10 (recession period). The Figure 4.4 depicts tax buoyancy from 1994/95 to 2013/14.



Source: National treasury and SARS, 2014 Tax Statistics.

Figure 4.4: The graph showing tax buoyancy from 1994/95 – 2013/14.

The improved tax compliance and strong economic growth result in buoyant tax revenue collections.

The following sections will deal will model development and analysis of the results thereof generated by three techniques discussed in this dissertation.

4.4 TEST FOR STATIONARITY FOR INDIVIDUAL TIME SERIES

The tests which are used in this dissertation to assess the stationarity of the time series variables are ADF and PP tests (discussed in chapter 3, section 3.3.1). The individual series are also plotted in various section of this chapter to visually investigate the possibility of stationarity. The ADF and PP tests results for each of the variables used in this study are reported in Table 4.1.

Table 4.1: ADF and PP tests results for stationarity of individual time series.

Variables	Augmented Dicky-Fuller		Phillips-Perron	
	Constant	Constant & trend	Constant	Constant & trend
Results for unit roots in levels				
LCIT	-1.788	-0.565	-1.680	-4.986***
LPIT	1.682	-2.197	0.352	-7.691***
LVATP	-0.610	-1.768	-0.271	-6.977***
LTTR	-1.031	-2.224	-0.021	-5.807***
LGOS	-2.880*	-1.260	-3.418***	-1.769
LRandol	-1.829	-2.163	-1.898	-2.211
LCoE	-0.480	-1.759	-0.745	-5.531***
LEmpl	-0.878	-1.566	-0.878	-1.561
LGDE	-1.018	-0.750	-1.891	-3.957
LPCE	-1.841	-0.615	-1.672	-6.286***
LGDP	-2.581	-0.087	-3.375**	0.873
LCPI	-1.500	-6.073***	-1.164	-6.263***
Variables	**Augmented Dicky-Fuller**		**Phillips-Perron**	
	Constant	**Constant & trend**	**Constant**	**Constant & trend**
Results for unit roots in first differences				
DLCIT	-13.795***	-14.068***	-23.127*	-35.102***
DLPIT	-3.712*	-4.294*	-29.596*	-34.100***
DLVATP	-5.298***	-5.279***	-31.128*	-33.350***
DLTTR	-2.880**	-2.956	-18.072*	-18.308***
DLGOS	-1.339	-2.954	-9.090*	-13.108***
DLRandol	-6.535***	-6.483***	-6.549*	-6.496***
DLCoE	-2.897**	-2.865	-18.665*	-18.450***
DLEmpl	-8.760***	-8.707***	-8.759*	-8.707***
DLGDE	-3.838***	-3.959**	-17.495*	-18.332***
DLPCE	-1.680	-2.415	-24.018*	-30.784***
DLGDP	-1.495	-2.896	-8.657*	-10.345***
DLCPI	-6.473***	-6.566***	-23.146*	-29.491***

***; **; * denotes stationarity at 1%, 5% and 10% respectively.

The stationary series are used in the development of models for various tax types, that is, CIT, PIT, VATP and finally TTR.

4.5 PROPOSED MODELS TO FORECAST QUARTERLY CIT

4.5.1 CIT ARIMA Models

4.5.1.1 Testing CIT Series for Stationarity

To Identifying an appropriate ***ARMA*** model a time series to be used must be stationary. An ***ARMA*** models to be stationary there is a need for roots modulus of the ***AR*** polynomial be bigger than unity, and for the ***MA*** process to be invertible it is also required that the roots of the ***MA*** polynomial lie outside the unit circle. To investigate the stationarity of the CIT, the graph of the quarterly CIT is plotted as shown in the Figure 4.5.

Figure 4.5: The graph showing Corporate Income Tax trend from 1998 – 2014.

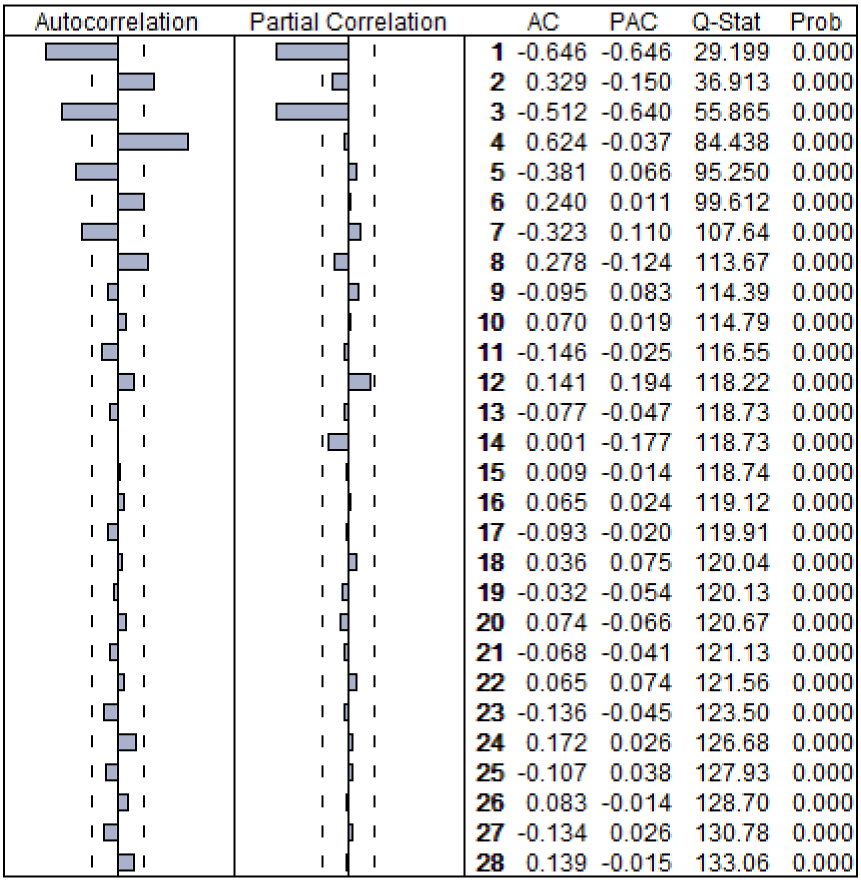
It is evident that the CIT time series displays a nonstationary pattern. Observing the CIT correlogram (See Appendix B, Figure 4-B1) it is clear that the quarterly data does not have seasonal pattern. The coefficients of autocorrelation start with a high value and slowly declines suggesting a non-stationary series. The Ljung-Box Q-statistic at the 28th lag has a probability value of 0.000 which is less than 0.05, also this confirm the non-stationarity of CIT data series. The Augmented Dickey Fuller test discussed in chapter 3 section 3.3.1 is used to explore the stationarity. And further to investigate the possibility of a unit root Phillip Perron test is also used and was discussed in section 3.3.2 of chapter 3. Both tests suggest that CIT series at the level is non-stationary as reflected by the unit root tests in Table 4.1.

The data became stationary after it was first differenced, the p-value of the differenced data is less than the significance level of 0.01, these means that the null hypothesis of nonstationarity is rejected suggesting that the data is stationary at first level (See Table 4.1). To further smoothen the data the sample logarithm was taken, normally this is done to lessen the severity of the data. The findings from unit root tests suggest that LCIT variable is also non-stationary at level, and then it became stationary at first difference (See Table 4.1).

4.5.1.2 CIT ARMA Model Identification and Estimation

Now that the correct order of differencing required to make the CIT series stationary has been determined, we now try to find an appropriate ARMA form to model the stationary CIT series. The Box-Jenkins techniques is the traditional and most commonly used methodology which involves examining plots of the sample autocorrelation and partial autocorrelation from a correlogram. Besides Box-Jenkins methodology, there are a number of other methods for identification suggested in the literature. These alternative methods include the Corner method proposed by Beguin, Gourieroux & Monfort (1980), the R and S Array method developed by Gray, Kelly & McIntire (1978) and canonical correlation methods by Tsay and Tiao (1985).

The correlogram of the stationary data (DLCIT) is plotted 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			-0.646	-0.646	29.199	0.000
2			0.329	-0.150	36.913	0.000
3			-0.512	-0.640	55.865	0.000
4			0.624	-0.037	84.438	0.000
5			-0.381	0.066	95.250	0.000
6			0.240	0.011	99.612	0.000
7			-0.323	0.110	107.64	0.000
8			0.278	-0.124	113.67	0.000
9			-0.095	0.083	114.39	0.000
10			0.070	0.019	114.79	0.000
11			-0.146	-0.025	116.55	0.000
12			0.141	0.194	118.22	0.000
13			-0.077	-0.047	118.73	0.000
14			0.001	-0.177	118.73	0.000
15			0.009	-0.014	118.74	0.000
16			0.065	0.024	119.12	0.000
17			-0.093	-0.020	119.91	0.000
18			0.036	0.075	120.04	0.000
19			-0.032	-0.054	120.13	0.000
20			0.074	-0.066	120.67	0.000
21			-0.068	-0.041	121.13	0.000
22			0.065	0.074	121.56	0.000
23			-0.136	-0.045	123.50	0.000
24			0.172	0.026	126.68	0.000
25			-0.107	0.038	127.93	0.000
26			0.083	-0.014	128.70	0.000
27			-0.134	0.026	130.78	0.000
28			0.139	-0.015	133.06	0.000

Figure 4.6: Correlogram of CIT (Logarithmic Form, 1st Differenced) (DLCIT).

The limits for ACF and PACF are $\pm\frac{2}{\sqrt{65}}=\pm0.2481$. Looking at Figure 4.4, it can be observed that the ACF cuts off at lag 4 $(q=4)$ and the PACF cuts off at lag 3 $(p=3)$. Now the ranges of models are explored $\{ARMA(p,q):0\le p\le 3,\ 0\le q\le 4\}$ and the best model is selected based on AIC and SIC as mentioned in chapter 3, section 3.4. After identifying the parameters, the automatic ***ARIMA*** forecasting was performed using Eviews and twenty models were generated and the best top five ***ARIMA*** models are shown in the Table 4.2.

Table 4.2: Top five CIT ARIMA models based on AIC.

Model	**LogL**	**AIC***	**BIC**	**HQ**
(4,1,1)	1.6012	0.0500	0.1585	0.0920
$(3,1,0)(1,0,1)_4$	2.7467	0.1142	0.3292	0.1977
$(4,1,0)(1,0,0)_4$	2.4979	0.1229	0.3379	0.2065
$(3,1,0)(0,0,1)_4$	1.2703	0.1309	0.3101	0.2005
$(2,1,0)(1,0,1)_4$	1.1029	0.1367	0.3160	0.2064

The models are selected based on AIC; the model with the smallest AIC is the best model. The appropriate model selected is $ARMA(4,1,1)$ with AIC of 0.05. The competing models are $ARMA(3,1,0)(1,0,1)_4$ with AIC of 0.1142; $ARMA(4,1,0)(1,0,0)_4$ with AIC of 0.1229; $ARMA(3,1,0)(0,0,1)_4$ with AIC of 0.1309 and $ARMA(2,1,0)(1,0,1)_4$ with AIC of 0.1367.

4.5.1.3 The best Model

The Table 4.2 indicate that the best model is $ARMA(4,1,1)$ based on AIC of 0.0500. The model output may be seen in Appendix B, Table 4-B1). The model is stationary at first difference $(d=1)$. The identified model is given as:

$$\hat{y}_t = \beta_1 \hat{y}_{t-1} + \beta_2 \hat{y}_{t-2} + \beta_3 \hat{y}_{t-3} + \beta_4 \hat{y}_{t-4} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \qquad where \qquad \hat{y}_t = \hat{Z}_t - \hat{Z}_{t-1}$$

or

$$DLCIT_t - DLCIT_{t-1} = 0.7228\, DLCIT_{t-4} + \varepsilon_t + 0.6901\, \varepsilon_{t-1}$$

The CIT ARIMA model is chosen based on meeting the prerequisites which are well 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 by Figure 4-B4 in Appendix B.

4.5.1.4 Diagnostic Checking of the CIT ARIMA Model

Diagnostic checking assists to check that the estimated model is statistically sound and acceptable. This is based on some statistical tests which are done to check whether the residuals of the models are not auto-correlated and are normally distributed. The Q-statistics test (Ljung-Box) explained in chapter 3 sections 3.1 is used to check autocorrelation and for normality test Jarque-Bera test (1980) is used. The Figure 4-B5 and 4-B6 in Appendix B, shows the autocorrelation and normality test of the residuals of the $ARMA(4,1,1)$ model. The Ljung-Box Q-statistics indicate that values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected, therefore, there is no autocorrelation detected in the residuals series. Also the normality test confirms that the residual of $ARMA(4,1,1)$ model follows a normal distribution. Since p-value 0.1275 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected. Therefore, the residuals follow a normal distribution.

4.5.1.5 CIT ARIMA Model Forecasts

One of the objectives of this study is to forecasts CIT into the future for 12 quarters ahead, from second quarter of 2012 to first quarter of 2015 using the best selected model $(ARMA(4,1,1))$.

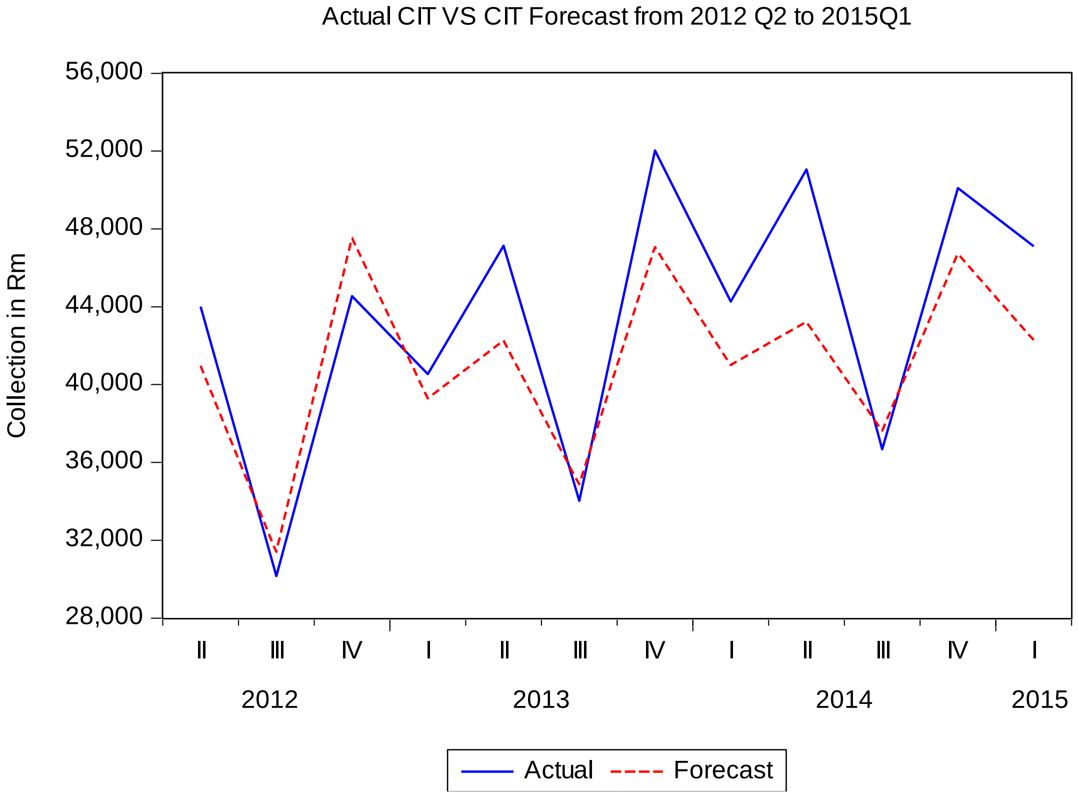


Figure 4.7: The graph showing actual collection and forecasts for h=12 quarters ahead with an $ARIMA(4,1,1)$ *model.*

Figure 4.7 shows the diagrammatic representation of the quarterly actual CIT collection in million rand and its forecasts. Checking the measures of forecast accuracy, the RMSE is 3847.81, MAE of 3286.44, MAPE is 7.12 while Theil statistics is 0.05.

4.5.2 CIT Error, Trend, Seasonal Models

4.5.2.1 CIT ETS Model Selection

The ETS models was performed by using Eviews Automatic Forecast tools which produces 30 models and select the best model based on AIC. Out of the 30 model specifications the Multiplicative error, Additive trend, Additive season (M, A, A) model was selected as the best model and represented by the equation:

$$\begin{aligned} y_t &= (\ell_{t-1} + b_{t-1} + s_{t-m})(1+\varepsilon_t) \\ \ell_t &= \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \\ b_t &= b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \\ s_t &= s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \end{aligned} \quad or \quad \begin{aligned} y_t &= (\ell_{t-1} + b_{t-1} + s_{t-m})(1+\varepsilon_t) \\ \ell_t &= \ell_{t-1} + b_{t-1} + 0.35(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \\ b_t &= b_{t-1} + 0(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \\ s_t &= s_{t-m} + 0.8(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t \end{aligned}$$

The selected model based on AIC has a level smoothing parameter estimate $\alpha = 0.35$, trend parameter $\beta = 0$ (zero indicate that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.80$. The output of the best model is shown in the Appendix B, Table 4-B2. The best five ETS model based on AIC is depicted in the Table 4.3.

Table 4.3: Top five CIT ETS models based on AIC

Model	Likelihood	AIC*	BIC	HQ
(M, A, A)	-545.1140	1174.9200	1191.2700	1181.2700
(M, M_D, A)	-544.2740	1175.2400	1193.6300	1182.3900
(M, A_D, A)	-545.1140	1176.9200	1195.3100	1184.0700
(M, M, A)	-546.2030	1177.1000	1193.4500	1183.4500
(M, A, M)	-551.7960	1188.2900	1204.6300	1194.6400

The four models which are competing with $ETS(M, A, A)$ as shown from Table 4.3 are $ETS(M, M_D, A)$ with AIC of 1175.2; $ETS(M, A_D, A)$ with AIC of 1176.9; $ETS(M, M, A)$ with AIC of 1177.1 and $ETS(M, A, M)$ with AIC of 1188.3.

4.5.2.2 CIT ETS Model Forecasts

The best $ETS(M, A, A)$ model is used to generate the CIT forecasts plotted in the graph (Figure: 4.8). By observing the graph, it can be seen that the forecast CIT series is closer to the actual series except in quarter four of 2012 and 2015.

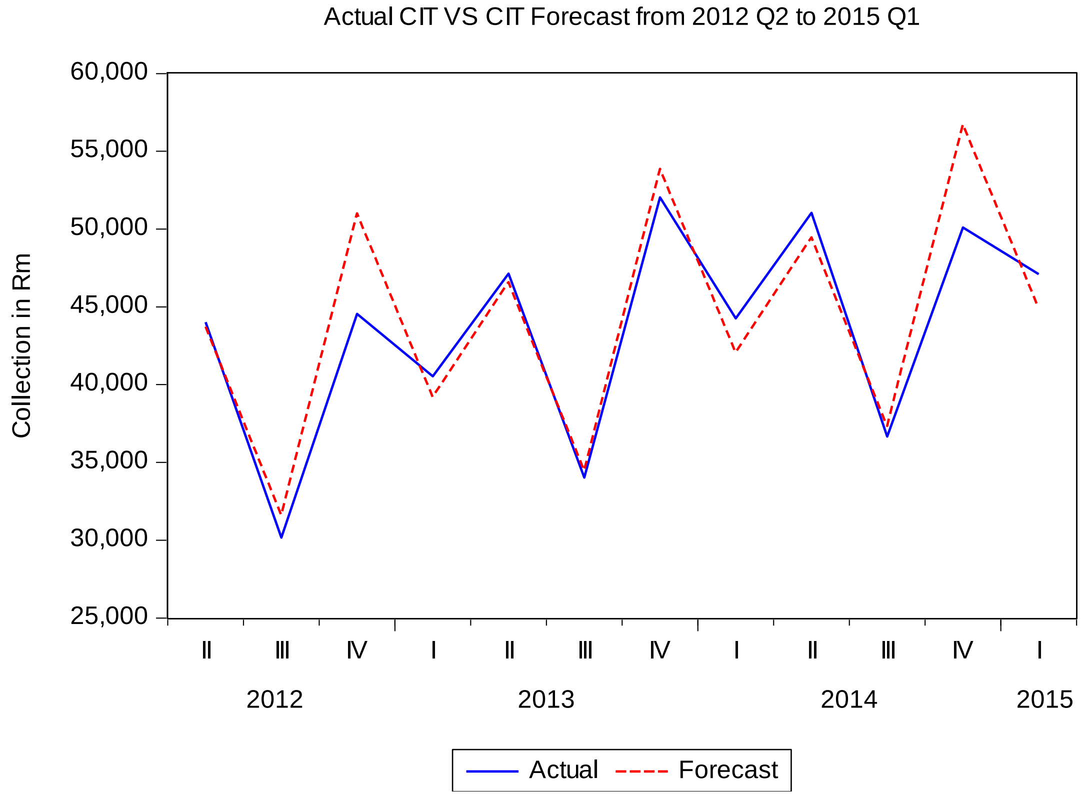


Figure 4.8: The graph showing actual collection and forecasts for h=12 quarters ahead with an $\boldsymbol{ETS}\ (\boldsymbol{M}, \boldsymbol{A}, \boldsymbol{A})$ *model.*

The closeness of the forecasting series to the actual series suggests that the selected model has better prediction power and is appropriate to forecast CIT. This is confirmed by the calculated accuracy measures, i.e. RMSE of 2975.36, MAE of 2134.91, MAPE of 4.75 and a Theil U statistics of 0.03.

4.5.3 Bayesian Vector Autoregressive (BVAR) Model for CIT

Selection of parameters in this study is a combination of the ones suggested in the literature and a search over a range of possible hyper-parameters to check which combination provides the best forecasting model with minimum RMSE. The different hyperparameters are used in order to obtain more robust results. Doan (2007) proposes that the priors should be selected as symmetric with an overall tightness of $\lambda_1 = 0.2$ and the relative weight $\lambda_2 = 0.1$ for small sized models. Caraiani (2010) estimate models with $\lambda_1 = 0.2$ and $\lambda_1 = 0.5$ with lag decay set to 1 and 2. In the study of Korobolis (2009), Kadiyala and Karlsson (1997) set the relative weight to 0.005. Sims-Zha (2007) propose $\lambda_0 = 1$, $\lambda_1 = 0.2$, $\lambda_3 = 1$ and 1 for unit root and trend dummies.

The CIT BVAR models are estimated using three priors, i.e. Minnesota prior, Normal-Wishart prior and Sims-Zha prior. The Minnesota prior parameters, $mu_1, \lambda_1, \lambda_2$ and λ_3 are set to 0.5; 0.5; 0.6 and 0.1 respectively. The parameters for Normal-Wishart, mu_1 and λ_3 are set to 0.5 and 0.01 respectively. The Sims-Zha parameters, λ_0, γ_1 and λ_3 are set to 1; 0.9 and 0.9 respectively. The λ_1 represents the overall tightness parameter and the range is [0,1]. It is the prior standard deviation of the coefficient of the first own lag, and basically controls the prior standard deviations of all the other lag coefficients. This prior determines how all the coefficients are concentrated around their prior means. When the tighter prior is desired λ_1 must be decreased. The λ_2 is the cross-variable weight tightness parameter, it represents the tightness of variable j in relation to variable i in equation i and the range is [0,1]. Own lags generally account for most of variation in a dependent variable, therefore the coefficients of cross lags are given smaller standard deviations than coefficients of own lags. The λ_3 ($\lambda_3 > 0$) is a decay factor that controls the tightness on lag l relative to lag 1. As coefficients of higher order lags are more likely to approach zeros than those of lower order lags, prior standard deviations of coefficients decrease as lag length l increases. The table 4.4 shows the results of CIT *BVAR* models using three priors.

Table 4.4: CIT BVAR models results with three priors

	Minnesota Prior			Normal Whishart Prior			Sims-Zha Prior		
	$DLCIT_t$	$DLGOS_t$	$DLRANDOL_t$	$DLCIT_t$	$DLGOS_t$	$DLRANDOL_t$	$DLCIT_t$	$DLGOS_t$	$DLRANDOL_t$
$DLCIT_{t-1}$	-0.631 (0.142)	-0.021 (0.019)	0.022 (0.045)	-0.778 (0.212)	-0.034 (0.157)	0.048 (0.163)	-0.670 (0.135)	-0.024 (0.021)	0.026 (0.046)
$DLCIT_{t-2}$	-0.283 (0.158)	-0.018 (0.020)	0.016 (0.048)	-0.435 (0.893)	-0.031 (0.661)	0.050 (0.689)	-0.295 (0.666)	-0.015 (0.102)	0.018 (0.227)
$DLCIT_{t-3}$	-0.342 (0.138)	-0.002 (0.018)	0.061 (0.043)	-0.452 (0.556)	-0.012 (0.412)	0.094 (0.429)	-0.336 (0.400)	0.000 (0.061)	0.060 (0.136)
$DLCIT_{t-4}$	0.163 (0.120)	-0.012 (0.016)	0.024 (0.038)	0.078 (0.241)	-0.024 (0.179)	0.048 (0.186)	0.139 (0.146)	-0.014 (0.022)	0.023 (0.050)
$DLGOS_{t-1}$	0.235 (0.659)	0.173 (0.101)	-0.082 (0.234)	0.364 (0.857)	0.219 (0.635)	-0.132 (0.661)	0.189 (0.636)	0.205 (0.098)	-0.095 (0.216)
$DLGOS_{t-2}$	1.716 (0.633)	-0.092 (0.099)	0.213 (0.226)	1.790 (0.558)	-0.094 (0.413)	0.195 (0.430)	1.846 (0.391)	-0.123 (0.060)	0.237 (0.133)
$DLGOS_{t-3}$	0.918 (0.677)	0.130 (0.105)	-0.060 (0.240)	1.181 (0.206)	0.161 (0.152)	-0.088 (0.159)	0.982 (0.122)	0.161 (0.019)	-0.088 (0.041)
$DLGOS_{t-4}$	-0.076 (0.688)	0.717 (0.106)	-0.115 (0.241)	0.165 (0.929)	0.697 (0.688)	-0.173 (0.716)	-0.086 (0.674)	0.629 (0.104)	-0.104 (0.229)
$DLRANDOL_{t-1}$	-0.027 (0.374)	0.112 (0.055)	0.380 (0.144)	0.028 (0.524)	0.145 (0.388)	0.363 (0.404)	-0.040 (0.346)	0.135 (0.053)	0.390 (0.118)
$DLRANDOL_{t-2}$	0.352 (0.372)	-0.015 (0.055)	-0.188 (0.145)	0.436 (0.177)	-0.028 (0.131)	-0.199 (0.137)	0.389 (0.104)	-0.023 (0.016)	-0.173 (0.036)
$DLRANDOL_{t-3}$	-0.173 (0.349)	-0.020 (0.052)	0.156 (0.136)	-0.160 (0.956)	-0.016 (0.708)	0.156 (0.737)	-0.206 (0.653)	-0.019 (0.100)	0.140 (0.222)
$DLRANDOL_{t-4}$	-0.136 (0.344)	-0.031 (0.051)	-0.173 (0.134)	-0.170 (0.515)	-0.040 (0.381)	-0.192 (0.397)	-0.156 (0.322)	-0.030 (0.050)	-0.129 (0.110)
RMSE	2690.400			2874.310			3416.080		

The best CIT *BVAR* model was selected by comparing the RMSE of the out-sampling forecasts accuracy and the smallest is that of ***BVAR*** Minnesota prior with RMSE of 2690.40, compared to BVAR Normal Whishart prior and Sims-Zha prior with RMSE of 2874.31 and 3416.08 respectively.

