FEDERAL TAX REVENUE FORECASTING IN PAKISTAN: ISSUES AND ALTERNATIVE METHODS EVALUATION

Ph.D. Dissertation for partial fulfillment of the degree of Doctor of Philosophy in Econometrics

Pakistan Institute of Development Economics

By

Syeda Um Ul Baneen
PIDE2016PHDETS08

Supervisor
Dr. Ahsan Ul Haq Satti
Co-Supervisor
Dr. Mahmood Khalid

Ph.D. Econometrics

PIDE School of Economics

Pakistan Institute of Development Economics, Islamabad

2025

Author's Declaration

I Ms. Syeda Um Ul Baneen hereby state that my PhD thesis titled **"Federal Tax Revenue Forecasting in Pakistan: Issues and Alternative Methods Evaluation"** is my own work and has not been submitted previously by me for taking any degree from **Pakistan Institute of Development Economics, Islamabad'** or anywhere else in the country/world.

At any time if my statement is found to be incorrect even after my Graduation the university has the right to withdraw my PhD degree.

Syeda Ba

Ms. Syeda Um Ul Baneen

PIDE2016FPHDETS08

Plagiarism Undertaking

I solemnly declare that research work presented in the thesis titled **"Federal Tax Revenue Forecasting in Pakistan: Issues and Alternative Methods Evaluation"** is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the **HEC** and **Pakistan Institute of Development Economics, Islamabad** towards plagiarism. Therefore I as an Author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of PhD degree, the University reserves the rights to withdraw/revoke my PhD degree and that HEC and the University has the right to publish my name on the HEC/University Website on which names of students are placed who submitted plagiarized thesis.

Students/Author Signature:________________

Ms. Syeda Um Ul Baneen

PIDE2016FPHDETS08

Certificate of Approval

This is to certify that the research work presented in this thesis, entitled: **"Federal Tax Revenue Forecasting in Pakistan: Issues and Alternative Methods Evaluation"** was conducted by **Ms. Syeda Um ul Baneen** under the supervision of **Dr. Ahsan ul Haq.** No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Econometrics from **Pakistan Institute of Development Economics, Islamabad.**

Student Name: Ms. Syeda Um Ul Baneen
PIDE2016FPHDETS08
Signature: ____________

Examination Committee:

a) **External Examiner: Dr. Naseem Faraz**
DG Macro-Fiscal Policy Unit/
Deputy Economic Advisor,
Ministry of Finance, Islamabad
Signature: ____________

b) **Internal Examiner: Dr. Abdul Jalil**
Professor,
NDU, Islamabad
Signature: ____________

Supervisor: Dr. Ahsan ul Haq
Assistant Professor,
PIDE, Islamabad
Signature: ____________

Co-Supervisor: Dr. Mahmood Khalid
Senior Research Economist,
PIDE, Islamabad
Signature: ____________

Dr. Iftikhar Ahmad
Head, PIDE School of Economics (PSE)
PIDE, Islamabad
Signature: ____________

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to Almighty Allah for His countless blessings and guidance throughout this journey. Without His help, the completion of this dissertation would not have been possible.

I extend my heartfelt gratitude to my thesis supervisors, Dr. Ahsan Ul Haq Satti and Dr. Mahmood Khalid. Dr. Ahsan Ul Haq Satti's profound expertise in econometrics laid the foundation for the analytical rigor of this research. His insightful guidance and unwavering support were instrumental in shaping this work. I am equally thankful to my co-supervisor, Dr. Mahmood Khalid, whose depth in applied economic research consistently steered me in the right direction. His constructive feedback and corrections significantly improved the quality of this dissertation. I remain sincerely indebted to both of them.

I am also thankful to my dear friends, Uzma Zia, Ambreen Fatima, and Sidra Bibi, for their valuable discussions, encouragement, and thoughtful suggestions, which contributed meaningfully to this work.

A special note of gratitude goes to my mother, Syeda Kalsoom Fatima, for her unconditional love, prayers, and steadfast support throughout this academic journey. Though my father, Syed Raza Hussain, is no longer with us, his belief in my potential and his dream of seeing me earn a doctorate have been a guiding light and a constant source of motivation.

Finally, I would like to express my deep appreciation to my husband Syed Alishan for his patience, encouragement, and unwavering belief in me. His emotional support and understanding during the most challenging phases of this journey have been invaluable, and I am truly grateful to have him by my side.

Syeda Um Ul Baneen
School of Economics,
PIDE, Islamabad,
Pakistan.

ABSTRACT

Every year, forecasting errors occur in Pakistan's government tax revenues. From 1970 to 2020, the federal government made massive projection errors across all tax categories. Using data from 1970 to 2014, two-thirds of the time, overall federal tax collections were overstated by more than 5 percent (Qasim & Khalid, 2016). The consequences of inaccurate tax revenue projections could be deceptive tax revenue targets, the unfair distribution of resources, and an increase in the debt burden. The reasons for tax revenue forecasting errors can be administrative, political, or methodological. Unfortunately, the political and organizational issues can't be measured accurately due to the qualitative and dispersed nature of the problems.

A significant part of these errors occurs due to wrong choices of forecasting methods, data discrepancy, and parameter issues. This study digs out the said problems in detail. The first objective is to document the current tax revenue forecasting mechanism and its impact on the economy. Forecasting errors are transferred into the economy through budget estimates. The study highlights the impacts specifically in terms of development expenditure, GDP, and debt accumulation.

The second objective is the decomposition of forecasting errors in terms of data, parameters, and estimation methods. FBR's default forecasting method is the Buoyancy approach. This method uses buoyancy estimates of tax revenues along with tax bases to forecast future tax revenue. FBR uses provisional data for forecasting purposes. This research has used different combinations of provisional, revised, final, and real-time data to check the effect of data discrepancies. The results show that the data regime has no significant impact on improving the forecasts.

The Buoyancy method uses the nominal GDP targets provided by the Ministry of Finance. The second part of the second objective is to check the effectiveness of the GDP targets revision. In different scenarios, the thesis has used provisional, revised and real-time targets of nominal GDP, and its components i.e., real GDP and inflation rate. The findings suggest that the real GDP target significantly improves the forecasts compared to provisional or revised targets. The inflation target regime has an insignificant impact on improving the projections.

The third part of the last objective is a check of the usefulness of the buoyancy approach and alternative forecasting methods. The forecasting is done using alternative theoretical, statistical, and machine learning methods. The mean absolute error suggests marginal tax rate approach is the best tax revenue forecasting method. On the other hand, the root mean square error suggests that LASSO (Least Absolute Shrinkage and Selection Operator)/Elastic Net is the best forecasting method for Total Tax, Custom Duties and FED. At the same time, the Box Jenkins methodology provides the best forecast for Direct Tax and Sales Tax. The study also considers the feasibility of the proposed methodology in terms of budget, time, and organizational terms.

Keywords: Tax Revenue Forecasting, Buoyancy Approach, LASSO, Marginal Tax Rate **JEL Classification*:*** B26; B41;C13; H20

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CBR	Central Board of Revenue
CD	Customs Duty
CFY	Current Fiscal Year
ECM	Error Correction Mechanism
ETR	Effective Tax Rate
FAA	Forecast Based on Actual Real GDP Growth and Actual Inflation Rate Targets
FAP	Forecast Based on Actual Real GDP Growth and Provisional Inflation Rate Targets
FAR	Forecast Based on Actual Real GDP Growth and Revised Inflation Rate Targets
FBR	Federal Board of Revenue
FED	Federal Excise Duty
FFF	Final Tax and Final Base Leading Forecast
FPA	Forecast Based on Provisional Real GDP Growth and Actual Inflation Rate Targets
FPR	Forecast Based on Provional Real GDP Growth and Revised Inflation Rate Targets
FPF	Final Tax and Provisional Base Leading Forecast
FRA	Forecast Based on Revised Real GDP Growth and Actual Inflation Rate Targets
FREF	Final Tax and Real Time Leading Forecast
FRF	Final Tax and Revised Base Leading Forecast
FRP	Forecast Based on Revised Real GDP Growth and Provisional Inflation Rate Targets
FY	Fiscal Year
GDP	Gross Domestic Product
LASSO	Least Absolute Shrinkage and Selection Operator
LSM	Large Scale Manufacturing

ME	Margin of Error
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MTR	Marginal Tax Rate
OBR	The Office for Budget Responsibility
PFF	Provisional Tax and Final Base Leading Forecast
PPF	Provisional Tax and Provisional Base Leading Forecast
PRF	Provisional Tax and Revised Base Leading Forecast
PREF	Provisional Tax and Real Time Base Leading Forecast
REREF	Real Time Tax and Real Time Base Leading Forecast
RMSE	Root Mean Square Error
RPF	Revised Tax and Provisional Base Leading Forecast
RREF	Revised Tax and Real Time Leading Forecast
RRF	Revised Tax and Revised Base Leading Forecast
SARS	South African Revenue Service
SITM	Scottish Income Tax Micro-Simulation Model
SPRS	Strategic Planning Reforms & Statistics
ST	Sales Tax
VAR	Vector Auto-Regression
VECM	Vector Error Correction Mechanism

CHAPTER 1

INTRODUCTION

"There are two kinds of forecasters: those who don't know, and those who don't know they don't know." John Kenneth Galbraith

“Forecasting" means assessing the value that a quantity will assume at some future point in time (Marriott, 2002). It has perpetually led the charge in policy-making and strategic planning. The utility of predictability is enormous. It is used in numerous scenarios: determining whether to construct another power generation plant in the next few years needs projections of upcoming demand; recruiting staff in a call center next month needs forecasts of call volumes; stocking an inventory involves predictions of stock supplies (Hyndman & Athanasopoulos, 2018).

The uncertainty surrounding the future is stimulating for individuals, organizations, and economies pursuing to minimize risks and maximize utilities. Forecasts lead to choices in all areas of economics and finance, and their worth can only be judged in relation to, and in the context of such choices (Elliott & Timmermann, 2008). Despite vast utility, forecasts are prone to some statistical, methodological and administrative issues. The predictableness of an incident or a quantity depends on many things including: factors that affect it, data availability, and the forecast’s effect on the future. For smooth implementation, these issues must be considered.

For decades, forecasting has been used to predict different economic factors. One of them is tax revenue forecasting. Tax receipts are vital to run an economy. Tax revenue[1] analysis and forecasting are critical to stabilizing tax and expenditure policy.

[1] Tax revenue is the revenue gathered through taxes on any income, profit, social security contributions, goods, and services. It also includes payroll taxes, taxes on ownership, property transfer, and miscellaneous taxes (OECD, 2022).

They play an essential part in annual budget preparation. Predictability provides policymakers and fiscal managers with the data needed to lead borrowing, use accumulated assets, or identify monetary measures to balance the budget. It also notifies about what fiscal arrangements are justifiable and how to balance fiscal policy to curb the complications in the balance of payments and foreign debt.

To augment the timely and compelling analysis of the revenue aspects of the fiscal policy, economies have gradually turned toward domestic tax policy entities rather than depending on tax experts from outside. For example, in the United Kingdom, the Office for Budget Responsibility (OBR), in Germany's Federal Ministry of Finance, and in Pakistan the Federal Board of Revenue (FBR) are in charge for projecting the annual estimates of different federal taxes.

There are many practices to estimate tax revenue projections. Each economy employs the forecasting strategy according to its economic, social, and feasibility conditions. For example, New Zealand and Ireland use the elasticity approach to forecast their federal tax revenues. In Scotland, a micro-simulation model is used for income tax forecasting known as the SITM or Scottish Income Tax Micro-Simulation Model. On the other hand, Pakistan practices the buoyancy approach to forecast federal tax revenues. FBR's Strategic Planning Reforms & Statistics Wing publishes head-wise tax revenue forecasts using nominal GDP forecasts. Annually, the ministry of Finance provides nominal GDP forecasts for the next three years.

FBR deals with the smooth functioning of tax collection, introducing and implementing tax rules along with predicting tax revenue estimates. Federal Taxes collected by FBR consist of two components Direct and Indirect Taxes. The Direct Taxes include: Income Tax, Capital Value Tax (CVT), Workers Welfare Fund

(WWF) and Workers Profits Participation Fund (WPPF). Indirect Taxes include: Sales Tax, Customs Duty, and Federal Excise Duty (FED). Moreover, all Domestic Taxes have also been categorized as Inland Revenue (FBR, 2021). Table 1.1 shows targets versus the net collection of federal taxes.

Table 1.1: Tax Targets Versus Net Collection (Rs. Billion)

Year	Targets	Collection	Growth	Target Achieved (%)
2016-17	3521	3368	8.2	95.7
2017-18	3935	3844	14.1	97.7
2018-19	4150	3828	-0.4	92.2
2019-20	3908	3997	4.4	102.3
2020-21	4691	4745	18.7	101.2

Source: FBR Year Book (2020-21)

Federal tax revenue forecasting has consistently been affected by errors, spanning from 1970 to 2020. Data analysis from 1970 to 2014 showed that two-thirds of the time, total federal taxes were overstated by more than five percent (Qasim & Khalid, 2016). These inaccuracies in tax revenue predictions have direct implications for budget formulation. They can lead to misleading revenue targets, misallocation of resources, reductions in development expenditures, and heightened debt burdens (Ademmer & Hogrefe, 2022). Moreover, the repercussions can extend to long-term monetary and fiscal effects.

Forecasting errors are transferred into the economy through budget estimates every year. This study is focused on measuring the development expenditure losses due to tax revenue errors. Following Ademmer & Hogrefe (2022), the fiscal costs of inaccurate forecasting are evaluated in terms of a fiscal quadratic loss function. In the first step, development expenditure losses are calculated. The GDP growth losses are derived through development expenditure losses. These measures estimate the economy's financial losses due to tax revenue errors.

The reasons for tax revenue forecasting errors can be administrative, political, or methodological. Unfortunately, the political and organizational issues can't be measured accurately due to the qualitative and dispersed nature of the problems. On the other hand, it is hypothesized that a significant number of these errors may occur due to wrong forecasting methods, data, and parameter choices. This study has dug out these problems in detail.

FBR uses the buoyancy approach for forecasting. The technique involves the calculation of tax buoyancies with respect to their bases. A product of tax buoyancy and nominal GDP growth target provides the tax revenue estimates. FBR uses the base year tax revenues and tax bases to calculate the buoyancies of all tax heads. Provisional data is used for that purpose. There data revision may improve the forecasts. The statement leads to the investigation of which data is more appropriate for forecasting purposes. What if tax revenues or bases or both are revised? Data discrepancy issues can be a possible reason for tax revenue forecasting errors.

The Ministry of Finance provides estimates of the GDP growth. This raises some questions; Are the choices of provisional parameters i.e. real GDP growth target and inflation rate target correct? What would happen if we use revised or real time parameters to forecast the tax revenues? Parameter choice may improve the forecasts. The effect of revision of nominal GDP growth target or real GDP growth target or inflation rate target can be checked. The results may tell about the suitability of parameters.

The choice of forecasting method plays a vital role in the efficiency of tax revenue forecasting. The inappropriate method would lead to misleading forecasts.The Federal Board of Revenue (FBR) employs the Buoyancy approach to predict tax revenues (FBR, 2021). This method involves assessing the proportional change in tax revenue relative to the proportional change in the tax base. These estimates are subsequently multiplied by nominal GDP growth targets.

As reported earlier, most of the time, forecasting errors were relatively higher (more than 5 percent). One probable explanation for these repeating errors might be the use of the same ineffective forecasting approach over time. There are several forecasting methods available. Some substitutes to the buoyancy approach can be applied for federal tax revenue forecasts. The comparison can tell about the strengths and weaknesses of the buoyancy approach.

In this study, theoretical, stataistical and machine learning methods are used as alternatives of Buoyancy approach for federal tax revenue forecasting of Pakistan. Theoretical methods are further devided into buoyancy approach, effective tax rate approach, marginal tax rate approach and structural VAR method. Statistical methods category used Box Jenkins methodology while machine learning methods include LASSO, elastic net and ridge regression. These methods with their implications and justifications to use in this study are given in table 1.2.

Table 1.2: Methods Used and Justifications

Category	Method	Description	Application	Justification
Theoretical/ Orthodox Methods	Buoyancy Approach	Measures tax revenue responsiveness to changes in GDP without policy changes.	Tax revenue forecasting, policy impact analysis.	Chosen because it helps in understanding the natural growth of tax revenues in response to economic changes without policy distortions.
	Effective Tax Rate Approach	Evaluates the actual tax burden on income or economic activity.	Tax policy assessment, fiscal studies.	Selected as it provides a more realistic view of the tax system's impact on individuals and businesses.
	Marginal Tax Rate Approach	Assesses the additional tax paid on an extra unit of income.	Tax efficiency analysis, behavioral tax modeling.	Included because it's essential for analyzing tax progressivity and economic incentives.
	Structural Vector Autoregression (SVAR) Method	Analyzes dynamic relationships between economic variables with structural shocks.	Macroeconomic policy evaluation, fiscal impact assessment.	Used for its ability to model economic shocks and policy impacts systematically.
Statistical Methods	Box-Jenkins Method	Time-series forecasting method based on ARIMA models.	Economic trend forecasting, revenue prediction.	Chosen due to its effectiveness in handling time-series data, making it ideal for revenue projections.
Machine Learning Methods	LASSO	Regression method that performs variable selection and regularization.	Feature selection, high-dimensional tax modeling.	Preferred because it reduces overfitting and selects the most relevant predictors in complex economic datasets.
	Ridge Regression	Addresses multicollinearity by penalizing large coefficients.	Predictive modeling, robust economic forecasting.	Used because it stabilizes predictions when dealing with correlated economic variables.
	Elastic Net	Hybrid of LASSO and Ridge Regression, balancing selection and shrinkage.	Complex tax revenue modeling, economic variable selection.	Chosen for its ability to handle highly correlated variables while maintaining model simplicity.

Although this study emphasizes on methodological and statistical reasons for forecasting problems, administrative constraints can't be ignored entirely. Pakistan has many constraints in terms of forecasting. Every method, data, and parameter is not feasible in all circumstances. Keeping in view the constraints, the study also considers the feasibility in terms of budget, time, and organizational terms. The empirically suggested methodologies are also evaluated on feasibility criteria considering

Pakistan's condition. The optimal recommendation is based on statistical errors and feasibility criteria.

1.1. Research Gap

There are many studies on the decomposition of tax revenue forecasting errors. Most of them studied either data issues or parameter issues or method issues. A comprehensive study of Pakistan's federal tax revenue forecasting errors was needed. There should be a research which evaluates the method/data/parameter discrepancy issues in the existing federal tax revenue forecasting mechanism and suggest better alternatives. This study will fill this gap by suggesting alternative advanced methodologies, data and parameter revisions where appropriate.

1.2. Research Questions

The key research questions of the proposed study are:

1. What are the existing FBR tax revenue forecasting methods, and how efficient are they?
2. How can different forecasting errors be decomposed in terms of data, parameters, and methodology?
3. What can be the best practice to improve Pakistan's federal tax revenue forecasting?

1.3. Objectives of the Study

The study will try to fulfill the following objectives:

- **To Analyze FBR's tax revenue forecasting methods**

The study documents the existing tax revenue forecasting practices at FBR. Their efficiency is descripted in terms of lousy optimal development expenditures and optimal debt.

- **To decompose tax revenue forecasting errors in terms of data discrepancy issues, parameter issues, and method choice**

  Forecasting errors are decomposed to find the real reason for these errors. The potential causes are incorrect forecasting method, incorrect parameters usage, or data discrepancy issues. Different data, parameter and methodology options are applied one by one on all federal taxes to see how errors react.

- **To Suggest a Revision of Data or/and Parameter or/and Method to Forecast Federal Tax Revenues**

  The study recommends alternative approaches to predict the federal tax revenues of Pakistan. The suggestions are based on the outcomes of objective 2.

1.4. Organization of the Thesis

In this study, chapter 2 explores the existing literature on the topic. Here literature on theoretical issues, forecasting methods, and Pakistan's federal tax revenue forecasting structure are discussed in detail. Chapter 3 provides methodology and data details. Chapter 4 analyzes the data and parameter issues in the existing tax revenue forecasting structure. It also provides a detailed comparison of Buoyancy and alternative tax revenue methods. The most suitable method is suggested in that chapter. Finally, the conclusion and policy recommendations are provided in chapter 5. References to studies used in this thesis are furnished at the end.

CHAPTER 2

LITERATURE REVIEW

Pakistan's tax revenue forecasting is a critical area of research due to the country's persistent fiscal constraints, low tax-to-GDP ratio, and structural inefficiencies in revenue mobilization. As a developing economy, Pakistan faces significant challenges, including a narrow tax base, high reliance on indirect taxation, widespread tax evasion, and a substantial informal sector, all of which hinder accurate revenue projections. Given the country's recurrent fiscal deficits and dependence on external financial assistance, precise tax revenue forecasting is imperative for effective budget formulation, fiscal discipline, and macroeconomic stability. Furthermore, Pakistan’s engagement with international financial institutions, such as the International Monetary Fund (IMF) and the World Bank, necessitates credible revenue estimates to fulfill reform commitments and ensure sustainable fiscal management. Thus, a comprehensive analysis of tax revenue forecasting methods and challenges in Pakistan is essential for informing evidence-based policymaking and enhancing revenue mobilization strategies.

The prevailing method for tax revenue forecasting in Pakistan is the tax buoyancy approach, which measures the responsiveness of tax revenue to changes in national income. While this method is widely used due to its simplicity and ease of implementation, it has several inherent flaws. The buoyancy approach assumes a stable relationship between tax revenue and GDP, overlooking structural inefficiencies in tax administration, policy changes, and shifts in economic composition. Additionally, it does not account for discretionary tax measures, compliance rates, or the effects of tax exemptions and incentives, leading to potential overestimations or underestimations of revenue projections. Beyond buoyancy, other forecasting methodologies, such as time series econometric models, computable general equilibrium models, and artificial intelligence-based techniques, offer more sophisticated analytical capabilities. However, these approaches also face challenges, including data limitations, economic volatility, and political influences. The instability of key macroeconomic variables, such as inflation, exchange rates, and sectoral growth, further complicates revenue estimation, reducing the reliability of traditional forecasting models. These challenges underscore the need for a rigorous evaluation of forecasting methodologies to identify the most effective approaches that can enhance predictive accuracy, support fiscal planning, and contribute to the overall stability of Pakistan's economy.

There is an enormous literature available on different dimensions of tax revenues. Since this study is focused upon methodological, data, and parametric issues in tax revenue forecasting structure, this chapter explored them in detail. Methodology issues are addressed in many studies but relatively fewer studies can be found on data and parametric issues in tax revenue forecasting. Along with these statistical problems, administrative issues and losses (statistical and fiscal) are also presented. More specifically, the chapter reviews the literature on the following topics. 1) Tax revenue forecasting methods 2) Studies on data and parameter issues in tax revenue forecasting 3) Studies on administrative issues in tax revenue forecasting 4) Statistical and fiscal losses due to tax revenue forecast errors and lastly 5) Studies on tax revenue forecasting of Pakistan.

2.1. Tax Revenue Forecasting Methods

There are a large number of methods suggested and implemented for tax revenue forecasting. Each economy forecasts taxes considering its economic, social, and organizational indicators. Moreover, academics and researchers have tried many methods which provide better estimates than classical methods. This section reviews the different methods suggested/used for tax revenue forecasting. Forecasting techniques are divided into groups according to their structure and usage feasibility.

2.1.1. Theoretical/ Orthodox Methods

Theoretical or orthodox methods of tax revenue forecasts refer to the methods build specifically for tax forecasting. These are practiced by most economies and also recommended by international economic and financial organizations.

Jenkins et al. (2000) provided a brief guide on the importance and practice of tax revenue forecasting. They covered forecasting methods involving bases like the buoyancy approach and the elasticity method. Following these, they also mentioned GDP involved methods and micro simulation techniques for predicting tax revenues at an economy level. IMF, in its paper IMF (2013), presented four methods for tax revenue forecasting; the effective tax rate method, the marginal tax rate method, the elasticity approach, and the regression approach (Firdawss & Karim, 2018). FBR uses the buoyancy method for forecasting taxes. These five approaches are categorized as theoretical/Orthodox methods to forecast tax revenues.

Bayer (2013) described the elementary technique for predicting tax revenues in terms of the macroeconomic methodology and from a microeconomic viewpoint. The microeconomic methodology used micro simulation approaches based on tax elasticities. The macroeconomic approach explained the basic features of regression analysis and time series analysis. It also described methods for evaluating the quality of predictions and the distribution of basic statistical error estimates.

The theoretical methods are also adopted by some researchers in Pakistan too. Although official tax revenue forecasting institutes, i.e., FBR itself, practice the buoyancy approach, a theoretical method to forecast. Qasim & Khalid (2016) did an accuracy analysis of tax receipt forecasts of Pakistan. The paper identified that there was a substantial gap between the budget estimates of Federal Revenue receipts and actual revenue receipts. This gap also holds for different constituents of revenue receipts. They also found that revenue forecasts were as bad after 1990 reforms as before them. There were no improvements found.

Firdawss & Karim (2018) applied the ETR approach to forecasting the federal tax revenue of Morocco for the years 2017 and 2018. When calculating the ETR, the authors observed that this rate was stable during the period, so it could be used to predict future tax revenues. Gumbo & Dhliwayo (2018) used ETR to forecast the VAT revenue of Zimbabwe for 2012 and 2013. The study found that Exponential Smoothing had relatively lower forecast errors than ETR. Firdawss & Karim (2018) applied the MTR approach to forecasting the tax revenue of Morocco for the years 2017 and 2018. When calculating the MTR, they noticed that this rate was not established during the period, so it could not be used to predict future tax revenues.

Gumbo & Dhliwayo (2018) used the Elasticity Approach to forecast the VAT revenue of Zimbabwe for 2012 and 2013. The study found that Exponential Smoothing had lower forecast errors than the Elasticity Approach. Firdawss & Karim (2018) applied an elasticity approach to forecast the tax revenue of Morocco for the years 2017 and 2018. When calculating the elasticity, they notice that this rate is stable during the period, which means it can be used to predict future tax revenues.

IMF suggested regression models were also used by some studies in the past. However, most of them are history now except Box-Jenkins, VAR, Exponential smoothing etc. Some old studies are mentioned here. While a large number of regression studies are discussed in univariate and multivariate literature sections. Reed (1983) used two regression models to forecast Indiana's tax revenues. The first model was estimated using two-stage least squares as the estimation technique for those model components that exhibit simultaneity. The second model was an adaptation of the first model. The only significant difference came in the way the behavioral equations for the sales and income tax bases were estimated. The variables in the second stage equations of the 2SLS model were earned over directly to the behavioral equations estimated using ordinary least squares. Grizzle & Klay (1994) used the linear regression method to forecast the tax revenues of 28 US states**.** The study found that the regression approach provides more accurate forecasting when there is a trend in a time series than a series with no trend. Firdawss & Karim (2018) applied a regression approach to forecast the federal tax revenue of Morocco for the years 2017 and 2018.

2.1.2. Statistical Methods

2.1.2.1. Univariate Approaches

Grizzle & Klay (1994) used the moving average method to forecast revenue using three years' data. The performance of the moving average method was quite bad, producing an overall mean absolute percentage error (MAPE) of 15.34 percent, about three times as large as the benchmark forecast. Grizzle & Klay (1994) applied the average change method to forecast the revenue of 28 US States. The method over casted the forecast by 4.4 % with a MAPE of 15.34. Grizzle & Klay (1994) used the curve fitting method to forecast the tax revenue of 28 US states. It was the most resource-intensive forecast but produced the lowest MAPE (6.84) in comparison to other extrapolative methods.

Bernard & Dent (1979) demonstrated an econometric model for tax forecasting using ARIMA methods to forecast several exogenous personal income modules. This approach was validated for the State of Iowa. Urrutia, Mingo, & Balmaceda (2015) established an ARIMA(0,1,0) model to estimate the Income Tax Revenue of the Philippines for the years 2014 to 2020. Forecasts attained from the model with a 99% coefficient of determination exposed that there was no significant difference between the actual values of Income Tax Revenue inspected through Paired T-Test.

Makananisa (2015) used different time series models and forecasted a few central taxes in the South African Revenue Service (SARS). The data used was monthly from January 1995 to March 2010 for Personal Income Tax (PIT), Value Added Tax (VAT), and Total Tax Revenue (TTAXR). The CIT data was transformed from monthly to quarterly frequency to cope with higher instability and negative values. The findings revealed that the SARIMA model performed well in modeling and projecting PIT and VAT. On the other hand, the SARIMA model underperformed in modeling and estimating the volatile CIT and TTAXR.

Molapo (2018) and Molapo et al. (2019) tried to forecast South Africa's significant tax revenues using the Autoregressive Moving Averages (ARIMA) model with data from 1998Q1 to 2012Q1 and compared the findings with a class of forecasting methods. While reaching on the basis of root mean square errors, the results did not prove the precision of the ARIMA model for forecasting significant tax revenues.

Streimikiene et al. (2018) forecasted the total tax revenue of Pakistan for the FY2017 using ARIMA models. The monthly data used for this study was from July 1985 to December 2016. The results confirmed that the ARIMA model did better tax revenue forecasting than other proposed models.

Ofori, Fumey, & Nketiah-Amponsah (2020) forecasted Ghana's Value Added Tax Revenue by comparing two methods. The first method was ARIMA with Intervention and the second method was the Holt linear trend method. Monthly VAT revenue data from 2002 to 2019 was used in the investigation. The findings showed that ARIMA with the Intervention method outpaced the Holt linear trend model in precision.

Zaw et al. (2020) used ARMA to forecast tax collection forecasts of Myanmar. The predictions were then compared with the official forecasts of the Myanmar Internal Revenue Department (IRD). The results were more accurate than official forecasts. Reed (1983) used the Box-Jenkins methodology to forecast the state tax revenue of Indiana. The models were estimated based on data through the second quarter of 1980. Then the model was used to prepare two series of forecasts. The first was a series of eight one-quarter ahead forecasts, while the second was a series of five four-quarter ahead forecasts. These forecasts were then equated with the actual data values from 1980Q3 through 1982Q2. The four-quarter ahead predictions of total taxes were all less than the substantial revenues realized. The errors ranged from $24,293 million to $133,691 million.

Tasi'u et al.(2024) focused on modeling and forecasting Nigeria's tax revenue using SARIMA and Holt-Winters models. Results showed the SARIMA model outperformed the Holt-Winters model in minimizing evaluation criteria, with an RMSE of 0.1654 and MAE of 0.0816. The study recommended applying the SARIMA model for accurate tax revenue forecasting.

Nandi et al. (2014) used the Box Jenkins methodology to forecast the total tax revenue of Bangladesh. The dataset consisted of one hundred and one observations from July 2004 to November 2012. The only strange discrepancy in one-step forecasting occurred in June 2008. During this period, an enormous surplus in income tax had contributed a larger share of this surplus in revenue as the provisional caretaker government of Bangladesh obligated many taxpayers to return due to tax payments.

Grizzle & Klay (1994) used the exponential smoothing method to forecast the tax revenue of 28 US states. The exponential smoothing method underforecasted the overall tax revenue by 90 percent. On the other hand, the linear, exponential smoothing method tended to over-forecast about as often as it under-forecast.

Cirincione et al. (1999) scrutinized the effect of the time series estimation method selection, the data length, and the data frequency on forecasting accurateness. The study forecasted non-tax revenue for six Connecticut municipalities. The study concluded that exponential smoothing models were the most precise. They also figured that local governments should use bi-monthly data for more accurate results.

Gumbo & Dhliwayo (2018) used the exponential smoothing method to forecast the VAT of Zimbabwe for the years 2012 and 2013. Exponential Smoothing had the lowest absolute forecast errors relative to other forecasting methods. Molapo (2018) and Molapo et al. (2019) tried to forecast South Africa's main tax revenues using the State Space exponential smoothing (ETS) model. The study used quarterly data from 1998Q1 to 2012Q1 and compared the results with the Bayesian vector autoregression (BVAR) method. The ETS appeared to be a suitable method for total tax revenue forecasting, as it overtook the BVAR method. Pavlik (2011) focused on forecasting income tax in cases where the time series was short. He used GARCH, EGARCH, and ARIMA. He applied these methodologies on an annual quarterly basis. EGARCH proved to be the best prediction method.

Mukherjee & Bhattacharya (2023) aimed to fill the gap in revenue forecasting literature in India by developing a conditional model for Corporate Income Tax (CIT) collection. The model addressed seasonality issues using a univariate Seasonal Auto Regressive Integrated Moving Average (SARIMA) model. The VAR model outperforms the univariate SARIMA model in terms of out-of-sample RMSE and Theil Inequality Index. Both VAR-based and univariate models had similar predictive powers, with potential for further improvement in future research.

2.1.2.2. Multivariate Approaches

Krol (2010) used the BVAR model to predict the tax revenue of California using the three most considerable taxes: sales tax, income tax, and corporate tax revenue. The BVAR model outperformed this study's VAR and random walk models based on root mean square errors. Yu et al. (2015) used an error correction mechanism (ECM) to improve accuracy using a single forecasting model. The study used 30-year data on Chinese tax revenue. It forecasted tax revenue using the LS-SVR model with and without ECM. The findings showed that the model with ECM gets superior results to those without ECM.

Streimikiene et al. (2018) used Vector Autoregression (VAR) model to predict the federal tax revenue of Pakistan for the years 2016 and 2017. The VAR forecasts were inferior to those calculated by the ARIMA model. Groening et al. (2019) used co-integration analysis to recognize factors that affect company tax revenue in Swaziland. Moreover, it used ARIMA and VAR to discover a forecasting method that will yield the least forecasting variance for CIT revenue. Combined forecasting was proven to cause the minor variance for one year ahead of the company tax revenue forecast.

Molapo (2018) and Molapo et al. (2019) tried to forecast South Africa's primary tax revenue using Bayesian Vector Autoregressive (BVAR) model for the data from 1998Q1 to 2012Q1. Based on RMSE, the results proved the correctness of the Bayesian Vector Autoregressive (BVAR) model for forecasting significant tax revenues. Carmody & Wiipongwii (2018) applied BSTS and compared it with the Bayesian VAR model. BVAR showed better results than BSTS.

Ajisola & Segun (2023) used Nigerian monthly data from the Federal Inland Revenue Service (FIRS) between 2010 and 2021 to analyze three models: Multivariate Linear Regression (MLR), seasonal Autoregressive integrated Moving Average (SARIMA), and Multi-variate Long Short-Term Memory Networks (LSTM). Both LSTM and MLR models performed better due to the ability to predict multiple independent variables. The LSTM model had a R2 score of 98.9% and an adjusted R2 score of 98.8%.

Alexi et al. (2024) presented an agent-based model using genetic algorithm-based decision-making to simulate an economy with taxation policy dynamics. The model assessed state tax revenues from regions or cities, introducing unique consumers and businesses. The model's efficacy was demonstrated on a small village, resulting in a mean relative error of 4% over 4 years. It also evaluated the effect of different taxation policies on economic activity and tax revenues.

2.1.3. Machine Learning Methods

This section reviews machine learning methods used for forecasting, including tax revenues and other variables.

Zhilou & Hua (2010) proposed an intelligent forecast system. They built a framework of different machine-learning methods. Their system could use noise data, get real-time prediction and model learning on line. Moreover, they established a forecasting model with learning function. So the model could be accustomed timely according to the data in the past to improve prediction accuracy. The predictions were quite accurate.

Ticano et al. (2017) predicted income tax revenues of Brazil. This work presented a hybrid model based on Genetic Algorithms (GAs) and Neural Networks (NNs) for a multi-step prediction of tax revenue collection. The outcomes were more precise than the findings that the Federal Revenue of Brazil's Secretariat (RFB) projected with the indicators method. The forecast results using dependent and independent variables were divided into two parts: (i) in 2013 (validation period), there was obtained a mean absolute percentage error (MAPE) of 2.37 percent and a decline of the relative error of 11.38 percent to 0.49 percent; (ii) in 2014 (testing data set) a decline of relative error of 10.82 percent to 3.51 percent was attained.

Simonov & Gligorov (2020) compared statistical, machine learning, and ensemble methods to forecast the Republic of North Macedonia customs. The study found that neural network autoregression (NNAR) predicted more accurately than statistical and ensemble methodology. Carmody & Wiipongwii (2018) used Support vector regression (SVR) to find the best parameters for the model accounting for epsilon and associated a cooverfitting in the model. Carmody & Wiipongwii (2018) also used the neural network process by implementing a method called 10-fold cross-validation. The model showed a promisingly small RMSE than SVR. Jang (2019) presented a supplementary tax forecast system that was based on artificial neural network. The system can help experts to forecast tax revenues proficiently.

Machine learning shrinkage methodologies, i.e., LASSO regression, elastic net, and ridge regression, were adopted by several researchers to forecast many macroeconomic factors. Waciko & Ismail (2020) used these shrinkage methods to forecast and compare the GDP of India and Indonesia. Under the MSE criteria, elastic net performed better than LASSO and Ridge Regression. Nakroji & Aminu (2022) used different machine learning techniques for inflation forecasting of Nigeria. They found ridge regression and artificial neural networks to be the best methods compared to LASSO, elastic net, and PLS.

Despite mounting anticipations and governments' growing use of artificial intelligence, academic research on AI applications in public administration has lagged. These gaps were rectified by Chung, Williams, and Do (2022), who concentrated on local governments' revenue estimates. They examined the accuracy of revenue forecasting by multiple ML algorithms and traditional forecasters, as well as the effectiveness of several ML models in predicting income for local governments. Their findings showed that, generally, traditional statistical forecasting techniques performed better than ML systems. However, KNN, one of the ML algorithms, was better at forecasting property tax income. The main benefit of this outcome was that it allowed public managers in local governments to deal with impending budgetary difficulties by making more precise income projections.

Noor et al. (2022) evaluated the forecasting performance of linear regression, random forest, and Feed Forward Neural Networks (FFNN) for Malaysian tax revenue projections. Following a series of tests, it was discovered that random forest and FFNN achieved the maximum accuracy. In terms of precision, linear regression did not do well. As a result, it was determined that the federal government revenue dataset was not a viable fit for this strategy.

Larson & Overton (2024) explored the question: What was the most effective method for forecasting sales tax revenue? Previous research primarily compared machine learning methods with traditional forecasting approaches, often overlooking the critical role of data pre-processing before applying forecasting models. This study revealed that machine learning did not always enhance forecasting accuracy. Instead, while modeling decisions influenced outcomes, their impact was smaller than previously believed. In contrast, pre-processing played the most crucial role in improving accuracy, highlighting the need for forecasters to understand the distinct characteristics of time series data. The key takeaway was that practitioners should have prioritized data pre-processing over model selection to achieve better results in sales tax revenue forecasting.

2.2. Studies on Data and Parameter Issues in Tax Revenue Forecasting

Along with other issues, data revision and parameter selection also affect the quality of tax revenue forecasts. This part provides a survey of these issues in detail.

Keene and Thomson (2007) examined the New Zealand Treasury's tax revenue errors for aggregate and disaggregated individual taxes. The study focused mostly on yearly one-year budget estimates, which are often based on comparing previous tax collections to growth rates in associated macroeconomic indicators such as GDP. The primary goal of the research was to identify the major drivers of tax revenue estimate inaccuracy and to identify opportunities for methodological improvements. Individual tax revenue projection mistakes were initially divided into two categories: errors caused by forecasting the macroeconomic factors used as a proxy for the tax base and errors caused by forecasting the tax ratio, or the ratio of tax revenue to proxy tax base. The tax ratio was further broken down into a component error resulting from anticipating the tax ratio trend and random error. According to the paper, the biggest source of tax revenue under-forecasting is an underestimation of the macroeconomic factors utilised as tax base proxies. The tax ratio projections were typically impartial, although less exact than the macroeconomic forecasts.

Mikesell (2018) found the boundaries of econometrics, the importance of disaggregation, the significance of tax administrators, the utility of causal models, the issue of data problems, the need to understand tax structure, the importance of consensus forecasts, the terror of recessions, and the reality of being wrong. In the Indiana consensus system, the experience provided tremendous respect for public servants seeking to make a sustainable fiscal system function. It contributed to making the revenue forecast binding in the budget process.

Göttert & Lehmann (2021) disentangled tax revenue forecast errors into influences originating from wrong macroeconomic assumptions and false predictions of the elasticities linking the tax base to its corresponding tax type. Across six tax types and the overall tax sum for Germany, they found a heterogeneous degree of the relative significance of both sources. Whereas wrong macroeconomic assumptions matter most for profit-related taxes and the wage tax, inaccurate estimates of the elasticities mainly drive the forecast errors of the energy and sales taxes. For the overall tax sum, more than two-third of the error was due to wrong macroeconomic predictions and approximately one-third to false assumptions on the elasticity. The results suggested that outsourcing the macroeconomic predictions to an independent forecaster and methodological improvements can reduce tax revenue forecast errors.

Stinson (2006) examined some of the causes for the large forecast variances from forecasts made for the state of Minnesota. Key factors identified include an inaccurate estimate for U.S. economic growth; inadequate, untimely, and inaccurate data; imperfect models; and unrecognized changes in the economy's structure. These factors came together and reinforced each other, ultimately producing a larger reduction in state revenues than anticipated.

Hannon et al. (2016) investigated the Department of Finance's tax revenue forecasting performance from 1997 to 2014. While the overall forecasting framework follows normal international practice, international standards discovered that predicting mistakes were rather substantial. An analysis of forecast mistakes shows significant contributions from sources other than errors in macroeconomic forecasting or calculating the previous year's revenue outturn. This showed that a formal examination of certain procedures and assumptions by the Department may lead to additional advances in predicting performance.

Mikesell (2018) found parametric and data issues along with administrative issues in the Indiana tax forecasting system. Göttert & Lehmann (2021) discovered certain parametric difficulties, such as incorrect macroeconomic assumptions and inaccurate forecasts of the elasticities that link the tax base to the relevant tax type. They discovered that the proportional relevance of both sources varied between six tax categories and the overall tax amount for Germany. They suggested outsourcing of forecasting for accurate predictions.

2.3. Administrative Issues in Tax Revenue Forecasting

This section reviews the development of tax revenue forecasting theory and the issues economists and econometricians face during forecasting.

Firstly, Shkurti (1990) briefly surveyed five tax revenue forecasting issues. The study also reviewed the status of revenue forecasting in the United States. In conclusion, it discussed the limits of revenue forecasting and good revenue forecasting practices. Bretschneider & Gorr (1992) explored the elements that influence fixed bias when estimating state sales tax revenue. An exploratory empirical research using survey data revealed indications of bias in predicting predictions due to political and policy influence. There was additional evidence that institutional improvements associated with 'excellent management' practices influence prediction bias.

Auerbach (1996) considered the problem of implementing dynamic revenue estimation methods. The study described the contemporary revenue estimation procedure and how these estimates were used in the development of policy. It then evaluated this procedure as an exercise in forecasting, bearing in mind its methodology and its performance. Finally, Auerbach (1996) related the choice among various revenue estimation procedures to the political process and the revenue estimates used.