4.5.3.1 CIT BVAR Forecasts

The best model was used to generate CIT ($BVAR_{Minne}$) quarterly forecasts as represented by Figure 4.9.

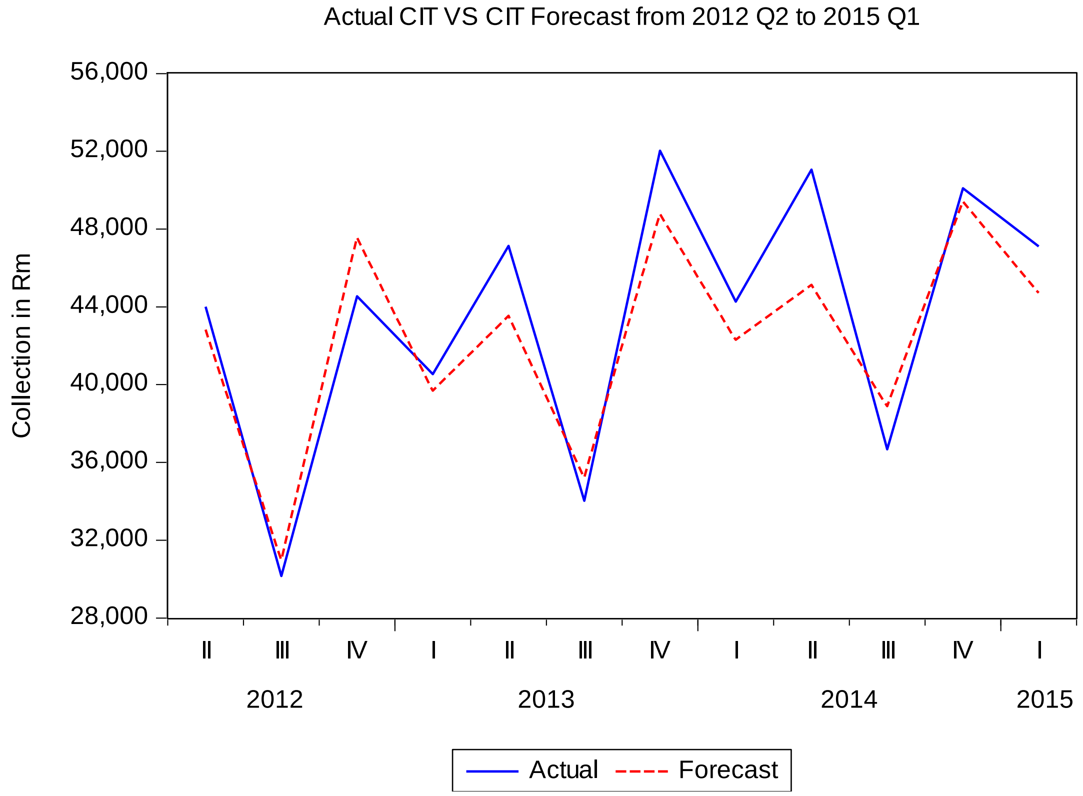


Figure 4.9: The graph showing actual collection and forecasts for h=12 quarters ahead with a CIT BVAR model.

4.5.4 The Corporate Income Tax (CIT) Results discussion

The CIT models were developed based on three techniques selected. Firstly, five different $ARIMA$ models were selected based on AIC and $ARIMA(4,1,1)$ was found to have the lowest AIC of 0.05 as compared to other four $ARIMA$ models in Table 4.2. In terms of forecast accuracy $ARIMA(4,1,1)$ was found to have a minimum RMSE of 3847.81. The results conclude that $ARIMA(4,1,1)$ is the appropriate model to fit CIT data better than other competing $ARIMA$.

Secondly, the best ETS model was found to be of a specification multiplicative error, additive trend, and additive season (M, A, A) with error parameter $\boldsymbol{\alpha = 0.35}$, trend parameter $\boldsymbol{\beta = 0}$ and seasonal parameter $\boldsymbol{\gamma = 0.80}$. The AIC of the (M, A, A) model was 1174.92 smaller than the other four competing models in Table 4.3. The forecasts evaluation was performed for out-of-sample period starting from second quarter of 2012 to first quarter of 2015 and (M, A, A) was found to have a minimum RMSE of 2975.36. These results suggest that $ETS(M, A, A)$ performs better than the other four CIT ETS models in Table 4.3.

The third approach used to project CIT series was Bayesian Vector Autoregression (BVAR) with three different priors. BVAR output does not have AIC information; therefore the best model was selected based on RMSE as it is also the case in the literature. Based on evaluation of forecasts accuracy, BVAR with Minnesota priors $BVAR_{\text{minne}}$ was the appropriate model with RMSE of 2690.40. The results suggest that $BVAR_{\min ne}$ was performing better than Bayesian vector autoregression with normal Wishart prior $(BVAR_{nw})$ and Bayesian vector autoregression with Sims and Zha priors $(BVAR_{sz})$

The final best computing model was selected based on RSME as depicted in Table 4.5. In conclusion $BVAR_{\min ne}$ was superior than selected $ARIMA(4,1,1)$ and $ETS(M, A, A)$ in handling the CIT series well, therefore is an appropriate technique that may be used to forecasts corporate income tax. The Table 4.5 shows the RMSE of the three approaches.

Table 4.5: RMSE for best CIT models

CIT Models	RMSE
$BVAR_{minne}$	**2 690.40**
ETS (M,A,A)	2 975.36
ARIMA (4,1,1)	3 847.81

4.6 PROPOSED MODELS TO FORECAST QUARTERLY PIT

4.6.1 PIT ARIMA Models

4.6.1.1 Testing PIT Series for Stationarity

The graph of PIT in Figure 4.10 indicates that the time series is not stationary, it keeps on shooting up.

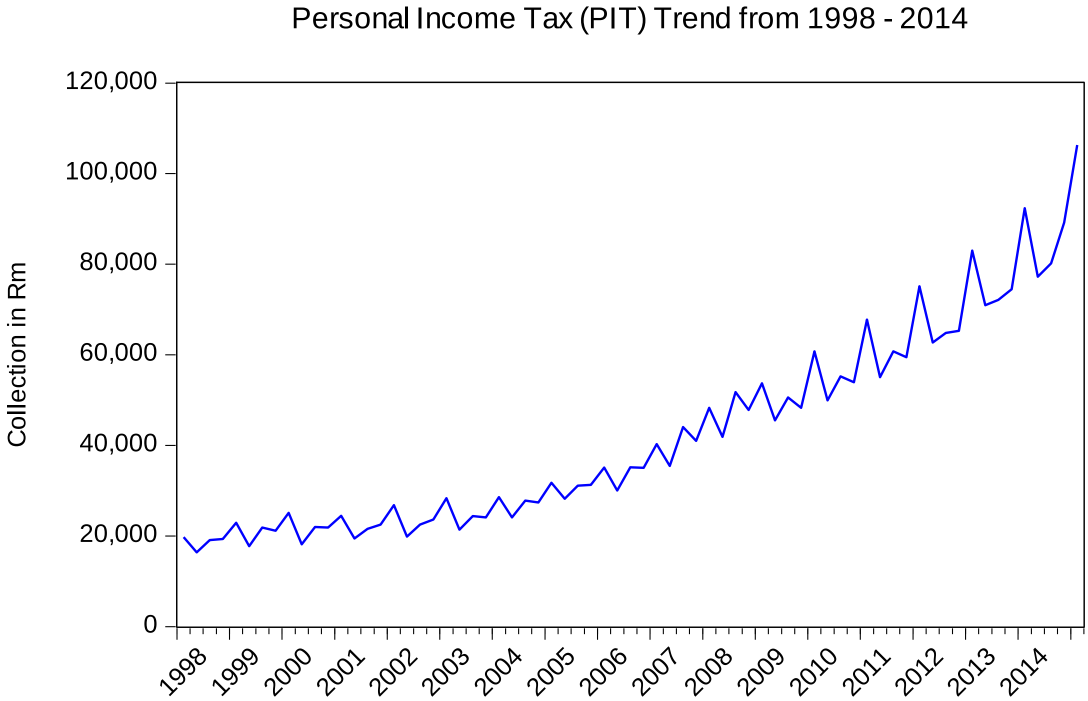


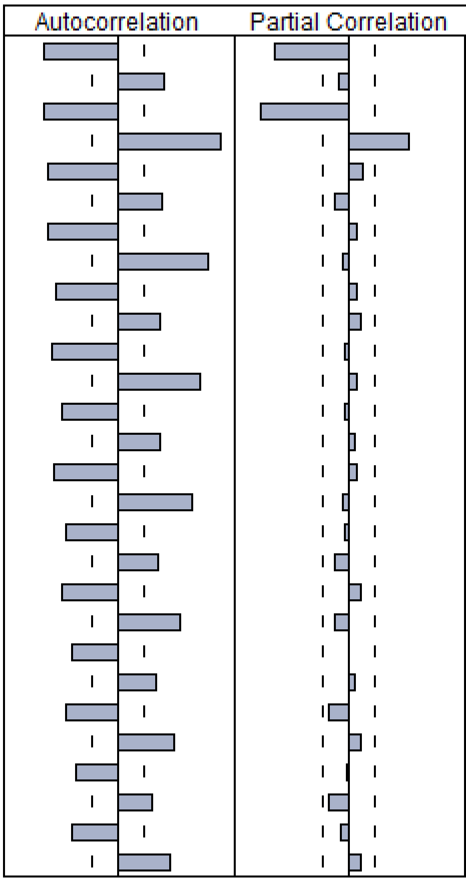
Figure 4.10: The graph showing Personal income tax trend from 1998 – 2014

Examining the PIT correlogram in Appendix C, Figure 4-C1 it is clear that the quarterly data shows some seasonal pattern although is not well distinct. Furthermore, the coefficients of autocorrelation start with a high value and slowly declines suggesting a non-stationary series. The Ljung-Box Q-statistic at the 28th lag has a probability value of 0.000 which is less than 0.05, also this confirm the non-stationarity of PIT data series. Resolving nonstationarity issue the data is made stationary by taking the first difference. The Augmented Dickey Fuller and Phillip Perron tests are used to explore the stationarity. The data became stationary after it was logged and differenced, the p-value of the differenced data is less than the significance level of 0.05, these means that the null hypothesis of nonstationarity is rejected suggesting that the data is stationary at logged first differenced (See Table 4.1).

4.6.1.2 PIT ARIMA Model Identification and Estimation

The graph of stationary series is plotted in Appendix C, Figure 4-C2 and the correlogram of the stationary data (DLPIT) is plotted in Figure 4.11. 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.



	AC	PAC	Q-Stat	Prob
1	-0.675	-0.675	32.391	0.000
2	0.409	-0.086	44.468	0.000
3	-0.676	-0.803	77.889	0.000
4	0.911	0.536	139.59	0.000
5	-0.624	0.128	169.03	0.000
6	0.394	-0.125	180.94	0.000
7	-0.635	0.081	212.38	0.000
8	0.815	-0.065	265.11	0.000
9	-0.551	0.080	289.57	0.000
10	0.379	0.111	301.38	0.000
11	-0.600	-0.028	331.48	0.000
12	0.734	0.072	377.29	0.000
13	-0.502	-0.049	399.13	0.000
14	0.378	0.054	411.75	0.000
15	-0.569	0.072	440.81	0.000
16	0.658	-0.050	480.41	0.000
17	-0.470	-0.046	500.99	0.000
18	0.367	-0.119	513.81	0.000
19	-0.510	0.107	539.05	0.000
20	0.563	-0.132	570.49	0.000
21	-0.415	-0.002	587.96	0.000
22	0.347	0.053	600.43	0.000
23	-0.462	-0.177	622.97	0.000
24	0.494	0.111	649.35	0.000
25	-0.372	-0.012	664.69	0.000
26	0.311	-0.179	675.66	0.000
27	-0.425	-0.073	696.63	0.000
28	0.471	0.113	723.05	0.000

Figure 4.11: Correlogram of PIT (Logarithmic Form, 1st Differenced) (DLPIT)

Examining the Figure 4.11, the ACF shows more spikes and PACF has less spikes. Now the range of models is explored and the best model is selected based on AIC. After identifying the parameters, the automatic $ARIMA$ forecasting was performed using Eviews and the best top five $ARIMA$ models are shown in Table 4.6.

Table 4.6: Top five PIT ARIMA models based on AIC

Model	**LogL**	**AIC***	**BIC**	**HQ**
(4,1,0)	94.9748	-3.1219	-2.9069	-3.0383
(3,1,0)	83.9210	-2.7692	-2.5899	-2.6995
(1,1,0)	45.3359	-1.4855	-1.3779	-1.4437
(2,1,0)	45.5132	-1.4566	-1.3132	-1.4009
(0,1,0)	24.4301	-0.7870	-0.7153	-0.7592

As shown in Table 4.6 the competing models are $ARIMA(4,1,0)$ with AIC of -3.1219; $ARIMA(3,1,0)$ with AIC of -2.7692, $ARIMA(1,1,0)$ with AIC of -1,4855, $ARMA(2,1,0)$ with AIC of -1.4566 and $ARIMA(0,1,0)$ with AIC of -0.7870.

4.6.1.3 PIT ARIMA Model Estimation

The best model selected is $ARMA(4,1,0)$ based on AIC of -3.1219. The model is stationary at first difference $(d = 1)$. The selected model ($ARMA(4,1,0)$) may be presented in a formula form as:

$$\hat{y}_t = \beta_0 + \beta_1 \hat{y}_{t-1} + \beta_2 \hat{y}_{t-2} + \beta_3 \hat{y}_{t-3} - \beta_4 \hat{y}_{t-4} + \varepsilon_t \qquad where \quad y_t = \hat{Z}_t - \hat{Z}_{t-1}$$

or

$$DLPIT_t - DLPIT_{t-1} = 0.0233 - 0.3709.DLPIT_{t-1} - 0.3276.DLPIT_{t-2} - 0.3574DLPIT_{t-3} + 0.5835.DLPIT_{t-4} + \varepsilon_t$$

The results of $ARIMA(4,1,0)$ model are shown in Appendix C, Table 4-C1. The roots of AR and MA characteristics polynomials for the $ARIMA(4,1,0)$ model indicate that the chosen ARIMA model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle (See Appendix C, Figure 4-C3).

4.6.1.4 Diagnostic Checking of the PIT ARIMA Model

Diagnostic checking is performed based on Ljung-Box and Jarque-Bera test to check whether the residuals of the models are not auto-correlated and are normally distributed. The Figure 4-C4 and 4-C5 in Appendix C shows the autocorrelation and normality test of the residuals of the $ARMA(4,1,0)$ model. The Ljung-Box Q-statistics values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected, therefore, therefore is no autocorrelation detected in the residuals series. The normality test confirms that the residuals follow a normal distribution. Since p-value 0.5520 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected. Therefore, the residuals follow a normal distribution.

4.6.1.5 PIT ARIMA Model Forecasts

The selected model $ARMA(4,1,0)$ is used to generate PIT forecasts into the future for 12 quarters ahead, from second quarter of 2012 to first quarter of 2015 as show in Figure 4.12.

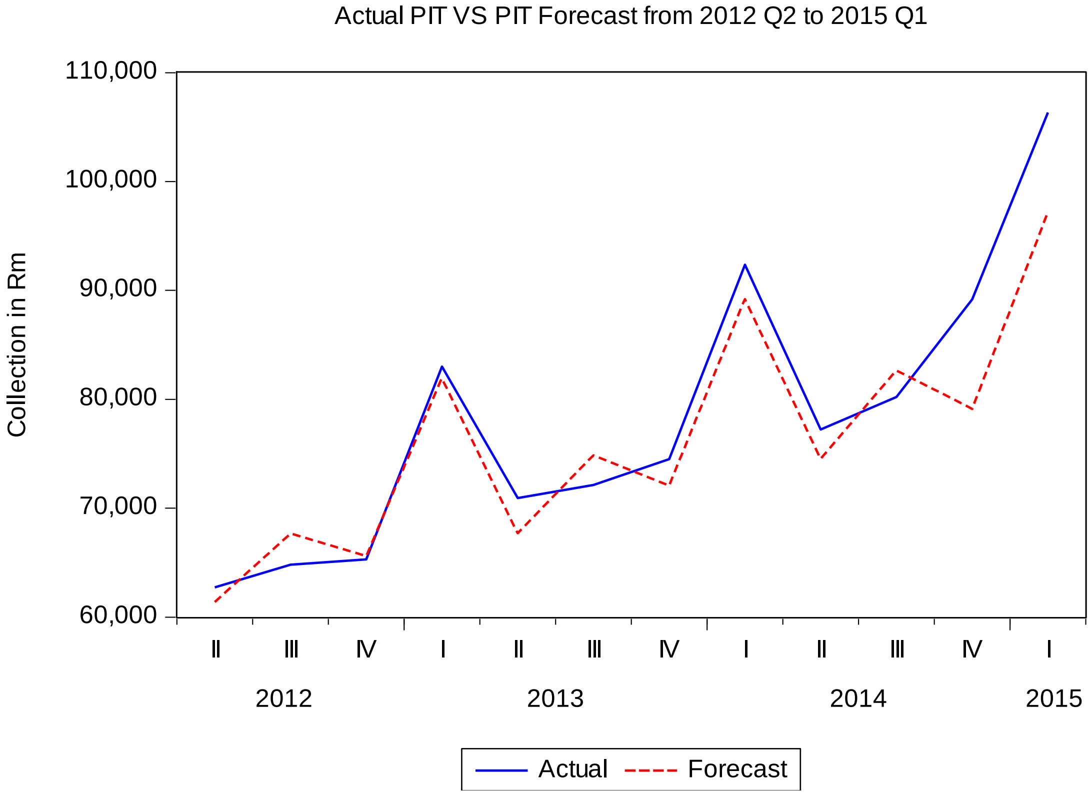


Figure 4.12: The graph showing actual collection and forecasts for h=12 quarters ahead with an $\boldsymbol{ARIMA(4,1,0)}$ *model*

Figure 4.12 shows the plotted representation of the quarterly actual PIT collection in million rand and its forecasts. The results of accuracy measures showed that RMSE is 4509.41, MAE is 3457.69 and MAPE is 4.15. The Theil's U statistic is 0.03.

4.6.2 PIT Error, Trend, Seasonal Models

4.6.2.1 PIT ETS Model Selection

The ETS models for PIT were performed and the best model was chosen based on AIC. Out of the 30 model specifications the additive error, multiplicative trend, additive seasonal (A, M, A) model was selected as the appropriate model and is represented by equations,

$$
\begin{aligned}
y_t &= \ell_{t-1}b_{t-1} + s_{t-m} + \varepsilon_t \\
\ell_t &= \ell_{t-1}b_{t-1} + \alpha\varepsilon_t \\
b_t &= b_{t-1} + \beta\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + \gamma\varepsilon_t
\end{aligned}
\quad \text{or} \quad
\begin{aligned}
y_t &= \ell_{t-1}b_{t-1} + s_{t-m} + \varepsilon_t \\
\ell_t &= \ell_{t-1}b_{t-1} + 0.31\varepsilon_t \\
b_t &= b_{t-1} + 0\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + 0.91\varepsilon_t
\end{aligned}
$$

The selected model has a level smoothing parameter estimate $\alpha = 0.31$, trend parameter $\beta = 0$ (zero indicate that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.91$. The output of the model may be seen in Appendix C, Table 4-C2. The best five ETS model based on AIC is represented by the Table 4.7.

Table 4.7: Top five PIT ETS models based on AIC

Model	**Likelihood**	**AIC***	**BIC**	**HQ**
(A, M, A)	-500.0870	1084.8700	1101.2100	1091.2200
(A, M_D, A)	-500.0870	1086.8700	1105.2600	1094.0200
(A, A, A)	-503.0750	1090.8500	1107.1900	1097.2000
(A, A_D, A)	-503.0750	1092.8500	1111.2300	1099.9900
(A, N, A)	-514.0730	1108.8400	1121.1000	1113.6000

The five models which are competing as shown in table 4.7 are $ETS(A, M, A)$ with AIC of 1084.87, $ETS(A, M_D, A)$ with AIC of 1086.9; $ETS(A, A, A)$ with AIC of 1090.9, $ETS(A, A_D, A)$ with AIC of 1092.9 and $ETS(A, N, A)$ with AIC of 1108.8.

4.6.2.2 PIT ETS Model Forecasts

The $ETS(A, M, A)$ model with the additive error and exhibiting seasonal pattern was used to generate PIT forecasts from second quarter of 2012 to first quarter of 2015 as depicted in the Figure 4.13. The PIT forecast series is closer to the PIT actual series suggesting the better fit of the data.

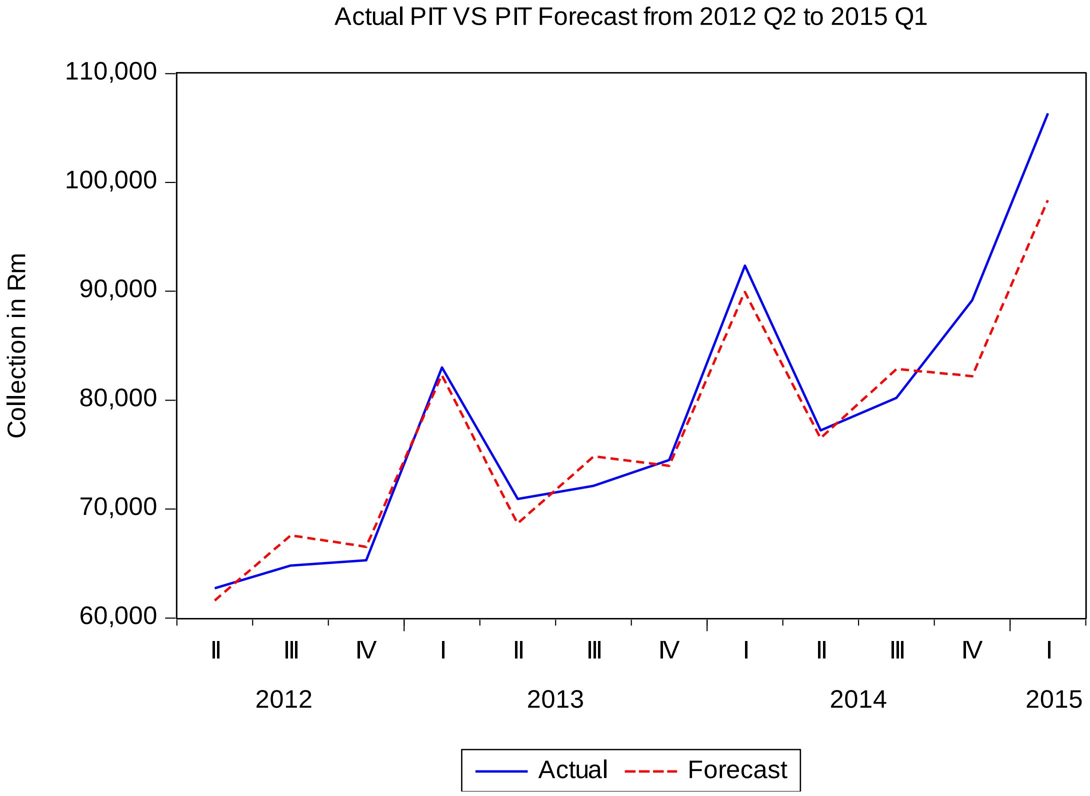


Figure 4.13: The graph showing actual collection and forecasts for h=12 quarters ahead with an $ETS(A, M, A)$ *model*

To assess the forecasting ability of the PIT $ETS(A, M, A)$ model the measures of accuracy of the model were determined which gives the following results, RMSE of 3526.07, MAE of 2667.40, MAPE of 3.21 and Theil U statistics of 0.02.

4.6.3 Bayesian Vector Autoregressive (BVAR) Model for PIT

The PIT BVAR models are estimated using three priors, i.e. Minnesota prior, Normal-Wishart prior and Sims-Zha prior. The Minnesota prior parameters for AR (1) coefficient (mu_1) is set at 0.5, the prior that controls overall tightness of the coefficients (λ_1) is set at 0.9 and the cross-variable weight tightness (λ_2) is set at 0.7 and λ_3 is set at 0.1. The parameters for Normal-Wishart, mu_1 and λ_3 are set to zero and 0.1 respectively. The Sims-Zha parameters, λ_0, γ_1 and λ_3 are set to 0.5; 0.7 and 0.9 respectively while mu_5 and mu_6 are both set at 1. The results of PIT BVARs with three different priors are depicted in the Table 4.8.

Table 4.8: PIT BVAR models results with three priors

	Minnesota Prior			Normal Whishart Prior			Sims-Zha Prior		
	$DLPIT_t$	$DLCoE_t$	$DLEMPL_t$	$DLPIT_t$	$DLCoE_t$	$DLEMPL_t$	$DLPIT_t$	$DLCoE_t$	$DLEMPL_t$
$DLPIT_{t-1}$	-0.453	0.037	-0.044	-0.281	-0.019	0.061	-0.225	0.039	-0.002
	(0.138)	(0.043)	(0.144)	(0.343)	(0.333)	(0.340)	(0.126)	(0.038)	(0.123)
$DLPIT_{t-2}$	-0.447	-0.012	-0.093	-0.173	-0.056	-0.026	-0.235	-0.019	-0.051
	(0.146)	(0.046)	(0.152)	(0.576)	(0.560)	(0.571)	(0.321)	(0.098)	(0.314)
$DLPIT_{t-3}$	-0.328	0.021	-0.414	-0.241	0.127	-0.118	-0.147	0.040	-0.310
	(0.149)	(0.047)	(0.155)	(0.450)	(0.437)	(0.446)	(0.144)	(0.044)	(0.141)
$DLPIT_{t-4}$	0.441	-0.020	0.005	0.573	0.001	0.017	0.574	-0.029	0.085
	(0.139)	(0.044)	(0.145)	(0.348)	(0.338)	(0.345)	(0.131)	(0.040)	(0.128)
$DLCoE_{t-1}$	0.727	0.081	-0.366	0.217	-0.024	-0.083	0.741	0.151	-0.459
	(0.309)	(0.101)	(0.334)	(0.581)	(0.565)	(0.576)	(0.329)	(0.101)	(0.322)
$DLCoE_{t-3}$	0.155	-0.144	0.561	-0.018	-0.134	0.002	-0.131	-0.170	0.475
	(0.320)	(0.105)	(0.346)	(0.447)	(0.434)	(0.443)	(0.139)	(0.042)	(0.136)
$DLCoE_{t-3}$	0.395	0.095	0.094	-0.051	-0.018	0.023	0.266	0.132	0.003
	(0.323)	(0.106)	(0.348)	(0.349)	(0.339)	(0.346)	(0.132)	(0.040)	(0.129)
$DLCoE_{t-4}$	-0.179	0.841	1.074	-0.062	0.176	0.108	-0.324	0.767	0.841
	(0.320)	(0.105)	(0.345)	(0.580)	(0.563)	(0.575)	(0.327)	(0.100)	(0.321)
$DLEMPL_{t-1}$	0.006	-0.052	-0.029	-0.013	-0.007	-0.045	-0.031	-0.052	0.135
	(0.137)	(0.045)	(0.151)	(0.447)	(0.435)	(0.444)	(0.138)	(0.042)	(0.135)
$DLEMPL_{t-2}$	0.102	0.008	0.074	0.042	-0.005	0.046	0.099	0.010	0.068
	(0.132)	(0.043)	(0.145)	(0.345)	(0.335)	(0.342)	(0.126)	(0.038)	(0.123)
$DLEMPL_{t-3}$	-0.129	0.042	0.108	-0.035	-0.009	0.014	-0.122	0.028	0.065
	(0.132)	(0.043)	(0.146)	(0.580)	(0.563)	(0.575)	(0.324)	(0.099)	(0.317)
$DLEMPL_{t-4}$	0.003	0.012	-0.188	-0.001	0.034	-0.034	0.020	0.018	-0.138
	(0.128)	(0.042)	(0.142)	(0.444)	(0.431)	(0.440)	(0.134)	(0.041)	(0.131)
C	0.000	0.000	0.000	0.023	0.023	0.011	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.049)	(0.047)	(0.048)	(0.000)	(0.000)	(0.000)
RMSE	3201.300			5209.404			3656.993		

The best $BVAR$ model is selected based on RMSE and is found to be $BVAR$ with Minnesota priors ($BVAR_{Minne}$) with RMSE of 3201.30. The RMSE of $BVAR$ with Normal Whishart prior and Sims-Zha prior are 5209.40 and 3656.99 respectively.

4.6.3.1 PIT $BVAR_{Minne}$ Forecasts

The best PIT BVAR model was used to generate the forecasts as shown in the Figure 4.14.

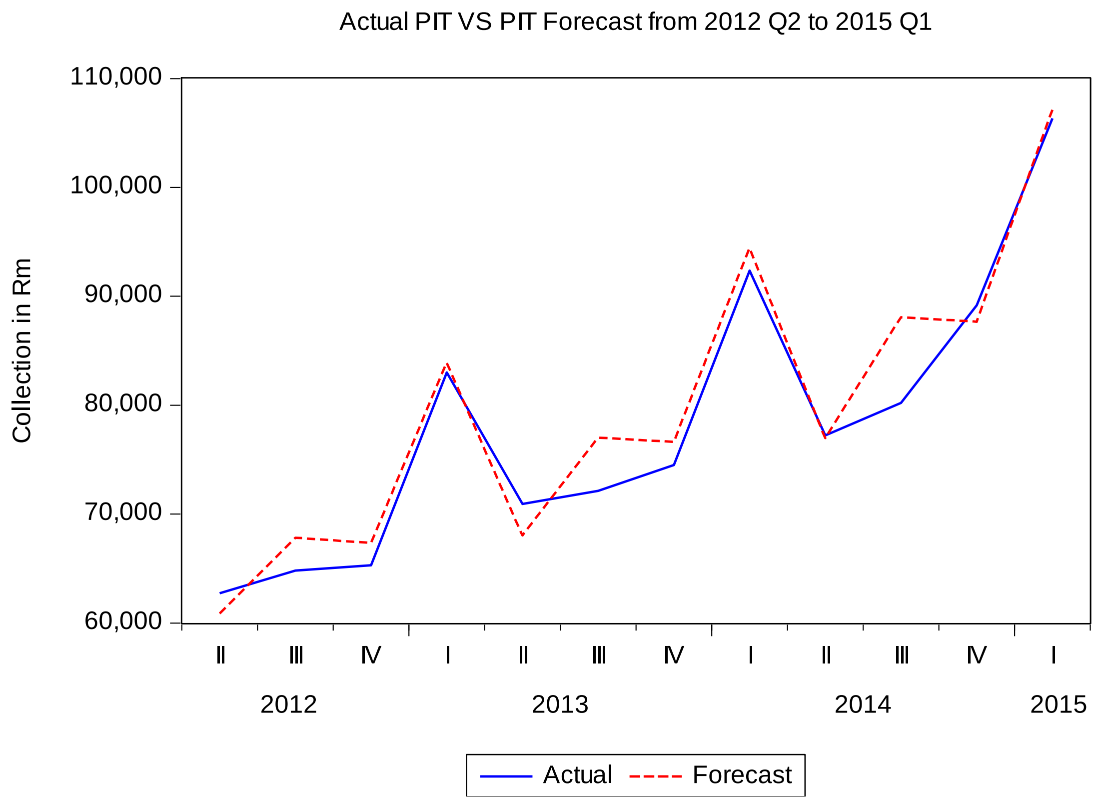


Figure 4.14: The graph showing actual collection and forecasts for h=12 quarters ahead with an $\boldsymbol{BVAR_{\min ne}}$ *model.*

4.6.4 The Personal Income Tax (PIT) Results discussion

The five appropriate PIT $ARIMA$ models selected were $ARIMA(4,1,0)$, $ARIMA(3,1,0)$, $ARIMA(1,1,0)$, $ARIMA(2,1,0)$ and $ARIMA(0,1,0)$. Based on AIC $ARIMA(4,1,0)$ was found to the best model with minimum AIC of -3.1219 as compared to other four $ARIMA$ models. In terms of forecast accuracy $ARIMA(4,1,0)$ was found to have a minimum RMSE of 4509.41. The results conclude that $ARIMA(4,1,0)$ is the appropriate model to fit PIT data better than other competing $ARIMA$ in Table 4.6.

The best ETS model was found to be of a specification additive error, multiplicative trend, and additive season (A, M, A) with error parameter $\boldsymbol{\alpha = 0.31}$, trend parameter $\boldsymbol{\beta = 0}$ and seasonal parameter $\boldsymbol{\gamma = 0.91}$. The AIC of the (A, M, A) model was 1084.87 smaller than the other four competing models. The forecasts evaluation was performed for out-of-sample period starting from second quarter of 2012 to first quarter of 2015 and (A, M, A) was found to have a minimum RMSE of 3526.07. These results show the superiority of $ETS(A, M, A)$ over the other four CIT ETS models in Table 4.7.

The results of PIT Bayesian Vector Autoregression with three different priors are observed. Based on evaluation of forecasts accuracy, $BVAR$ with Minnesota priors $BVAR_{\text{minne}}$ was the appropriate model with RMSE of 3201.30. The results suggest that $BVAR_{\min ne}$ was performing better than Bayesian vector autoregression with normal Wishart prior $(BVAR_{nw})$ and Bayesian vector autoregression with Sims and Zha priors $(BVAR_{sz})$

In conclusion, based on RMSE it was observed that Bayesian VARS with Minnesota priors perform better in comparison with $ARIMA(4,1,0)$ and $ETS(A,M,A)$ models shown in Table 4.9. Therefore $BVAR_{\min ne}$ is an appropriate technique that may be used to forecasts personal income tax. The table 4.9 shows the RMSE of the three approaches.

Table 4.9: RMSE for best PIT models

PIT Models	RMSE
BVAR$_{\text{minne}}$	**3 201.30**
ETS (A,M,A)	3 526.07
ARIMA (4,1,0)	4 509.41

4.7 PROPOSED MODELS TO FORECAST QUARTERLY VATP

4.7.1 VATP ARIMA Models

4.7.1.1 Testing VATP Series for Stationarity

Observing the plotted VATP series in Figure 4.15, it indicates that the time series reveals a nonstationarity trend and the data does not show any sign of seasonality.

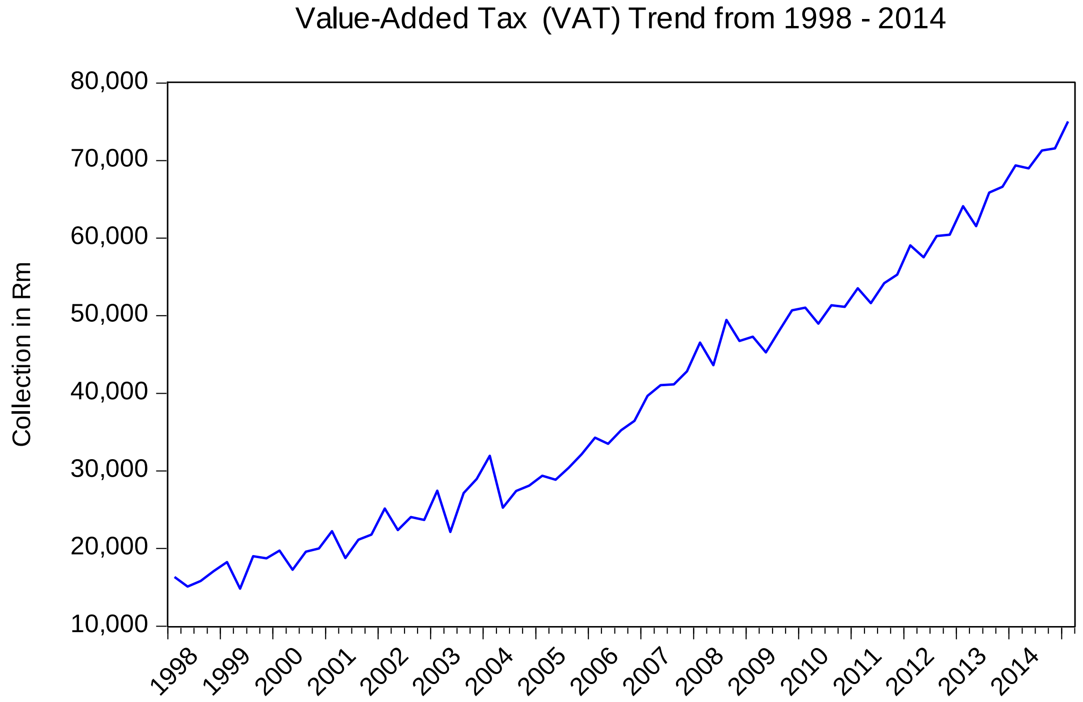
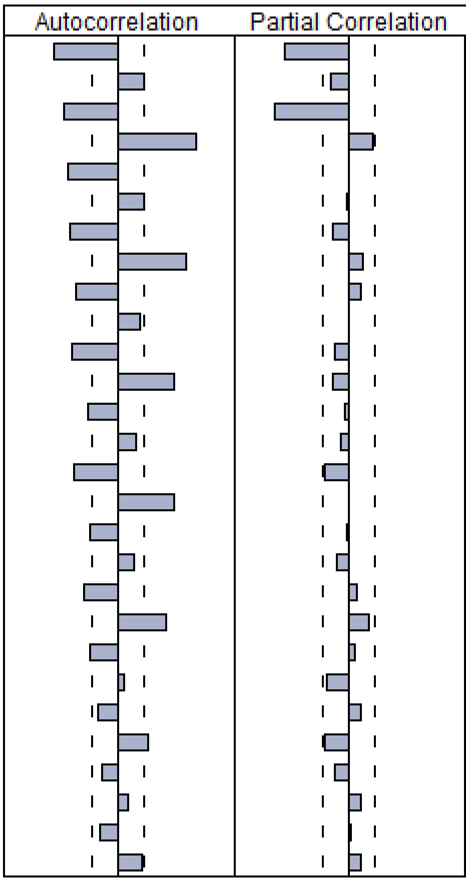


Figure 4.15: The graph showing Value Added-tax (VATP) trend from 1998 – 2014.

The VATP correlogram was plotted with the 28 lags at level (See Appendix D, Figure 4-D1). It can be seen that the coefficients of autocorrelation (ACF) starts with a high value and declines slowly, suggesting the non-stationary series. The non-stationary is further confirmed by the Q-statistic of Ljung-Box (1978) at the 28th lag with a probability value of 0.000 which is smaller than 0.05, so we cannot reject the null hypothesis that the VATP series is non-stationary. Therefore, the series must be log transformed and differenced as shown in Appendix D, Figure 4-D2. The Figure 4.16 depicts VATP correlogram plot.



	AC	PAC	Q-Stat	Prob
1	-0.578	-0.578	23.762	0.000
2	0.229	-0.159	27.528	0.000
3	-0.493	-0.666	45.356	0.000
4	0.707	0.213	82.525	0.000
5	-0.449	0.005	97.785	0.000
6	0.235	-0.022	102.04	0.000
7	-0.440	-0.145	117.12	0.000
8	0.615	0.119	147.09	0.000
9	-0.379	0.106	158.68	0.000
10	0.196	0.002	161.84	0.000
11	-0.416	-0.118	176.28	0.000
12	0.498	-0.155	197.40	0.000
13	-0.272	-0.037	203.81	0.000
14	0.168	-0.074	206.29	0.000
15	-0.406	-0.227	221.08	0.000
16	0.492	-0.007	243.28	0.000
17	-0.252	-0.022	249.22	0.000
18	0.134	-0.098	250.94	0.000
19	-0.313	0.079	260.43	0.000
20	0.437	0.179	279.41	0.000
21	-0.251	0.059	285.81	0.000
22	0.064	-0.194	286.23	0.000
23	-0.174	0.106	289.45	0.000
24	0.271	-0.210	297.42	0.000
25	-0.143	-0.138	299.67	0.000
26	0.089	0.109	300.56	0.000
27	-0.165	0.012	303.73	0.000
28	0.214	0.106	309.19	0.000

Figure 4.16: Correlogram of VATP (Logarithmic Form, 1st Differenced) (DLVAT).

The transformed data is tested for stationarity using ADF test and PP test both of which proved the series to be stationary (See Table 4.1). Therefore, the null hypothesis of that DLVATP series is stationary cannot be rejected.

4.7.1.2 VATP ARIMA Model Identification and Estimation

Various *ARIMA* models were tested to find the best fitting model. The Eviews automatic *ARIMA* tool was used to generate models; five models were selected as presented in the Table 4.10.

Table 4.10: Top five VATP ARIMA models based on AIC.

Model	LogL	AIC*	BIC	HQ
(3,1,0)	81.2777	-2.6764	-2.4972	-2.6068
$(3,1,0)(2,0,0)_4$	82.7016	-2.6562	-2.4053	-2.5587
$(3,1,0)(1,0,0)_4$	81.4052	-2.6458	-2.4307	-2.5622
$(1,1,0)(2,0,0)_4$	80.1939	-2.6384	-2.4592	-2.5687
$(2,1,0)(2,0,0)_4$	81.0432	-2.6331	-2.4180	-2.5495

The models are selected based on minimum AIC. The best model selected is $ARIMA(3,1,0)$ with AIC of -2.6764. The competing models are; $ARIMA(3,1,0)(2,0,0)_4$ with AIC of -2.6562; $ARIMA(3,1,0)(1,0,0)_4$ with AIC of -2.6458; $ARIMA(1,1,0)(2,0,0)_4$ with AIC of -2.6384 and $ARIMA(2,1,0)(2,0,0)_4$ with AIC of -2.6331.

4.7.1.3 VATP ARIMA Best Model

The appropriate best model selected is $ARIMA(3,1,0)$. The model is stationary at first difference $(d = 1)$. The model is formulated as indicated by the equation,

$$\hat{y}_t = \beta_0 + \beta_1 \hat{y}_{t-1} + \beta_2 \hat{y}_{t-2} + \beta_3 \hat{y}_{t-3} + \varepsilon_t \qquad where \quad \hat{y}_t = \hat{Z}_t - \hat{Z}_{t-1}$$

$$or$$

$$DLVATP_t - DLVATP_{t-1} = 0.0232 - 0.7823 DLVATP_{t-1} - 0.6187 DLVATP_{t-2} - 0.6577 DLVATP_{t-3} + \varepsilon_t$$

The results of $ARIMA(3,1,0)$ model are shown in Appendix D, Table 4-D1. The inverse roots of AR characteristics polynomials for the stability of the selected $ARIMA$ model are presented in Appendix D, Figure 4-D3. It is clearly that the chosen $ARIMA$ model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle.

4.7.1.4 Diagnostic Checking of the VATP ARIMA Model

Diagnostic checking was performed to check that the estimated model is statistically sound and acceptable. This is based on some statistical tests which are done to check whether the residuals of the models are not auto-correlated and are normally distributed. The Q-statistics test confirmed that there is no autocorrelation and

Jarque-Bera test (1980) confirmed that the residuals are normally distributed as depicted in Appendix D, Figure 4-D4 and 4-D5 respectively. The Ljung-Box Q-statistics values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected, therefore, there is no autocorrelation detected in the residuals series. The normality test confirms that the residual of the model follows a normal distribution since p-value 0.5078 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected.

4.7.1.5 VATP ARIMA Model Forecasts

The VATP series was projected into the future for 12 quarters ahead, from second quarter of 2012 to first quarter of 2015 using the best selected model $ARIMA(3,1,0)$ as depicted by the Figure 4.17.

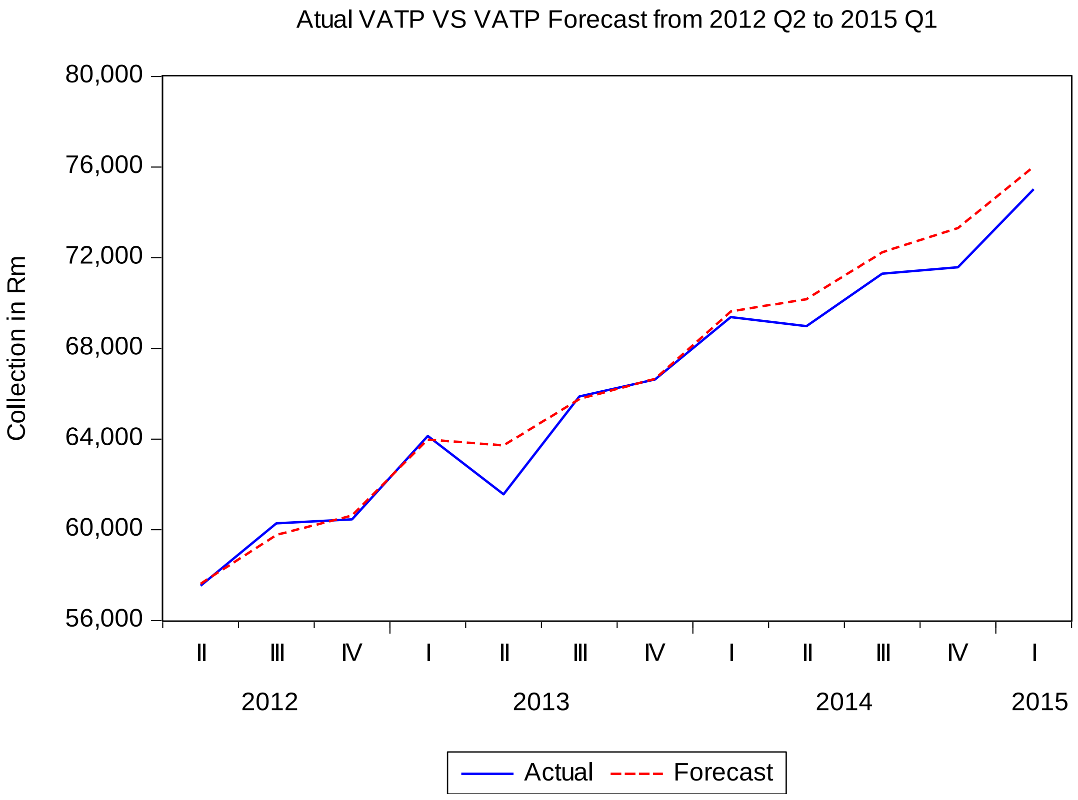


Figure 4.17: The graph showing actual collection and forecasts for h=12 quarters ahead with an $ARIMA(3,1,0)$ *model.*

Figure 4.17 shows the diagrammatic representation of the quarterly actual VATP collection in million rand and its forecasts. The results of accuracy measures showed that RMSE is 972.16, MAE is 692.91 and MAPE is 1.0307. The Theil's U statistic is 0.007.

4.7.2 VATP Error, Trend, Seasonal Models

4.7.2.1 VATP ETS Model Selection

The best ETS model selected for VATP is specified as Multiplicative error, Multiplicative trend, Additive season (M, M, A) model based on minimum AIC. The model is represented by the following equation:

$$
\begin{aligned}
y_t &= (\ell_{t-1}b_{t-1} + s_{t-m})(1+\varepsilon_t) \\
\ell_t &= \ell_{t-1}b_{t-1} + \alpha(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t \\
b_t &= b_{t-1} + \beta(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + \gamma(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t
\end{aligned}
\quad or \quad
\begin{aligned}
y_t &= (\ell_{t-1}b_{t-1} + s_{t-m})(1+\varepsilon_t) \\
\ell_t &= \ell_{t-1}b_{t-1} + 0.45(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t \\
b_t &= b_{t-1} + 0(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + 0(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t
\end{aligned}
$$

The selected model has a level smoothing parameter estimate $\alpha = 0.46$, trend (β) and seasonal (γ) parameters are equal to zero, indicate that the trend components and seasonal components do not change from its starting value. The output of the model may be seen in Appendix D, Table 4-D2. The best five ETS model based on AIC is represented by the Table 4.11.

Table 4.11: Top five ETS VATP models based on AIC.

Model	Likelihood	AIC*	BIC	HQ
(M, M, A)	-493.0450	1070.7800	1087.1300	1077.1400
(M, M_D, A)	-492.8160	1072.3300	1090.7100	1079.4700
(M, A, A)	-494.4290	1073.5500	1089.9000	1079.9000
(M, A_D, A)	-494.4290	1075.5500	1093.9400	1082.7000
(M, N, A)	-505.0860	1090.8700	1103.1300	1095.6300

The five competing models are shown in Table 4.11 and are $ETS(M, M, A)$ with AIC of 1070.78, $ETS(M, M_D, A)$ with AIC of 1072.3; $ETS(M, A, A)$ with AIC of 1073.6; $ETS(M, A_D, A)$ with AIC of 1075.6 and $ETS(M, N, A)$ with AIC of 1090.9.

4.7.2.2 VATP ETS Model Forecasts

The $ETS(M, M, A)$ with multiplicative error was used to generate VATP forecasts series as indicated in the Figure 4.18. The model generally fit the VATP data well as the forecast series is closer to the actual series except for the fourth quarter of 2014 and first quarter of 2015.

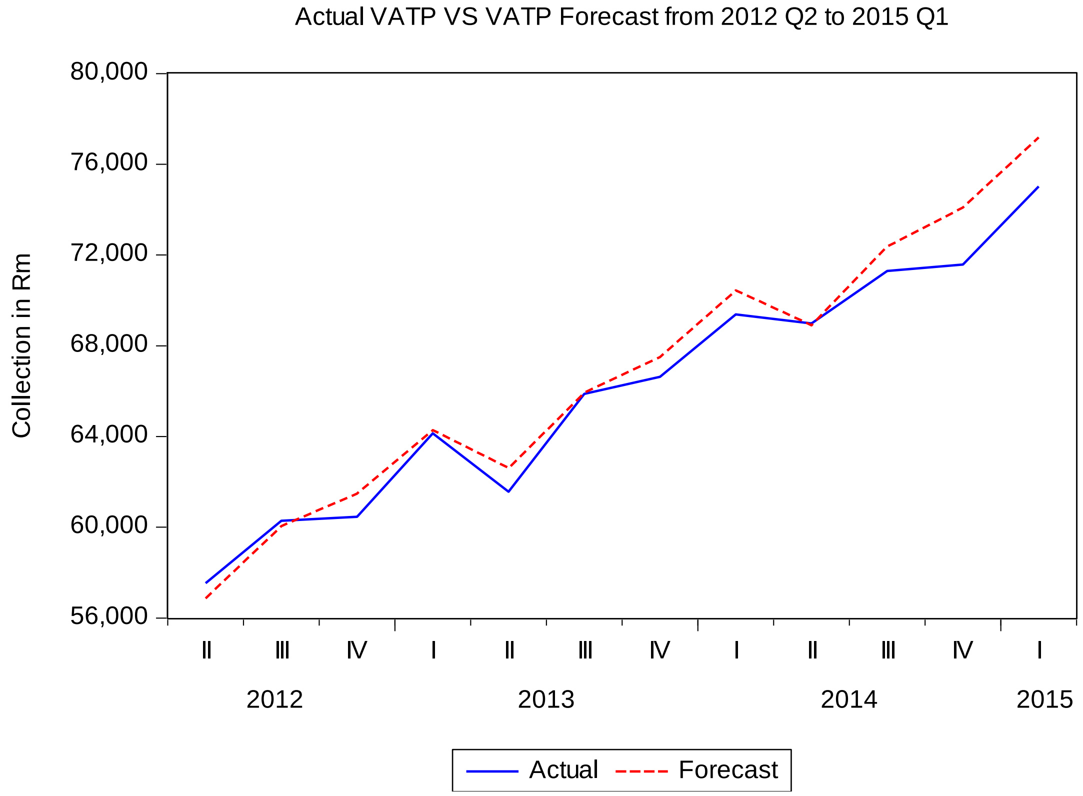


Figure 4.18: The graph showing actual collection and forecasts for h=12 quarters ahead with an $\boldsymbol{ETS\,(M, M, A)}$ *model.*

The closeness of the VATP forecast series to the actual series imply that the selected model has better prediction power and is suitable to forecast VATP series. This is confirmed by the calculated accuracy measures, i.e. RMSE of 1179.92, MAE of 909.30 MAPE of 1.34 and a Theil U statistics of 0.01.