Rubin et al. (1999) focused on revenue forecasting and estimation. The study described institutional responsibilities, techniques, problems, and issues associated with forecasting and estimation. In addition to revenue forecasting, the study also estimated the impact of proposed tax changes. Penner (2008) found that forecasts of federal revenue play a vital role in shaping the national debate over upcoming spending and tax policy. The forecasts of federal tax revenue are often incorrect. According to the study, the main issue is the complexity of the forecasting process. The forecasting procedure involves many steps and inaccuracies that are likely to happen. Many minor errors occurring in the same direction can make a forecast look useless.

Danninger & Kyobe (2005) examined interference and timeliness in the revenue-forecasting process using data from low-income nations. The data overwhelmingly supported the concept, and the findings were resilient to alternative explanations. The article also created three indices: transparency, formality, and organisational simplicity. More transparent and straightforward forecasting techniques result in early budget restrictions, whereas formality has no discernible influence.

Couture & Imbeau (2009) sought to establish and explain the difference between speech and action by comparing revenue predictions stated in Budget Speeches to actual revenues reported in provincial public accounts in Canada from 1986 to 2004. The regression analysis revealed that the mistakes were due to uncertainty. Furthermore, reliance on federal funds has ambiguous consequences. It resulted in underestimating during fiscal liberalism and overestimation during budgetary restriction. They also discovered that revenue projection is susceptible to political influence. Revenue is consistently overstated during election years, and right-wing administrations have recently greatly understated their revenue. Finally, when anti-deficit legislation exists, revenue forecasting mistakes are reduced.

Buettner & Kauder (2010) examined the practice and performance of revenue forecasting in many OECD nations. Cross-country disparities in revenue forecasting performance were shown to be connected with uncertainty regarding macroeconomic fundamentals. Differences in anticipated time were found to be significant. However, revenue forecasting's independence from potential government manipulation has a large favourable influence on revenue prediction accuracy.

Glenday (2013) first outlined the tenacities of tax revenue forecasting. Second, it discussed the elementary ideas and concerns involved with tax revenue, estimation, and growth. Third, it described the kinds of models used for analyzing and forecasting many types of tax revenues. Finally, it discussed the administrative preparations needed to support tax revenue forecasting.

Buettner and Kauder (2015) investigated if and how political biases emerge in an institutional framework in which revenue forecasting is entrusted to a group of experts from organizations that are not part of nor controlled by the government. The empirical investigation focuses on the performance of German federal tax revenue estimates, for which external experts' opinion has long been employed as an institutional safeguard. While income estimates were generally impartial, the outcomes revealed that the government had some impact. Specifically, optimism/pessimism in the government's GDP prediction helps to explain why the revenue forecast is overly optimistic/pessimistic. Furthermore, governmental projections of the revenue implications of tax-law changes were discovered to lead to prediction errors.

Mikesell (2018) reviewed the limits of econometrics, the implication of tax administrators, the utility of causal models, the importance of disaggregation, the issue of data problems, the need to understand tax structure, the significance of consent forecasts, and the reality of being wrong. Shirinov (2021) pointed out the issues in the "forecasting tax revenues" concept. He offered his own definition of this concept. Based on the study of the concept of forecasting tax revenues, Shirinov (2021) made theoretical conclusions.

Some studies highlighted administrative issues in Pakistan’s tax revenue scenario.Vazquez (2015) found that Pakistan's tax system needs improvement, especially in raising cash. Tax fraud is common, because personal and corporate income taxes and the GST have narrow bases. In 2014-15, the tax-to-GDP ratio had dropped dramatically. Most political party manifestos propose big expenditures and revenue increases. Pakistani officials needed a comprehensive policy reform agenda to raise money. The study proposed a complete tax structure and administration review.

Amin et al. (2014) empirically established Pakistani direct and indirect tax collecting factors. Empirical testing found that tax collection was affected by tax to GDP ratio, corruption, political instability, trade openness, real per capita income, and inflation. Regression analysis employed 1980–2010 time series data. International Country Risk Guide (ICRG), Handbook of Statistics, and Polity to Project IMF provided all variables. The Autoregressive Distributed Lagged (ARDL) model was used to determine short- and long-term model dynamics. Corruption, inflation, and political instability reduced tax revenue, but trade openness and real per capita income increased it.

Zafar (2017) indicated that Pakistan's tax system needs major modernization due to its broken processes, confusing legislation, and arbitrariness and discretion. Tax policy and administration have the biggest room for improvement. The restricted tax base has to be broadened and made more egalitarian by removing special treatment for some sectors of society, which results in minimal income from agricultural, immovable property, and capital gains taxes. Tax administration must also eliminate corruption and inefficiencies. These factors have long impeded tax reform in Pakistan.

Munir and Sultan (2018) discovered that direct taxes have a long-run beneficial relationship with economic development. Sales tax, tariffs on foreign commerce, and other indirect taxes have a favourable long-term and short-term influence on Pakistan's economic growth. The government should raise direct taxes by broadening the tax base. Indirect taxes often have a negative impact after one to two years; as a result, the government should reduce its reliance on indirect taxes. The government should encourage tax knowledge among the general public to boost tax morale and expand the tax base.

2.4. Statistical and Fiscal Losses due to Tax Revenue Forecasting Errors

This section is an overview of the studies that explored the impact of tax revenue forecast errors in terms of statistical and fiscal losses.

Gentry (1989) investigated whether state revenue predictions include all relevant economic and political data. The study formalised the idea of rationality and then evaluated it using time series data from New Jersey's key revenue sources. The findings revealed that most revenue sources' projections had a negative bias and did not properly use the information provided to the forecaster. Some research suggested that the magnitude of this bias was positively connected to the uncertainty of the revenue source.

Voorhees (2006) found that underestimating revenue forecasts leads to less loss for forecasters than overestimations. Furthermore, with the increased use of national forecasting organisations that produce economic projections based on revenue estimates, a secondary source of forecaster bias may exist in many state-level forecasts. This theory was backed by a rise in the number of states that use such organisations, as well as a reduction in the yearly mean percentage state prediction error standard deviation.

Becker & Buettner (2007) empirically evaluated tax revenue forecasts in Germany. An empirical examination of predictions from 1971 to 2006 investigated their unbiasedness and efficiency. The findings led to impartial and typically efficient tax revenue predictions in Germany. Uncertainty over the growth rate accounts for a significant portion of projection mistakes. However, variations between the government's growth estimates and the forecasts of independent research organisations had no substantial consequences. Only the electoral cycle's outcomes revealed the possibility of progress.

Krol (2013) investigated the accuracy of state revenue forecasts using a flexible loss function. Using Californian data, the study came to the same result of rejecting rationality. However, the rejection might be due to an asymmetric loss function. After accounting for the asymmetry of the loss function with a flexible loss function, the study discovered that under-forecasting California's revenues is less expensive than over-forecasting. The research also discovered that the prediction errors that account for this imbalance were independent of the information available at the time of forecasting. These results revealed that a failure to adjust for probable asymmetry in the loss function in prior studies might have caused misleading results.

Chatagny & Siliverstovs (2013) added to the empirical literature on budget forecasts by presenting fresh data on Swiss cantons. Using data from 26 Swiss cantons from 1944 to 2010, the study used the approach Elliott et al. (2005) to examine the reasonableness of direct tax revenue estimates. The study principally discovered that 1) when considering the percent prediction inaccuracy, loss functions were asymmetric in a majority of cantons, and 2) allowing for asymmetric losses, findings of rationality tests were significantly altered in the sense that more cantons turned out to provide reasonable forecasts. 3) When evaluating projections of growth rates, absolutely little evidence of an asymmetric loss function was discovered. Finally 4) Forecasts of tax revenue growth rate proved to be sensible in a larger number of cantons than tax revnue levels.

Chatagny (2015) investigated the link between the finance minister's ideology and tax revenue prediction inaccuracies, and analysed how the strictness of fiscal laws affected this relationship. The study employed a panel dataset of twenty-six Swiss cantons from 1980 to 2007, as well as a novel dataset of ninety-nine canton-level finance ministers. The empirical technique considered the fact that the distribution of departments to elected politicians is a random procedure from the voters' perspective. The article found a counterintuitive beneficial influence of the finance minister's ideology on tax revenue forecasting mistakes, with a more left-wing finance minister producing comparatively more conservative estimates. They also discovered that fiscal laws mitigated the impact of ideology on tax revenue prediction inaccuracies. These findings showed that left-wing finance ministers need to reduce deficits more than right-wing finance ministers in order to convey the same degree of competence to voters.

Wawire (2007) sought to identify the determinants of Value Added Tax income and evaluate yield responses to changes in bases and reforms. The major conclusions of the study were derived from Paul Samuelson's (1955) work "Fundamental General Equilibrium Analysis of the Public Sector". The estimation findings revealed that Kenya’s GDP elasticity is lower than elasticities with regard to monetary GDP, implying the presence of an underground economy. It was also shown that tax yields responded with significant delays to changes in their drivers, whereas revenues were sensitive to reforms and unique conditions. The study revealed that Kenya's VAT income was highly responsive to changes in its determinants, particularly foreign commerce.

Ademmer & Boysen-Hogrefe (2022) evaluated the effect of errors in medium-term tax revenue predictions on the ultimate budget balance. The research was based on budget statistics from all German states. They discovered that forecast errors at different forecast horizons had a significant impact on the ultimate budget balance, implying that expenditure plans were only modestly changed when revenue predictions were altered. As a result, inaccuracies in tax revenue predictions had a significant impact on the growth of the public debt. The findings revealed that a considerable portion of German states' overall debt is the consequence of overly optimistic forecasts.

2.5. Studies on Federal Tax Revenue of Pakistan

This section is a review of the methods used for tax revenue forecasting in Pakistan. The Federal Board of Revenue (FBR) is the official body that forecasts the tax revenues of Pakistan. The Strategic Planning Reforms & Statistics (SPRS) wing of FBR forecasts the head-wise tax revenue estimates based on buoyancy estimates from past years on a rolling basis. Buoyancy estimates are found using the past collection and relevant base data(FBR, 2021).

Qasim & Khalid (2016) investigated the accuracy of Pakistan's federal revenue receipt budget estimates. The analysis found a significant difference between budget predictions of federal revenues and actual revenue receipts. This disparity also exists for several components of revenue collections. Additional research was conducted on changes implemented before to and following the 1990s. For a variety of factors, revenue forecasts did not improve following the revisions. The revisions did not include revenue forecasting as one of their aims. Secondly, the estimations were strongly dependent on the accuracy of macroeconomic projections and buoyancy indicators. This investigation produced policy recommendations for strengthening the predictability of fiscal tools.

Streimikiene et al. (2018) forecasted the federal tax revenues of Pakistan for FY2017 using three time series methodologies. The study used the Autoregressive Integrated

Moving Average model (ARIMA), the Autoregressive model (AR with seasonal dummies), and the Vector Autoregression (VAR) model. The study also analyzed the effect of indirect taxes on the working class and the time series models' efficiency. This study used monthly data from July 1985 to December 2016. The results of this study discovered that among the techniques used, the ARIMA model gave the best-forecasted values for the tax revenues of Pakistan.

Altaf et al. (2021) employed multiple linear and nonlinear models to estimate Pakistan's tax-to-GDP ratio. The data from 1991 to 2013 was analysed in this context, and forecasts were made until 2025. The models' performance was confirmed by comparing the levels of the forecasting errors. The investigation showed that the nonlinear model was adequate for forecasting the tax-to-GDP ratio. Furthermore, there was no significant difference between the actual values and those predicted by the model. As a result, the study pushed the policymakers and government to scrutinize Pakistan's tax revenue.

Hussain et al. (2014) forecasted the federal revenue collections and long-term total indirect tax revenues of Pakistan using different forecasting models. The Autoregressive Integrated Moving Average model proved to produce minimum forecast errors based on nonparametric Mean Absolute Deviation (MAD) test. The study also simulated tax data for next ten years i.e., 2011-2020.

Hussain & Siddiqui (2014) employed Minitab 16 to analyse many ARIMA models in order to determine the best forecast models for direct income tax and indirect federal revenue reports. It has been discovered that the ARIMA(1,1,1) and ARIMA(1,1,0) models are suitable forecasting models for federal indirect revenue and direct income tax, respectively.

Table 2.1: Studies on Federal Tax Revenue of Pakistan

Author name & Year	**Title**	**Model or Method used**	**Main contribution**
Hussain & Siddiqui (2014)	Time Series Forecast Models for Pakistan's Economic Activities in Past Four Decades	Autoregressive Integrated Moving Average model (ARIMA)	Found that ARIMA(1,1,1) and ARIMA(1,1,0) models are suitable forecasting models for federal indirect revenue and direct income tax, respectively.
Hussain et al. (2014)	Forecast simulation of tax collection in Pakistan	A number of Time series models	Autoregressive Integrated Moving Average model (ARIMA) proved to be best option to forecast indirect taxes. Simulated the data for next 10 years.
Qasim & Khalid (2016)	Accuracy of Revenue Forecast: Analysis of Pakistan's Federal Revenue		The analysis found a significant difference between budget predictions of federal revenues and actual revenue receipts.
Streimikiene et al. (2018)	Forecasting Tax Revenues using Time Series Techniques – a Case of Pakistan	Autoregressive Integrated Moving Average model (ARIMA), Autoregressive model (AR with seasonal dummies), and Vector Autoregression (VAR) model	Analyzed the effect of indirect taxes on the working class and the time series models' efficiency.
Altaf et al. (2021)	A mathematical forecasting model to estimate tax to GDP ratio	Linear and Non-linear models	No-linear model predicted most accurately.
FBR Reports	Evidence-Based Revenue Forecasting	Buoyancy approach	Did forecasts of all federal tax heads

	FY 2021-22, FY 2022-23, FY 2023-24		through Buoyancy approach

Source: Followed the pattern of Bhatti & Do (2019)

2.6. Conclusion

Tax revenue forecasting is a multifaceted task with consequences for fiscal management and policymaking. Researchers continue to investigate new methodologies and data sources in order to enhance prediction accuracy, particularly in the face of economic uncertainty and complicated tax systems. Accurate forecasting is still required for efficient governance and sensible budgetary planning.

This chapter reviewed the literature on Tax revenue forecasting methods, studies on data and parameter issues in tax revenue forecasting, studies on administrative issues in tax revenue forecasting, statistical and fiscal losses due to tax revenue forecast errors, and lastly tax revenue forecasting in Pakistan. The literature survey showed that many tax revenue studies of data, parameter, and method issues are available. Most of these researches are for advanced economies. An inclusive study of Pakistan’s federal tax revenue frecasting errors was needed. There should be a research that appraises the method, data and parameter discrepancy issues in the existing federal tax revenue forecasting mechanism. The current system needs better alternatives too. This thesis will fill this gap.

CHAPTER 3

RESEARCH METHODOLOHY

Pakistan is a federation in which the resources are collected at the center and then shared through the revenue sharing formula (NFC award). So any forecast error for revenue stream would result in mismanaged fiscal operations, mis-targeting targets, or additional debt burdens. Therefore, it is imperative to evaluate current forecasting practices for federal taxes in Pakistan.

As studies have noted that revenue forecasting errors are significant. This may be due to a large number of factors. Hannon et al. (2016) found significant contributions from sources other than flaws in macroeconomic forecasting revealed by an investigation of forecast errors. This demonstrated how the Department could make further progress in performance prediction by formally reviewing certain practices and presumptions. Göttert & Lehmann (2021) have shown that tax revenue forecast errors can be decreased by methodological changes and outsourcing the macroeconomic projections to a third-party forecaster.

For this study, we have focused on evaluating data, parameters, and methods issues. This chapter includes 4 parts. Firstly, the theoretical framework is discussed. In the second part, methodology is explained. In the third part evaluation criteria are discussed. While, the last part includes data and sources.

3.1. Theoretical Framework

As discussed in chapter 1, prevailing method of tax revenue forecasting is Buoyancy Approach. It mainly consists of following two steps.

Step 1: Calculation of Buoyancy

$$Buoyancy = \frac{\Delta T/_{T}}{\Delta GDP/_{GDP}} \quad (3.1)$$

Here,

$\Delta T/_{T}$ =Ratio of change in tax to previous year tax revenue

$\Delta GDP/_{GDP}$=Ratio of tax base(mostly GDP) to Previous year tax base(GDP)

Step 2: Calcualtion of tax revenue forecast through Buoyancy:

$$Tax_f = GDP\ grwoth\ target * Tax\ buoyancy \quad (3.2)$$

Here,

$$Tax_f = Tax\ rewvenue\ forecast\ for\ future\ year$$

$GDP\ grwoth\ target = GDP\ growth\ target\ calculated\ by\ \ Ministry\ of\ Finance$

$Tax\ buoyancy = Tax\ Buoyancy\ calculated\ in\ step$ 1

In the following sections, data, parameter and method discrepancies are checked by making ammendments to this frame work.

3.2. Methodology for Checking Data Discrepancy

As discussed in the last chapter, FBR uses provisional data on taxes, bases, and GDP growth targets. The study intends to check whether data revision can help improve different heads' tax revenue forecasts. If data revision has any impact on forecast improvement real time data should predict more accurately. On the other hand, provisional data should perform the worst.

The current FBR forecasting method (Buoyancy approach) is applied to the following data combinations for within-sample forecasts.

- **Case 1**

These forecasts use provisional taxes, provisional bases, and actual nominal GDP growth targets. According to data revision theory, these predictions should have the maximum errors.

- **Case 2**

Case 2 uses revised taxes, revised bases, and actual nominal GDP growth targets.

- **Case 3**

Case 3 utilizes final taxes, final bases, and actual nominal GDP growth targets.

- **Case 4**

Case 4 has the tax revenue forecasts using real time taxes, real-time bases, and actual nominal GDP growth targets. If tax revision affects forecasting, these forecasts should be more accurate.

- **Case 5**

These forecasts use provisional taxes, real time bases, and actual nominal GDP growth targets.

- **Case 6**

This case uses revised taxes, real time bases, and actual nominal GDP growth targets.

- **Case 7**

Case 7 consists of final taxes, real time bases, and actual nominal GDP growth targets.

- **Case 8**

Case 8 uses Provisional taxes, final bases, and actual nominal GDP growth targets.

- **Case 9**

Case 9 uses revised taxes, final bases, and actual nominal GDP growth targets.

- **Case 10**

Case 10 uses Provisional taxes, revised bases, and actual nominal GDP growth targets.

- **Case 11**

Case 11 has the tax revenue predictions using Final taxes, revised bases, and actual nominal GDP growth target.

- **Case 12**

Case 12 has the tax predictions using revised taxes, provisional bases, and actual nominal GDP growth targets.

- **Case 13**

Case 13 uses final taxes, provisional bases, and actual nominal GDP growth targets.

3.3. Methodology for Checking Parametric Issues

The thesis intends to check whether parameter revision can help improve different heads' tax revenue forecasts. Both real GDP growth targets and inflation rate targets are used as provisional, revised, and actual data along with real-time taxes and bases. If parameter revision has any impact on forecast improvement, then actual nominal GDP growth or actual real GDP growth or actual inflation rate targets should predict most accurately. On the other hand, provisional targets should perform the worst.

- **Case 1**

These forecasts use real-time taxes, real-time bases, and actual nominal GDP growth targets. If parameter revision has any impact, these forecasts must be optimum.

- **Case 2**

  Case 2 uses real time taxes, real time bases and provisional nominal GDP growth targets. If parameter revision hypothesis is true, these forecasts should have the most inaccurate predictions.

- **Case 3**

  Case 3 uses real time taxes, real time tax bases and revised nominal GDP growth targets.

- **Case 4**

  Case 4 uses real time taxes, real time bases, actual real GDP growth targets and provisional inflation targets.

- **Case 5**

  These forecasts use real time taxes, real time bases, actual real GDP growth targets and revised inflation targets.

- **Case 6**

  Case 6 utilizes real time taxes, real time tax bases, provisional real GDP growth targets and actual inflation targets.

- **Case 7**

  Case 7 consists of real time taxes, real time tax bases, revised real GDP growth targets and actual inflation targets.

- **Case 8**

  Case 8 is estimated using real-time taxes, real time tax bases, provisional real GDP growth targets and revised inflation targets.

- **Case 9**

  Case 9 uses real time taxes, real time tax bases, real GDP growth targets and provisional inflation targets.

3.4. Methodology for Method Selection

As previously stated, in Pakistan, forecasting errors were primarily on the high side, resulting in fiscal mismanagement. A tax revenue projection error causes over/under-booking of government spending, which, if reversed, can result in waste or under-provisioning. If not changed, it will result in further funding, placing pressure on the monetary side as well. Furthermore, given Pakistan's federative structure, any revenue estimate error would be carried over to the provinces via the revenue sharing channel.

One probable explanation for these frequent failures is the use of an incorrect forecasting approach over time. There are several alternative forecasting approaches available to improve tax revenue stream prediction accuracy. In this section, we propose potential alternatives to the buoyancy approach for forecasting federal tax income. These strategies were evaluated to determine which method produced the highest accurate predictions for each federal tax head.

Different forecasting methods are used on federal tax revenues data to check the method's accuracy. The methods are grouped as theoretical, statistical, and machine learning methods.