4.7.3 Bayesian Vector Autoregressive (BVAR) Model for VATP

The VATP BVAR models are estimated using three priors, i.e. Minnesota prior, Normal-Wishart prior and Sims-Zha prior. The Minnesota prior parameters, $mu_1, \lambda_1, \lambda_2$ and λ_3 are set to 0; 0.5; 0.5 and 1 respectively. The parameters for Normal-Wishart, mu_1 and λ_3 are set to 0 and 0.1 respectively. The Sims-Zha parameters, λ_0, γ_1 and λ_3 are set to 1; 0.9 and 1 respectively with mu_5 and mu_6 both set to 1. The Table 4.12 depicts the results of BVARs with different priors.

Table 4.12: VATP BVAR models results with three priors.

	Minnesota Prior			Normal Whishart Prior			Sims-Zha Prior		
	$DLVATP_t$	$DLGDE_t$	$DLPCE_t$	$DLVATP_t$	$DLGDE_t$	$DLPCE_t$	$DLVATP_t$	$DLGDE_t$	$DLPCE_t$
$DLVATP_{t-1}$	-0.209 (0.090)	0.027 (0.019)	0.000 (0.010)	-0.310 (0.378)	0.144 (0.367)	-0.006 (0.367)	-0.443 (0.141)	0.190 (0.071)	0.040 (0.044)
$DLVATP_{t-2}$	-0.036 (0.073)	0.000 (0.010)	-0.002 (0.005)	-0.192 (0.520)	0.061 (0.504)	-0.022 (0.504)	-0.359 (0.261)	0.180 (0.131)	0.083 (0.081)
$DLVATP_{t-3}$	-0.106 (0.056)	-0.001 (0.007)	0.002 (0.004)	-0.189 (0.501)	-0.065 (0.485)	0.122 (0.486)	-0.290 (0.300)	-0.065 (0.151)	0.072 (0.093)
$DLVATP_{t-4}$	0.106 (0.045)	-0.001 (0.005)	0.000 (0.003)	0.284 (0.386)	-0.042 (0.374)	0.020 (0.374)	0.198 (0.140)	-0.073 (0.070)	0.050 (0.043)
$DLGDE_{t-1}$	0.055 (0.075)	-0.182 (0.096)	-0.008 (0.020)	0.144 (0.523)	-0.188 (0.507)	-0.044 (0.508)	0.253 (0.272)	-0.427 (0.137)	-0.045 (0.084)
$DLGDE_{t-2}$	-0.019 (0.039)	0.069 (0.077)	0.000 (0.011)	-0.135 (0.502)	0.067 (0.487)	-0.047 (0.487)	-0.172 (0.294)	0.093 (0.148)	-0.097 (0.091)
$DLGDE_{t-3}$	0.008 (0.026)	-0.042 (0.057)	-0.001 (0.007)	0.046 (0.384)	-0.035 (0.372)	0.003 (0.373)	-0.005 (0.136)	-0.001 (0.069)	-0.070 (0.042)
$DLGDE_{t-4}$	-0.003 (0.020)	0.024 (0.045)	0.001 (0.005)	-0.003 (0.523)	0.016 (0.507)	0.022 (0.508)	-0.011 (0.244)	-0.007 (0.123)	-0.023 (0.076)
$DLPCE_{t-1}$	0.362 (0.110)	-0.028 (0.054)	-0.391 (0.053)	0.221 (0.504)	0.003 (0.489)	-0.169 (0.489)	0.318 (0.292)	0.128 (0.147)	-0.248 (0.091)
$DLPCE_{t-2}$	-0.082 (0.068)	0.014 (0.034)	-0.240 (0.052)	-0.150 (0.371)	-0.031 (0.359)	-0.065 (0.360)	-0.197 (0.124)	-0.086 (0.062)	-0.221 (0.038)
$DLPCE_{t-3}$	0.003 (0.047)	-0.013 (0.023)	-0.327 (0.044)	-0.056 (0.508)	-0.053 (0.492)	-0.181 (0.493)	-0.127 (0.201)	-0.107 (0.101)	-0.294 (0.062)
$DLPCE_{t-4}$	-0.005 (0.036)	0.005 (0.018)	0.372 (0.040)	-0.028 (0.509)	0.088 (0.494)	0.393 (0.494)	-0.034 (0.303)	0.223 (0.153)	0.728 (0.094)
C	0.020 (0.010)	0.032 (0.006)	0.040 (0.004)	0.031 (0.059)	0.030 (0.057)	0.025 (0.057)	0.043 (0.032)	0.027 (0.016)	0.027 (0.010)
RMSE	645.690			968.658			2485.677		

The best VATP BVAR model is selected based on RMSE and is found to be BVAR of Minnesota priors ($BVAR_{Minne}$) with RMSE of 645.69. BVAR with Normal Whishart prior RMSE is 968.66 while that of BVAR Sims-Zha prior is 2485.68.

4.7.3.1 VATP $BVAR_{Minne}$ Forecasts

The best PIT $BVAR$ model was used to generate the forecasts as shown in the Figure 4.19.

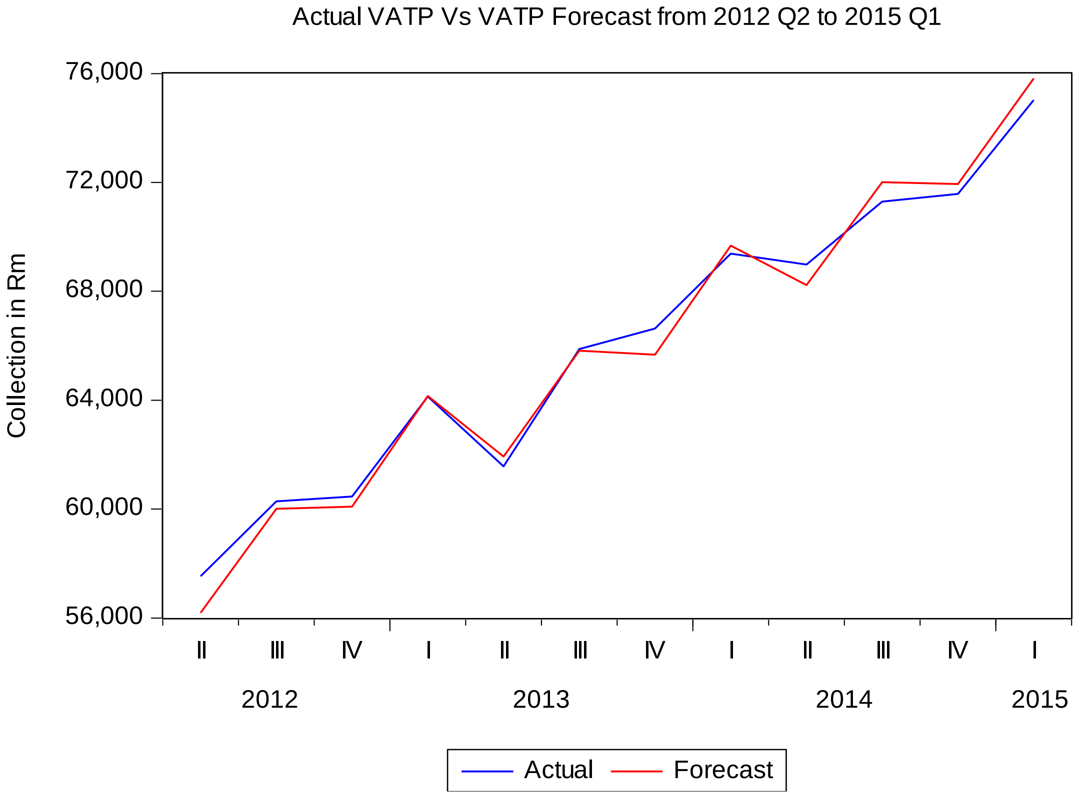


Figure 4.19: The graph showing actual collection and forecasts for h=12 quarters ahead with an VATP BVAR model.

4.7.4 The Value-Added Tax Payments (VATP) Results discussion

Five different VATP $ARIMA$ models were selected based on AIC and $ARIMA(3,1,0)$ was found to have the lowest AIC of -2.6764 as compared to other four $ARIMA$ models in Table 4.10. In terms of forecast accuracy $ARIMA\ (3,1,0)$ was found also to have a minimum RMSE of 972.16. The results suggest that $ARIMA(3,1,0)$ is the appropriate model to fit VATP data better than other competing $ARIMA$.

With regards to ETS models, the best ETS model was found to be of a specification multiplicative error, multiplicative trend, and additive season (M, M, A) with error parameter $\boldsymbol{\alpha = 0.46}$, trend parameter $\boldsymbol{\beta = 0}$ and seasonal parameter $\boldsymbol{\gamma = 0.}$. The AIC of the (M, M, A) model was 1070.78 smaller than the other four competing models in Table 4.11. The forecasts evaluation was performed for out-of-sample period starting from second quarter of 2012 to first quarter of 2015 and (M, M, A) was found to have a minimum RMSE of 1179.92. These results suggest that $ETS\,(M, M, A)$ performs better than the other four VATP ETS models.

The Bayesian Vector Autoregression with three different priors was performed. Based on evaluation of forecasts accuracy, BVAR with Minnesota priors $BVAR_{\text{minne}}$ was the appropriate model with RMSE of 645.69. The results suggest that $BVAR_{\min ne}$ was performing better than Bayesian vector autoregression with normal Wishart prior $(BVAR_{nw})$ and Bayesian vector autoregression with Sims and Zha priors $(BVAR_{sz})$

In conclusion VATP $BVAR_{\min ne}$ was better than selected $ARIMA(3,1,0)$ and $ETS(M,M,A)$ in handling the VATP series well, therefore is an appropriate technique that may be used to forecasts value-added tax. The Table 4.13 shows the RMSE of the three approaches.

Table 4.13: RMSE for best VATP models $BVAR$

VATP Models	RMSE
$BVAR_{minne}$	**645.69**
ARIMA (3,1,0)	972.16
ETS (M,M,A)	1 179.92

4.8 PROPOSED MODELS TO FORECAST QUARTERLY TOTAL TAX REVENUE (TTR)

4.8.1 TTR ARIMA Models

4.8.1.1 Testing TTR Series for Stationarity

The graph of TTR in Figure 4.20 reveals that the time series is non-stationary with an upward trend and with the signs of seasonality. The non-stationarity is confirmed by ADF test and PP test in Table 4.1.

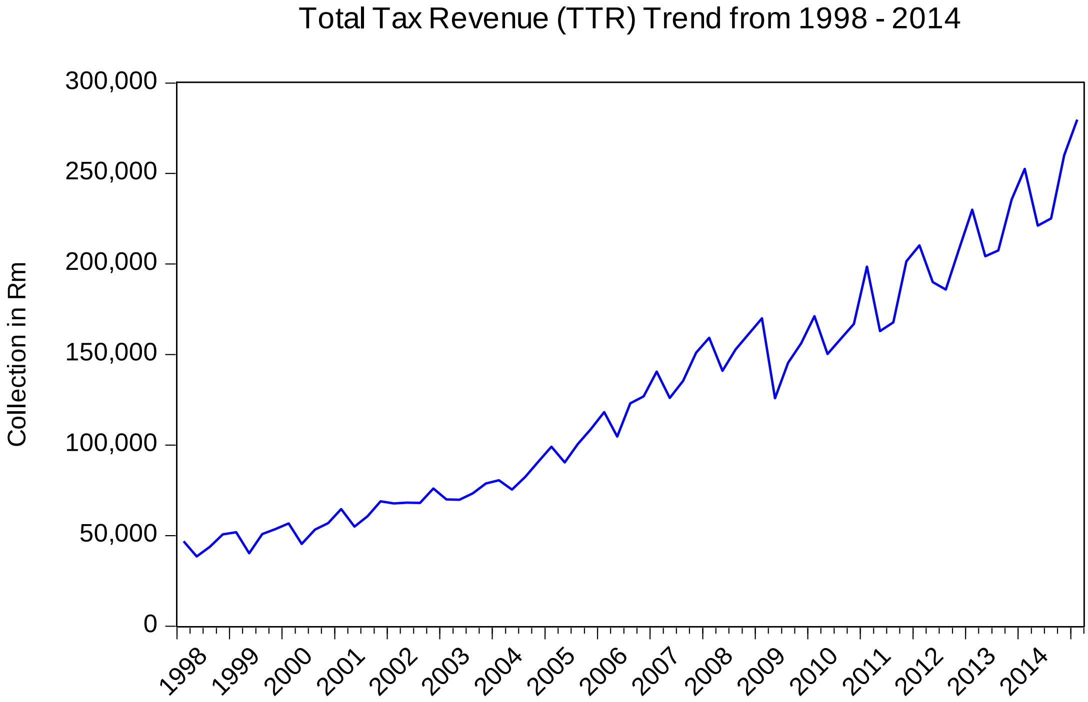
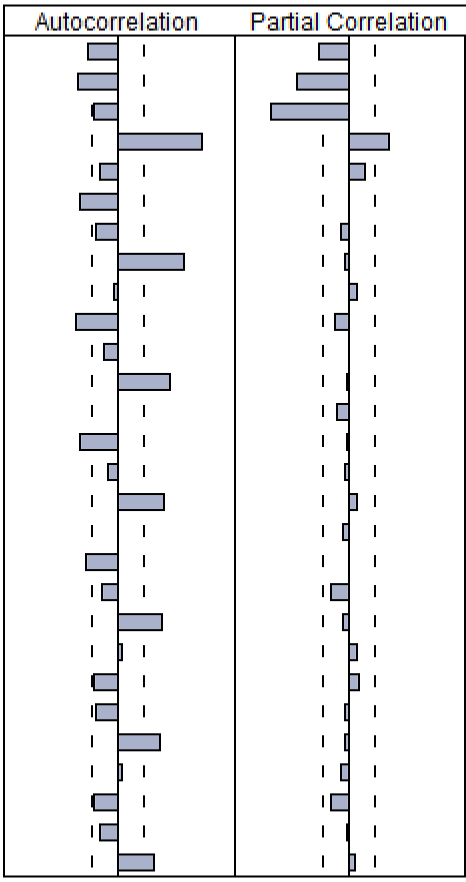


Figure 4.20: The graph showing Total Tax Revenue (TTR) trend from 1998 - 2014

Figure 4-E1 in Appendix E depicts the correlogram of the TTR series with the 28 lags at level. It can be seen that the coefficients of autocorrelation (ACF) starts with a high value and declines slowly, suggesting the non-stationary series. The Q-statistic of Ljung-Box (1978) at the 28th lag has a probability value of 0.000 which is less than 0.05, so we cannot reject the null hypothesis that the TTR series is non-stationary. Therefore, the series is log transformed and first-differenced. The correlogram is shown by the Figure 4.21.



	AC	PAC	Q-Stat	Prob
1	-0.276	-0.276	5.3994	0.020
2	-0.361	-0.473	14.787	0.001
3	-0.224	-0.698	18.466	0.000
4	0.761	0.357	61.590	0.000
5	-0.157	0.142	63.463	0.000
6	-0.350	-0.002	72.868	0.000
7	-0.191	-0.066	75.721	0.000
8	0.598	-0.049	104.05	0.000
9	-0.037	0.074	104.16	0.000
10	-0.384	-0.125	116.29	0.000
11	-0.125	-0.007	117.60	0.000
12	0.469	-0.019	136.33	0.000
13	0.006	-0.113	136.34	0.000
14	-0.343	-0.012	146.69	0.000
15	-0.098	-0.037	147.56	0.000
16	0.403	0.070	162.45	0.000
17	-0.003	-0.064	162.45	0.000
18	-0.280	-0.002	169.94	0.000
19	-0.143	-0.170	171.91	0.000
20	0.394	-0.061	187.27	0.000
21	0.028	0.078	187.35	0.000
22	-0.229	0.090	192.78	0.000
23	-0.197	-0.028	196.86	0.000
24	0.373	-0.032	211.88	0.000
25	0.030	-0.078	211.98	0.000
26	-0.221	-0.168	217.51	0.000
27	-0.164	-0.018	220.65	0.000
28	0.335	0.050	234.03	0.000

Figure 4.21: Correlogram of Total Tax Revenue (Logarithmic Form, 1st Differenced) (DLTTR).

The transformed data is tested for stationarity using ADF test (at 5% and 10%) and PP test both of which proved the series to be stationary (See Table 4.1). Therefore, the null hypothesis of that DLTTR series is stationary cannot be rejected.

4.8.1.2 TTR ARIMA Model Identification and Estimation

Various *ARIMA* models were tested to find the best fitting model. The Eviews automatic *ARIMA* tool was used to generate models, and the five best models are presented in the Table 4.14.

Table 4.14: Top five TTR ARIMA models based on AIC.

Model	LogL	AIC*	BIC	HQ
(4,1,1)	76.0672	-2.6095	-2.5010	-2.5675
(4,1,2)	82.2289	-2.6045	-2.3178	-2.4931
(4,1,0)	78.4038	-2.5405	-2.3254	-2.4569
(3,1,1)	76.2278	-2.4641	-2.2491	-2.3806
(4,0,0)	71.7819	-2.3432	-2.1640	-2.2736

The best model was selected by using minimum AIC. The competing models are $ARIMA(4,1,1)$ with AIC of -2.6095; $ARIMA(4,1,2)$ with AIC of -2.6045; $ARIMA(4,1,0)$ with AIC of -2.5405; $ARIMA(3,1,1)$ with AIC of -2.4641and $ARIMA(4,0,0)$ with AIC of -2.3432.

4.8.1.3 TTR ARIMA Best Model

The best model selected is $ARIMA(4,1,1)$. The model is stationary at first difference $(d=1)$. The model is represented by the following equation:

$$\hat{y}_t = \beta_1\hat{y}_{t-1} + \beta_2\hat{y}_{t-2} + \beta_3\hat{y}_{t-3} + \beta_4\hat{y}_{t-4} + \varepsilon_t - \theta_1\varepsilon_{t-1} \qquad where \qquad \hat{y}_t = \hat{Z}_t - \hat{Z}_{t-1}$$

$$or$$

$$DLTTR_t - DLTTR_{t-1} = 0.8549DLTTR_{t-4} - 0.488.\varepsilon_{t-1} + \varepsilon_t$$

The results of $ARIMA(4,1,1)$ model are shown in Appendix E, Table 4-E1. The inverse roots of AR and MA characteristics polynomials for the stability of the selected $ARIMA$ model are depicted in Appendix E, Figure 4-E3. It is clearly the selected $ARIMA$ model is stable as the inverse roots of the characteristic polynomials are not outside the unit circle.

4.8.1.4 Diagnostic Checking of the TTR ARIMA Model

Diagnostic checking was performed to check that the estimated model is statistically sound and acceptable. This is based on checking whether the residuals of the models are not auto-correlated and are normally distributed. The Q-statistics test (Ljung-Box) confirmed that there is no autocorrelation and Jarque-Bera test (1980) confirmed that the residuals are normally distributed as seen in Figure 4-E4 and 4-E5 in Appendix E respectively. The Ljung-Box Q-statistics values for all the 24 lags are greater than 0.05 meaning that the null hypothesis of no autocorrelation cannot be rejected, therefore, there is no autocorrelation detected in the residuals series. The normality test confirms that the residual of model follows a normal distribution since p-value 0.3354 is greater than 0.05, the null hypothesis for the Jarque-Bera test is not rejected.

4.8.1.5 TTR ARIMA Model Forecasts

The TTR series was projected into the future for 12 quarters ahead, from second quarter of 2012 to first quarter of 2015 using the selected $ARIMA(4,1,1)$ model and forecasts plot is depicted by the Figure .4.22.

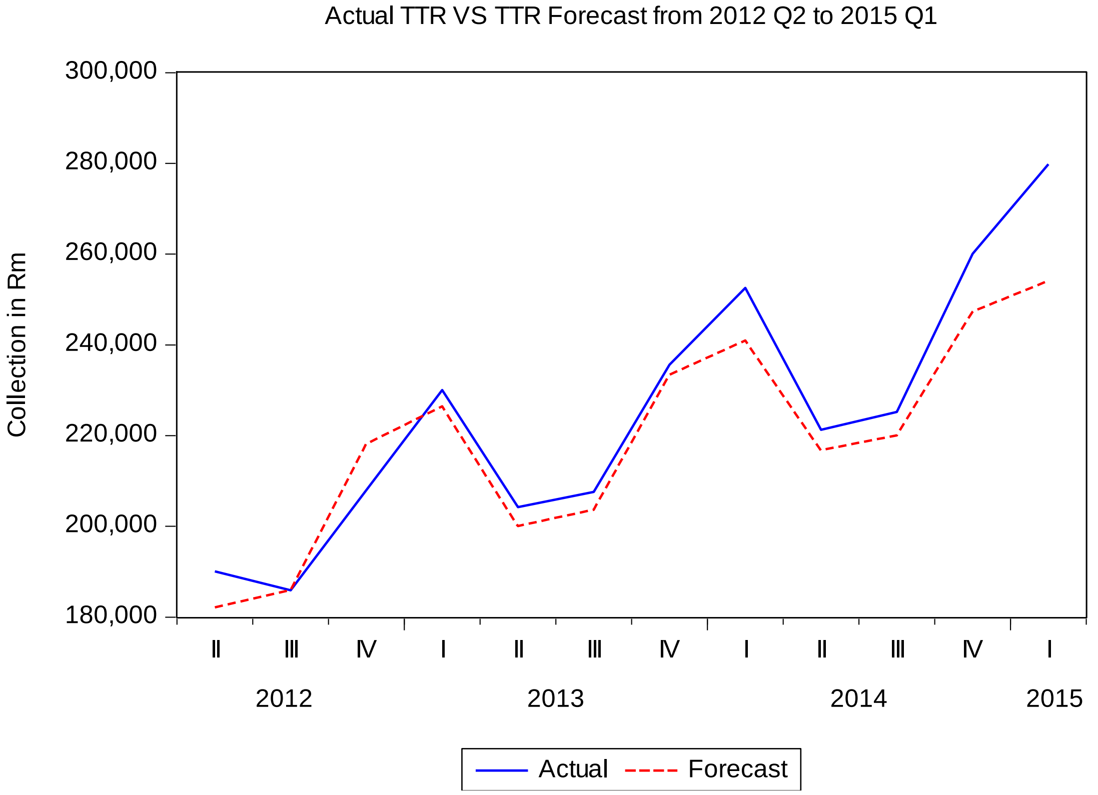


Figure 4.22: The graph showing actual collection and forecasts for h=12 quarters ahead with an $\boldsymbol{ARIMA(4,1,1)}$ *model.*

Figure 4.22 shows the graph representation of the quarterly actual TTR collection in million rand and its forecasts. A time series graph which shows actual TTR values, fitted TTR value and residual values may be seen in Appendix E Figure 4-E6. The results of accuracy measures showed that RMSE is 9999.92, MAE is 7548.62 and MAPE is 3.16 while Theil's U statistic is 0.02

4.8.2 TTR Error, Trend, Seasonal (ETS) Models

4.8.2.1 TTR ETS Model Selection

The ETS models selected out of the 30 model was the Multiplicative error, Multiplicative-dampened trend, Additive season (M, M_d, A) model. The model was selected based on AIC and is given by equations,

$$
\begin{aligned}
y_t &= (\ell_{t-1} b^{\phi}{}_{t-1} + s_{t-m})(1+\varepsilon_t) \\
\ell_t &= \ell_{t-1} b^{\phi}{}_{t-1} + \alpha(\ell_{t-1} b^{\phi}{}_{t-1} + s_{t-m})\varepsilon_t \\
b_t &= b^{\phi}{}_{t-1} + \beta(\ell_{t-1} b^{\phi}{}_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + \gamma(\ell_{t-1} b^{\phi}{}_{t-1} + s_{t-m})\varepsilon_t
\end{aligned}
\quad or \quad
\begin{aligned}
y_t &= (\ell_{t-1} b^{0.99}{}_{t-1} + s_{t-m})(1+\varepsilon_t) \\
\ell_t &= \ell_{t-1} b^{0.99}{}_{t-1} + 0.39(\ell_{t-1} b^{0.99}{}_{t-1} + s_{t-m})\varepsilon_t \\
b_t &= b^{0.99}{}_{t-1} + 0(\ell_{t-1} b^{0.99}{}_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1} \\
s_t &= s_{t-m} + 0.46(\ell_{t-1} b^{0.99}{}_{t-1} + s_{t-m})\varepsilon_t
\end{aligned}
$$

The selected model has a level smoothing parameter estimate $\alpha = 0.37$, trend parameter $\beta = 0$ (zero indicate that the trend components do not change from its starting value) and the seasonal parameter $\gamma = 0.46$. The results of the model are shown in Appendix E, Table 4-E2. The best five ETS model based on AIC is represented by the Table 4.15.

Table 4.15: Top five TTR ETS models based on AIC.

Model	Likelihood	AIC*	BIC	HQ
(M, M_D, A)	-564.4130	1215.5200	1233.9100	1222.6700
(M, M_D, M)	-565.3140	1217.3200	1235.7100	1224.4700
(A, M_D, M)	-573.0910	1232.8800	1251.2600	1240.0200
(A, M_D, A)	-579.4510	1245.6000	1263.9800	1252.7400
(M, M_D, N)	-594.6970	1268.0900	1278.3000	1272.0600

The five models which are competing are shown in Table 4.15 and are $ETS(M, M_D, A)$ with AIC of 1215.52, $ETS(M, M_D, M)$ with AIC of 1217.4; $ETS(A, M_D, M)$ with AIC of 1232.9; $ETS(A, M_D, A)$ with AIC of 1245.6 and $ETS(M, M_D, A)$ with AIC of 1268.1.

4.8.2.2 TTR ETS Model Forecasts

To generate TTR forecasts the selected $ETS(M, M_d, A)$ with multiplicative error, dumped additive trend and additive seasonal was used. The Figure 4.23 shows the generated TTR forecasts which are much closer to the actual TTR series suggesting a better fit and predictive accuracy.

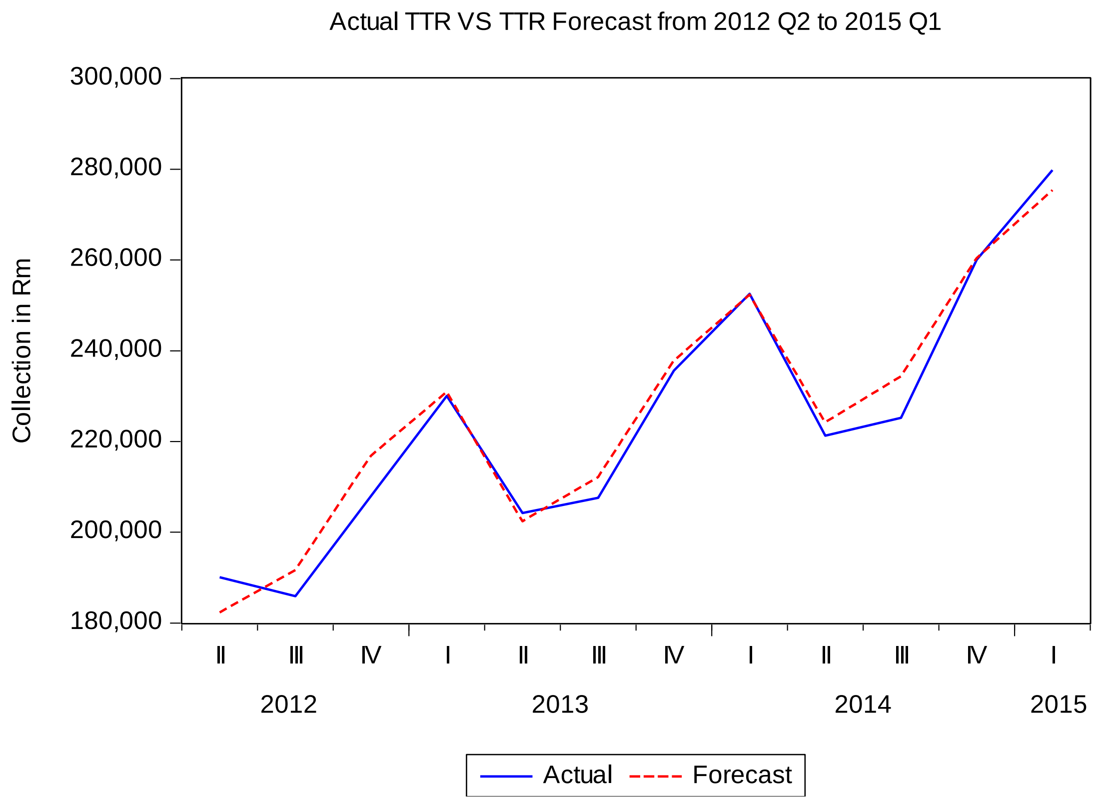


Figure 4.23: The graph showing actual collection and forecasts for h=12 quarters ahead with an $ETS\ (M, M_D, A)$ *model.*

The calculated measures of accuracy confirm the predictive power of the selected model and the results are as follows, RMSE is 4976.50, MAE is 3969.07, MAPE is 1.87 and Theil U statistics is 0.01.

4.8.3 Bayesian Vector Autoregressive (BVAR) Model for TTR

The TTR BVAR models are estimated using three priors, i.e. Minnesota prior, Normal-Wishart prior and Sims-Zha prior. The Minnesota prior parameters for AR (1) coefficient (mu_1) is set at 0, the prior that controls overall tightness of the coefficients (λ_1) is set at 1 and the cross-variable weight tightness (λ_2) is set at 1 and λ_3 is set at 0.5. The parameters for Normal-Wishart, mu_1 and λ_3 are set to zero and 0.01 respectively. The Sims-Zha parameters, λ_0, γ_1 and λ_3 are set to 1; 0.9 and 1 respectively while mu_5 and mu_6 are both set at 1. The results of TTR BVARs with three different priors are depicted in the Table 4.16.

Table 4.16: TTR BVAR models results with three priors

	Minnesota Prior			Normal Whishart Prior			Sims-Zha Prior		
	$DLTTR_t$	$DLGDP_t$	$DLCPI_t$	$DLTTR_t$	$DLGDP_t$	$DLCPI_t$	$DLTTR_t$	$DLGDP_t$	$DLCPI_t$
$DLTTR_{t-1}$	-0.633 (0.144)	-0.020 (0.035)	0.022 (0.035)	-0.564 (0.460)	-0.036 (0.443)	0.027 (0.443)	-0.146 (0.091)	-0.056 (0.022)	0.016 (0.019)
$DLTTR_{t-2}$	-0.404 (0.158)	0.046 (0.038)	0.041 (0.038)	-0.404 (1.484)	0.040 (1.431)	0.050 (1.431)	0.143 (0.516)	0.029 (0.124)	0.024 (0.107)
$DLTTR_{t-3}$	-0.253 (0.166)	0.098 (0.040)	-0.025 (0.040)	-0.288 (1.531)	0.102 (1.476)	-0.013 (1.475)	0.261 (0.615)	0.075 (0.148)	-0.044 (0.128)
$DLTTR_{t-4}$	0.242 (0.152)	-0.017 (0.037)	0.007 (0.037)	0.328 (0.501)	-0.023 (0.483)	0.007 (0.483)	0.656 (0.092)	-0.052 (0.022)	0.003 (0.019)
$DLGDP_{t-1}$	1.057 (0.619)	-0.097 (0.149)	0.013 (0.150)	0.557 (1.429)	-0.012 (1.378)	0.001 (1.377)	0.980 (0.509)	0.379 (0.122)	-0.053 (0.106)
$DLGDP_{t-2}$	0.881 (0.568)	-0.203 (0.137)	0.012 (0.138)	0.486 (1.547)	-0.094 (1.492)	-0.003 (1.492)	0.733 (0.694)	0.166 (0.167)	-0.016 (0.144)
$DLGDP_{t-3}$	0.408 (0.569)	-0.258 (0.137)	0.038 (0.138)	0.195 (0.521)	-0.113 (0.502)	0.021 (0.502)	-0.083 (0.088)	0.123 (0.021)	0.017 (0.018)
$DLGDP_{t-4}$	-1.382 (0.553)	-0.002 (0.134)	0.055 (0.134)	-0.905 (1.429)	0.025 (1.378)	0.031 (1.377)	-1.641 (0.490)	0.321 (0.118)	0.052 (0.102)
$DLCPI_{t-1}$	0.185 (0.655)	0.184 (0.158)	0.294 (0.158)	0.110 (1.495)	0.083 (1.442)	0.130 (1.441)	-0.116 (0.627)	0.047 (0.151)	0.595 (0.131)
$DLCPI_{t-2}$	0.267 (0.684)	0.103 (0.165)	-0.119 (0.165)	0.115 (0.481)	0.049 (0.463)	-0.038 (0.463)	0.164 (0.095)	-0.082 (0.023)	0.067 (0.020)
$DLCPI_{t-3}$	-0.595 (0.634)	0.072 (0.153)	-0.042 (0.153)	-0.209 (1.421)	0.051 (1.370)	-0.042 (1.370)	-0.512 (0.443)	0.004 (0.106)	0.123 (0.092)
$DLCPI_{t-4}$	0.297 (0.614)	0.315 (0.148)	0.000 (0.149)	0.187 (1.487)	0.136 (1.433)	0.000 (1.433)	0.454 (0.529)	0.034 (0.127)	0.207 (0.110)
C	0.027 (0.042)	0.034 (0.010)	0.010 (0.010)	0.043 (0.116)	0.030 (0.112)	0.012 (0.112)	-0.003 (0.022)	0.002 (0.005)	0.001 (0.005)
@ trend	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)	0.000 (0.002)	0.000 (0.002)	0.000 (0.002)	0.000 (0.001)	0.000 (0.000)	0.000 (0.000)
RMSE	5738.990			6105.545			5951.652		

The best TTR $BVAR$ model is selected based on RMSE and is found to be $BVAR$ of Minnesota priors ($BVAR_{Minne}$) with RMSE of 5739.00. The competing models were $BVAR$ with Normal Whishart prior with RMSE of 6105.55 and $BVAR$ with Sims-Zha prior with RMSE of 5951.65.

4.8.3.1 TTR $BVAR_{Minne}$ Forecasts

The best TTR $BVAR$ model was used to generate the forecasts as shown in the Figure 4.24.

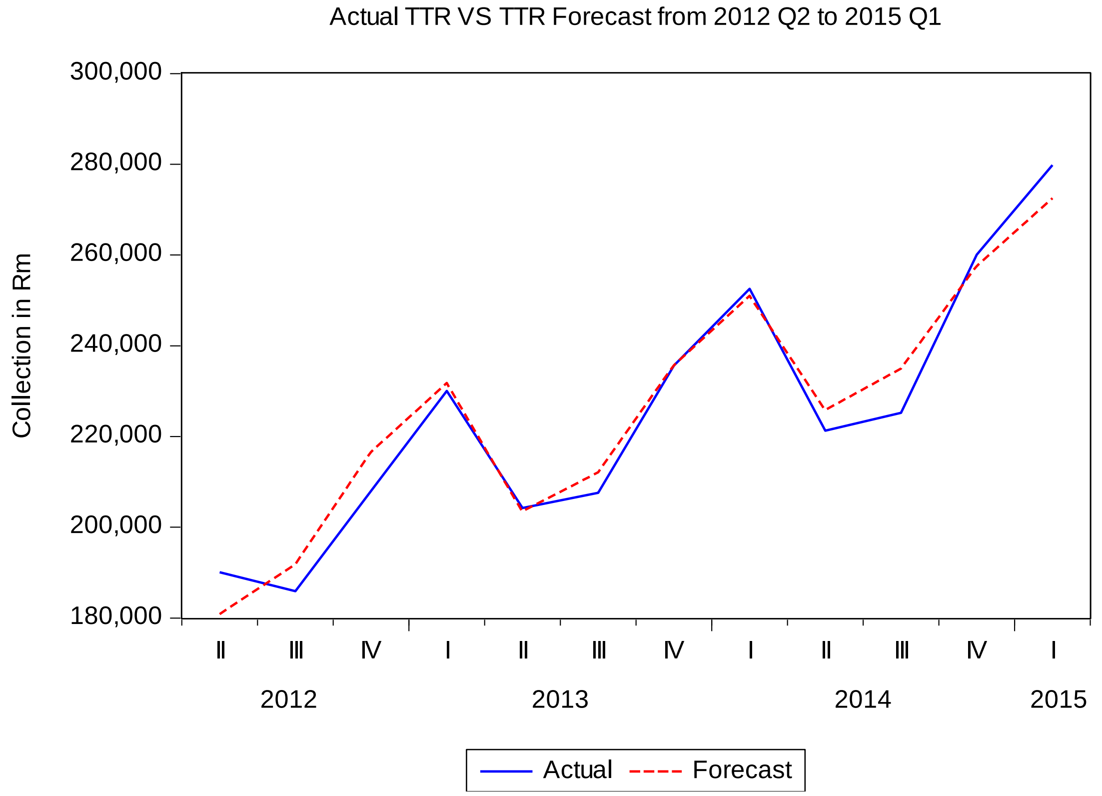


Figure 4.24: The graph showing actual collection and forecasts for h=12 quarters ahead with TTR BVAR model.

4.8.4 The Total Tax Revenue (TTR) Results discussion

The developed TTR $ARIMA$ models were selected based on AIC and $ARIMA(4,1,1)$ was found to have the lowest AIC of -2.6095 as compared to other four $ARIMA$ models in Table 4.14. In terms of forecast accuracy $ARIMA(4,1,1)$ was found also to have a minimum RMSE of 999.92. The results conclude that $ARIMA(4,1,1)$ is the appropriate model to fit total tax revenue data better than other competing $ARIMA$.

The best ETS model for TTR was found to be of a specification multiplicative error, multiplicative-dampened trend, and additive season (M, M_D, A) with error parameter $\alpha = \mathbf{0.37}$, trend parameter $\beta = \mathbf{0}$ and seasonal parameter $\gamma = \mathbf{0.46}$. The AIC of the (M, M_D, A) model was 1115.52 smaller than the other four competing models in Table 4.15. The forecasts evaluation was performed for out-of-sample period starting from second quarter of 2012 to first quarter of 2015 and (M, M_D, A) was found to have a minimum RMSE of 4976.50. These results suggest that $ETS(M, M_D, A)$ performs better than the other four ETS models.

Bayesian Vector Autoregression of TTR with three different priors shows that $BVAR$ with Minnesota priors ($BVAR_{minne}$) was the appropriate model with RMSE of 5738.99. The results suggest that TTR $BVAR_{\min ne}$ was performing better than Bayesian vector autoregression with normal Wishart prior $(BVAR_{nw})$ and Bayesian vector autoregression with Sims and Zha priors $(BVAR_{sz})$

In conclusion TTR $BVAR_{\min ne}$ was not performing better than selected $ARIMA(4,1,1)$ and $ETS(M,M_D,A)$. In this case $ETS(M,M_D,A)$ seems to handle the TTR series well, therefore is an appropriate technique that may be used to total tax revenue. The Table 4.17 shows the RMSE of the three approaches.

Table 4.17: RMSE for best TTR models

TTR Models	**RMSE**
ETS (M,M_D,A)	**4 976.50**
$BVAR_{minne}$	5 738.99
ARIMA (4,1,1)	9 999.92

4.9 DISCUSSION SUMMARY

In selecting the best model, the first step was to choose the appropriate model within specific tax type based on AIC. Five $ARIMA$ models and five ETS models were generated for each tax types. From each five models of specific tax type, one was selected based on minimum AIC. There were forty models generated in total, twenty $ARIMA$ and twenty ETS. Finally, there was eight best model selected, four $ARIMA$ and ETS, meaning $ARIMA$ for CIT, PIT, VATP and TTR, together with ETS for CIT, PIT, VATP and TTR.

In the case of $BVAR$, Models for each tax types were chosen based on minimum root mean squared error of the out-sample forecasts. BVAR output does not have AIC information; therefore the best model was selected based on RMSE and this is in line with the literature. The three priors were used, i.e. Minnesota prior, Normal-Wishart and Sims-Zha, therefore only twelve models were generated. The best four $BVAR$ models were selected for CIT, PIT, VATP and TTR.

Finally, the best selected models for *ARIMA*, ***ETS*** and *BVAR* for each tax types were competing. The models forecasting performance are evaluated by determining the root mean squared error for the out-of-sample forecasts between second quarter of 2012 and first quarter of 2015. The Figure 2.25 shows the out of sample forecasts generated by three techniques (*ARIMA, ETS and BVAR*).

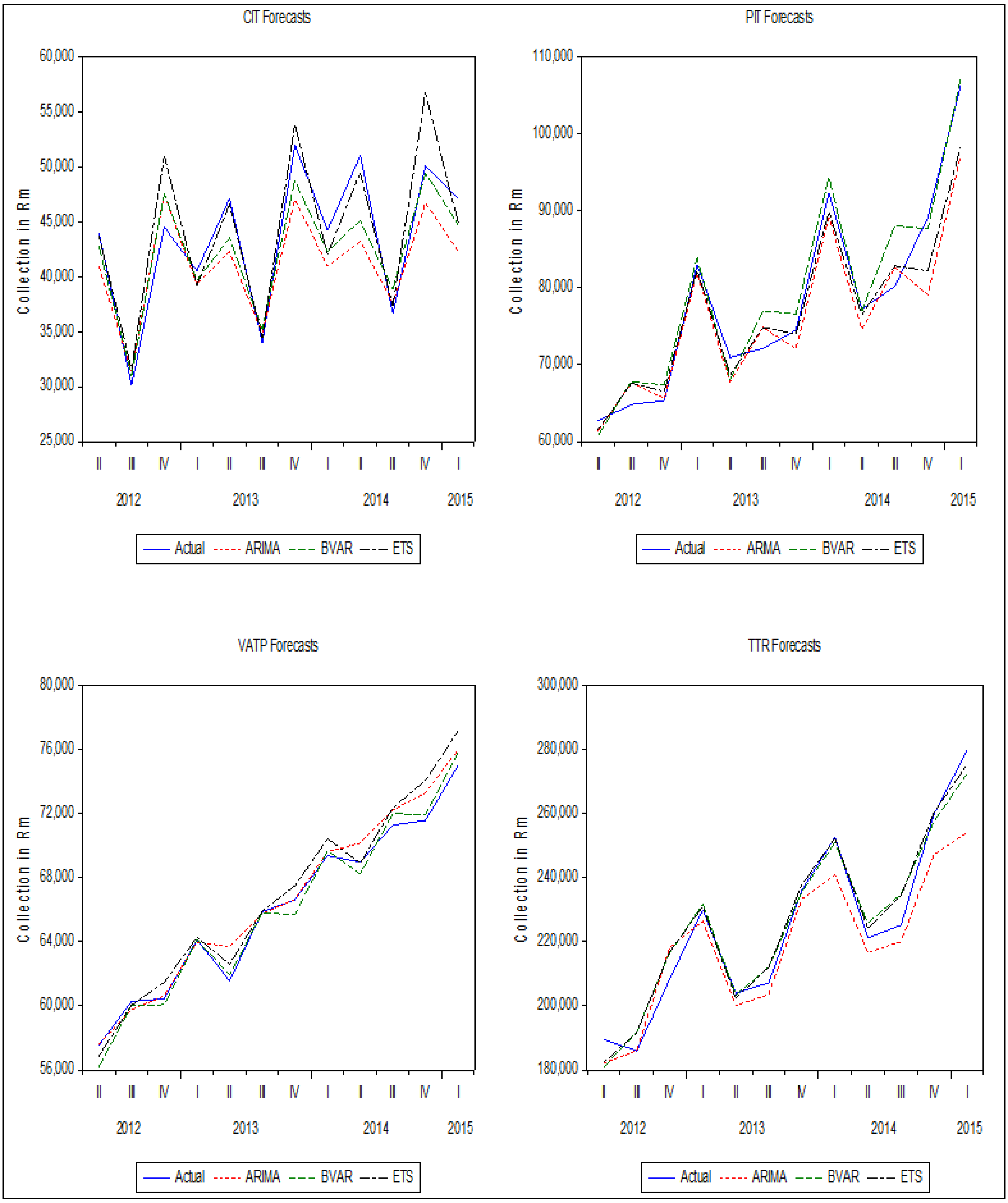


Figure 4.25: The graphs showing forecasts generated by ARIMA, ETS and BVAR method

Bayesian Vector Autoregression with Minnesota priors seems to perform well at individual tax level; it outperforms both *ARIMA* and *ETS* models. In the case of aggregated tax (TTR), *BVAR* fails to outperform both *ARIMA* and ***ETS*** models. The best model for TTR was $ETS(M, M_D, A)$ which also outperforms *ARIMA*.

Comparing the CIT BVAR with ETS and ARIMA, the finding is that, BVAR with Minnesota prior has the lowest RMSE of 2690.40, while ETS and ARIMA have RMSE of 2975.36 and 3847.81 respectively. These suggest that for CIT BVAR is more accurate than the two other techniques. ETS is the second best and performs better than ARIMA with RMSE of 2975.36.

With regards to PIT, BVAR also outperforms ETS and ARIMA with a RMSE of 3201.30. The ETS is the second best with RMSE of 3526.07 compared to ARIMA with RMSE of 4509.41. Based on RMSE, PIT BVAR forecasts are more accurate than ETS and ARIMA forecasts.

In the case of VATP, BVAR is also superior to the ETS and ARIMA with RMSE of 645.69. ETS and ARIMA RMSE are 1179.92 and 972.16, unlike CIT and PIT, ARIMA model is the second best outperforming ETS. VATP BVAR is considered to be the best model to forecasts VATP series.

TTR is a total revenue collection incorporating a number of taxes that contribute approximately 20% including CIT, PIT and VATP contributing around 80%. The model which was best suitable to fit TTR series was ETS with a RMSE of 4976.50, outperforming BVAR and ARIMA with RMSE of 5738.99 and 9999.92 respectively. BVAR is the second best outperforms ARIMA. ETS is the best suited for generating forecasts of TTR. The Table 4.18 shows the RMSE results.

Table 4.18: The RMSE of ARIMA, ETS and BVAR models.

Models	**CIT**	**PIT**	**VATP**	**TTR**
BVAR	**2 690.40**	**3 201.30**	**645.69**	5 738.99
ETS	2 975.36	3 526.07	1 179.92	**4 976.50**
ARIMA	3 847.81	4 509.41	972.16	9 999.92

The reason associated with *BVAR* not performing against *ETS* model may be associated to the misspecification of TTR model, total revenue collection includes various tax components (may be influenced by various variables) including CIT, PIT and VATP.

The results of this study are comparable to the results in the literature although different hyperparameters were used. The BVAR performs better in many instances except on TTR. In Krol (2010) the BVAR models performs better than VAR models for Sales Tax Revenue, Corporate Tax Revenue and Total Tax Revenue except for Income Tax Revenue, the only multivariate case where a BVAR did not produce the lowest RMSE. Krol (2010) follows the literature in specifying the parameter values. The results are also in line with the economic studies where BVAR outperforms the other comparative models being the econometric or time series models.

CHAPTER 5

CONCLUSION

This study attempts to forecast revenue collection for Corporate Income tax, Personal Income Tax, Value-Added Tax and Total Tax Revenue by comparing the outcome of three methodologies. The three methodologies are Autoregressive Moving Averages (ARIMA); State space smoothing methods, i.e. Error, Trend, seasonal models (ETS) and Bayesian Vector Autoregression (BVAR). ARIMA and ETS models uses the historical data for specific taxes while BVAR uses selected variables such as Gross Domestic Product (GDP) and Consumer Price Index (CPI) as explanatory variables for total tax revenue, Gross Operating Surplus (GOS) and rand/dollar (randol) exchange rate as explanatory variables for Corporate Income Tax (CIT). The chosen explanatory variables for Value-Added tax are Gross Domestic Expenditure (GDE) and Private Consumption Expenditure (PCE) while Personal income tax is Compensation of Employee (CoE) and Employment (Empl). The variables are chosen based on economic theory and literature. The in-sample period of the data series starts from 1998Q4 to 2012Q1. The out-of-sample is for the period 2012Q2 to 2015Q1.

Akaike Information Criterion (AIC) is used to compare models derived from same techniques except in the case of BVAR where root mean squared error (RMSE) was used. The root mean squared error was used to compare models from different approaches (ARIMA, ETS and BVAR). BVAR models are estimated using three priors, Minnesota prior, Normal-Wishart prior and Sims-Zha priors. The forecast performances of the selected three techniques are evaluated based on the minimum RMSE.

The results confirm the accuracy of Bayesian Vector Autoregression for predicting tax data. BVAR using Minnesota priors performs better than ARIMA and ETS in all taxes under consideration except for total tax revenue. The total tax revenue was best fitted by ETS models; the ETS models also outperform ARIMA/SARIMA in forecasting total tax revenue.

BVAR models in this dissertation may be improved by selecting more appropriate variables that explains various taxes and also by including more variables. Heidari

(2009) stated that in practice a VAR model with four variables and three lags is more common than a VAR model with four variables and one lag. Forecasting is not an easy task, Gürkaynak; Kisacikoğlu & Rossi (2013) have alluded to the fact that there is no absolute best forecasting methods. To the best of my knowledge BVAR models has not been used for forecasting tax revenue variables of South Africa and this is the unique study that employs BVAR for the purpose of forecasting tax revenue.

This study reveals the accuracy of BVAR models in forecasting tax revenue and it is recommended that BVAR models may be used in forecasting tax revenues of South Africa. For future studies or researches BVAR techniques may be extended to other small taxes in South Africa to investigate whether it will fit these taxes accurately as it is in the case of major taxes.

It was also observed that ETS models are powerful tools in forecasting tax revenues; this study reveals that ETS models in most of the instances are the second best to BVAR and in forecasting aggregate tax collection were superior to BVAR. Also more economic variables which have an impact to tax revenue should be explored in forecasting revenue with BVAR techniques. Therefore, future research should focus on suitability of ETS methods to forecast tax revenue as an alternative approach to the existing tax models. ETS methods are not used adequately (if not at all) in tax revenue forecasting, we have not come across a single study where ETS were used for forecasting tax purposes. The commonly used methodology is simple exponential smoothing.

REFERENCES

Auerbach, A. J. (1999). On the Performance and Use of Government Revenue Forecasts. *National Tax Journal, 52*(4), 765-782.

Australian Fiscal Advisory Council. (2014). Forecasting principle and methods.

Bako, H. Y. (2014). *Forecasting pelagic fish in Malaysia using ETS state space approach.* Doctoral dissertation, Universiti Tun Hussein Onn Malaysia.

Beguin, J. M., Gourieroux, C., & Monfort, A. (1980). *"Identification of a Mixed Autoregressive Moving Average Process: The Corner Method", in O.D. Anderson.* North-Holland: Amsterdam.

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

Brown, R. J. (1959). *Statistical forecasting for inventry control.* New York: McGraw-Hill.

Burnham, K., & Anderson, D. (2002). *Model Selection and Multimodal Inference.* New York (Print).: Springer.

Caraiani, P. (2010). "Forecasting Romanian GDP using a BVAR model". *Romanian journal of economic forecasting, 13*(4), 76-87.

Caraiani, P. (2010). "Forecasting Romanian GDP using a BVAR model". *Romanian journal of economic forecasting, Volume 13*(4), 76-87.

Carreiro, A., & Mumtaz, H. (2012, September). *"Forecatsing scalor variables using Bayesian vector autoregression with constant and drifting coeficients".* Retrieved January 21, 2015, from Webmeets.com: http://www.webmeets.com/files/papers/res/2013/945/FiscForc_vers2.pdf

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

Ciccarelli, M., & Rebucci, A. (2003). *"Measuring contagion with a bayesian time-varying coefficient model".* Working Paper, International Monetary Fund, Research Department, WP/03/171.

Contreras, J., Espínola, R., Nogales, F., & Conejo, A. J. (2000). ARIMA models to predict next-day electricity prices. *IEEE Trans. Power Syst., 18*(3), 1014–1020.

Dadzie, J. E. (2013). *Time series analysis of Value Added Tax revenue collection in Ghana.* College of Science, Department of Mathematics.

Danninger, S. (2005). *Revenue forecasts as performance targets.* Working paper, International Monetary Fund WP/05/14, Fiscal Affairs Department.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association , 74*(366), 427–431.

Doan, T. (2007). RATS User's Manual version 7.0, ESTIMA.

Doan, T., Litterman, R. B., & Sims, C. A. (1984). *"Forecasting and conditional projection using realistic prior distributions".* Econometric Reviews, 3, 1: 1-100.

Doan, T., Litterman, R. B., & Sims, C. A. (1986). Forecasting and Conditional Projection Using Realistic Prior Distribution. Tech. rep.

Enders, W. (2003). *Applied econometric time series* (Third ed.). United States of America: Wiley.

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.

Gamboa, J. (2002). *Development of tax forecasting models: Corporate and individual income taxes.* Philippine Institute for Development Studies.

Gardner Jr, E. S. (1985). Exponential smoothing: The state of the art. *Journal of forecasting, 4*, 1-28.

Giannone, D., Lenza, M., & Primiceri, G. (2012). *Prior selection for vector autoregressions.* CEPR Discussion Paper, 8755.