3.4.1. Theoretical/Orthodox Methods

Theoretical tax revenue forecasting methods are those that involve some fiscal or revenue theory/ logic in their estimation. In the current study, effective tax rate approach, marginal tax rate approach, and Buoyancy approach are used as theoretical forecasting methods. Following are brief reviews of these methods.

i. Buoyancy Approach

Tax revenue buoyancy means the total response of tax revenue to change in the tax base or proxy tax base. Buoyancy is given as:

$$Buoyancy = \frac{\Delta T / T}{\Delta GDP / GDP} \tag{3.1}$$

And tax revenue forecast through Buoyancy:

$$Tax_f = GDP\ grwoth\ target * Tax\ buoyancy \tag{3.2}$$

A tax is said to be buoyant if the proportionate change in tax revenue is more than the proportional change in the tax base or GDP. Buoyancy is a desirable property of any tax. It shows the revenue productivity of the tax system. As total income increases, tax revenue automatically follows it (Gumbo & Dhliwayo, 2018).

ii. Effective Tax Rate Approach

The effective tax rate (ETR) is the quantity of tax revenue as a percentage of the tax base. While using the ETR, one must consider some factors, like tax evasion and tax exemptions. We can say that a relationship occurs between the tax base and the tax revenue if the ETR is constant over time. Once ETR is established, we can utilize it to forecast tax revenue by multiplying the tax base with the tax rate. But, the forecast is controlled by the trouble of determining the tax base because we need detailed evidence to measure the development of different tax bases. Mainly since these data are not always accessible or published. Even if it is probable to get the tax base for several years, it is not possible to forecast it in every case.

For these causes, IMF (2013) suggests a tax base replacement to examine the tax revenue and forecast tax revenue. This tax base can be any economic or financial variable strictly associated with the actual tax base and for which data are obtainable. So, to forecast tax revenue, we first determine the ETR, which is defined as the tax revenue divided by the proxy tax base:

$$Effective\ tax\ rate = \frac{Tax\ revenue}{Proxy\ tax\ Base} \qquad (3.3)$$

The forecast of tax revenue is attained by applying the formula given below:

$$Tax(f) = Taxbase(f) * (ETR) \qquad (3.4)$$

Once we determine that the ETR is constant, we can forecast tax revenue by multiplying the forecasted tax base with the tax rate. If the effective tax rate is unstable, it can be replaced by the MTR, i.e., the marginal tax rate (Firdawss & Karim, 2018).

iii. Marginal Tax Rate Approach

The marginal tax rate (MTR) is the ratio of the change in tax revenue to the change in the tax base:

$$Marginal\ tax\ rate = \frac{\Delta Tax\ revenue}{\Delta Proxy\ tax\ base} \qquad (3.5)$$

If the MTR is constant, we forecast the revenues by multiplying the MTR by the forecasted tax base.

$$Tax(f) = MTR * Taxbase(f) \qquad (3.6)$$

iv. Structural Vector Autoregression (SVAR) Method

Christopher Sims proposed Vector Autoregression in 1980. A VAR model is the general/multivariate form of the univariate autoregressive model. It can tackle time series vectors. If the series is stationary, we forecast them using a VAR model to the data straight. If the series is non-stationary, differences of the data are taken to make them stationary. VAR is fitted after that. In these cases, the models are estimated using the OLS. Another option is that the series is non-stationary and cointegrated. A VAR with ECM term would be included in this case, and alternative estimation methods would be used. Forecasts are recursively produced from the VAR. The VAR model forecasts for all variables involved in the system. The fundamental VAR model is represented in the following reduced form equation based on Lütkepohl and Kratzig (2004).

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t \qquad (3.7)$$

Where

$y_t = N\text{x}1\ Vector\ of\ Variables$

$A_i = N\text{xN Matrix of Autoregression Coefficients}$

$\varepsilon_t = Unobserved\ error\ term$

To convert reduced form VAR to Structural VAR, Several parameter constraints are conceivable. In Cholesky's writing, the short-term limitation may be used to explain the sequence of shocks, While in Blanchard-Quah's long-term restriction can be used to describe the appearance of the shock.

In case of Pakistan, Streimikiene et al. (2018) forecasted the tax revenue of Pakistan for the fiscal year 2016–17 using VAR model. The result was worse than ARIMA while they compared the statistical errors.

3.4.2. Statistical Methods

Statistical forecasting methods use statistics to forecast future values based on historical data. These methods involve data and don't care about the economic/ fiscal theory. Regressions are few of those statistical methods. In this chapter, a univariate regression technique i.e., Box-Jenkins methodology and a multivariate approach i.e., Vector-Autoregression method, are used. The IMF sometimes uses the multivariate regression method to estimate the effect of tax base variables on tax revenue. The accuracy of this method depends on the strength of the relationship between the independent variables (Firdawss & Karim, 2018). Below is the description of both univariate and multivariate methods.

i. Box-Jenkins Method

The Box-Jenkins method was presented by G. Box and G. Jenkins in their combined book "Time Series Analysis: Forecasting and Control" in 1970. The method starts with the hypothesis that the process that produced the time series can be estimated using ARMA modeling if it is stationary or an ARIMA setting if it is non-stationary. This method has the following three steps;

a. Identification: In this step, the properties of the data are examined using statistical tests and visual aids like time series plots and autocorrelation plots. The objective is to determine the values of three parameters in order to find the right ARIMA model.

b. Estimation: In this stage, statistical methods such as maximum likelihood estimation are used to estimate the model's coefficients.
c. Diagnostic Checking: Now the fitted model is assessed to verify that it appropriately represents the data patterns. This includes checking the residuals (errors) for randomness and independence. If the model proves insufficient, improvements are performed, and the process loops back to identification.

Following Molapo (2018); Streimikiene et al. (2018); Molapo et al. (2019) and Ofori et al. (2020) we have applied the Box-Jenkins method on all federal tax heads.

3.4.3. Machine Learning Methods

The term'machine learning' was originally used in 1959 by Arthur Samuel, who was at the time working at IBM, and characterised it as the field of research that allows computers to learn without being explicitly programmed (Gutierrez, 2015). In this section, we have used three machine learning techniques to forecast federal tax heads of Pakistan.

i. LASSO

"LASSO" is the acronym for Least Absolute Shrinkage and Selection Operator. It is a statistical formula to regularize data models and feature choice. It is utilized in regression methods for more accurate predictions. It uses the shrinkage technique. Shrinkage is the point where values are shrunk to the mean. The LASSO is an extension of OLS, which adds a penalty to the Residual Sum of Squares (RSS) equal to the sum of the absolute values of the non-intercept beta coefficients multiplied by parameter λ that slows or accelerates the penalty. LASSO reduces the overfitting in regressions by shrinking the parameters to 0.

This regression is appropriate for models with high multicollinearity levels or when you want to mechanize some parts of model selection, like variable choice/parameter elimination (Lau, 2017). The mathematical equation of this regression is given as

$$\sum_{i=1}^{n}(y_i - \sum_{j} x_{ij}\beta_j) + \lambda \sum_{j=1}^{p} |\beta_j | \qquad (3.8)$$

Here

λ is showing the shrinkage.

- λ = 0 shows that all features are considered and it is equal to the linear regression.
- λ = ∞ shows that no feature is considered.

We have used all potential predictors of each tax head revenue. The LASSO automatically chooses the most relevant predictors while minimizing the variance.

ii. Ridge Regression

Ridge regression is also one of the machine learning shrinkage methods. Hoerl and Kennard (1970) proposed Ridge regression. The main difference between LASSO and Ridge regression is that LASSO shrinks the slope to zero while Ridge regression just asymptotically minimizes it close to zero. It is given as follows:

$$R = \sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p} \beta_j^2 \qquad (3.9)$$

Equations 6.8 and 6.9 show that LASSO regression takes the absolute magnitude of the coefficients, while ridge regression takes the squares. For comparison purposes, we have used ridge regression to get the tax revenue forecasts using all potential regressors.

iii. Elastic Net

The Elastic Net Regression method is a convex amalgamation of the Ridge Regression Penalty and Least Absolute Shrinkage and Selection Operator (LASSO) Penalty. Zou and Hastie developed it in 2005. As the name suggests, it is flexible in adopting the λ value. The Elastic Net = Ridge Regression when λ =1 and The Elastic Net = LASSO when λ = 0. This study follows Waciko & Ismail (2020) in using elastic net to forecast tax revenues.

Table 3.1: Comparison of LASSO, Ridge, and Elastic Net Assumptions

Method	**Assumes Sparsity**	**Handles Multicollinearity**	**Feature Selection**	**Assumes Linearity**	**Error Independence**
LASSO	Yes	Partially (selects one feature from correlated groups)	Yes	Yes	Yes
Ridge	No	Yes (shrinks all coefficients)	No	Yes	Yes
Elastic Net	Partial	Yes (keeps grouped features)	Yes	Yes	Yes

3.5. Evaluation Criteria

All combinations and forecasting methods are judged on the bases of statistical errors, statistical loss functions and feasibility to use.

3.5.1. Statistical Errors

These forecasts' performance is judged based on the root mean square error (RMSE) and mean absolute error (MAE). A brief description of these errors is given below.

Table 3.2: Errors Description and Formulas

Description	**Formula**
Root Mean Square Error The root mean square error (RMSE) is used as a standard statistical metric to measure model	$RMSE = \sqrt{\frac{\sum\left(Y_t^p - Y_t^A\right)^2}{n}}$

performance in different social sciences, including Economics.	Here $Y_t^A = Actual\ Data$ $Y_t^p = Predicted\ Data$
Mean Absolute Error(MAE) The mean absolute error (MAE) is another useful measure widely used in model evaluations. Many researchers recommended this evaluation measure due to the absolute nature of the error. Some prominent ones are Willmott and Matsuura (2005) and Willmott et al. (2009).	$MAE = \frac{\sum \lvert Y_t^p - Y_t^A \rvert}{n}$ Here $Y_t^A = Actual\ Data$ $Y_t^p = Predicted\ Data$

RMSE and MAE were selected as evaluation criteria due to their effectiveness in measuring prediction accuracy and error magnitude.

RMSE is particularly useful because it penalizes larger errors more heavily due to the squared term. This makes it ideal when significant deviations from actual values are critical. RMSE is widely used in regression tasks where minimizing large prediction errors is a priority.

MAE provides a straightforward measure of the average error magnitude without emphasizing large deviations. It is useful when all errors should be treated equally and provides a more interpretable metric, especially when dealing with datasets that may contain outliers.

By using both RMSE and MAE, we balance sensitivity to large errors (RMSE) with a robust measure of overall accuracy (MAE), ensuring a comprehensive evaluation of model performance.

3.5.2. Statistical Loss Functions

Numerous macroeconomic projections have been assessed using the rational expectations method. It is also known as symmetric loss function. Usually, this method works on the assumption that the forecast loss function is symmetric and quadratic. It is well-liked in the literature on prediction evaluation because it has the appealing feature that the conditional expectation is the best or rational forecast, implying that forecasts are objective (Elliott et al., 2005; Elliott, Komunjer, & Timmermann, 2008; Elliot,2013).

The statistical loss function would check the rationality of the econometric methods. It will check whether the method is biased or unbiased and comprehensive or not. The two parts are given as follows:

i.Unbiasness

Loss Function is given as

$$R_{t+h} = \propto + \beta F_h^t + \mu_t \qquad (3.12)$$

Null Hypothesis : $\propto$=0 and β=1 (unbiased forecast)

Here

R_{t+h} = The percentage change in tax revenues from period t to period t + h

F_h^t = The forecasted h-period-ahead percentage change in tax revenues in period t

Rejecting this joint hypothesis rejects the idea that the forecast is unbiased.

ii. Method Comprehension

Information loss is calculated as

$$\in_t = \gamma + \eta_1 X_t + \eta_2 X_{t-1} + v_t \qquad (3.13)$$

Null Hypothesis: $\eta_1=\eta_2=0$ = Insignificant (all information used)

X_t and X_{t-1} represent information available to the analyst at times t and t – 1.

Rejecting this null hypothesis means that information available to the analyst was not used and could have reduced the error if used.

3.5.3. Feasibility Analysis

All methods are evaluated on the bases of time, budget, and complication criteria. The skill sets of concerned department, organizational structure and sludges are also considered to suggest a forecasting method for each tax head.

3.6. Variables and Data Sources

Table 3.3 is the complete list of variables which are used in this study.

Table 3.3: Variables Description

Variable	Description
Total federal tax	Total federal tax revenue collection by FBR (million Rs.)
Direct tax	Yearly direct tax revenue collection by FBR (million Rs.)
Custom duties	Yearly custom duties revenue collection by FBR (million Rs.)
Federal excise duty	FBR Yearly federal excise duty revenue collection (million Rs.)
Sales tax	Yearly Sales tax revenue collection by FBR (million Rs.)
Nominal GDP (CFC)	Nominal GDP measured at current factor cost (million Rs.)
Non-Agr GDP (CFC)	Non-agricultural GDP at current factor cost (million Rs.)
Large Scale Manufacturing	Large Scale manufacturing Value addition in GDP (million Rs.)
Dutiable Imports	Imports on which duty is leviable (million Rs.)
Total Consumption	Total Consumption expenditure (million Rs.)

Real GDP growth Target	Real GDP growth estimate in percentage
Inflation Rate Target	Inflation rate estimate in percentage
Consumer Price Index	Consumer Price Index taking base price of 2010
Control Variables	A large number of control variables which affect tax revenue forecasting according to literature

Source: FBR (2021), Economic Survey of Pakistan, World Development Indicators, Annual plans of the Ministry of Planning, Development, and Special Initiatives

The data is secondary and yearly from 1984-85 to 2018-19. The data for nominal GDP growth targets is real-time. On the other hand, taxes and bases data are collected on different revisions i.e., provisional, revised, final, and real-time. The data is extracted from the official website of FBR, different issues of the Economic Survey of Pakistan, budget documents, and annual plans issued by the Ministry of Planning, Development, and Special Initiatives.

3.6. Conclusion

This chapter is a review of the methodology used in the dissertation. It consists of 3 parts. Firstly, the Methodology is explained. The methodology is divided into 3 subparts. In the first subpart, the methodology to check data discrepancies is discussed. Thirteen unique cases of data alternatives are found in this part. In the second subpart, the methodology to check parameter discrepancies is explained. In this part nine cases of alternative parameters are found. In the third subpart, different forecasting methods are discussed.

In the second part evaluation criteria are discussed. Three main criteria are used. These include statistical errors, loss functions and feasibility criteria.The last part consists of data and sources. In the following chapter, the dicussed methodologies are implemented on Pakistan's data to gain insights.

CHAPTER 4

RESULTS AND DISCUSSION

Tax revenue forecasting works as a base for budget estimates: the more accurate the tax revenue forecasts, the better fiscal decisions. Pakistan's economy has been suffering from poor fiscal decisions since independence. The tax revenue forecast errors were a stimulus for wrong budget allocations. The system of data collection and forecasting needs a critical diagnosis. Once real reasons for these loopholes are identified, solutions can be suggested accordingly.

This study focuses on three potential resaons of tax revenue forecast errors of Pakistan; data discrepancy, parametric issues and methodological issues. This chapter is the practical testing of these issues. The succeeding sections are detailed analyses with reference to different heads of federal tax revenue of Pakistan. In the first section, data discrepancies are tested. In the second section parameter discrepancies and in the last section, method selection errors are checked.

4.1. Checking Data Discrepancy

The current FBR forecasting method (Buoyancy approach) is applied to the different data combinations for within-sample forecasts.

4.1.1. Total Tax

Our hypothesis that data revision increases forecasting accuracy is not valid for total tax. As Figure 4.1 shows, all data types show similar patterns. It means provisional, revised and final data are quite similar to the real time data. According to table 4.1, RMSE, and MAE suggest different data combinations. Interestingly, they all consist of provisional, revised, and final data sets. RMSE is minimum i.e., 111538.3 for

FREF and MAE's optimum value of 65333.92 suggests PPF to be the best combination for total tax forecasts.

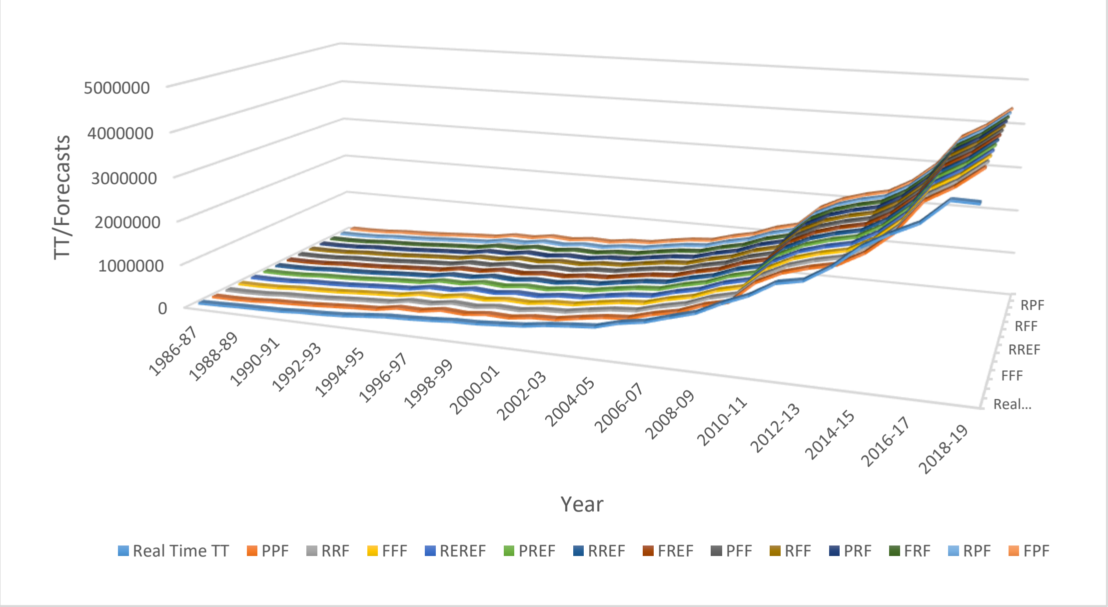


Figure 4.1:Real Time Total Tax and Forecasts

Table 4.1: Statistical Errors for Total Federal Tax Revenue Forecasts

CASE	RMSE	MAE
Case 1: PPF	113254	65333.9
Case 2: RRF	146079	81377.7
Case 3: FFF	111676	66710
Case 4: REREF	130979	71486.4
Case 5: PREF	112915	65171.8
Case 6: RREF	145948	81401.7
Case 7: FREF	111538	66691.9
Case 8: PFF	113223	65408.5
Case 9: RFF	146070	81421.1
Case 10: PRF	113228	65342.7
Case 11: FRF	111687	66666.3

Case 12: RPF	146086	81312.4
Case 13: FPF	111693	66600.8

Source: Authors' calculations

4.1.2. Direct Tax

Figure 4.2 depicts that all forecasts are slightly higher than real time direct tax but are in harmony with one another. Table 4.2 is showing that both errors are minimized at different combinations. RMSE with a minimum value of 51666.32 depicts RPF as the best forecasting combination while MAE is the minimum, i.e., 34488.85 for PREF.

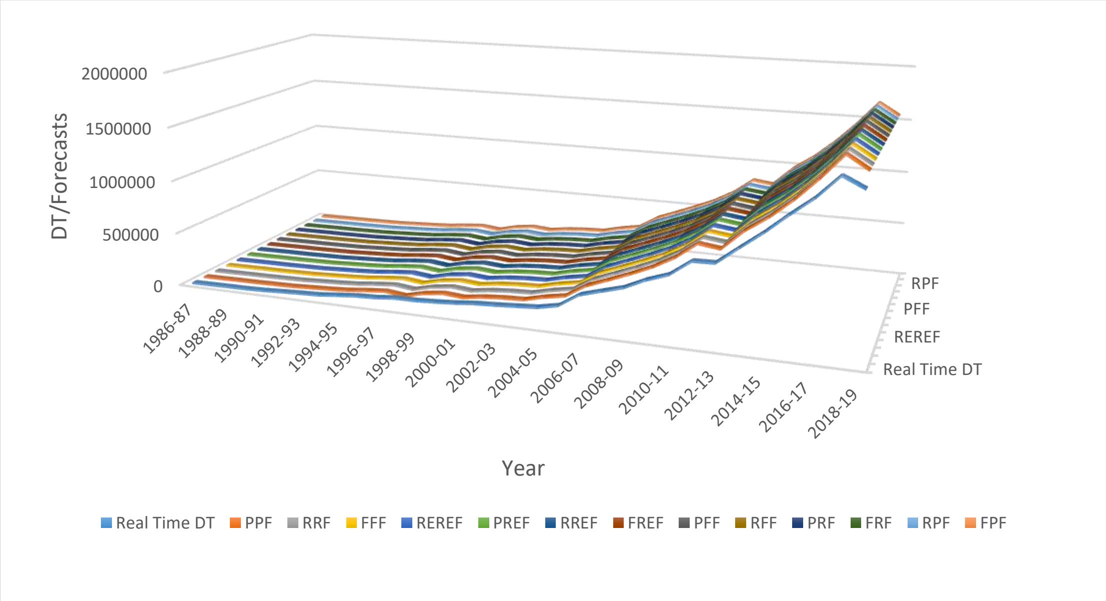


Figure 4.2: Real-Time Direct Tax and Forecasts

Table 4.2: Statistical Errors for Direct Tax Revenue Forecasts

CASE	RMSE	MAE
Case 1: PPF	55956.1	34724.4
Case 2: RRF	51810.1	35589.5
Case 3: FFF	52252.1	35430
Case 4: REREF	52677.7	35701.6

Case 5: PREF	55895.6	34488.9
Case 6: RREF	52405.1	35959.6
Case 7: FREF	52700.7	35727
Case 8: PFF	55978.6	34715.4
Case 9: RFF	51961.5	35664.1
Case 10: PRF	55928	34684.1
Case 11: FRF	52102.5	35354.8
Case 12: RPF	51666.3	35509.7
Case 13: FPF	51959.2	35274.8

Source: Authors' calculations

4.1.3. Sales Tax

Figure 4.3 shows that forecasts fluctuate around the real time sales tax. All predictions move together. In table 4.3, RMSE is in favor of data revision accuracy as the minimum RMSE, i.e., 46927.69 is for REREF. But MAE is not supporting it. It is minimum i.e., 29574.96 for FREF.