Golosov, M., & King, J. (2002). *Tax revenue forecasts in IMF supported programs.* Working Paper, International Monetary Fund, Fiscal Affairs Department.

Gray, H., Kelley, G., & McIntire, D. (1978). "A New Approach to ARMA Modelling", Communications in Statistics,. *B7*(1), 1 - 77.

Guizzi, G., Silvestri, C., Romano, E., & Revetria, E. (2015). *A comparison of forecast models to predict weather parameteres.* University of Naples Federico II, Naples.

Gürkaynak R, S., Kisacikoğlu, B., & Rossi, B. (2013). Do DSGE models forecast more accurately out-of-sample than VAR models? VAR models in macroeconomics-New developments and applications: Essays in Honor of Christopher A. Sims. *Advancees in Econometrics., 13*, 27 - 79.

Hannon, A., Leahy, E., & O'Sullivan, R. (2015). *An analysis of tax forecasting errors in Ireland*. Retrieved from www.fiscalcouncil.ie

Heidari, H. (2012). An evaluation of alternative BVAR models for forecasting Iranian inflation. *Iranian Journal of Economic Research, 17*(50), 65 - 81.

Holt, C. C. (1957). *Forecasting seasonal and trends by exponentially weighted moving averages.* Office of Naval Research, Research Memorandum No. 52.

Hurvich, C. M., & Tsai, C. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika, 76*, 297-307.

Huselius, E., & Walled, L. (2014). *Forecasting inflation in Sweden: A univariate approach.* Stockholms Universitet, Department of Statistics, Nr 2014: 11, Stockholms.

Hyndman, R. J., & Athanasopoulos, G. (2012). *Forecasting: principles and practice.* [Online]. Available: http://otexts.com/fpp/.

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

Hyndman, R. J., Koehle, A. B., Snyder, R. D., & Grose, S. (2002). *A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods.* International Journal of Forecasting, 18(3), 439 - 454.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2005). Prediction intervals for exponential smoothing using two new classes of state space models. *Journal of Forecasting, 24*, 17–37.

Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008b). *Forecasting with exponential smoothing: The state space approach.* Springer Science & Business Media.

Jarque, C., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters, 6*(3), 255-259.

Jenkins, G. P., & Kuo, C. (1995). *A VAT Revenue Simulation Model for Tax Reform in Developing Countries.* Harvard Institute for International Development.

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

Keene, M., & Thompson, P. (2007). *An analysis of tax revenue forecast errors.* Working Paper 07/02, New Zealand Treasury.

Koirala, T. P. (2011). Government Revenue Forecasting in Nepal. *NBR Economic Review, 4*(2).

Krol, R. (2010). *"Forecasting state tax revenue: A Bayesian vector autoregression approach".* California State University, Department of Economics.

Legeida, N., & Sologoub, D. (2003). *Modeling Value Added Tax (VAT) revenues in a transition economy: case of Ukraine.* Working paper / Institute for Economic Research and Policy Consulting ; 22 .

Litterman, R. B. (1980). *A Bayesian Procedure for Forecasting with Vector Autoregression.* Working paper, Massachusetts Institute of Technology, Department of Economics.

Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregression-Five Years of experience. *Journal of Business & Economic Statistics, 4*(1), 25 - 38.

Ljung, G. M., & Box, G. E. (1978). On a measure of a lack of fit in time series models. *Biometrika, 65*(2), 297-303.

Lu, Y. (2009). *Modeling and forecasting China's GDP data with time series models.* D-level Essay, Statistics Department of Economics and Society, Högskolan Darlana, Sweden.

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.

Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting methods and applications* (Third ed.). United States of America: John Wiley & Sons, Inc.

Mehmood, S., & Ahmad, Z. (2012). Forecasting Pakistan's Exports to SAARC: An Application of Univiriate ARIMA Model. *Journal of Contemporary Issues in Business Research, 1* (3), 41-54.

Meylar, Aidan, Kenny, G., & Quinn, T. (1998). *Forecasting Irish Inflation Using ARIMA Models.* Munich Personal RePEs Archive (MPRA) Paper No 11359 (December).

Ministry of Finance. (2013). Budget speech. Cape Town, South Africa: National Treasury, ISBN: 978-0-621-41456-1 RP: 345/12. February 2013.

Nandi, B. K., Chaudhury, M., & Hasan, G. Q. (2014). *Univariate time series forecasting: A study of monthly tax revenue of Bangladesh.* EWUCRT Working Paper No.9. , East West University, Department of economics, Dhaka, Bangladesh.

Nazmi, N., & Leuthold, J. H. (1985). *Forecasting State Income Tax Receipts: A Time Series Approach.* College of Commerce and Business Administration, University of Illinois at Urbana-Champaign, Bureau of Economic and Business Research. Urbana, Ill.

Ord, J. K., Koehler, A. B., & Snyder, R. D. (1997). *Estimation and Prediction for a Class of Dynamic Nonlinear Statistical Models.* Journal of the American Statistical Association, 92, 1621 - 1629.

Pegels, C. C. (1969). Exponential forecasting: Some new variations. *Management Science, 15*(5), 311-315.

Phillips, P. C., & Perron, P. (1988). Testing for unit roots in time series regression. *Biometrika, 75*, 335 - 346.

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.

SARS. (2004). SARS-AR-06-Annual report 2003-2004. Pretoria, Gauteng, South Africa.

SARS. (2012, October). Quarterly bulletin-revenue Planning, Analysis and Reporting. (03).

SARS. (2015). SARS-AR-20-Annual Report 2014-2015. Pretoria, Gauteng, South Africa.

Sevinc, V., & Ergun, G. (2009). Usage of different prior distributions in Bayesion vector autoregressive models. *Hacettepe Journal of Mathematics and Statistics, 38*(1), 85 - 93.

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

Spulbăr, C., Niţoi, M., & Stanciu, C. (2012). Monetary policy analysis in Romania: A Bayesian VAR approach. *African Journal of Business Management, 6*(36), 9957-9968.

Taylor, J. W. (2003). Exponential smoothing with a damped multiplicative trend. *International Journal of Forecasting, 19*, 715-725.

Treasury, S. a. (2015). Tax statistics publication.

Tsay, R., & Tiao, G. (1985). “Use of Canonical Analysis in Time Series Model Identification”, . *Biometrika, Vol. 72,*(No. 2,), 299 -315.

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.

Van Heerden, Y., & Schoeman, N. J. (2010). *An empirical dissemination of the personal income tax regime in South Africa using a Microsimulation Tax Model.* Working Paper 25, University of Pretoria, Department of Economics, Pretoria.

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science , 6*, 324–342.

Yang, D., Sharma, V., Ye, Z., Lim, L. H., Zhao, L., & Aryaputera, A. W. (2015). Forecasting of global horizontal irradiance by exponential smoothing, using decompositions. *Energy, 81*, 111 - 119.

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

Yusof, F., & Kane, I. L. (2012). Modelling monthly rainfall time series using ETS state space and SARIMA models. *International Journal of Current Research, 4*(9), 195 - 200.

Yusof, F., & Kane, I. L. (2012). Modelling monthly rainfall time series using ETS state space and SARIMA models. *International Journal of Current Research, 4*(09), 195 - 200.

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

APPENDIX A: ETS Formulae.

Table 3-A1: Formulae for recursive calculations and point forecasts

Trend	Seasonal		
	N	A	M
N	$\hat{y}_{t+h\|t} = \ell_t$ $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}$	$\hat{y}_{t+h\|t} = \ell_t + s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}) + (1-\gamma)s_{t-m}$	$\hat{y}_{t+h\|t} = \ell_t s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1}$ $s_t = \gamma(y_t / \ell_{t-1}) + (1-\gamma)s_{t-m}$
A	$\hat{y}_{t+h\|t} = \ell_t + hb_t$ $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$	$\hat{y}_{t+h\|t} = \ell_t + hb_t + s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - b_{t-1}) + (1-\gamma)s_{t-m}$	$\hat{y}_{t+h\|t} = (\ell_t + hb_t)s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t /(\ell_{t-1} - b_{t-1})) + (1-\gamma)s_{t-m}$
A_d	$\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t$ $\ell_t = \alpha y_t + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$	$\hat{y}_{t+h\|t} = \ell_t + \phi_h b_t + s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1} - \phi b_{t-1}) + (1-\gamma)s_{t-m}$	$\hat{y}_{t+h\|t} = (\ell_t + \phi_h b_t)s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)(\ell_{t-1} + \phi b_{t-1})$ $b_t = \beta^*(\ell_t - \ell_{t-1}) + (1-\beta^*)\phi b_{t-1}$ $s_t = \gamma(y_t /(\ell_{t-1} - \phi b_{t-1})) + (1-\gamma)s_{t-m}$
M	$\hat{y}_{t+h\|t} = \ell_t b_t^h$ $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$	$\hat{y}_{t+h\|t} = \ell_t b_t^h + s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}) + (1-\gamma)s_{t-m}$	$\hat{y}_{t+h\|t} = \ell_t b_t^h s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1} + b_{t-1}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}$ $s_t = \gamma(y_t /(\ell_{t-1} - b_{t-1})) + (1-\gamma)s_{t-m}$
M_d	$\hat{y}_{t+h\|t} = \ell_t b_{t-1}^{\phi h}$ $\ell_t = \alpha y_t + (1-\alpha)\ell_{t-1}b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$	$\hat{y}_{t+h\|t} = \ell_t b_t^{\phi h} + s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t - s_{t-m}) + (1-\alpha)\ell_{t-1}b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$ $s_t = \gamma(y_t - \ell_{t-1}b_{t-1}^{\phi}) + (1-\gamma)s_{t-m}$	$\hat{y}_{t+h\|t} = \ell_t b_t^{\phi h} s_{t-m+h_m^+}$ $\ell_t = \alpha(y_t / s_{t-m}) + (1-\alpha)\ell_{t-1} + b_{t-1}^{\phi}$ $b_t = \beta^*(\ell_t / \ell_{t-1}) + (1-\beta^*)b_{t-1}^{\phi}$ $s_t = \gamma(y_t /(\ell_{t-1} - b_{t-1}^{\phi})) + (1-\gamma)s_{t-m}$

Source: Hyndman, R. J. and Athanasopoulos, G. Forecasting: principles and practice.

Table 3-A2: Formula for Additive error models

ADDITIVE ERROR MODELS			
Trend		**Seasonal**	
	N	**A**	**M**
N	$y_t = \ell_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$	$y_t = \ell_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \alpha\varepsilon_t / s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t / \ell_{t-1}$
A	$y_t = \ell_{t-1} + b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha\varepsilon_t / s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t / s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t /(\ell_{t-1} + b_{t-1})$
Ad	$y_t = \ell_{t-1} + \phi b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$	$y_t = \ell_{t-1} + \phi b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t$ $b_t = \phi b_{t-1} + \beta\varepsilon_t$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha\varepsilon_t / s_{t-m}$ $b_t = \phi b_{t-1} + \beta\varepsilon_t / s_{t-m}$ $s_t = s_{t-m} + \gamma\varepsilon_t /(\ell_{t-1} + \phi b_{t-1})$
M	$y_t = \ell_{t-1}b_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t / \ell_{t-1}$	$y_t = \ell_{t-1}b_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t$ $b_t = b_{t-1} + \beta\varepsilon_t / \ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}b_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha\varepsilon_t / s_{t-m}$ $b_t = b_{t-1} + \beta\varepsilon_t /(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t /(\ell_{t-1}b_{t-1})$
M$_d$	$y_t = \ell_{t-1}b^{\phi}{}_{t-1} + \varepsilon_t$ $\ell_t = \ell_{t-1}b^{\phi}{}_{t-1} + \alpha\varepsilon_t$ $b_t = b^{\phi}{}_{t-1} + \beta\varepsilon_t / \ell_{t-1}$	$y_t = \ell_{t-1}b^{\phi}{}_{t-1} + s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1}b^{\phi}{}_{t-1} + \alpha\varepsilon_t$ $b_t = b^{\phi}{}_{t-1} + \beta\varepsilon_t / \ell_{t-1}$ $s_t = s_{t-m} + \gamma\varepsilon_t$	$y_t = \ell_{t-1}b^{\phi}{}_{t-1}s_{t-m} + \varepsilon_t$ $\ell_t = \ell_{t-1}b^{\phi}{}_{t-1} + \alpha\varepsilon_t / s_{t-m}$ $b_t = b^{\phi}{}_{t-1} + \beta\varepsilon_t /(s_{t-m}\ell_{t-1})$ $s_t = s_{t-m} + \gamma\varepsilon_t /(\ell_{t-1}b^{\phi}{}_{t-1})$

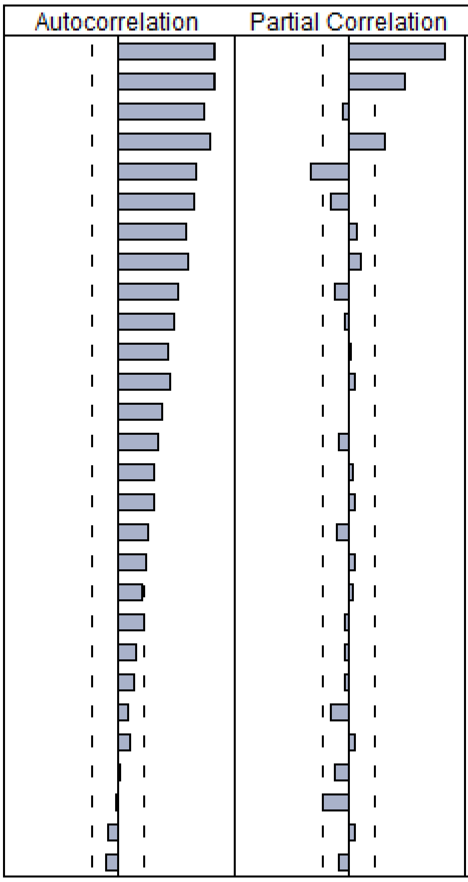
Source: Forecasting: principles and practice, Rob J Hyndman &George Athanasopoulos

Table 3-A3: Formulae for multiplicative error models

MULTIPLICATIVE ERROR MODELS			
Trend		**Seasonal**	
	N	**A**	**M**
N	$y_t = \ell_{t-1}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$	$y_t = (\ell_{t-1} + s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1} + \alpha(\ell_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}s_{t-m}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}(1+\alpha\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A	$y_t = (\ell_{t-1} + b_{t-1})(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1} + s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1} + b_{t-1} + \alpha(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + b_{t-1})s_{t-m}(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1} + b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = b_{t-1} + \beta(\ell_{t-1} + b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
A_d	$y_t = (\ell_{t-1} + \phi b_{t-1})(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1} + s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1} + \phi b_{t-1} + \alpha(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$ $s_t = s_{t-m} + \gamma(\ell_{t-1} + \phi b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = (\ell_{t-1} + \phi b_{t-1})s_{t-m}(1+\varepsilon_t)$ $\ell_t = (\ell_{t-1} + \phi b_{t-1})(1+\alpha\varepsilon_t)$ $b_t = \phi b_{t-1} + \beta(\ell_{t-1} + \phi b_{t-1})\varepsilon_t$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
M	$y_t = \ell_{t-1}b_{t-1}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}(1+\beta\varepsilon_t)$	$y_t = (\ell_{t-1}b_{t-1} + s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b_{t-1} + \alpha(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b_{t-1} + \beta(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1}$ $s_t = s_{t-m} + \gamma(\ell_{t-1}b_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}b_{t-1}s_{t-m}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b_{t-1}(1+\beta\varepsilon_t)$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$
M_d	$y_t = \ell_{t-1}b^{\phi}_{t-1}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b^{\phi}_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b^{\phi}_{t-1}(1+\beta\varepsilon_t)$	$y_t = (\ell_{t-1}b^{\phi}_{t-1} + s_{t-m})(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b^{\phi}_{t-1} + \alpha(\ell_{t-1}b^{\phi}_{t-1} + s_{t-m})\varepsilon_t$ $b_t = b^{\phi}_{t-1} + \beta(\ell_{t-1}b^{\phi}_{t-1} + s_{t-m})\varepsilon_t / \ell_{t-1}$ $s_t = s_{t-m} + \gamma(\ell_{t-1}b^{\phi}_{t-1} + s_{t-m})\varepsilon_t$	$y_t = \ell_{t-1}b^{\phi}_{t-1}s_{t-m}(1+\varepsilon_t)$ $\ell_t = \ell_{t-1}b^{\phi}_{t-1}(1+\alpha\varepsilon_t)$ $b_t = b^{\phi}_{t-1}(1+\beta\varepsilon_t))$ $s_t = s_{t-m}(1+\gamma\varepsilon_t)$

Source: Forecasting: principles and practice, Rob J Hyndman &George Athanasopoulos

APPENDIX B: Corporate Income Tax Tables and Figures



	AC	PAC	Q-Stat	Prob
1	0.854	0.854	52.570	0.000
2	0.866	0.503	107.35	0.000
3	0.784	-0.065	152.96	0.000
4	0.828	0.327	204.63	0.000
5	0.709	-0.340	243.06	0.000
6	0.687	-0.165	279.78	0.000
7	0.616	0.074	309.81	0.000
8	0.637	0.111	342.41	0.000
9	0.540	-0.118	366.18	0.000
10	0.514	-0.049	388.09	0.000
11	0.453	0.019	405.42	0.000
12	0.467	0.046	424.14	0.000
13	0.394	-0.008	437.73	0.000
14	0.361	-0.092	449.33	0.000
15	0.318	0.035	458.51	0.000
16	0.328	0.051	468.44	0.000
17	0.264	-0.102	475.02	0.000
18	0.246	0.062	480.85	0.000
19	0.211	0.026	485.23	0.000
20	0.222	-0.037	490.15	0.000
21	0.168	-0.042	493.03	0.000
22	0.146	-0.046	495.24	0.000
23	0.091	-0.170	496.13	0.000
24	0.100	0.044	497.21	0.000
25	0.024	-0.136	497.27	0.000
26	-0.025	-0.233	497.34	0.000
27	-0.083	0.056	498.15	0.000
28	-0.101	-0.087	499.37	0.000

Figure 4-B1: Correlogram of CIT at level.

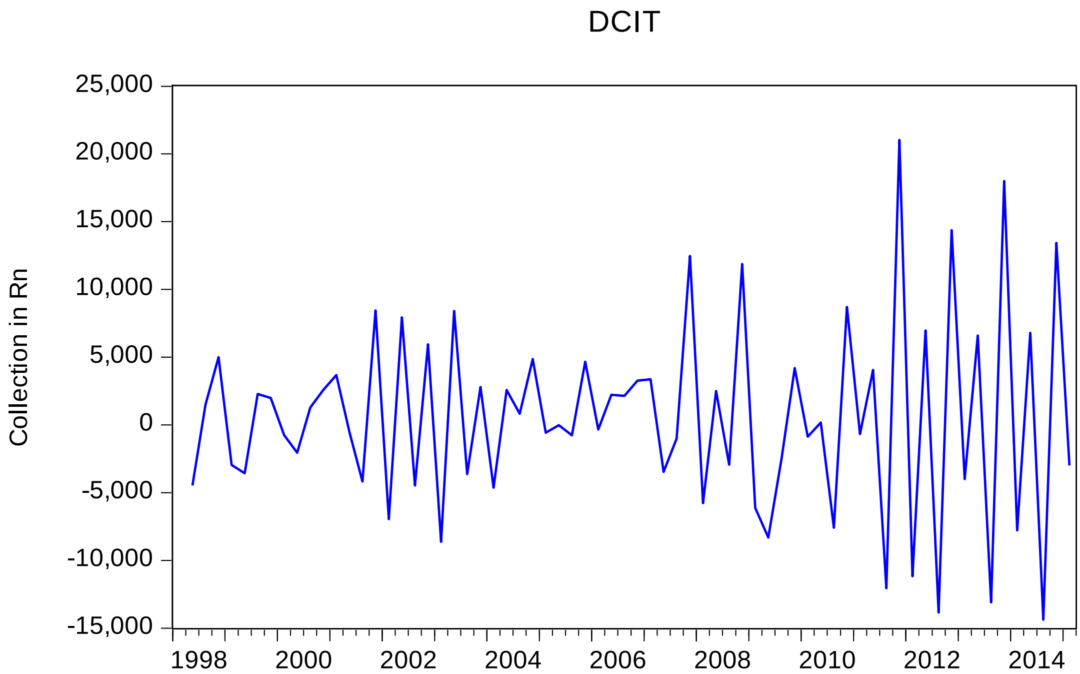


Figure 4-B1: Graph of CIT at first difference (DLCIT).

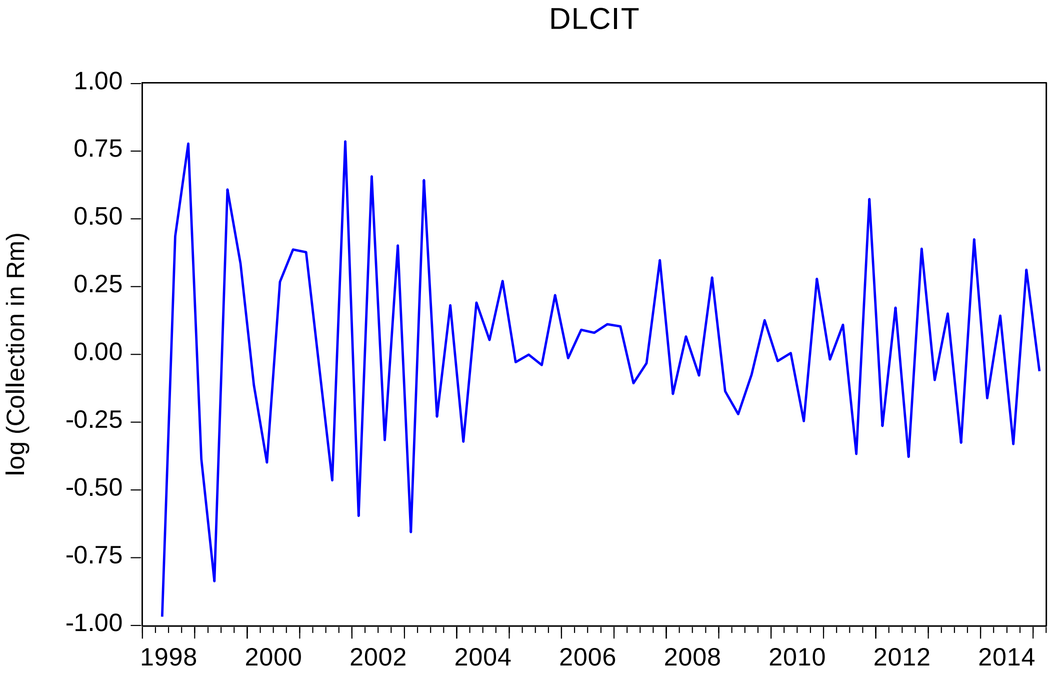


Figure 4-B3: Graph of CIT (Logarithmic Form, 1st Differenced) (DLCIT).

Table 4-B1: Estimation results of **ARIMA (4,1,1)** .

Dependent Variable: DLOG(CIT)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.722793	0.066527	10.86465	0.0000
MA(1)	-0.690059	0.065876	-10.47512	0.0000
SIGMASQ	0.052183	0.007727	6.753020	0.0000
R-squared	0.619863	Mean dependent var		0.029180
Adjusted R-squared	0.605518	S.D. dependent var		0.373857
S.E. of regression	0.234811	Akaike info criterion		0.049958
Sum squared resid	2.922230	Schwarz criterion		0.158459
Log likelihood	1.601175	Hannan-Quinn criter.		0.092024
Durbin-Watson stat	1.932358			
Inverted AR Roots	.92	.00-.92i	-.00+.92i	-.92
Inverted MA Roots	.69			

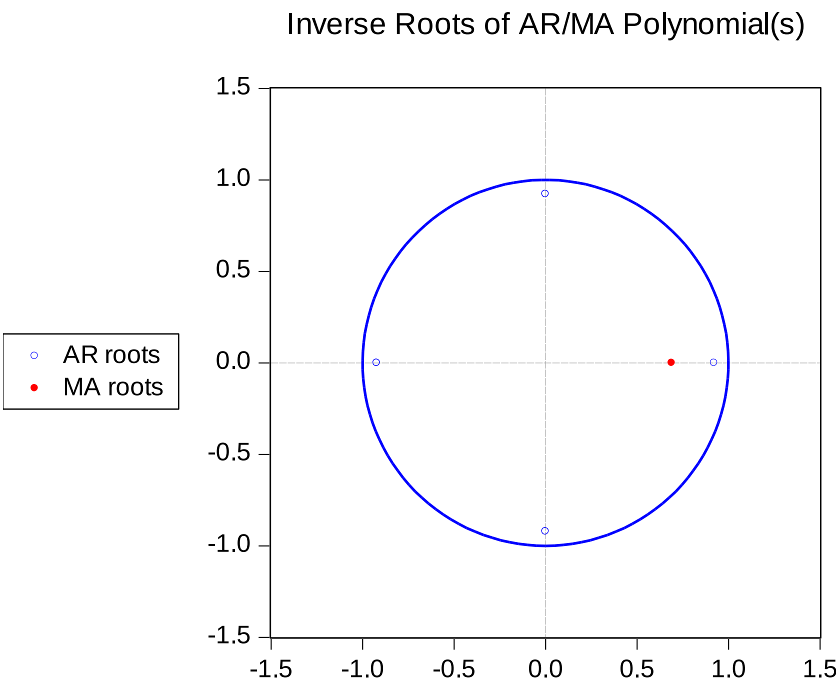
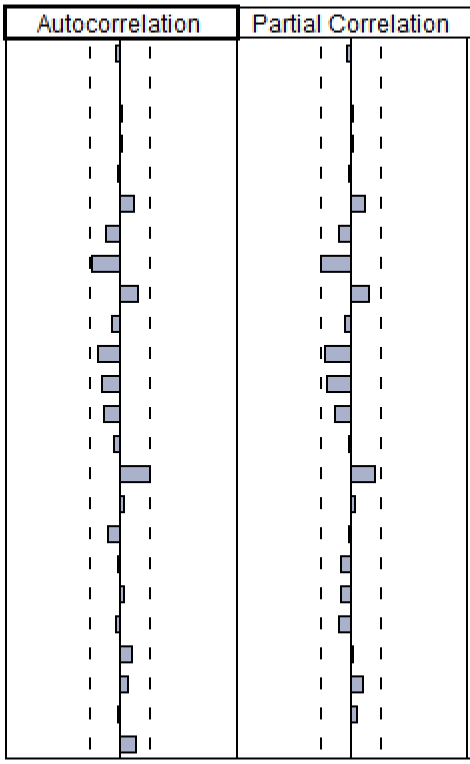


Figure 4-B4: Inverse Roots of AR and MA Process of ***ARIMA*** **(4,1,1)** .



	AC	PAC	Q-Stat	Prob
1	-0.045	-0.045	0.1207	
2	0.004	0.002	0.1216	
3	0.012	0.012	0.1301	0.718
4	0.016	0.017	0.1454	0.930
5	-0.022	-0.020	0.1749	0.982
6	0.128	0.126	1.2367	0.872
7	-0.118	-0.109	2.1632	0.826
8	-0.255	-0.272	6.5651	0.363
9	0.170	0.161	8.5537	0.286
10	-0.067	-0.055	8.8671	0.354
11	-0.202	-0.233	11.818	0.224
12	-0.167	-0.214	13.884	0.178
13	-0.141	-0.155	15.395	0.165
14	-0.061	-0.025	15.685	0.206
15	0.270	0.207	21.449	0.065
16	0.025	0.039	21.502	0.089
17	-0.104	-0.013	22.402	0.098
18	-0.009	-0.096	22.409	0.130
19	0.034	-0.093	22.511	0.166
20	-0.030	-0.098	22.591	0.207
21	0.107	0.010	23.656	0.210
22	0.068	0.111	24.099	0.238
23	-0.016	0.058	24.123	0.287
24	0.155	-0.003	26.577	0.228

Figure 4-B5: Correlogram Residuals of CIT Model ***ARIMA*** **(4,1,1)** .

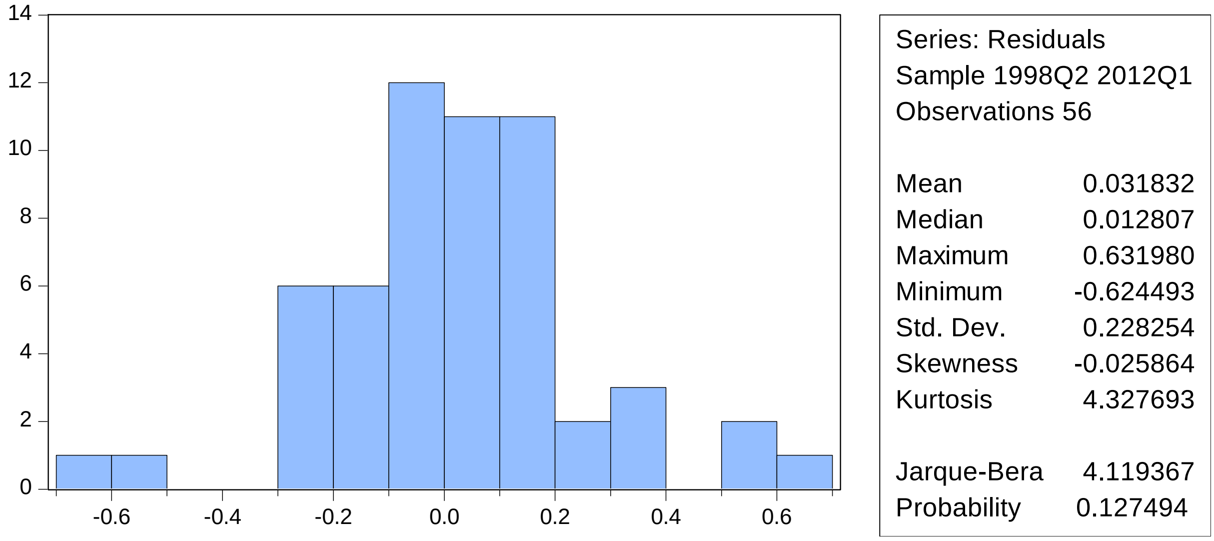


Figure 4-B6: Histogram of the residuals $\boldsymbol{ARINMA}(4,1,1)$.

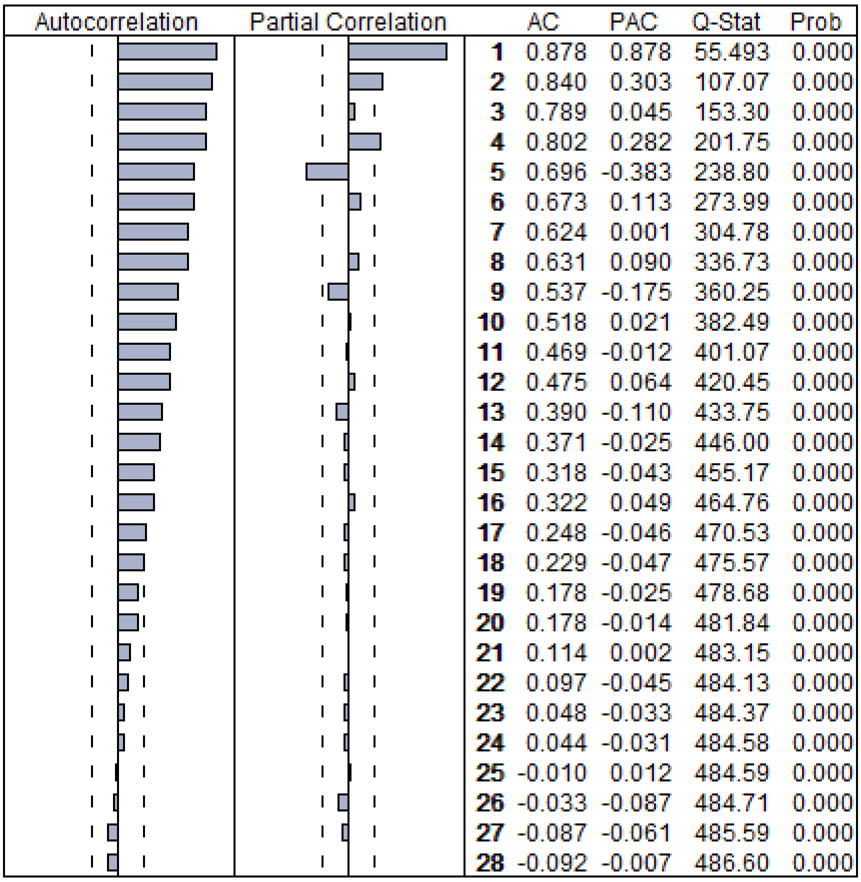
Table 4-B2: Estimation results of CIT $\boldsymbol{ETS}(M, A, A)$.

CIT ETS Smoothing Model: M,A,A - Multiplicative Error, Additive Trend, Additive Season	
Parameters	
Alpha:	0.346961
Beta:	0.000000
Gamma:	0.803269
Initial Parameters	
Initial level:	5185.268
Initial trend:	716.0343
Initial state 1:	5680.516
Initial state 2:	-2872.243
Initial state 3:	-3948.290
Initial state 4:	1140.016
Compact Log-likelihood	-579.4613
Log-likelihood	-545.1138
Akaike Information Criterion	1174.923
Schwarz Criterion	1191.267
Hannan-Quinn Criterion	1181.275
Sum of Squared Residuals	2.202027
Root Mean Squared Error	0.196550
Average Mean Squared Error	17089418

Table 4-B3: Estimation results of CITBVAR.

Bayesian VAR Estimates			
Prior type: Litterman/Minnesota			
Hyper-parameters: Mu: 0.5, λ_1 : 0.5, λ_2 : 0.6, λ_3 : 0.1			
	D(LOG(CIT))	**D(LOG(GOS))**	**D(LOG(RANDOL))**
D(LOG(CIT(-1)))	-0.631 (0.142) [-4.44699]	-0.021 (0.019) [-1.13928]	0.022 (0.045) [0.49381]
D(LOG(CIT(-2)))	-0.283 (0.158) [-1.79144]	-0.018 (0.020) [-0.88288]	0.016 (0.048) [0.32070]
D(LOG(CIT(-3)))	-0.342 (0.138) [-2.48820]	-0.002 (0.018) [-0.10886]	0.061 (0.043) [1.42221]
D(LOG(CIT(-4)))	0.163 (0.120) [1.36232]	-0.012 (0.016) [-0.77908]	0.024 (0.038) [0.62401]
D(LOG(GOS(-1)))	0.235 (0.659) [0.35661]	0.173 (0.101) [1.70995]	-0.082 (0.234) [-0.35217]
D(LOG(GOS(-2)))	1.717 (0.633) [2.71246]	-0.092 (0.099) [-0.93587]	0.213 (0.226) [0.93988]
D(LOG(GOS(-3)))	0.918 (0.677) [1.35556]	0.130 (0.105) [1.23186]	-0.060 (0.240) [-0.24970]
D(LOG(GOS(-4)))	-0.076 (0.688) [-0.11090]	0.717 (0.106) [6.75656]	-0.115 (0.241) [-0.47549]
D(LOG(RANDOL(-1)))	-0.027 (0.374) [-0.07346]	0.112 (0.055) [2.03927]	0.380 (0.144) [2.64192]
D(LOG(RANDOL(-2)))	0.352 (0.372) [0.94796]	-0.015 (0.055) [-0.27027]	-0.188 (0.145) [-1.29315]
D(LOG(RANDOL(-3)))	-0.173 (0.350) [-0.49456]	-0.020 (0.052) [-0.39467]	0.156 (0.136) [1.14809]
D(LOG(RANDOL(-4)))	-0.137 (0.345) [-0.39624]	-0.031 (0.051) [-0.61520]	-0.173 (0.134) [-1.29484]
R-squared	0.705	0.667	0.244
Adj. R-squared	0.623	0.575	0.036
Sum sq. resids	1.712	0.037	0.215
S.E. equation	0.207	0.031	0.073
F-statistic	8.670	7.284	1.171
Mean dependent	0.034	0.028	0.005
S.D. dependent	0.337	0.047	0.075

APPENDIX C: Personal Income Tax Tables and Figures



Lag	AC	PAC	Q-Stat	Prob
1	0.878	0.878	55.493	0.000
2	0.840	0.303	107.07	0.000
3	0.789	0.045	153.30	0.000
4	0.802	0.282	201.75	0.000
5	0.696	-0.383	238.80	0.000
6	0.673	0.113	273.99	0.000
7	0.624	0.001	304.78	0.000
8	0.631	0.090	336.73	0.000
9	0.537	-0.175	360.25	0.000
10	0.518	0.021	382.49	0.000
11	0.469	-0.012	401.07	0.000
12	0.475	0.064	420.45	0.000
13	0.390	-0.110	433.75	0.000
14	0.371	-0.025	446.00	0.000
15	0.318	-0.043	455.17	0.000
16	0.322	0.049	464.76	0.000
17	0.248	-0.046	470.53	0.000
18	0.229	-0.047	475.57	0.000
19	0.178	-0.025	478.68	0.000
20	0.178	-0.014	481.84	0.000
21	0.114	0.002	483.15	0.000
22	0.097	-0.045	484.13	0.000
23	0.048	-0.033	484.37	0.000
24	0.044	-0.031	484.58	0.000
25	-0.010	0.012	484.59	0.000
26	-0.033	-0.087	484.71	0.000
27	-0.087	-0.061	485.59	0.000
28	-0.092	-0.007	486.60	0.000

Figure 4-C1: Correlogram of PIT at level.

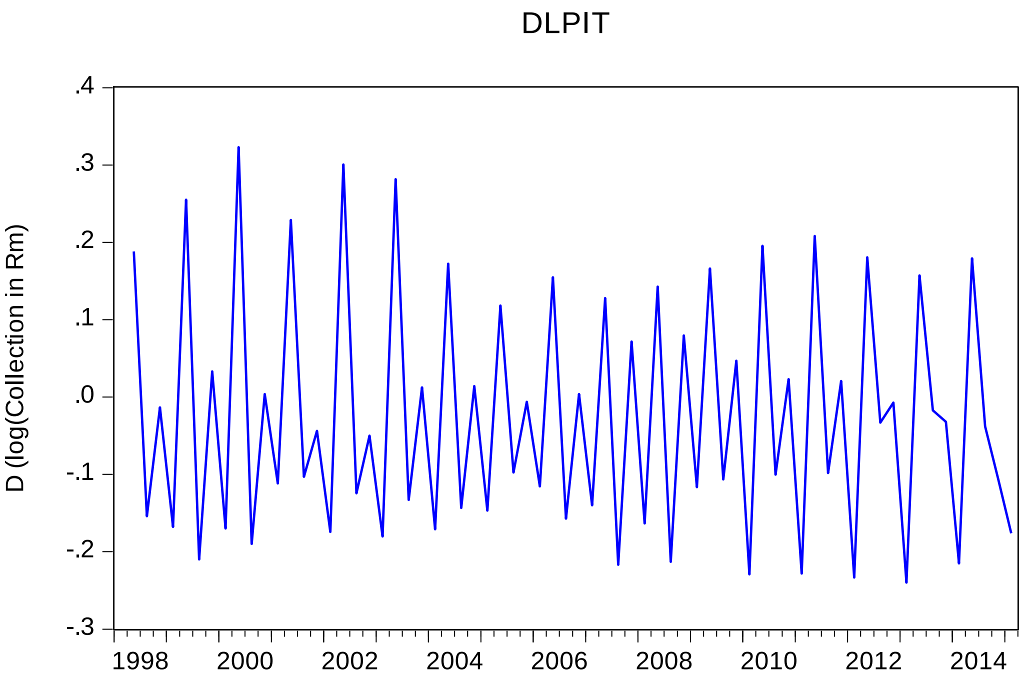


Figure 4-C2: Graph of PIT (Logarithmic Form, 1st Differenced) (DLPIT).

Table 4-C1: Estimation Output of ***ARIMA*** **(4,1,0)** .

Dependent Variable: DLOG(PIT)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.023305	0.004236	5.501938	0.0000
AR(1)	-0.370883	0.139828	-2.652430	0.0107
AR(2)	-0.327574	0.131005	-2.500462	0.0157
AR(3)	-0.357430	0.117929	-3.030891	0.0039
AR(4)	0.583468	0.130358	4.475886	0.0000
R-squared	0.928567	Mean dependent var		0.023840
Adjusted R-squared	0.921423	S.D. dependent var		0.157839
S.E. of regression	0.044245	Akaike info criterion		-3.177672
Sum squared resid	0.097880	Schwarz criterion		-2.960670
Log likelihood	94.97482	Hannan-Quinn criter.		-3.093541
F-statistic	129.9906	Durbin-Watson stat		2.155283
Prob(F-statistic)	0.000000			
Inverted AR Roots	.62	.00-.97i	.00+.97i	-1.00

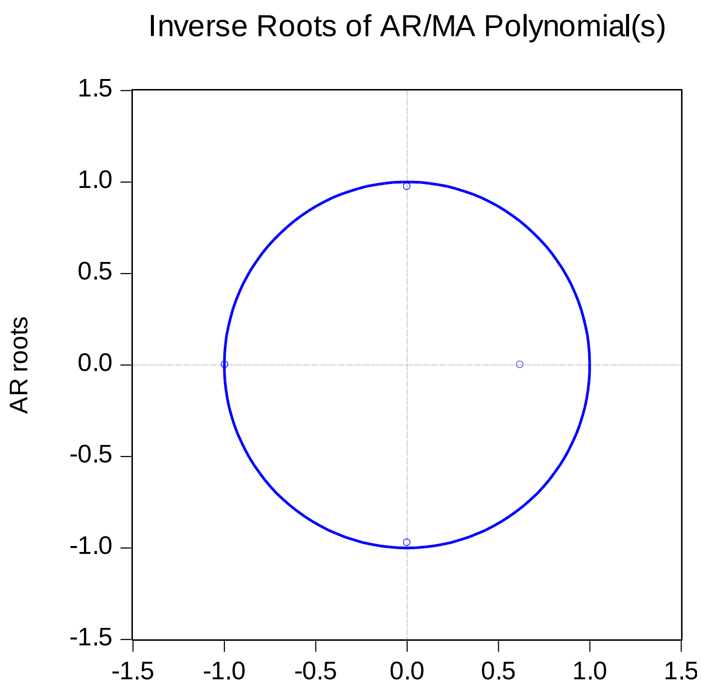
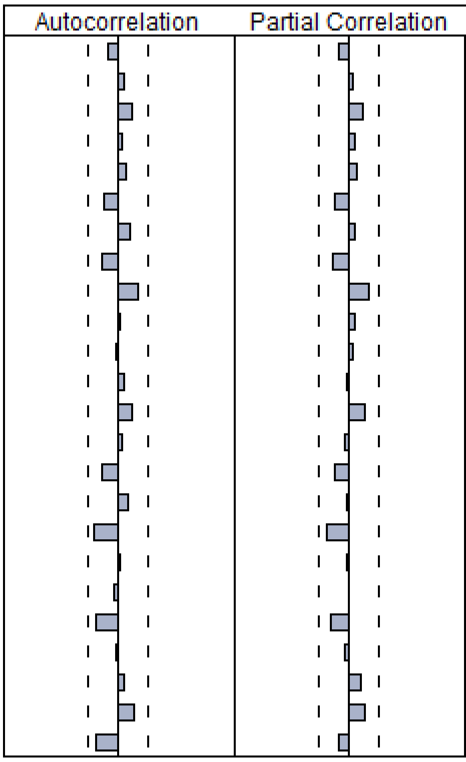


Figure 4-C3: Inverse Roots of AR and MA Process of ***ARIMA*** **(4,1,0)** .



	AC	PAC	Q-Stat	Prob
1	-0.097	-0.097	0.5501	
2	0.045	0.037	0.6745	
3	0.118	0.127	1.5268	
4	0.040	0.063	1.6261	
5	0.078	0.079	2.0103	0.156
6	-0.116	-0.125	2.8903	0.236
7	0.107	0.065	3.6488	0.302
8	-0.146	-0.149	5.0985	0.277
9	0.177	0.181	7.2559	0.202
10	0.020	0.044	7.2854	0.295
11	-0.012	0.036	7.2958	0.399
12	0.055	-0.009	7.5229	0.481
13	0.119	0.150	8.5879	0.476
14	0.030	-0.032	8.6597	0.565
15	-0.155	-0.119	10.558	0.481
16	0.091	-0.014	11.231	0.509
17	-0.208	-0.191	14.828	0.318
18	0.015	-0.011	14.847	0.389
19	-0.028	0.000	14.915	0.458
20	-0.200	-0.164	18.540	0.293
21	-0.017	-0.046	18.565	0.354
22	0.064	0.109	18.961	0.394
23	0.143	0.137	20.961	0.339
24	-0.200	-0.098	25.025	0.200

Figure 4-C4: Correlogram Residuals of PIT Model $\boldsymbol{ARIMA}(4,1,0)$.

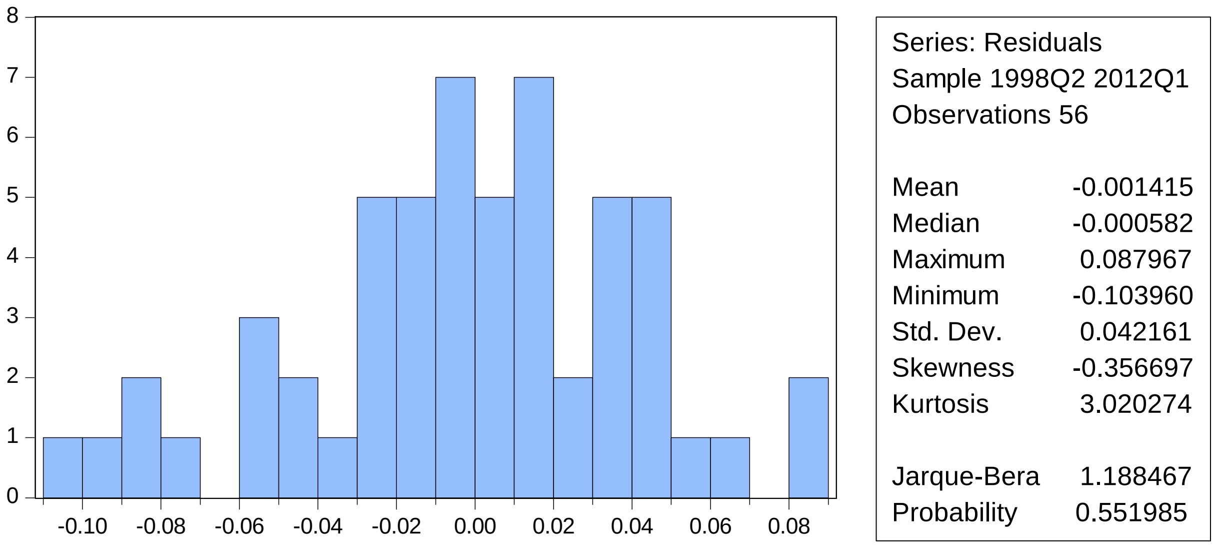


Figure 4-C5: Histogram of the Residuals of Model $\boldsymbol{ARIMA}(4,1,0)$.

Table 4-C2: Estimation Results of PIT $\mathbf{ETS}(\boldsymbol{A},\boldsymbol{M},\boldsymbol{A})$.

PIT ETS Smoothing Model: A,M,A - Additive Error, Multiplicative Trend, Additive Season	
Parameters	
Alpha:	0.309350
Beta:	0.000000
Gamma:	0.905182
Initial Parameters	
Initial level:	17298.92
Initial trend:	1.025161
Initial state 1:	-68.90486
Initial state 2:	224.7097
Initial state 3:	-1938.715
Initial state 4:	1782.910
Compact Log-likelihood	-534.4346
Log-likelihood	-500.0872
Akaike Information Criterion	1084.869
Schwarz Criterion	1101.214
Hannan-Quinn Criterion	1091.221
Sum of Squared Residuals	1.39E+08
Root Mean Squared Error	1563.247
Average Mean Squared Error	3010502.

Table 4-C3: Estimation Results of PIT $BVAR$.

Bayesian VAR Estimates			
Prior type: Litterman/Minnesota			
Hyper-parameters: Mu: 0.5, λ_1 : 0.9, λ_2: 0.7, λ_3 : 0.1			
	D(LOG(PIT))	**D(LOG(COE))**	**D(LOG(EMPL))**
D(LOG(PIT(-1)))	-0.453 (0.138) [-3.29141]	0.037 (0.044) [0.84080]	-0.048 (0.145) [-0.33142]
D(LOG(PIT(-2)))	-0.448 (0.146) [-3.06872]	-0.012 (0.046) [-0.25526]	-0.096 (0.153) [-0.62455]
D(LOG(PIT(-3)))	-0.328 (0.149) [-2.19497]	0.022 (0.047) [0.45619]	-0.419 (0.156) [-2.67963]
D(LOG(PIT(-4)))	0.440 (0.140) [3.15554]	-0.020 (0.044) [-0.45139]	0.002 (0.147) [0.01683]
D(LOG(COE(-1)))	0.731 (0.310) [2.35550]	0.081 (0.102) [0.80142]	-0.368 (0.335) [-1.09766]
D(LOG(COE(-2)))	0.154 (0.321) [0.47915]	-0.145 (0.105) [-1.37358]	0.568 (0.347) [1.63628]
D(LOG(COE(-3)))	0.399 (0.324) [1.23078]	0.095 (0.106) [0.89782]	0.094 (0.349) [0.27036]
D(LOG(COE(-4)))	-0.181 (0.321) [-0.56398]	0.841 (0.105) [7.99878]	1.085 (0.346) [3.13063]
D(LOG(EMPL(-1)))	0.006 (0.137) [0.04141]	-0.053 (0.045) [-1.17928]	-0.029 (0.151) [-0.19348]
D(LOG(EMPL(-2)))	0.103 (0.132) [0.77581]	0.008 (0.043) [0.18124]	0.074 (0.146) [0.50983]
D(LOG(EMPL(-3)))	-0.130 (0.133) [-0.98154]	0.042 (0.043) [0.97241]	0.109 (0.146) [0.74638]
D(LOG(EMPL(-4)))	0.003 (0.129) [0.02356]	0.012 (0.042) [0.28211]	-0.189 (0.142) [-1.33671]
@TREND	0.000 (0.000) [1.03475]	0.000 (0.000) [0.40779]	0.000 (0.000) [-0.65928]
R-squared	0.945	0.906	0.298
Adj. R-squared	0.928	0.877	0.083
Sum sq. resids	0.071	0.008	0.084
S.E. equation	0.043	0.014	0.046
F-statistic	55.649	31.281	1.383
Mean dependent	0.023	0.025	0.010
S.D. dependent	0.159	0.040	0.048

APPENDIX D: Value-Added Tax Tables and Figures

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.945	0.945	64.397	0.000
		2	0.904	0.096	124.16	0.000
		3	0.856	-0.074	178.50	0.000
		4	0.820	0.079	229.13	0.000
		5	0.769	-0.130	274.45	0.000
		6	0.726	-0.002	315.39	0.000
		7	0.680	-0.011	351.93	0.000
		8	0.648	0.076	385.65	0.000
		9	0.599	-0.143	414.94	0.000
		10	0.559	0.018	440.90	0.000
		11	0.515	-0.024	463.32	0.000
		12	0.483	0.028	483.34	0.000
		13	0.439	-0.074	500.17	0.000
		14	0.401	-0.006	514.50	0.000
		15	0.362	0.001	526.41	0.000
		16	0.335	0.029	536.78	0.000
		17	0.296	-0.070	545.04	0.000
		18	0.263	-0.014	551.68	0.000
		19	0.225	-0.035	556.65	0.000
		20	0.196	-0.003	560.50	0.000
		21	0.157	-0.072	563.00	0.000
		22	0.116	-0.076	564.41	0.000
		23	0.082	0.053	565.13	0.000
		24	0.057	0.019	565.48	0.000
		25	0.023	-0.065	565.54	0.000
		26	-0.014	-0.089	565.56	0.000
		27	-0.059	-0.095	565.96	0.000
		28	-0.090	0.028	566.92	0.000

Figure 4-D1: Correlogram of VATP at level.