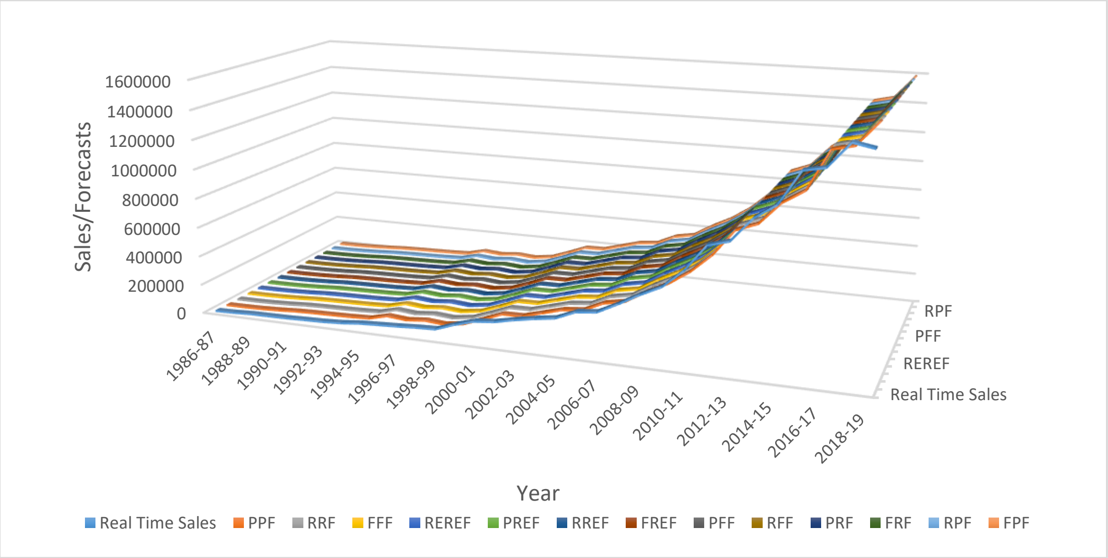


Figure 4.3:Real Time Sales Tax and Forecas

Table 4.3: Statistical Errors for Sales Tax Revenue Forecasts

CASE	RMSE	MAE
Case 1: PPF	47651.9	30687.9
Case 2: RRF	47145.5	29758.4
Case 3: FFF	47117	29715.3
Case 4: REREF	46927.7	29306.5
Case 5: PREF	47517.3	30503.1
Case 6: RREF	47007.2	29601.2
Case 7: FREF	47008.4	29575
Case 8: PFF	47627.8	30653.8
Case 9: RFF	47115.9	29741.5
Case 10: PRF	47656.9	30669.9
Case 11: FRF	47146.6	29732.2
Case 12: RPF	47140	29774.7
Case 13: FPF	47141	29748.5

Source: Authors' calculations

4.1.4. Custom Duties

Figure 4.4 and table 4.4 show that the data revision strategy is partially applicable for custom duty forecasts. All residuals suggest that accuracy is highest while tax base data is real-time, but base year tax data is provisional. As it is shown that minimum RMSE i.e., 27602.67, and MAE i.e., 11.48988 suggest PREF as the most accurate combination.

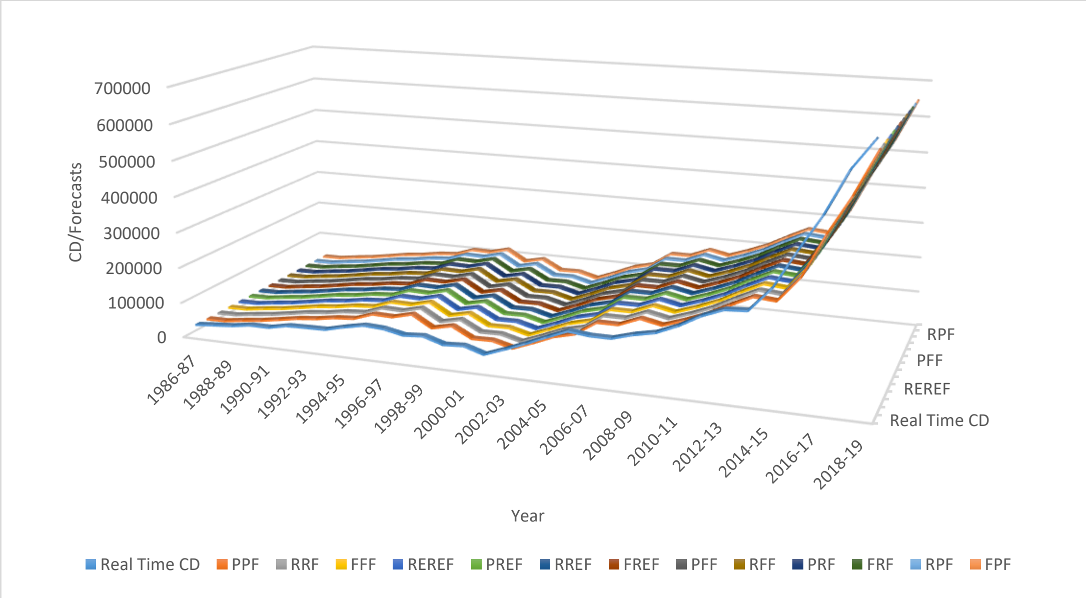


Figure 4.4: Real-Time Custom Duties and Forecasts

Table 4.4: Statistical Errors for Custom Duties Forecasts

CASE	RMSE	MAE
Case 1: PPF	27705.9	17682.1
Case 2: RRF	28227.6	19188.2
Case 3: FFF	28454.9	19668.8
Case 4: REREF	28493.6	19950.9
Case 5: PREF	27602.7	17599.2
Case 6: RREF	28167.5	19199.4
Case 7: FREF	28433.8	19691.2
Case 8: PFF	27643.1	17638.6
Case 9: RFF	28190.5	19161
Case 10: PRF	27679.5	17659.7
Case 11: FRF	28491.9	19695.2
Case 12: RPF	28251.7	19201.7
Case 13: FPF	28515.3	19710.7

Source: Authors' calculations

4.1.5. Federal Excise Duty (FED)

According to figure 4.5 and table 4.5, real time data performs the best for FED forecasting according to MAPE i.e., 7048.247. RMSE and MAE are in conflict. RMSE is the minimum for RPF i.e., 9531.769, while MAE is the minimum for PPF i.e., 9.57102.

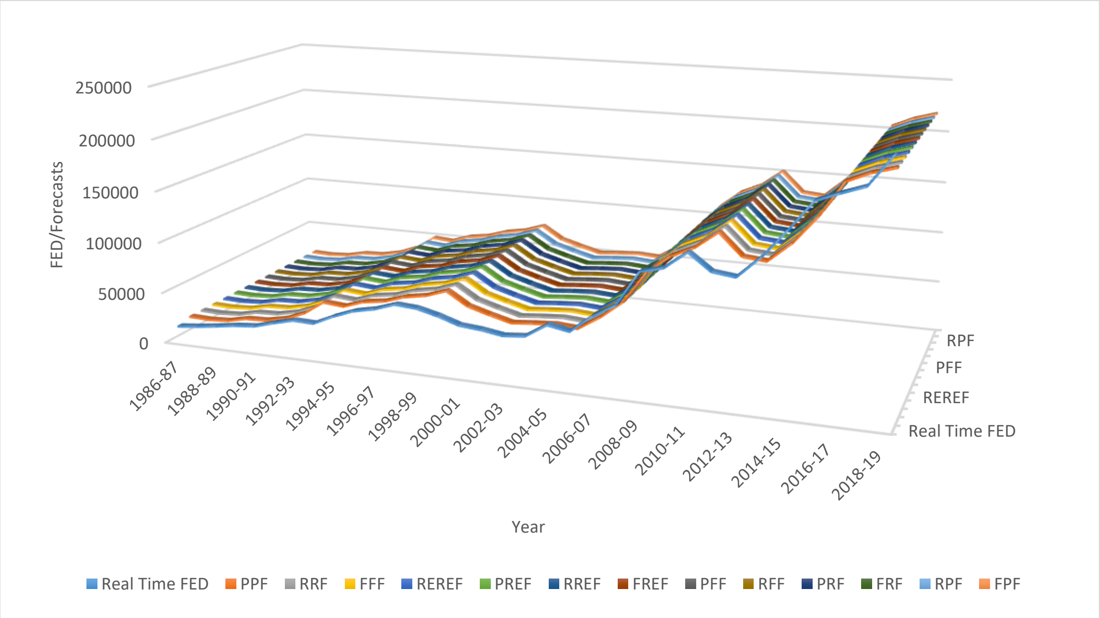


Figure 4.5: Real Time FED and Forecasts

Table 4.5: Statistical Errors for Federal Excise Duty Forecasts

CASE	RMSE	MAE
Case 1: PPF	9627.37	7211.62
Case 2: RRF	9538.48	7068.93
Case 3: FFF	9610.71	7106.04
Case 4: REREF	9618.7	7048.25
Case 5: PREF	9638.77	7214.68
Case 6: RREF	9561.16	7084.44
Case 7: FREF	9624.85	7116.36

Case 8: PFF	9640.83	7221.74
Case 9: RFF	9547.52	7074.13
Case 10: PRF	9631.22	7216.58
Case 11: FRF	9601.75	7100.84
Case 12: RPF	9531.77	7060.64
Case 13: FPF	9594.89	7092.55

Source: Authors' calculations

4.1.6. Conclusion

This section has used different combinations of provisional, revised, final, and real-time data to check the effect of data discrepancies. If the data regime matters, the real-time data results should give minimum forecasting errors. Comparing RMSE and MAE, results show that the data revision has no significant impact on the forecasts of any federal tax. Our first hypothesis that data discrepancy is affecting the accuracy is not accepted. From this chapter, we can conclude that federal tax revenue forecast errors are due to some reason other than data discrepancy issues.

4.2. Checking Parameter Discrepancy

The buoyancy approach is applied to the following data sets for within-sample forecasts. The taxes and bases data is real-time in all sets. On the other hand, the parameters, i.e., real GDP growth target and inflation rate are revised. Forecasts are calculated and compared on the basis of RMSE and MAE.

4.2.1. Total Tax

According to figure 4.6 , all forecasts move almost together. They move along with real time total tax till 2012-13. In 2013-14, real time total tax exceeded the forecasts.

In 2014-15, back to harmony and in 2015-16, forecasts under predicted the total tax. Fluctuation persisted and in 2018-19, there was over prediction.

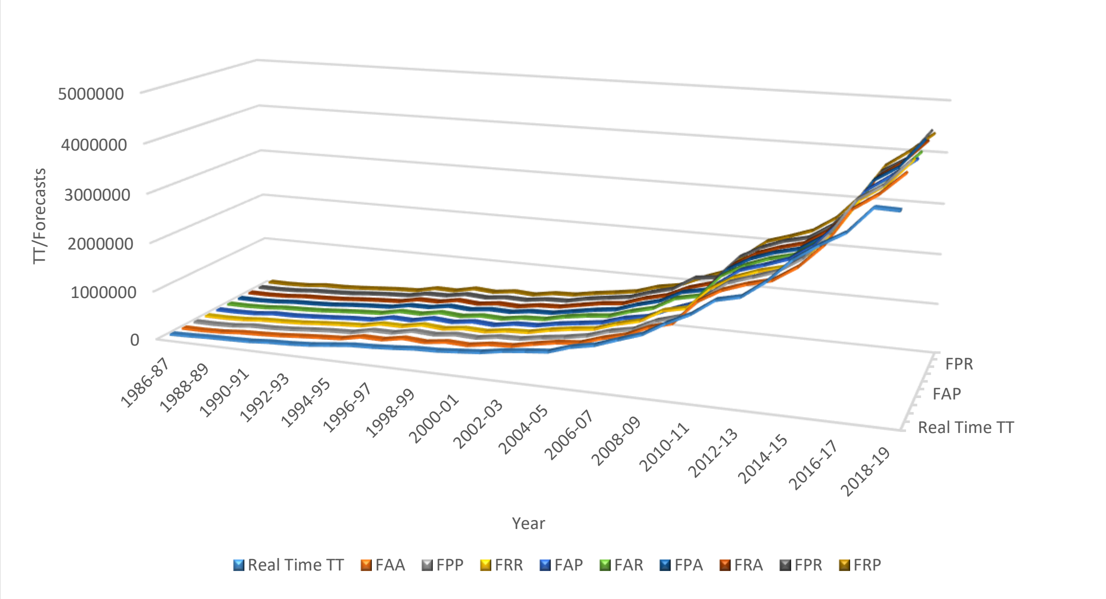


Figure 4.6 : Real-Time Total Tax and Forecasts

According to table 4.6, RMSE and MAE are minimized when actual real GDP growth and provisional inflation rate targets are used. RMSE is 126881.8, and MAE is 69350.28 when minimized. The results suggest that real GDP growth targets revision affects the quality of the forecasts.

Table 4.6: Statistical Errors for Total Federal Tax Forecasts

CASE	RMSE	MAE
Case 1: FAA	130979	71486.4
Case 2: FPP	149384	74115.8
Case 3: FRR	135080	72291.4
Case 4: FAP	126882	69350.3
Case 5: FAR	129426	71352.5
Case 6: FPA	153663	78762.6

Case 7: FRA	136413	72580.4
Case 8: FPR	152721	78180.2
Case 9: FRP	132227	69564.2

Source: Authors' calculations

4.2.2. Direct Tax

According to figure 4.7, all forecasts move almost together. They move along with real time direct tax till 2007-08. Since 2008-09, all forecasts have under predicted the direct tax.

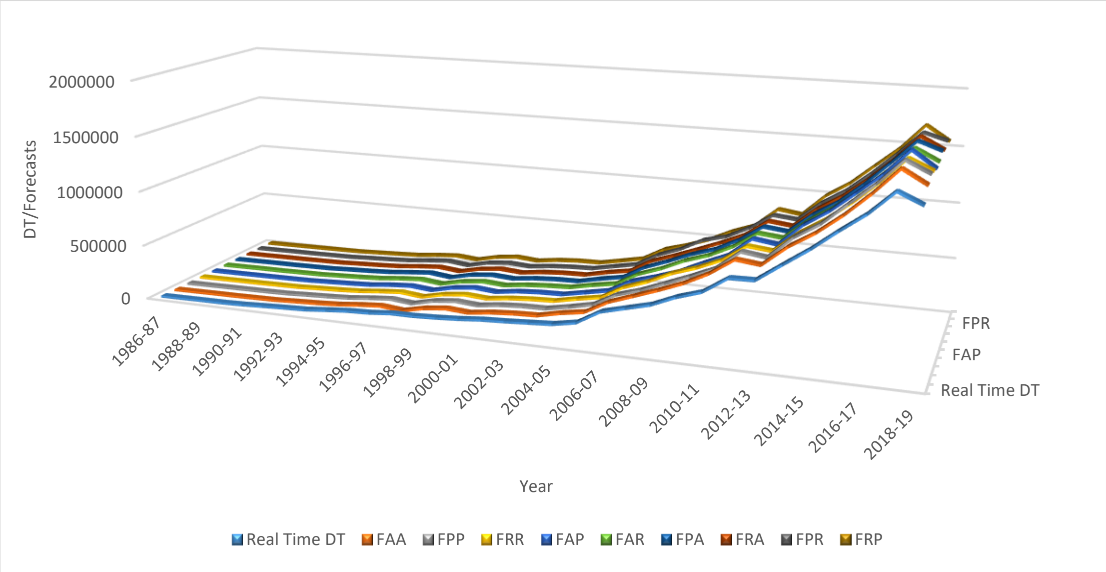


Figure 4.7: Real Time Direct Tax and Forecasts

According to table 4.7, RMSE and MAE are minimized when actual real GDP growth and actual inflation rate targets are used. RMSE is 52677.7 and MAE is 35701.5 at optimum forecasts. The errors suggest that real GDP growth targets revision plays a role in forecasts improvement.

Table 4.7: Statistical Errors for Direct Tax Forecasts

CASE	RMSE	MAE
Case 1: FAA	52677.7	35701.5

Case 2: FPP	61778.5	40183.6
Case 3: FRR	54120.5	36953.4
Case 4: FAP	55906.1	36677.8
Case 5: FAR	52742.6	36094.2
Case 6: FPA	59125.9	39252.7
Case 7: FRA	53994.4	36591.8
Case 8: FPR	59332	39614.3
Case 9: FRP	57059.8	37586.7

Source: Authors' calculations

4.2.3. Sales Tax

According to figure 4.8, sales tax forecasts fluctuated around the real time data. They under predicted in 2011-12 and 2015-16 and over predicted in 2016-17 and 2018-19.

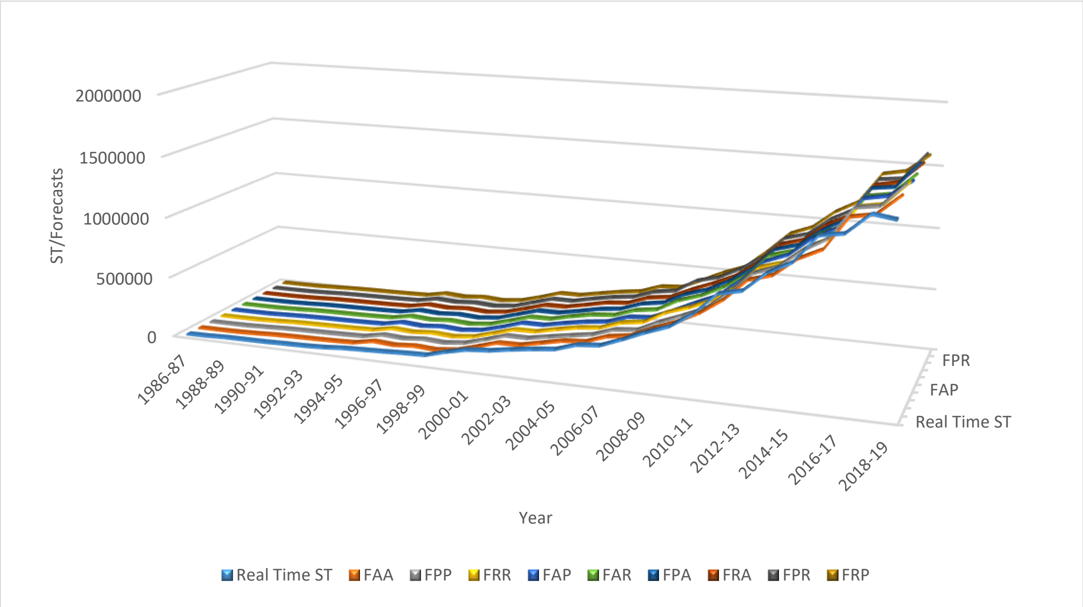


Figure 4.8: Real Time Sales Tax and Forecasts

According to table 4.8, RMSE and MAE are minimum for sales tax when actual real GDP growth and provisional inflation rate targets are used. At optimum forecasts, RMSE is 44180.0, and MAE is 28261.4. The results suggest that real GDP growth targets revision affects while inflation rate targets does not affect the quality of the forecasts.

Table 4.8 : Statistical Errors for Sales Tax Forecasts

CASE	RMSE	MAE
Case 1: FAA	46927.6	29306.4
Case 2: FPP	48432.1	28867
Case 3: FRR	48224.1	28866.1
Case 4: FAP	44180	28261.4
Case 5: FAR	46927.8	28601.8
Case 6: FPA	51199.6	30059.6
Case 7: FRA	48169.5	29570.7
Case 8: FPR	51329.8	29890
Case 9: FRP	45387	28439.6

Source: Authors' calculations

4.2.4. Custom Duties

Figure 4.9 shows that there are fluctuations in custom duties forecasts. Since 2014-15, forecasts have under predicted the custom duties tax revenue. Following table 4.9, RMSE is the minimum i.e. 26424.6 for provisional nominal GDP growth targets. At the same time, MAE is optimized i.e. 18957.6 at provisional nominal GDP growth targets. For custom duties, neither real GDP targets nor inflation rate targets revisions have an impact on forecasts.

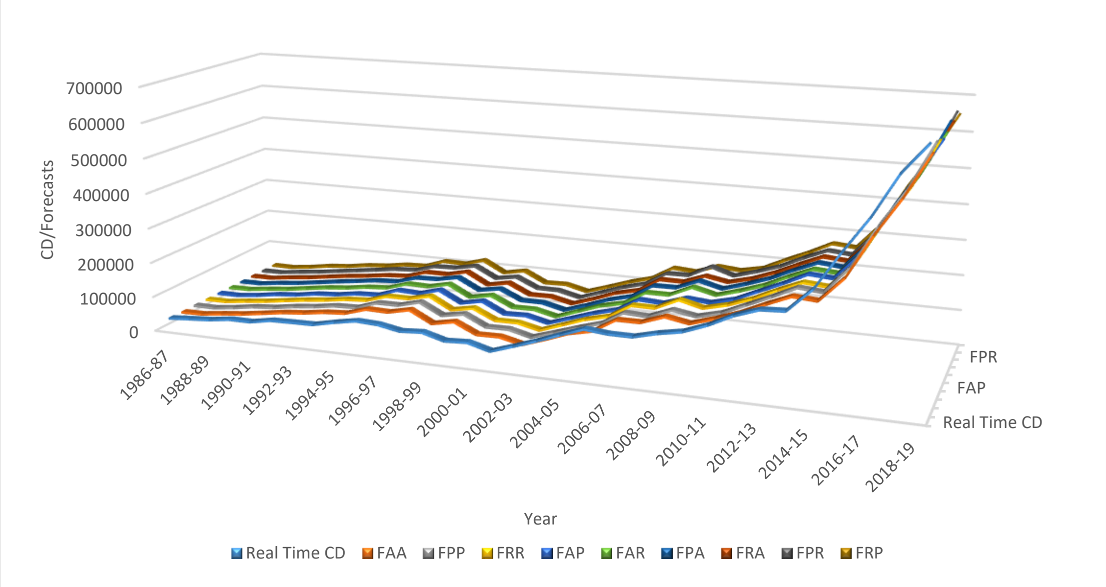


Figure 4.9: Real-Time Custom Duties and Forecasts

Table 4.9: Statistical Errors for Custom Duties Forecasts

CASE	RMSE	MAE
Case 1: FAA	28493.6	19950.9
Case 2: FPP	26424.6	18957.6
Case 3: FRR	28933.2	20347.5
Case 4: FAP	27106.2	19038.7
Case 5: FAR	28994.7	20214
Case 6: FPA	27953	19869.7
Case 7: FRA	28416.5	20084.3
Case 8: FPR	28504.4	20132.9

Case 9: FRP	26995.5	19172.2

Source: Authors' calculations

4.2.5. Federal Excise Duty (FED)

Figure 4.10 depicts that FED forecasts fluctuated around the real time tax. In 2018-19, forecasts underpredicted the federal excise duty.

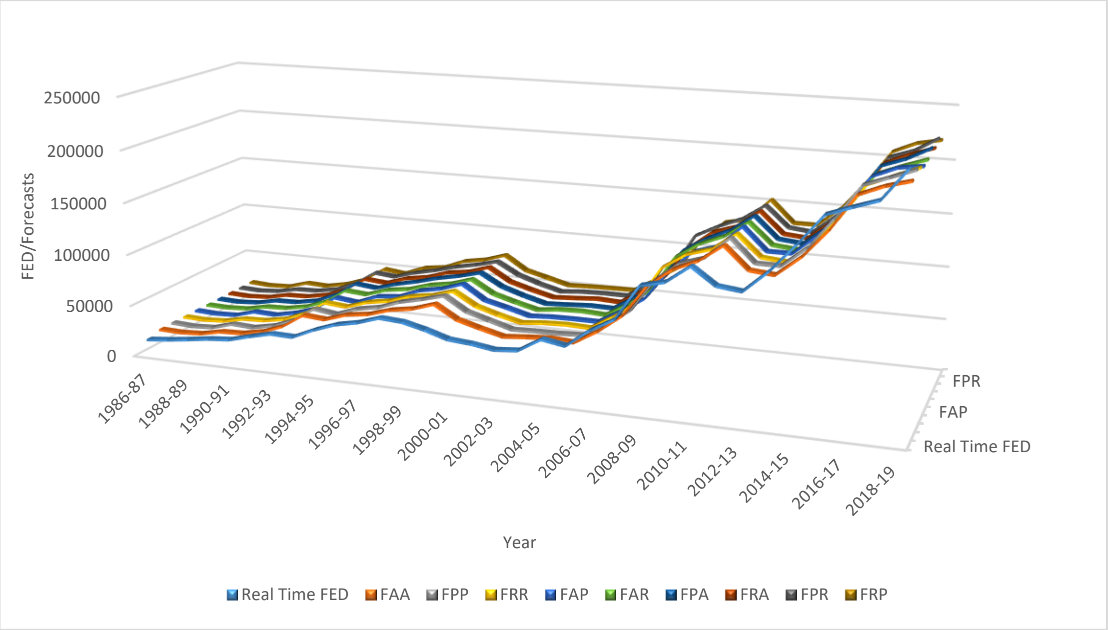


Figure 4.10: Real Time FED and Forecasts

According to table 4.10, RMSE is the minimum i.e., 9509.31 for revised nominal GDP growth targets. In contrast, MAE suggests actual real GDP growth and revised inflation rate targets for optimum forecasts. At optimal forecasts MAE is 6895.37. Following Willmott and Matsuura (2005) and Willmott et al.(2009), we take MAE as superior. So, the result suggests that real GDP growth target revision improves the forecasts.

Table 4.10 : Statistical Errors for Federal Excise Duty Forecasts

CASE	RMSE	MAE
Case 1: FAA	9618.7	7048.24
Case 2: FPP	9674.36	7337.06
Case 3: FRR	9509.31	6934.94
Case 4: FAP	9759.7	7365.61
Case 5: FAR	9526.06	6895.37
Case 6: FPA	9596.92	7045.63
Case 7: FRA	9586.96	7087.81
Case 8: FPR	9543.81	6953.54
Case 9: FRP	9717.45	7394.61

Source: Authors' calculations

4.2.6. Conclusion

This section provides some insights into parameter selection in federal tax revenue forecasting. Considering the values of RMSE, and MAE, most tax forecasts are optimal while actual real GDP growth targets are used. If we use the actual real GDP growth targets, the results would be more accurate than provisional and revised GDP growth. On the other hand, inflation rate target revisions have no significant impact on forecasting. We can use either provisional, revised, or actual inflation rate targets along with actual real GDP growth targets to get better results. This chapter accepts the null hypothesis that parameter revision affects Pakistan's federal tax revenue forecasts.

4.3. Method Selection

The root mean square error and mean absolute error for all eight forecasting methods are discussed in this section. Detailed forecasts and errors are given in appendix 2A.1 to appendix 2A.8.

4.3.1. Total Tax

In table 4.11, according to root mean square error (RMSE), Least Absolute Shrinkage and Selection Operator (LASSO), and Elastic Net provide the best total tax revenue forecasts with a minimum error of 0.03. On the other hand, mean absolute error (MAE) suggests Marginal Tax Rate (MTR) as the best forecaster with an error of 0.016. Following Willmott and Matsuura (2005) and Willmott *et al.* (2009), we take MAE as superior. The recommended tax revenue forecasting method for total tax is MTR or the marginal tax rate method.The results deny Streimikiene, Ahmed, V Veinhardt, & Pervaiz (2018). They found out that Box Jenkins provided better predictions for FY17 for Pakistan.

Table 4.11 : Statistical Errors for Total Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.108	0.093
Box Jenkins	0.05	0.036
SVAR Model	0.052	0.044
Elastic Net	0.031	0.025
LASSO	0.031	0.024
Ridge Regression	0.046	0.038
Effective Tax Rate	0.118	0.095
Marginal Tax Rate	0.751	0.016

Source: Authors' calculations

4.3.2. Direct Tax

According to table 4.12, root mean square error (RMSE), with a value of 0.089, suggests that Box Jenkins methodology provides the most accurate forecast for direct tax. On the other hand, mean absolute error (MAE) recommends Marginal Tax Rate (MTR) as the best method to forecast direct tax with an error value of 0.025. We take MTR as the recommended method.

Table 4.12: Statistical Errors for Direct Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.15	0.09
Box Jenkins	0.089	0.067
SVAR Model	0.101	0.073
Elastic Net	0.098	0.077
LASSO	0.098	0.077
Ridge Regression	0.102	0.075
Effective Tax Rate	0.149	0.128
Marginal Tax Rate	1.434	0.025

Source: Authors' calculations

4.3.3. Sales Tax

Table 4.13: Statistical Errors for Sales Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.196	0.159
Box Jenkins	0.088	0.073

SVAR Model	0.117	0.087
Elastic Net	0.102	0.079
LASSO	0.101	0.078
Ridge Regression	0.107	0.082
Effective Tax Rate	0.16	0.117
Marginal Tax Rate	0.896	0.019

Source: Authors' calculations

Table 4.13 shows that according to RMSE, Box Jenkins provides the most accurate sales tax forecasts with an error of 0.088. On the other hand, MAE suggests that the Marginal Tax Rate (MTR) approach is the most efficient technique, with an error of 0.019. The marginal tax rate approach is recommended for sales tax forecasting

4.3.4. Custom Duties

According to table 4.14, the custom duties of Pakistan are best forecasted through LASSO and Elastic Net with root mean square error (RMSE) of 0.128. In contrast, mean absolute error (MAE) value of Marginal Tax Rate (MTR) is minimum, i.e., 0.046 proving it the most appropriate method to forecast custom duties. Following MAE, the marginal tax rate (MTR) is recommended.

Table 4.14: Statistical Errors for Custom Duties Forecasting Methods

Estimation Methodology	**RMSE**	**MAE**
Buoyancy Approach	0.204	0.142
Box Jenkins	0.157	0.124
SVAR Model	0.141	0.107
Elastic Net	0.128	0.093

LASSO	0.128	0.093
Ridge Regression	0.13	0.097
Effective Tax Rate	0.248	0.174
Marginal Tax Rate	2.5	0.046

Source: Authors' calculations

4.3.5. Federal Excise Duty (FED)

According to table 4.15, Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net have a minimum root mean square error (RMSE) of 0.089. In contrast, Marginal Tax Rate (MTR) has a minimum mean absolute error (MAE) of 0.033 while forecasting the federal excise duty of Pakistan. Following MAE, the marginal tax rate method is suggested here.

Table 4.15: Statistical Errors for FED Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.121	0.090
Box Jenkins	0.115	0.101
SVAR Model	0.111	0.093
Elastic Net	0.089	0.072
LASSO	0.089	0.072
Ridge Regression	0.09	0.075
Effective Tax Rate	0.167	0.142
Marginal Tax Rate	1.609	0.033

Source: Authors' calculations

4.3.6. Feasibility

The statistical errors suggest LASSO, Elastic Net regression, Box Jenkins and MTR as suitable methods for forecasting different tax revenues of Pakistan. Marginal Tax Rate is crude method of forecasting which suggest that LASSO/Elastic Net and Box Jenkins should be adopted for tax revenue forecasts of Pakistan. LASSO/elastic net method is a machine learning method. It is time and cost effective procedure.

The marginal tax rate approach, and Box Jenkins are statistical methods consist of a number of steps. Moreover, LASSO/Elastic Net is a one step regression making it handy. They are convenient and can quickly be adopted by FBR. So, the time, budget, and complication criteria prove that the Box Jenkins and LASSO/Elastic Net are superior to other approaches.

4.3.7 Conclusion

Pakistan is a federation in which the resources are collected at the center and then shared through the revenue sharing formula (NFC award). So any forecast error for revenue stream would result in mismanaged fiscal operations, mistargeting targets, or additional debt burdens. Therefore it is imperative to evaluate and propose alternative methods of forecasts for federal taxes in Pakistan. As studies have noted (Qasim and Khalid, 2016) that revenue forecasting errors are significant. This may be due to other reasons, but for this study, we have focused on evaluating different methods of forecasting to propose alternative but better methods for different federal tax streams.

The forecasting is done using alternative theoretical, statistical, and machine-learning methods. The RMSE suggests that LASSO, and Elastic Net are the best choices to forecast Total Tax, Custom Duties, and FED. Box Jenkins predicts accurately while forecasting Sales Tax and Direct tax. On the other hand, MAE suggests the marginal tax rate approach as the most appropriate forecasting method for all taxes. According to the results, it is evident that method revision can play a vital role in improving Pakistan's federal tax revenue forecasting.

4.4. Statistical Loss Function

As mentioned in previous section, all forecasts and other variables are in differenced forms. In this section, statistical loss functions for all federal taxes are computed. To check statistical losses, we have applied biasedness and method comprehension criteria on both Buoyancy approach and LASSO regression. In the biasedness test, tax revenues are regressed upon their forecasts made in previous year. A test of $\propto=0$ and $\beta=1$ is done. If rejected, the tax forecasts are biased.

The second statistical loss is calculated through method comprehension approach. We check whether all relevant information is used by the given method. The test is done by regressing forecasting errors over a number of relevant variables. The null hypothesis of $\eta_1=\eta_2=0$ = all insignificant coefficients, is checked. If rejected, all relevant information is not used. It means the method is not comprehensive.

4.4.1. Total Tax

i.Biasedness

Table 4.16: Test Results for Unbiased Total Tax Revenue Forecasts

Method	$\propto$	β	R-Squared
Buoyancy Approach	0.118*** (0.017)	0.022 (0.106)	0.0014
LASSO	-0.005 (0.015)	1.046*** (0.1130	0.726

Elastic Net	-0.005	1.046***	0.726
	(0.015)	(0.1130	

S.E are in parenthesis. * means significant at 10%level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.16, the buoyancy approach does not accept the null hypothesis of $\propto$=0 and β=1. In contrast, LASSO and Elastic Net regression's coefficients are almost equal to the null hypothesis in absolute terms. We can conclude that LASSO/Elastic Net regression provides unbiased forecasts for total tax revenues of Pakistan.

ii. Method Comprehension

Table 4.17: Testing Information Efficiency using P-Values for Total Tax Forecasts

Method	GDP Growth Target	Total Consumption(L)	Dutiable Imports(L)	LSM(L)	Total Tax(L)	Nominal GDP(L)
Buoyancy Approach	0.57	0.81	0.25	0.86	0.02**	0.35
LASSO	0.55	0.42	0.18	0.20	0.90	0.71
Elastic Net	0.55	0.42	0.18	0.20	0.90	0.71

* means significant at 10% level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.17, for the Buoyancy approach, the null hypothesis that $\eta_1=\eta_2=0$ = all insignificant coefficients, is rejected. It means all relevant information was not used in Buoyancy Approach. On the other hand, in case of LASSO regression, all relevant information was used.

4.4.2. Direct Tax

i.Biasedness

Table 4.18: Test Results for Unbiased Direct Tax Revenue Forecasts

Method/Parameter	∝	β	R-Squared
Buoyancy Approach	0.119* (0.026)	0.217*** (0.124)	0.092
Box Jenkins Method	0.001 (0.084)	0.997* (0.558)	0.090

S.E are in parenthesis.* means significant at 10% level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.18, the buoyancy approach does not accept the null hypothesis of ∝=0 and β=1. In contrast, Box Jenkins's coefficients are almost equal to the null hypothesis in absolute terms. We can conclude that Box Jenkins Methodology provides unbiased forecasts for direct tax revenues of Pakistan.

ii.Method Comprehension

Table 4.19: Testing Information Efficiency using P-Values for Direct Tax Forecasts

Method	GDP Growth Target	CPI	Exchange Rate	Total Employed Labor	Education Expenditure	Direct Tax(L)	Non-agricultural GDP(L)
Buoyancy Approach	0.71	0.86	0.76	0.75	0.42	0.48	0.12
Box Jenkins Method	0.91	0.49	0.18	0.91	0.65	0.69	0.54

* means significant at 10%level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.19, the null hypothesis that $\eta_1 = \eta_2 = 0$ = all insignificant coefficients, is not rejected in cases of both the buoyancy approach and Box Jenkins methodology. It means all relevant information is used in both methods while forecasting direct tax.

4.4.3. Sales Tax

i.Biasedness

Table 4.20: Test Results for Unbiased Sales Tax Revenue Forecasts

Method	$\propto$	β	R-Squared
Buoyancy Approach	0.152*** (0.032)	0.096 (0.131)	0.0174
Box Jenkins Method	-0.001 (0.047)	1.013*** (0.261)	0.320

S.E are in parenthesis.* means significant at 10%level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.20, the buoyancy approach does not accept the null hypothesis of $\propto$=0 and β=1. In contrast, Box Jenkins's coefficients are almost equal to the null hypothesis in absolute terms. We can conclude that Box Jenkins methodology provides unbiased forecasts for sales tax revenues of Pakistan.

ii. Method Comprehension

According to table 4.21, for the Buoyancy approach, the null hypothesis that $\eta_1=\eta_2=0$ = all insignificant coefficients, is rejected. It means all relevant information was not used in Buoyancy Approach. On the other hand, in case of Box Jenkins methodology, all relevant information was used.

Table 4.21: Testing Information Efficiency using P-Values for Sales Tax Forecasts

Method	GDP Growth Target	Imports	CPI	Exchange Rate	Sales Tax(L)	Total Consumption(L)
Buoyancy Approach	0.08*	0.94	0.84	0.91	0.01***	0.26
Box Jenkins Method	0.70	0.92	0.37	0.56	0.28	0.14

* means significant at 10%level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

4.4.4. Custom Duties

i. Biasedness

Table 4.22: Test Results for Unbiased Custom Duties Tax Revenue Forecasts

Method	∝	β	R-Squared
Buoyancy Approach	0.085*** (0.030)	0.092 (0.161)	0.010
LASSO	-0.003 (0.045)	1.034** (0.396)	0.175
Elastic Net	-0.003 (0.045)	1.034** (0.396)	0.175

S.E are in parenthesis.* means significant at 10% level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.22, the buoyancy approach does not accept the null hypothesis of ∝=0 and β=1. In contrast, LASSO regression's coefficients are almost equal to the null hypothesis in absolute terms. We can conclude that LASSO regression provides unbiased forecasts for Custom Duties tax revenues of Pakistan.

ii. Method Comprehension

Table 4.23: Testing Information Efficiency using P-Values for Custom Duties Forecasts

Method	GDP Growth Target	LSM	CPI	Exchange Rate	Dutiable Imports(L)	Custom Duties(L)
Buoyancy Approach	0.86	0.90	0.88	0.68	0.97	0.01***
LASSO	0.39	0.61	0.97	0.77	0.42	0.63

Elastic Net	0.39	0.61	0.97	0.77	0.42	0.63

* means significant at 10% level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.23, in the Buoyancy approach, the null hypothesis that $\eta_1=\eta_2=0$ = all insignificant coefficients, is rejected. It means all relevant information was not used in Buoyancy Approach. On the other hand, in case of LASSO regression and Elastic Net, all relevant information was used.

4.4.5. FED

i. Biasedness

According to table 4.24, the buoyancy approach does not accept the null hypothesis of $\propto=0$ and $\beta=1$. In contrast, Elastic Net and LASSO regression's coefficients are almost equal to the null hypothesis in absolute terms. We can conclude that LASSO regression and Elastic Net provide unbiased forecasts for total tax revenues of Pakistan.

Table 4.24: Test Results for Unbiased FED Tax Revenue Forecasts

Method	$\propto$	$\boldsymbol{\beta}$	R-Squared
Buoyancy Approach	0.051**	0.440***	0.218
	(0.021)	(0.152)	
LASSO	-0.000	1.016***	0.357
	(0.024)	(0.240)	
Elastic Net	-0.000	1.016***	0.357
	(0.024)	(0.240)	

S.E are in parenthesis.* means significant at 10%l evel of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

ii.Method Comprehension

Table 4.25: Testing Information Efficiency using P-Values for FED Forecasts

Method	GDP Growth Target	Total Employed Labor	Education Expenditure	CPI	Exchange Rate	FED(L)	LSM(L)
Buoyancy Approach	0.67	0.96	0.34	0.80	0.69	0.02***	0.18
LASSO	0.53	0.75	0.60	0.52	0.39	0.44	0.97
Elastic Net	0.53	0.75	0.60	0.52	0.39	0.44	0.97

* means significant at 10%level of Significance.** means significant at 5% level of significance.*** means significant at 1% level of significance.

According to table 4.25, in the Buoyancy approach, the null hypothesis that $\eta_1=\eta_2=0$ = all insignificant coefficients, is rejected. It means all relevant information was not used in the buoyancy approach. On the other hand, in the case of LASSO regression ans Elastic Net, all relevant information was used.

4.4.6. Conclusion

The statistical loss functions provide evidence for the non-suitability of the buoyancy approach. LASSO regression provides unbiased forecasts for total tax, customs duties and FED while Box Jenkins methodolody provides unbiased forecasts for direct tax and sales tax. On the other hand, In case of method comprehension, LASSO uses all relevant information for total tax, custom duties and FED. Box Jenkins methodology uses all relevant information in case of sales tax but not in case of direct tax. On the other hand, the buoyancy approach does not uses all relevant information for any federal tax.

CHAPTER 5

CONCLUSION AND POLICY RECOMMENDATIONS

5.1. Conclusion

Tax revenue forecasting is a critical component of effective fiscal management, playing a central role in shaping government policies and ensuring economic stability. Accurate forecasts allow policymakers to allocate resources efficiently, design sound fiscal policies, and maintain macroeconomic balance. Conversely, inaccurate forecasts lead to fiscal mismanagement, unforeseen debt accumulation, and compromised developmental priorities. Forecasting errors can be particularly detrimental in federated systems where revenues are collected at the federal level and distributed to subnational governments, exacerbating fiscal imbalances across all levels of governance. Such systemic inefficiencies highlight the broader socio-economic ramifications of unreliable revenue forecasts, affecting poverty reduction strategies, infrastructure development, and social welfare programs.

This thesis set out to examine the tax revenue forecasting framework in Pakistan, addressing three overarching objectives: a comprehensive evaluation of the existing forecasting mechanism, identification of the key determinants of forecasting errors, and the recommendation of alternative forecasting methodologies to improve accuracy and economic outcomes.

The first objective sought to thoroughly analyze the current tax revenue forecasting structure employed by Pakistan's Federal Board of Revenue (FBR). The research identified significant weaknesses in the prevailing buoyancy approach, which assumes a static relationship between tax revenues and GDP growth. While simple to implement, this method fails to account for dynamic economic conditions, structural shifts in the tax base, and variations in compliance behavior. These limitations have led to persistent forecasting errors, causing substantial economic costs. For instance, errors in predicting tax revenues directly impact the fiscal deficit, necessitating either unplanned borrowing or abrupt expenditure cuts. Such fiscal disruptions undermine long-term investment projects, delay public service delivery, and increase reliance on external financing, ultimately constraining fiscal sovereignty and economic growth.