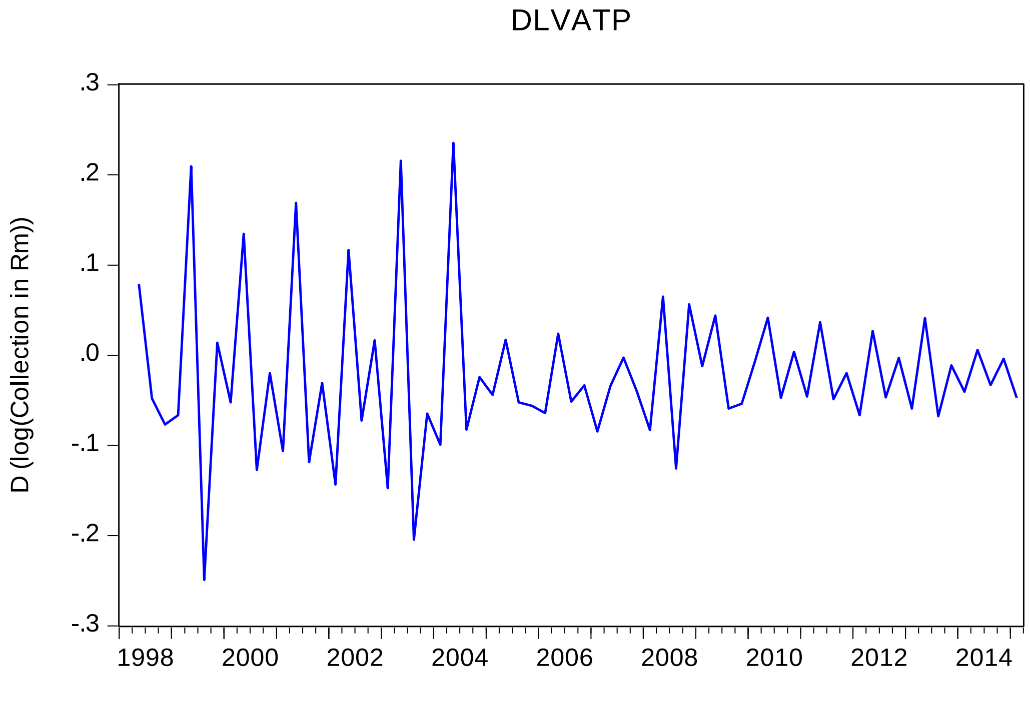


Figure 4-D2: Graph of VATP (Logarithmic Form, 1st Differenced) (DLVATP).

Table 4-D1: Estimation Results of $ARIMA(3,1,0)$.

Dependent Variable: DLOG(VATP)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.023216	0.002587	8.972431	0.0000
AR(1)	-0.782265	0.091996	-8.503275	0.0000
AR(2)	-0.618652	0.120169	-5.148183	0.0000
AR(3)	-0.657705	0.085951	-7.652068	0.0000
R-squared	0.649209	Mean dependent var		0.022972
Adjusted R-squared	0.621695	S.D. dependent var		0.094735
S.E. of regression	0.058268	Akaike info criterion		-2.724204
Sum squared resid	0.173155	Schwarz criterion		-2.543369
Log likelihood	81.27772	Hannan-Quinn criter.		-2.654095
F-statistic	23.59638	Durbin-Watson stat		1.707522
Prob(F-statistic)	0.000000			
Inverted AR Roots	.06+.85i	.06-.85i	-.90	

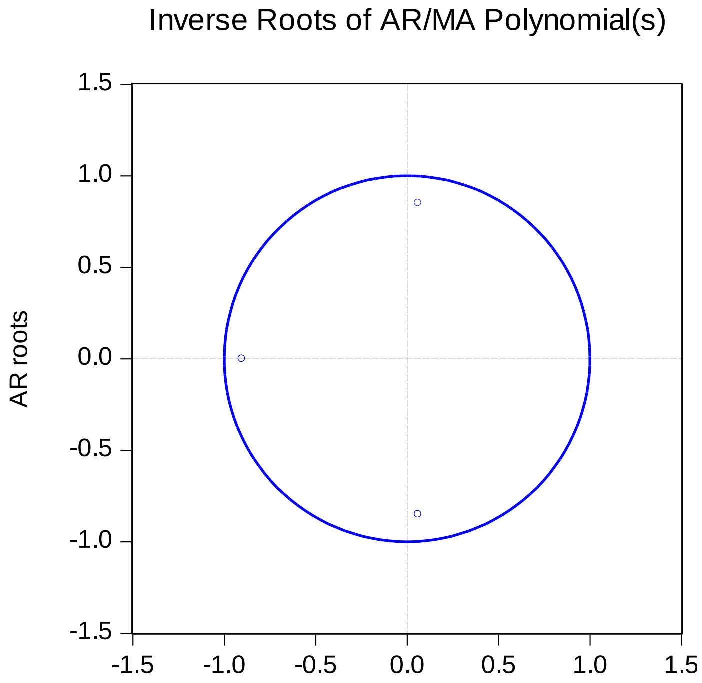
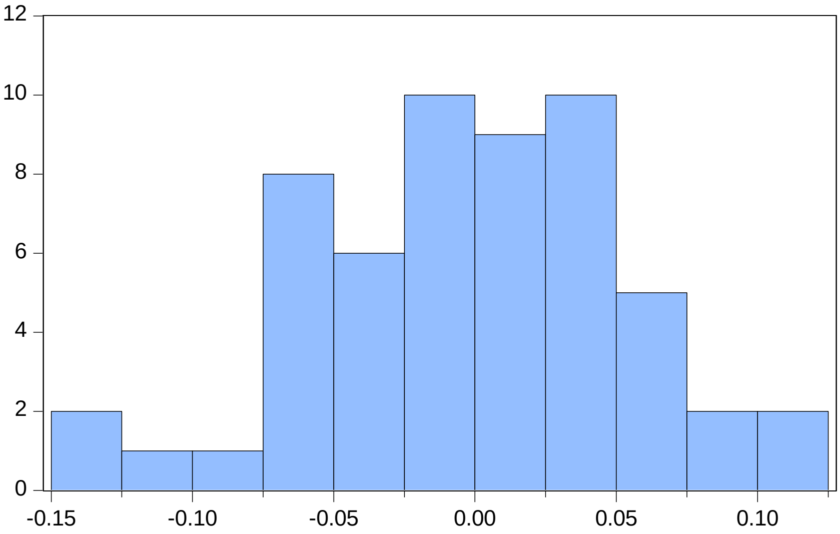


Figure 4-D3: Inverse Roots of AR and MA Process of $ARIMA(3,1,0)$.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.134	0.134	1.0619	
		2	0.031	0.013	1.1188	
		3	0.046	0.041	1.2489	
		4	-0.022	-0.035	1.2793	0.258
		5	-0.018	-0.012	1.2992	0.522
		6	0.021	0.024	1.3274	0.723
		7	0.020	0.017	1.3534	0.852
		8	0.113	0.110	2.2147	0.819
		9	-0.146	-0.185	3.6811	0.720
		10	-0.184	-0.152	6.0597	0.533
		11	-0.209	-0.183	9.2159	0.324
		12	-0.143	-0.079	10.717	0.296
		13	-0.066	-0.027	11.043	0.354
		14	-0.106	-0.108	11.912	0.370
		15	-0.170	-0.179	14.213	0.287
		16	0.144	0.178	15.885	0.255
		17	0.065	0.109	16.237	0.299
		18	-0.015	0.013	16.257	0.365
		19	-0.008	-0.038	16.263	0.435
		20	0.172	0.154	18.934	0.332
		21	-0.123	-0.245	20.343	0.314
		22	-0.125	-0.195	21.830	0.293
		23	0.096	0.074	22.729	0.302
		24	0.064	-0.100	23.145	0.336

Figure 4-D4: Correlogram Residuals of Model $ARIMA(3,1,0)$.



Series: Residuals
Sample 1998Q2 2012Q1
Observations 56

Mean	-0.000784
Median	0.004594
Maximum	0.117278
Minimum	-0.149365
Std. Dev.	0.056104
Skewness	-0.380230
Kurtosis	3.050567
Jarque-Bera	1.355333
Probability	0.507801

Figure 4-D5: Histogram of the Residuals of Model $ARIMA(3,1,0)$.

Table 4-D2: Estimation Results of $\boldsymbol{ETS(M,M,A)}$.

VATP ETS Smoothing Model: M,M,A - Multiplicative Error, Multiplicative Trend, Additive Season	
Parameters	
Alpha:	0.458603
Beta:	0.000000
Gamma:	0.000000
Initial Parameters	
Initial level:	15271.34
Initial trend:	1.023637
Initial state 1:	123.7055
Initial state 2:	106.8306
Initial state 3:	-1699.165
Initial state 4:	1468.629
Compact Log-likelihood	-527.3924
Log-likelihood	-493.0450
Akaike Information Criterion	1070.785
Schwarz Criterion	1087.129
Hannan-Quinn Criterion	1077.137
Sum of Squared Residuals	0.116389
Root Mean Squared Error	0.045188
Average Mean Squared Error	2899152.

Table 4-D3: Estimation Results of VATP BVAR.

Bayesian VAR Estimates			
Prior type: Litterman/Minnesota			
Hyper-parameters: Mu: 0, λ_1 : 0.5, λ_2 : 0.5, λ_3 : 1			
	D(LOG(VATP))	**D(LOG(GDE))**	**D(LOG(PCE))**
D(LOG(VATP(-1)))	-0.209 (0.090) [-3.40110]	0.027 (0.019) [2.37847]	0.000 (0.010) [0.02847]
D(LOG(VATP(-2)))	-0.036 (0.073) [-2.32401]	0.000 (0.010) [1.89718]	-0.002 (0.005) [0.64395]
D(LOG(VATP(-3)))	-0.106 (0.056) [-2.14614]	-0.001 (0.007) [-0.58274]	0.002 (0.004) [1.01814]
D(LOG(VATP(-4)))	0.106 (0.045) [1.50351]	-0.001 (0.005) [-0.56591]	0.000 (0.003) [0.40015]
D(LOG(GDE(-1)))	0.055 (0.075) [0.76710]	-0.182 (0.096) [-3.62779]	-0.008 (0.020) [-0.27875]
D(LOG(GDE(-2)))	-0.019 (0.039) [-0.64314]	0.069 (0.077) [-0.20820]	0.000 (0.011) [-0.61900]
D(LOG(GDE(-3)))	0.008 (0.026) [0.31728]	-0.042 (0.057) [-0.46400]	-0.001 (0.007) [-0.49958]
D(LOG(GDE(-4)))	-0.003 (0.020) [-0.07541]	0.024 (0.045) [0.01687]	0.001 (0.005) [-0.17200]
D(LOG(PCE(-1)))	0.362 (0.110) [2.02807]	-0.028 (0.054) [0.21052]	-0.391 (0.053) [-4.77905]
D(LOG(PCE(-2)))	-0.082 (0.068) [-0.57149]	0.014 (0.034) [-0.41022]	-0.240 (0.052) [-4.02067]
D(LOG(PCE(-3)))	0.003 (0.047) [-0.47517]	-0.013 (0.023) [-0.81421]	-0.327 (0.044) [-5.09446]
D(LOG(PCE(-4)))	-0.005 (0.036) [0.05205]	0.005 (0.018) [0.72524]	0.372 (0.040) [7.77089]
C	0.020 (0.010) [1.52970]	0.032 (0.006) [3.24586]	0.040 (0.004) [5.15856]
R-squared	0.713	0.541	0.921
Adj. R-squared	0.625	0.400	0.897
Sum sq. resids	0.137	0.037	0.013
S.E. equation	0.059	0.031	0.018
F-statistic	8.093	3.834	37.895
Mean dependent	0.023	0.028	0.025
S.D. dependent	0.097	0.040	0.057

APPENDIX E: Total Tax Revenue Tables and Figures

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.920	0.920	60.975	0.000
		2	0.854	0.048	114.27	0.000
		3	0.832	0.259	165.62	0.000
		4	0.819	0.107	216.18	0.000
		5	0.748	-0.314	259.02	0.000
		6	0.685	-0.010	295.46	0.000
		7	0.660	0.094	329.92	0.000
		8	0.646	0.054	363.39	0.000
		9	0.580	-0.180	390.83	0.000
		10	0.523	0.013	413.50	0.000
		11	0.500	0.037	434.62	0.000
		12	0.483	0.005	454.67	0.000
		13	0.425	-0.090	470.45	0.000
		14	0.367	-0.059	482.42	0.000
		15	0.344	0.041	493.16	0.000
		16	0.333	0.044	503.40	0.000
		17	0.276	-0.129	510.57	0.000
		18	0.233	0.088	515.80	0.000
		19	0.208	-0.100	520.04	0.000
		20	0.197	0.048	523.92	0.000
		21	0.151	-0.049	526.24	0.000
		22	0.107	-0.034	527.44	0.000
		23	0.083	-0.039	528.17	0.000
		24	0.076	0.055	528.79	0.000
		25	0.030	-0.109	528.89	0.000
		26	-0.019	-0.064	528.93	0.000
		27	-0.051	-0.059	529.24	0.000
		28	-0.071	-0.049	529.84	0.000

Figure 4-E1: Correlogram of TTR at level.

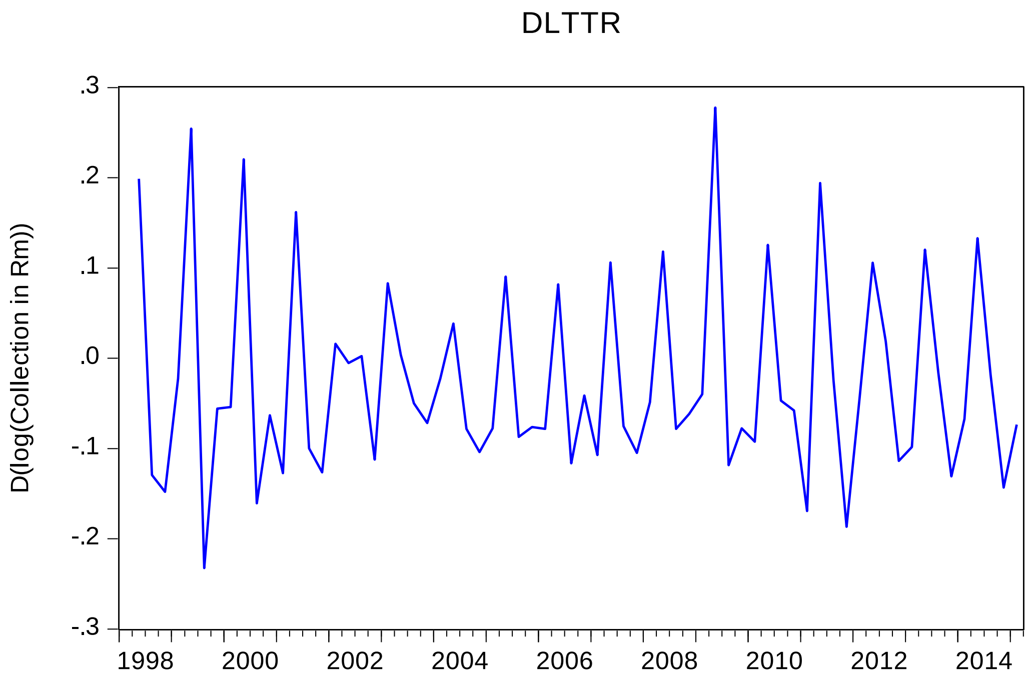


Figure 4-E2: Graph of Total Tax Revenue (Logarithmic Form, 1st Differenced) (DLTTR).

*Table 4-E1: Estimation Results of **ARIMA (4,1,1)** .*

Dependent Variable: DLOG(TTR)				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
AR(4)	0.854887	0.063352	13.49419	0.0000
MA(1)	-0.488816	0.159967	-3.055740	0.0035
R-squared	0.723053	Mean dependent var		0.026819
Adjusted R-squared	0.712602	S.D. dependent var		0.113639
S.E. of regression	0.060922	Akaike info criterion		-2.609544
Sum squared resid	0.196706	Schwarz criterion		-2.501043
Log likelihood	76.06724	Hannan-Quinn criter.		-2.567479
Durbin-Watson stat	2.006652			
Inverted AR Roots	.96	.00-.96i	.00+.96i	-.96
Inverted MA Roots	.49			

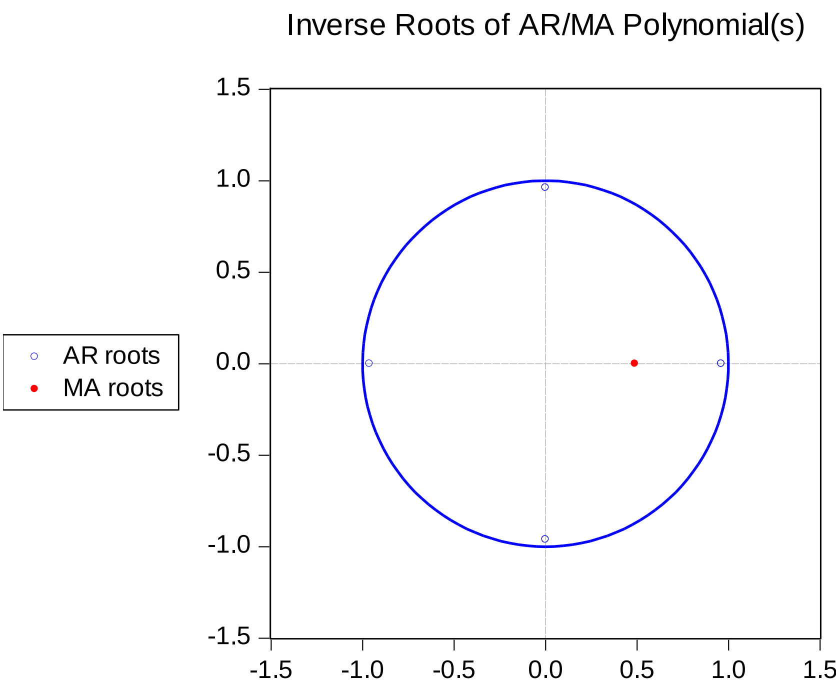
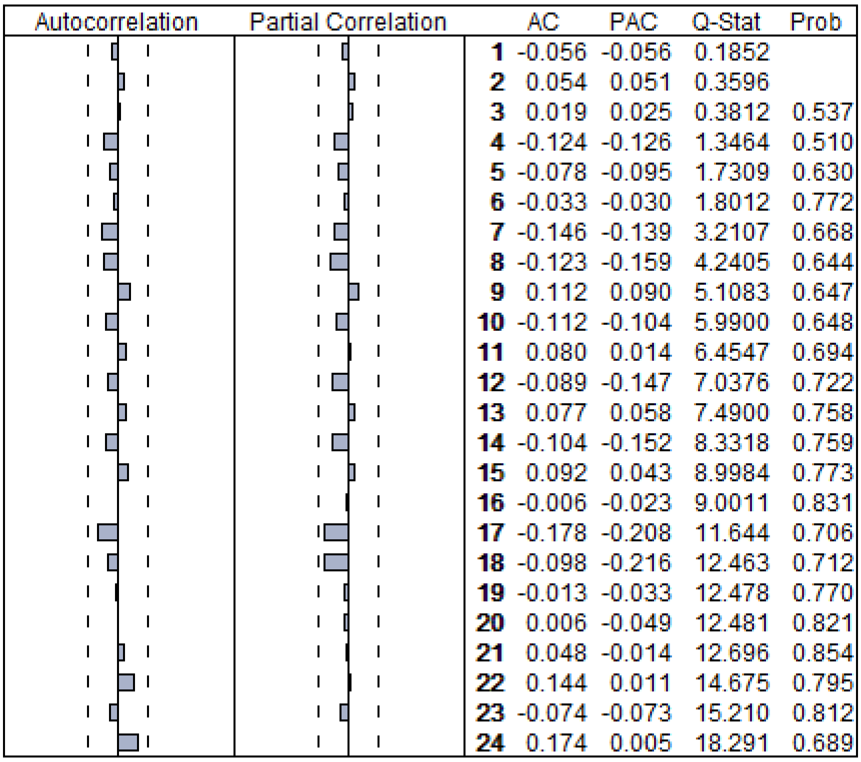


*Figure 4-E3: Inverse Roots of AR and MA Process of **ARIMA (4,1,1)** .*



	AC	PAC	Q-Stat	Prob
1	-0.056	-0.056	0.1852	
2	0.054	0.051	0.3596	
3	0.019	0.025	0.3812	0.537
4	-0.124	-0.126	1.3464	0.510
5	-0.078	-0.095	1.7309	0.630
6	-0.033	-0.030	1.8012	0.772
7	-0.146	-0.139	3.2107	0.668
8	-0.123	-0.159	4.2405	0.644
9	0.112	0.090	5.1083	0.647
10	-0.112	-0.104	5.9900	0.648
11	0.080	0.014	6.4547	0.694
12	-0.089	-0.147	7.0376	0.722
13	0.077	0.058	7.4900	0.758
14	-0.104	-0.152	8.3318	0.759
15	0.092	0.043	8.9984	0.773
16	-0.006	-0.023	9.0011	0.831
17	-0.178	-0.208	11.644	0.706
18	-0.098	-0.216	12.463	0.712
19	-0.013	-0.033	12.478	0.770
20	0.006	-0.049	12.481	0.821
21	0.048	-0.014	12.696	0.854
22	0.144	0.011	14.675	0.795
23	-0.074	-0.073	15.210	0.812
24	0.174	0.005	18.291	0.689

Figure 4-E4: Correlogram Residuals of TTR Model $\boldsymbol{ARIMA}(4,1,1)$.

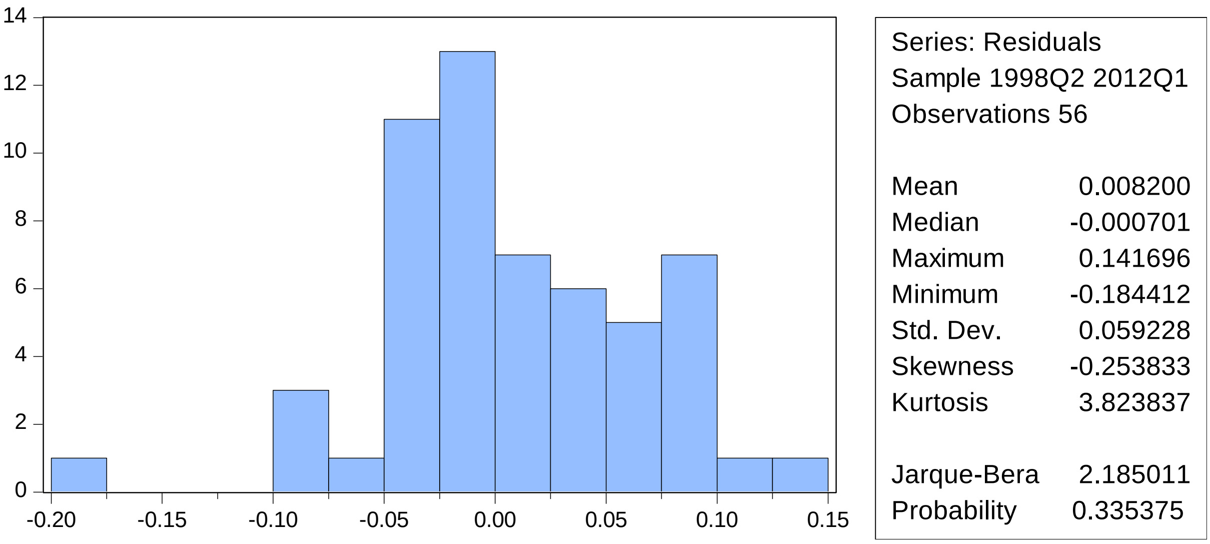


Figure 4-E5: Histogram of the Residuals of Model $\boldsymbol{ARIMA}(4,1,1)$.

Table 4-E2: Estimation Results of TTR $\boldsymbol{ETS(M, M_D, A)}$.

TTR ETS Smoothing	
Model: M,MD,A - Multiplicative Error, Multiplicative-Dampened Trend, Additive Season	
Parameters	
Alpha:	0.386361
Beta:	0.000000
Gamma:	0.463687
Phi:	0.995864
Initial Parameters	
Initial level:	41643.37
Initial trend:	1.031284
Initial state 1:	3656.653
Initial state 2:	-1114.836
Initial state 3:	-6560.208
Initial state 4:	4018.392
Compact Log-likelihood	-598.7603
Log-likelihood	-564.4128
Akaike Information Criterion	1215.521
Schwarz Criterion	1233.908
Hannan-Quinn Criterion	1222.667
Sum of Squared Residuals	0.150792
Root Mean Squared Error	0.051434
Average Mean Squared Error	49016413

Table 4-E3: Estimation Results of TTR $BVAR$.

Bayesian VAR Estimates			
Prior type: Litterman/Minnesota			
Hyper-parameters: Mu: 0, λ_1 : 1, λ_2 : 1, λ_3 : 0.5			
	D(LOG(TTR))	**D(LOG(GDP))**	**D(LOG(CPI))**
D(LOG(TTR(-1)))	-0.633 (0.144) [-4.38021]	-0.020 (0.035) [-0.56118]	0.022 (0.035) [0.62508]
D(LOG(TTR(-2)))	-0.404 (0.158) [-2.55233]	0.046 (0.038) [1.21509]	0.041 (0.038) [1.07354]
D(LOG(TTR(-3)))	-0.253 (0.166) [-1.52206]	0.098 (0.040) [2.45429]	-0.025 (0.040) [-0.62495]
D(LOG(TTR(-4)))	0.242 (0.152) [1.59876]	-0.017 (0.037) [-0.46224]	0.007 (0.037) [0.19603]
D(LOG(GDP(-1)))	1.057 (0.619) [1.70881]	-0.097 (0.149) [-0.64990]	0.013 (0.150) [0.08455]
D(LOG(GDP(-2)))	0.881 (0.568) [1.54948]	-0.203 (0.137) [-1.47881]	0.012 (0.138) [0.08394]
D(LOG(GDP(-3)))	0.408 (0.569) [0.71698]	-0.258 (0.137) [-1.87755]	0.038 (0.138) [0.27330]
D(LOG(GDP(-4)))	-1.382 (0.553) [-2.50028]	-0.002 (0.134) [-0.01863]	0.055 (0.134) [0.41223]
D(LOG(CPI(-1)))	0.185 (0.655) [0.28289]	0.184 (0.158) [1.16139]	0.294 (0.158) [1.85645]
D(LOG(CPI(-2)))	0.267 (0.684) [0.39066]	0.103 (0.165) [0.62166]	-0.119 (0.165) [-0.71771]
D(LOG(CPI(-3)))	-0.595 (0.634) [-0.93907]	0.072 (0.153) [0.46864]	-0.042 (0.153) [-0.27102]
D(LOG(CPI(-4)))	0.297 (0.614) [0.48360]	0.315 (0.148) [2.12166]	0.000 (0.149) [0.00259]
C	0.027 (0.042) [0.65738]	0.034 (0.010) [3.35153]	0.010 (0.010) [1.00252]
@TREND	0.000 (0.001) [-0.03697]	0.000 (0.000) [-1.35220]	0.000 (0.000) [-0.35044]
R-squared	0.814	0.693	0.286
Adj. R-squared	0.751	0.588	0.042
Sum sq. resids	0.128	0.007	0.007
S.E. equation	0.058	0.014	0.014
F-statistic	12.832	6.591	1.171
Mean dependent	0.027	0.027	0.015
S.D. dependent	0.116	0.022	0.014