The second objective investigated the root causes of forecasting inaccuracies by examining the effects of data revisions, parameter adjustments, and methodological choices:

- **Data Revisions**: Contrary to initial assumptions, data revisions exhibited negligible effects on tax revenue forecast accuracy across major tax heads, suggesting that the primary issue lies within model specification rather than data quality.
- **Parameter Revisions**: The analysis revealed that while inflation rate targets had limited impact on forecast errors, revisions in real GDP growth targets significantly influenced outcomes. Forecasts using actual GDP growth rates were notably more accurate compared to those based on provisional or revised figures, indicating the critical need for reliable macroeconomic projections.
- **Methodological Limitations**: The study found that the buoyancy approach's static assumptions rendered it unsuitable for dynamic tax environments. Its inability to adapt to economic shocks, policy changes, or structural transformations necessitated the adoption of more flexible and robust forecasting techniques.

The third objective proposed alternative methodologies tailored to Pakistan's fiscal landscape. Empirical evaluations identified several superior forecasting models:

- The Marginal Tax Rate (MTR) Method demonstrated the highest accuracy for forecasting all major tax categories, according to Mean Absolute Error (MAE) analysis. By reflecting incremental revenue changes relative to marginal tax adjustments, the MTR method offers a more nuanced and responsive approach.
- Least Absolute Shrinkage and Selection Operator (LASSO) Regression and Elastic Net emerged as optimal for predicting aggregate tax revenues, customs duties, and federal excise duties, as confirmed by Root Mean Square Error (RMSE) results. These methods' ability to handle multicollinearity and select significant predictors enhances their predictive robustness.
- The Box-Jenkins Methodology (also known as ARIMA models) was recommended for direct tax and sales tax revenue forecasts due to its proficiency in modeling time-series data patterns and accounting for seasonal fluctuations and autocorrelations.

5.2. Policy Recommendations

The study's findings call for a comprehensive reform of Pakistan's tax revenue forecasting practices. Key recommendations include:

- **Adopt the Marginal Tax Rate Method for Forecasting**: Transitioning to the MTR method would enable the Federal Board of Revenue (FBR) to produce more accurate and dynamic tax revenue forecasts. Unlike the buoyancy approach, the MTR method can better capture the effects of policy changes, tax rate adjustments, and evolving economic conditions, thereby improving fiscal planning and policy implementation.

- **Enhance Real GDP Growth Target Estimation**: The Ministry of Finance should prioritize refining its GDP forecasting processes. Accurate GDP projections are fundamental to reliable tax forecasts. Incorporating advanced econometric models, integrating macroeconomic indicators beyond simple production data, and utilizing real-time analytics could significantly enhance predictive accuracy.
- **Reform the Use of the Buoyancy Approach**: While the buoyancy approach remains widely used due to its simplicity, its predictive performance can be enhanced by coupling it with improved GDP estimates. By aligning GDP forecasts with more sophisticated modeling techniques, the method's inherent limitations can be partially mitigated.
- **Develop Capacity-Building Programs**: Strengthening the analytical capacity of forecasting units within the Ministry of Finance and the FBR is vital. Investing in specialized training on modern econometric software, machine learning tools, and advanced statistical methods would enhance the proficiency of analysts and reduce dependency on outdated methods.
- **Establish a Tax Forecasting Review Committee**: Creating an independent committee composed of economists, statisticians, and policy experts to regularly evaluate and update forecasting methodologies would ensure the adoption of best practices and facilitate continuous improvement.
- **Institutionalize Forecasting Transparency**: Publishing detailed reports on forecasting assumptions, methodologies, and error margins would enhance accountability and foster public trust. Transparent practices encourage

informed policy debates and provide stakeholders with a clearer understanding of fiscal challenges.

5.3. Future Research Directions

Although this study contributes significantly to the discourse on tax revenue forecasting in Pakistan, several avenues remain unexplored. Future research could:

- **Incorporate Political and Administrative Factors**: Political influences and bureaucratic inefficiencies often affect revenue collection and forecasting. Future studies could analyze how electoral cycles, governance quality, and administrative capacity impact forecast accuracy, providing a more comprehensive understanding of systemic challenges.
- **Evaluate GDP Forecasting Methodologies**: The link between GDP target inaccuracies and tax revenue forecast errors calls for an in-depth examination of the Ministry of Finance's GDP modeling techniques. Identifying methodological gaps and incorporating advanced forecasting models would yield more reliable macroeconomic indicators.
- **Quantify Fiscal Costs of Forecasting Errors**: Future research should empirically measure the economic consequences of forecasting inaccuracies, including their effects on budget deficits, public debt, and development spending. Quantifying these impacts would strengthen the case for methodological reforms and guide policy prioritization.
- **Explore Asymmetrical Loss Functions in Forecast Accuracy**: Unlike traditional MAE and RMSE metrics, asymmetrical loss functions consider the differing economic impacts of overestimation versus underestimation.

Incorporating these measures would offer a more nuanced assessment of forecast quality.

- **Leverage Machine Learning and Artificial Intelligence**: Emerging technologies hold the potential to revolutionize tax forecasting. Future research could compare the effectiveness of more advanced machine learning algorithms, neural networks, and AI-driven models against traditional approaches, determining their viability for Pakistan's fiscal environment.
- **Investigate Sector-Specific Forecasting Models**: Different sectors of the economy contribute variably to tax revenues. Future studies could develop sector-specific forecasting models that account for industry-specific growth patterns, compliance rates, and tax elasticity.

By addressing these research gaps and implementing robust policy reforms, Pakistan can establish a resilient, data-driven forecasting framework. Enhanced forecasting accuracy will not only strengthen fiscal policy but also support broader economic goals, including sustainable growth, poverty alleviation, and improved public service delivery.

REFERENCES

Aamir, M., Qayyum, A., Nasir, A., Hussain, S., Khan, K. I., & Butt, S. (2011). Determinants of tax revenue: A comparative study of direct taxes and indirect taxes of Pakistan and India. *International Journal of Business and Social Science*, *2*(19).

Abate, D. (2013). Analysis of Tax Revenue Forecasting in Ethiopia: An Autoregressive Distributed Lag Approach.

Acconcia, A., d'Amato, M., Martina, R., & Ratto, M. (2019). *The Response of Taxpayer Compliance to the Large Shock of Italian Unification*. https://doi.org/10.1016/j.ejpoleco.2021.102158

Ademmer, M., & Boysen-Hogrefe, J. (2022). The impact of forecast errors on fiscal planning and debt accumulation. *Jahrbücher für Nationalökonomie und Statistik*, *242*(2), 171-190. https://doi.org/10.1515/jbnst-2020-0054

Ajisola, A. S. (2023). *An Efficient Time Series Model for Tax Revenue Forecasting: A Case of Nigeria* (Doctoral dissertation, AUST).

Alexi, A., Lazebnik, T., & Shami, L. (2024). Microfounded tax revenue forecast model with heterogeneous population and genetic algorithm approach. *Computational Economics*, *63*(5), 1705-1734.

Alhassan, B. G., & Tela, N. M. (2019). An Improvement Forecast Using Vector Autoregression. *Journal of Engineering and Applied Scientific*, 1-34.

Alhassan, B. G., Yusof, F. B., Norrulashikin, S. M., & Kane, I. L. (2020). Dynamics of Heteroscedasticity Modelling and Forecasting of Tax Revenue in a

Developing Economy: A Review. *International Journal of Management and Humanities (IJMH)*, 34-4. https://doi.org/10.35940/ijmh.C31161.115320

Altaf, S., Khan, S. Y., Ali, M. W., & Khan, M. W. (2021). A mathematical forecasting model to estimate tax to GDP ratio. *International Journal of Advanced Research in Engineering and technology*, *12*(3). https://doi.org/10.34218/IJARET.12.3.2021.029

Amin, A., Nadeem, A. M., Parveen, S., Kamran, M. A., & Anwar, S. (2014). Factors affecting tax collection in Pakistan: An empirical investigation. *Journal of Finance and Economics*, *2*(5), 149-155.

Auerbach, A. J. (1996). Dynamic Revenue Estimation. *Journal of Economic Perspectives*, 141-157. https://doi.org/110.1257/jep.1210.1251.1141

Bagdigen, M. (2002). How Accurate is Revenue Forecasting in Turkey? An Empirical Analysis. *Yapi Kredi Economic Review*, *13*(2), 29-37.

Barnard, J. R., & Dent, W. T. (1979). State tax revenues-new methods of forecasting. *The Annals of Regional Science*, *13*(3), 1-14. https://doi.org/10.1007/BF01287742

Bayer, O. (2013). Research of Estimates of Tax Revenue: An Overview. *European Financial and Accounting Journal*, *8*(3-4), 59-73. https://doi.org/10.18267/j.efaj.107

Bedrynets, M. (2019). System of tax revenues forecasting in Ukraine: theory and reality. *Економічний вісник університету*, (42), 171-179.

Becker, I., & Buettner, T. (2007). Are German Tax-Revenue Forecasts Flawed. *Arbeitspapier präsentiert auf der Jahrestagung des Vereins für Socialpolitik.*

Berwick, M. D., & Malchose, D. C. (2012). *Forecasting North Dakota fuel tax revenue and license and registration fee revenue* (No. DP-249). Fargo, ND, USA: Upper Great Plains Transportation Institute.

Boyd, D. J., & Dadayan, L. (2014). State tax revenue forecasting accuracy. *Rockefeller Institute of Government*, 1-54.

Bretschneider, S., & Gorr, W. (1992). Economic, organizational, and political influences on biases in forecasting state sales tax receipts. *International Journal of Forecasting*, *7*(4), 457-466. https://doi.org/10.1016/0169-2070(92)90029-9

Brown, B. W., & Maital, S. (1981). What do economists know? An empirical study of experts' expectations. *Econometrica: Journal of the Econometric Society*, 491-504. https://doi.org/10.2307/1913322

Buettner, T., & Kauder, B. (2010). Revenue forecasting practices: Differences across countries and consequences for forecasting performance. *Fiscal Studies*, *31*(3), 313-340. https://doi.org/10.1111/j.1475-5890.2010.00117.x

Buettner, T., & Kauder, B. (2015). Political biases despite external expert participation? An empirical analysis of tax revenue forecasts in Germany. *Public Choice*, *164*(3), 287-307. https://doi.org/10.1007/s11127-015-0279-2

Cameron, S., Spyropoulos, N., Saied-Tessier, A., Gusanu, S., & Barkis, N. (2017). *A Review of Tax Revenue Forecasting Models for the Scottish Housing Market*.

Cangiano, M., & Pathak, R. (2019). Revenue Forecasting in Low-Income and Developing Countries: Biases and Potential Remedies. In *The Palgrave*

Handbook of Government Budget Forecasting (89-114). Switzerland: Springer Nature Switzerland AG https://doi.org/10.1007/978-3-030-18195-6_5

Carmody, K., & Wiipongwii, T. (2018). *Virginia Tax Revenue Forecasting Policy Research Seminar Working Paper*.

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)?–Arguments against avoiding RMSE in the literature. *Geoscientific model development*, *7*(3), 1247-1250. https://doi.org/10.5194/gmd-7-1247-2014, 2014

Charemza, W. W., & Deadman, D. F. (1997). *New directions in Econometric Practice*.

Chatagny, F., & Siliverstovs, B. (2013). Rationality of direct tax revenue forecasts under asymmetric losses: Evidence from Swiss cantons. https://dx.doi.org/10.2139/ssrn.2199004

Chatagny, F. (2015). Incentive effects of fiscal rules on the finance minister's behavior: Evidence from revenue projections in Swiss Cantons. *European Journal of Political Economy*, *39*, 184-200. https://doi.org/10.1016/j.ejpoleco.2015.04.012

Chung, I. H., Williams, D. W., & Do, M. R. (2022). For Better or Worse? Revenue Forecasting with Machine Learning Approaches. *Public Performance & Management Review*, 1-21.

Cirincione, C., Gurrieri, G. A., & Sande, B. v. d. (1999). Municipal Government Revenue Forecasting: Issues of Method and Data. *Public Budgeting and Finance*, 26-46 https://doi.org/10.1046/j.0275-1100.1999.01155.x

Couture, J., & Imbeau, L. M. (2009). Do governments manipulate their revenue forecasts? Budget speech and budget outcomes in the Canadian provinces. In *Do They Walk Like They Talk?* (155-166): Springer. New York, NY. https://doi.org/10.1007/978-0-387-89672-4_9

Davidson, J. E., Hendry, D. F., Srba, F., & Yeo, S. (1978). Econometrics modelling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. *The Economic Journal*, 661-692. https://doi.org/610.2307/2231972.

Danninger, S., & Kyobe, A. J. (2005). The political economy of revenue-forecasting experience from low-income countries.

DBE, D. S. R., Breedon, F., Smith, A., & Ulph, D. (2021). *How we forecast income tax*.

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. J. N. (2016). Mean absolute percentage error for regression models. *192,* 38-48. https://doi.org/10.1016/j.neucom.2015.12.114

Dept., I. M. F. E. (2016). *Germany: 2016 Article IV Consultation-Press Release; Staff Report; and Statement by the Executive Director for Germany*.

Elliott, G., Timmermann, A., & Komunjer, I. J. T. R. o. E. S. (2005). Estimation and testing of forecast rationality under flexible loss. *72*(4), 1107-1125. https://doi.org/10.1111/0034-6527.00363

Elliott, G., & Timmermann, A. (2008). Economic forecasting. *Journal of Economic Literature, 46*(1), 3-56. https://doi.org/10.1257/jel.46.1.3

FBR. (2021). *Evidence-Based Revenue Forecasting FY 2021-22*. Islamabad.

Finance Division, P. o. P. (2022-23). *Federal Budget*. Islamabad.

Firdawss, T., & Karim, M. (2018). Forecasting Moroccan Tax Revenues: An Analysis Using International Institutions Methodologies and VECM. *Journal of Economics and Public Finance*, 304-323.

Gamboa, A. M. S. J. (April 2002). Development of Tax Forecasting Models: Corporate and Individual Income Taxes. *PIDS Discussion Paper Series, No. 2002-06*.

Gardner, E. S., & Jr., E. M. (1985). Forecasting Trends in Time Series. *Management science*, 1199-1334. https://doi.org/1110.1287/mnsc.1131.1110.1237.

Gentzkow, M., & Kamenica, E. (2014). Costly Persuasion. *American Economic Review*, *104*(5), 457-462. https://doi.org/ 10.1257/aer.104.5.457

Gentry, W. M. (1989). Do state revenue forecasters utilize available information?. *National Tax Journal*, *42*(4), 429-439.

Gholami, E., & Mousavi Jahromi, Y. (2015). Forecasting of the Value Added Tax from Tobacco Consumption Using Neural Network Method. *Journal of Economic Modeling Research*, *5*(20), 55-72.

Glenday, G. (2013). Revenue Forecasting. In *The International Handbook of Public Financial Management* (435-452): Palgrave Macmillan, a division of Macmillan Publishers Limited.

Golosov, M., & King, J. (2002). *Tax Revenue Forecasts in IMF-Supported Programs.* Paper presented at the IMF Fiscal Affairs Department seminar.

Göttert, M., & Lehmann, R. (2021). Tax Revenue Forecast Errors: Wrong Predictions of the Tax Base or the Elasticity? https://dx.doi.org/10.2139/ssrn.3872387

Greoning, E., Zivanomoyo, J., & Tsaurai, K. (2019). Determinants of Company Tax Revenue in Swaziland (1990-2015). *Acta Universitatis Danubius: Oeconomica*, *15*(5).

Grizzle, G. A., & Klay, W. E. (1994). Forecasting State Sales Tax Revenues: Comparing the Accuracy of Different Methods. *State and Local Government Review*, 142-152.

Gumbo, V., & Dhliwayo, L. (2018). VAT Revenue Modelling: The Case of zimbabwe. 1-44.

Gursoy, D., & Rutherford, D. G. (2004). Host attitudes toward tourism: An Improved Structural Model. *Annals of Tourism research*, 465-516 https://doi.org/410.1016/j.annals.2003.1008.1008.

Haas, M. P., & Knittel, M. J. (1997). *Revenue Forecasting in Michigan.* Paper presented at the: Proceedings. Annual Conference on Taxation and Minutes of the Annual, Michigan.

Hambor, J. C., Norman, M. R., & Russell, R. R. (1974). A tax revenue forecasting model for the State of Hawaii. *Public Finance Quarterly, 2*(4), 432-450. https://doi.org/410.1177/109114217400200403.

Hannon, A., Leahy, E., & O'Sullivan, R. (2016). An analysis of tax forecasting errors in Ireland. *The Economic and Social Review*, *47*(3, Autumn), 391-423.

Hoerl, A. E., & Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *12*(1), 55-67.

Husain, F., Qasim, M. A., Khalid, M., & Ghani, E. (2010). The Relationship between Federal Government Revenues and Expenditures in Pakistan [with Comments]. *The Pakistan Development Review*, 641-649.

Hussain, M. A., Murshid, S. M., & Soomro, S. (2014). Forecast simulation of tax collection in Pakistan. *Journal of Business Strategies*, *8*(2), 31.

Hussain, M. A., & Siddiqui, A. A. (2014). Time Series Forecast Models for Pakistan's Economic Activities in Past Four Decades. *Journal of Business Strategies*, *8*(1), 1.

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. OTexts.

Ibragimov, M., Ibragimov, R., & Sirajiddinov, N. (2009). *Modelling and Forecasting Income Tax Revenue: The Case of Uzbekistan*. Retrieved from Tashken.

J.Jordan, E., A.Vogt, C., & P.DeShon, R. (2015). A stress and coping framework for understanding resident responses to tourism development. *Tourism Management*, 500-512 https://doi.org/510.1016/j.tourman.2015.1001.1002.

Jean-Baptiste, Hiriart-Urruty, & Lemaréchal, C. (2001). *Fundamentals of Convex Analysis*. Switzerland: Springer.

Jenkins, G. P., Kuo, C. Y., & Shukla, G. (2000). Tax analysis and revenue forecasting. *Cambridge, Massachusetts: Harvard Institute for International Development, Harvard University*.

Jonas, K., Rest, G. J., & Atkinson, T. (1992). Virginia's Revenue Forecasting Process and Models. *Public Budgeting and Finance*, 70-81.https://doi.org/10.1111/1540-5850.00939.

Fullerton Jr, T. M. (1989). A composite approach to forecasting state government revenues: Case study of the Idaho sales tax. *International Journal of Forecasting*, *5*(3), 373-380. https://doi.org/310.1016/0169-2070(1089)90040-X.

Jang, S. B. (2019, January). A design of a tax prediction system based on artificial neural network. In *2019 International Conference on Platform Technology and Service (PlatCon)* (pp. 1-4). IEEE. https://doi.org/10.1109/PlatCon.2019.8669416

Kallis, G., Kostakis, V., Lange, S., Muraca, B., Paulson, S., & & Schmelzer, M. (2018). Research on degrowth. *Annual Review of Environment and Resources*, 291-316 . https://doi.org/210.1146/annurev-environ-102017-025941.

Kauder, B. o. (2011). *Institutions and Public Sector Performance: Empirical Analyses of Revenue Forecasting and Spatial Administrative Structures*. M¨unchen: an der Ludwig-Maximilians-Universit¨at. https://doi.org/10.5282/edoc.13683

Keene, M., & Thomson, P. (2007). *An analysis of tax revenue forecast errors*. (No. 07/02). New Zealand Treasury Working Paper.

Kennedy, P. (2008). *A guide to modern econometrics*.

Khairina, D. M., Muaddam, A., Maharani, S., & Hatta, H. R. (2019). Forecasting of Groundwater Tax Revenue Using Single Exponential Smoothing Method. *E3S W eb of Conferences,* 1-5. https://doi.org/10.1051/e3sconf/201912523006

Koszegi, B., & Rabin, M. (2006). A Model of Reference-Dependent Preferences. *The Quarterly Journal of Economics, 121*(4), 1133-1165. https://doi.org/10.1093/qje/121.4.1133

Krol, R. (2010). Forecasting State Tax Revenue: A Bayesian Vector Autoregression Approach. 1-18.

Krol, R. (2013). Evaluating state revenue forecasting under a flexible loss function. *International Journal of Forecasting*, *29*(2), 282-289. https://doi.org/10.1016/j.ijforecast.2012.11.003

Krol, R. (2014). Forecast bias of government agencies. *Cato J.*, *34*, 99.

Larson, S., & Overton, M. (2024). Modeling approach matters, but not as much as preprocessing: Comparison of machine learning and traditional revenue forecasting techniques. *Public Finance Journal*, *1*(1), 29-48. https://doi.org/10.59469/pfj.2024.8

Lütkepohl, H., & Krätzig, M. (Eds.). (2004). *Applied time series econometrics*. Cambridge university press.

Makananisa, M. P. (2015). Forecasting Annual Tax Revenue of the South African Taxes Using Time Series Holt-Winters and ARIMA/SARIMA Models. South Africa: University of South Africa.

Malikov, T., Jalilov, S., Kobulov, K., & Sultonov, S. (2021). Methodological approaches to assessing and forecasting the tax potential of the region. *Turkish Journal of Computer and Mathematics Education*, 7056-7060.

Manan, A., Nawaz, Z., Ahmed, W., & Talib, M. N. A. (2022). Assessing the determinants of tax revenue: empirical evidence from Pakistan. *Pacific Business Review (International)*, *14*(11), 69-84.

Martinez-Vazquez, J. (2006). *Pakistan : A Preliminary Assessment of the Federal Tax System*. Retrieved from Atlanta.

Martinez-Vazquez, J. (2015). The role of taxation in Pakistan's revival. *OUP Catalogue*.

Mehmood, K., Ahmad, S., Mehmood, T., Mohsin, M., & Ishfaq, M. (2022). Does Laffer Curve Exist in Tax Structure of Pakistan? A Threshold Regression Analysis. *Journal of Economic Impact*, *4*(1), 145-149. https://doi.org/10.52223/jei4012217

Mikesell, J. L. (2018). Often Wrong, Never Uncertain: Lessons from 40 Years of State Revenue Forecasting. *Public Administration Review*, 795-802. https://doi.org/10.1111/puar.12954

Molapo, M. A. (2018). Employing Bayesian Vector Auto-Regression (BVAR) Method as an Alternative Technique for Forecasting Tax Revenue in South Africa (Doctoral dissertation, MSc dissertation. University of South Africa).

Molapo, M. A., Olaomi, J. O., & Ama, N. O. (2019). Bayesian Vector Auto-Regression Method as an Alternative Technique for Forecasting South African Tax Revenue. *Southern African Business Review*, 1-28 https://doi.org/10.25159/21998-28125/24416.

Mukherjee, S., & Bhattacharya, R. (2023). Revenue forecasting of corporate income tax (CIT) in India. *Indian Economic Review*, *58*(2), 329-349.

Munir, K., & Sultan, M. (2018). Are some taxes better for growth in Pakistan? A time series analysis. *International journal of social economics*, *45*(10), 1439-1452.

Nakorji, M., & Aminu, U. (2022). Forecasting inflation using machine learning techniques. *The Review of Finance and Banking*, *14*(1), 45-55.http//dx.doi.org/10.24818/rfb.22.14.01.04

Nandi, B. K., Chaudhury, M., & Hasan, G. Q. (2014). *Univariate Time Series Forecasting: A Study of Monthly Tax revenue of Bangladesh*.

Noor, N., Sarlan, A., & Aziz, N. (2022). *Revenue Prediction for Malaysian Federal Government Using Machine Learning Technique.* Paper presented at the 2022 11th International Conference on Software and Computer Applications. https://doi.org/10.1145/3524304.3524337

Ofori, M. S., Fumey, A., & Nketiah-Amponsah, E. (2020). Forecasting Value Added Tax Revenue in Ghana. *Journal of Economics and Financial Analysis*, 63-99. https://ojs.tripaledu.com/index.php/jefa/article/view/58

Pavlík, M. (2011). Forecasting of the Income Tax in the Slovak Republic. *Ekonometria*, (31), 125-131.

Penner, R. G. (2008). Federal Revenue Forecasting. In *Government Budget Forecasting Theory and Practice* (pp. 11-25). New York: CRC Press Taylor and Francis Group.

Pesaran, H. H., & Shin, Y. J. E. l. (1998). Generalized impulse response analysis in linear multivariate models. *58*(1), 17-29. https://doi.org/10.1016/S0165-1765(97)00214-0

Petropoulos, F., Apiletti, D., & Assimakopoulos, V. a., et. (2022). Forecasting: theory and practice. *International Journal of Forecasting*. https://doi.org/10.1016/j.ijforecast.2021.1011.1001.

Qasim, M. A., & Khalid, M. (2016). Accuracy of Revenue Forecast: Analysis of Pakistan's Federal Revenue. *Forman Journal of Economic Studies*, 41-56.

Reed, D. A. (1983). *A Simultaneous Equation Tax Revenue Forecasting Model for the State of indiana*. Indiana University.

Rehman, W., Khan, R. A., & Kousar, S. (2020). Government revenue and economic growth of Pakistan (1979–2017). *Studies of Applied Economics*, *38*(3). https://doi.org/10.25115/eea.v38i3.3034

Renner, N. A. (1986). *Forecasting quarterly sales tax revenues: a comparative study.* North Texas State University.

Robertson, M. (2018). Wrongful Conviction, Persuasion and Loss Aversion. *Centre for Research in Economic Theory and its Applications*, (48). https://dx.doi.org/10.2139/ssrn.3456756

Rubin, M. M., Peters, J. L., & Mantell, N. (1999). Revenue Forecasting and Estimation. In *Handbook on Taxation* (32). New York: Routledge Taylor & Francis Group https://doi.org/10.4324/9781315093161.

Sabaj, E., & Kahveci, M. (2018). *Forecasting tax revenues in an emerging economy: The case of Albania*. Retrieved from Istanbul.

Simonov, J., & Gligorov, Z. (2020). Customs Revenues Prediction Using Ensemble Methods (Statistical Modelling vs Machine Learning).

Satu, M. S., Abedin, M. Z., Khanom, S., Ouenniche, J., & Kaiser, M. S. (2020). *Application of Feature Engineering with Classification Techniques to Enhance Corporate Tax Default Detection Performance.* Paper presented at the Proceedings of International Conference on Trends in Computational and Cognitive Engineering Singapore. https://doi.org/10.1007/978-981-33-4673-4_5

Shirinov, S. (2021). The Concept of "Forecasting of Tax Revenues" and Approaches to its Definitions. *International Finance and Accounting*, 1-7.

Shkurti, W. J. (March 1990). A User's Guide to State Revenue Forecasting. *Public Budgeting and Finance*, 79-94 https://doi.org/10.1111/1540-5850.00853.

Špalek, J., Moravanský, D. J. C. J. o. E., & Finance. (2005). Evaluation of Tax-Revenue Forecasts in the Czech Republic (in Czech). *55*(3-4), 162-187.

Stock, & Watson. (2015). *Introduction to econometrics*.

Stock, J. H., & Watson, M. W. J. J. o. E. p. (2001). Vector autoregressions. *15*(4), 101-115.

Streimikiene, D., Ahmed, R. R., Vveinhardt, J., Ghauri, S. P., & Zahid, S. (2018). Forecasting tax revenues using time series techniques – a case of Pakistan. *Economic Research-Ekonomska Istraživanja*, 722–754 https://doi.org/710.1080/1331677X.1332018.1442236.

Stinson, T. F. (2006). Sources of error in state revenue forecasts or how can the forecast possibly be so far off. *Journal of Public Budgeting, Accounting & Financial Management*. https://doi.org/10.1108/JPBAFM-18-01-2006-B005.

Tasi'u, M., Usman, A. M., & Garba, H. D. (2024). Modelling and Forecasting Nigeria's Tax Revenue: A Comparative Analysis of SARIMA and Holt-Winters Models. *UMYU Scientifica*, *3*(3), 118-129.

Tolmacheva, E. A., & Podbornova, E. S. (2021). Comparative analysis of change in budget income from the increased VAT rates. *Vestnik of Samara University. Economics and Management*, *12*(1), 66-74. https://doi.org/10.18287/12542-10461-12021-18212-18281-18266-18274.

Ticona, W., Figueiredo, K., & Vellasco, M. (2017, August). Hybrid model based on genetic algorithms and neural networks to forecast tax collection: application using endogenous and exogenous variables. In *2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON)* (1-4). IEEE.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 297-323.

Tyler, C. (2017). Wanted: academics wise to the needs of government. *Nature*, https://doi.org/10.1038/d41586-41017-07744-41581.

Urrutia, J. D., Mingo, F. L. T., & Balmaceda, C. N. M. (2015). Forecasting Income Tax Revenue of the Philippines using Autoregressive Integrated Moving Average(ARIMA) Modeling: A Time Series Analysis. *American Research Thoughts*, 1938-1992 http://dx.doi.org/1910.6084/m1939.figshare.1469539

Voorhees, W. R. (2006). Consistent underestimation bias, the asymmetrical loss function, and homogeneous sources of bias in state revenue forecasts. *Journal of Public Budgeting, Accounting & Financial Management*. https://doi.org/10.1108/JPBAFM-18-01-2006-B003

Waciko, K. J., & Ismail, B. (2020). Performance of Shrinkage Methods for Forecasting GDP.

Wahid, U., & Wallace, S. (2008). *Incidence of taxes in Pakistan: primer and estimates* (No. paper0813). International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University.

Wang, D. (2018). Government Revenue Forecasting During Exceptional Times 1-125.

Wawire, N. (2017). Determinants of value added tax revenue in Kenya. *Journal of Economics Library*, *4*(3), 322-344. http://dx.doi.org/10.1453/jel.v4i3.1391

Willmott, C. J., & Matsuura, K. J. C. r. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *30*(1), 79-82.https/doi.org/10.3354/cr030079

Wong, J. D. (1995). Local Government Revenue Forecasting: Using Regression and Econometric Revenue Forecasting in a Medium-Sized City. *Journal of Public*

Budgeting, Accounting & Financial Management, 315-335. https://doi.org/10.1108/JPBAFM-07-03-1995-B001

Yu, Z., Yang, C., Zhang, Z., & Jiao, J. (2015). Error correction method based on data transformational GM(1,1) and application on tax forecasting *Applied Soft Computing*, 554-560. https://doi.org/510.1016/j.asoc.2015.1009.1001

Zafar, T. (2017). Taxation system of Pakistan and its impact on economy. *Punjab Economic Research Institute, Planning and Development Department. Lahore: Government of Punjab, Pakistan*.

Zaw, T., Kyaw, S. S., & Oo, A. N. (2020). *ARMA Model for Revenue Prediction.* Paper presented at the Proceedings of the 11th International Conference on Advances in Information Technology. https://doi.org/10.1145/3406601.3406617

Zhilou, Y., & Hua, J. (2010). *Research of tax revenue intelligent forecast system.* Paper presented at the 2010 International Forum on Information Technology and Applications. https://doi.org/10.1109/IFITA.2010.30

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, *67*(2),301-320.https://doi.org/10.1111/j.1467-9868.2005.00503.x

APPENDICES

Appendix 1A

Key Terms

1. Tax Revenue

Tax revenue is the revenue collected through taxes on any income, profit, social security contributions, goods, and services. It also includes payroll taxes, taxes on ownership, property transfer, and miscellaneous taxes (OECD,2022).

2. Tax Year

A tax year is a time period of 12 months ending on the 30th of June. It is represented by the year in which that last date falls i.e., 30th June. For example, the tax year for the period July 01, 2019, to June 30, 2020, will be declared the year 2020 (FBR,2021).

3. Forecasting

“Forecasting" is assessing the value that a quantity will assume at some future point in time(Marriott, 2002).

4. Forecast Error

Througout this thesis we denote to the difference between a forecast and actual results as a “forecast error.” This term is commonly used in forecasting whether they be forecasts of the inflation, the weather, or state tax revenue. It does not imply that the forecaster made an avoidable mistake or was somehow unprofessional. All forecasts of economic activity are wrong to some extent, regardless of the expertise of the forecaster or the quality of the tools used(Boyd & Dadayan, 2014).

5. Federal Board of Revenue (FBR)

The Federal Board of Revenue (FBR) collects, manages, and forecasts all federal taxes. The Federal Board of Revenue was created as the Central Board of Revenue (CBR) on the 01st of April, 1924, through the Central Board of Revenue act 1924. In 1944, the Revenue Division was created under the Ministry of Finance. In Pakistan, the arrangement continued till 1960, when CBR was created as an affiliated department of the Ministry of Finance. The CBR was converted into FBR by enacting the FBR act 2007.

6. Gross Domestic Product (GDP)

The GDP is the market value of all final goods and services, produced in the economy during a year. GDP is a better idea to visualize domestic production in the economy (PBS).

7. Federal Excise Duty (FED)

FED is the deduction of the amount of duty paid on goods used in the manufacture or production of other goods from the amount of duty payable on such other goods in the prescribed manner (FBR).

Appendix 2A.1

Buoyancy Approach Forecasts

Taxes

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85	9313	4674	15308	23371	52666
1985-86	9781	4928	15515	29343	59567
1986-87	10570	6409	15361	33300	78949
1987-88	11840	8743	17399	38000	88958
1988-89	13920	14700	20038	42362	105517
1989-90	14942	15575	23038	50665	114004
1990-91	19840	16909	24739	50520	124310
1991-92	28847	20799	30334	61821	156830
1992-93	36762	23516	35169	63225	171477
1993-94	43451	30379	34519	64240	199607
1994-95	61585	43571	43691	77652	248059
1995-96	78152	49869	51104	88908	293915
1996-97	85051	55868	55297	86094	310681
1997-98	103182	63942	62011	74496	338042
1998-99	103476	68680	60572	78654	375078
1999-00	112100	115767	55600	61600	386800
2000-01	124585	153565	49000	65047	422500
2001-02	142649	166561	47189	47818	459300
2002-03	151976	195139	44002	68836	534000
2003-04	164497	219167	45823	91045	583000
2004-05	176930	238537	58670	115374	624700
2005-06	215000	294798	55000	138384	766900
2006-07	334168	309396	71575	132299	852866
2007-08	387563	377430	83594	132200	1009902
2008-09	440271	451744	116055	148403	1158586
2009-10	528649	516348	121182	160272.7	1418007
2010-11	594689	633357	137207	184853	1634775
2011-12	731941	804899	122014	216906	1945697
2012-13	735758	842528	119453	239459	2048509
2013-14	884118	996382	139084	242810	2374540
2014-15	1029244	1087790	163969	306220	2811773
2015-16	1191602	1302371	190581	404572	3112472
2016-17	1343197	1328965	198570	496000	3367874
2017-18	1536636	1485306	205877	608373	3843755
2018-19	1445594	1459213	233591	685575	3828482

Forecasts

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86					
1986-87	10372.68	5629.391	18961.32	32462.39	71845.16
1987-88	12038.03	6197.692	18238.98	33005.13	70549.41
1988-89	14431.03	8667.04	19323.06	40460.55	84066.23
1989-90	15412.41	13784.62	23649.18	46838.31	119597.6
1990-91	18565.63	21249.37	24368.24	55889.8	134588.7
1991-92	21550.2	24689.08	28412.84	65259.92	156229.6
1992-93	30247.75	26873.08	35894.13	70923.83	189261.5
1993-94	38138.81	29881.31	49040.75	79994.79	219213.1
1994-95	49077.27	35395.79	47332.5	84310.86	243059.4
1995-96	69240.04	67198.43	54010.36	103120.4	328861.5
1996-97	84059	55640.66	56827.36	101672	322206.3
1997-98	63362.48	67900	63173.12	115397.4	399428.4
1998-99	107783.6	57137.05	66573.62	83211.41	368719.4
1999-00	130029.1	75439.1	73313.54	95554.15	425277
2000-01	114984.2	123167.8	60782.18	66657.19	408691.6
2001-02	140648.4	174725.5	54935.48	67675.41	476412.2
2002-03	155294.9	176690.8	49664.01	53252.48	495239.5
2003-04	162213.9	213684.2	51462.23	74996.55	584752.1
2004-05	207208.6	250670.6	52896.12	98725.61	663072.3
2005-06	240399.2	264131.7	51362.27	112632.8	702854
2006-07	353412.1	325725.3	64165.97	150489.8	848900.4
2007-08	426528.8	350722.4	80503.56	150695.8	960738.9
2008-09	500894	429293.1	109885.5	173432.5	1159310
2009-10	581680.5	497150.5	126502.6	162021.4	1283463
2010-11	685292.4	587222.8	135982.8	181683.9	1748674
2011-12	825393.6	707922.3	151672.1	204609.6	2002988
2012-13	801066.8	876847.8	131887.9	234416.7	2157748
2013-14	973387.9	928451.9	129891	261791	2271027
2014-15	1103407	1068804	147752.1	256723.3	2557985
2015-16	1263602	1153780	172561.9	323018	2994000
2016-17	1452461	1408987	204052.9	436670.1	3670715
2017-18	1664725	1439659	212879.7	534445.2	3971324
2018-19	1550608	1592947	218722.4	649641.8	4380153

Errors

	Direct Tax	Sales Tax	FED	Customs	Total
RMSE	52677.7	46927.69	9618.702	28493.63	130978.7
MAE	35701.56	29306.46	7048.247	19950.9	71486.38

Log Dif Taxes					
FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86	0.04903	0.052918	0.013432	0.227558	0.123132
1986-87	0.077578	0.26277	-0.00998	0.126503	0.2817
1987-88	0.113464	0.31055	0.124581	0.132029	0.119362
1988-89	0.161843	0.519594	0.141218	0.108666	0.170708
1989-90	0.070849	0.05782	0.139515	0.178984	0.077361
1990-91	0.283524	0.082179	0.071236	-0.00287	0.086545
1991-92	0.374306	0.207059	0.203888	0.201874	0.232384
1992-93	0.242459	0.122776	0.147896	0.022457	0.089287
1993-94	0.167169	0.256071	-0.01866	0.015926	0.151901
1994-95	0.348784	0.36064	0.235632	0.189611	0.217316
1995-96	0.238237	0.135008	0.156721	0.135365	0.169624
1996-97	0.084595	0.113592	0.078856	-0.03216	0.055476
1997-98	0.193243	0.134985	0.114593	-0.14469	0.084403
1998-99	0.002845	0.071482	-0.02348	0.054313	0.103964
1999-00	0.080052	0.522122	-0.08565	-0.2444	0.030774
2000-01	0.105597	0.282544	-0.12636	0.054448	0.088282
2001-02	0.135399	0.081238	-0.03766	-0.30771	0.083514
2002-03	0.063336	0.158351	-0.06993	0.364325	0.150692
2003-04	0.07917	0.116122	0.040551	0.279627	0.087791
2004-05	0.072862	0.08469	0.247142	0.236825	0.069084
2005-06	0.194884	0.211766	-0.0646	0.181853	0.205085
2006-07	0.441006	0.048332	0.263413	-0.04497	0.106246
2007-08	0.148235	0.198763	0.155226	-0.00075	0.169006
2008-09	0.127512	0.179731	0.328092	0.115616	0.137347
2009-10	0.182934	0.133665	0.043229	0.076945	0.202052
2010-11	0.117714	0.204253	0.124197	0.142684	0.142253
2011-12	0.207661	0.239683	-0.11735	0.159903	0.174115
2012-13	0.005201	0.04569	-0.02121	0.098918	0.051492
2013-14	0.183689	0.167724	0.152155	0.013897	0.147692
2014-15	0.151989	0.087773	0.164599	0.232025	0.169012
2015-16	0.146474	0.180038	0.1504	0.278526	0.101602
2016-17	0.119754	0.020214	0.041064	0.203746	0.078864
2017-18	0.134543	0.11122	0.036137	0.204212	0.132168
2018-19	-0.06108	-0.01772	0.126293	0.11947	-0.00398

Log Dif Forecasts					
FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86					
1986-87					
1987-88	0.148896	0.096176	-0.03884	0.016581	-0.0182
1988-89	0.18131	0.33535	0.057738	0.203664	0.175292
1989-90	0.065792	0.464026	0.202029	0.146374	0.352528
1990-91	0.186139	0.432774	0.029952	0.17668	0.11809
1991-92	0.149073	0.150034	0.15356	0.154996	0.149104
1992-93	0.339037	0.084764	0.233733	0.083228	0.191802
1993-94	0.231811	0.106108	0.312078	0.120355	0.146915
1994-95	0.252164	0.16936	-0.03545	0.052549	0.103262
1995-96	0.344183	0.641057	0.131979	0.201387	0.302331
1996-97	0.19394	-0.18874	0.050842	-0.01415	-0.02044
1997-98	-0.28265	0.199122	0.105861	0.12663	0.214842
1998-99	0.531253	-0.17258	0.05243	-0.327	-0.08
1999-00	0.187633	0.277873	0.096437	0.138309	0.142705
2000-01	-0.12296	0.490222	-0.18745	-0.36013	-0.03978
2001-02	0.201468	0.349669	-0.10114	0.01516	0.153323
2002-03	0.099063	0.011185	-0.10088	-0.23968	0.038758
2003-04	0.04359	0.190098	0.035568	0.342398	0.166146
2004-05	0.24481	0.159641	0.027482	0.274902	0.125696
2005-06	0.148575	0.052308	-0.02943	0.131788	0.058265
2006-07	0.385334	0.209606	0.222569	0.289763	0.188793
2007-08	0.188045	0.073941	0.226828	0.001368	0.123761
2008-09	0.160715	0.202145	0.311138	0.140525	0.187878
2009-10	0.149527	0.146753	0.140824	-0.06806	0.101737
2010-11	0.163924	0.166512	0.072266	0.11454	0.309296
2011-12	0.186015	0.18693	0.109193	0.118836	0.135782
2012-13	-0.02992	0.213999	-0.13977	0.135996	0.074425
2013-14	0.194838	0.057185	-0.01526	0.110446	0.051167
2014-15	0.125375	0.140777	0.12884	-0.01955	0.118988
2015-16	0.135563	0.076503	0.15522	0.229709	0.15739
2016-17	0.139293	0.199828	0.167624	0.30147	0.203776
2017-18	0.1364	0.021536	0.042348	0.202051	0.078713
2018-19	-0.07101	0.101179	0.027076	0.195192	0.097984

Errors					
	Direct Tax	Sales Tax	FED	Customs	Total
RMSE	0.150851	0.196116	0.121269	0.204486	0.10793
MAE	0.090527	0.159996	0.090234	0.142784	0.092808

Appendix 2A.2

Box-Jenkins Forecasts

Total Tax

Forecast: DLTTF	
Actual: DLTT	
Forecast sample: 1 35	
Adjusted sample: 12 35	
Included observations: 24	
Root Mean Squared Error	0.050592
Mean Absolute Error	0.036871
Mean Abs. Percent Error	156.7217
Theil Inequality Coef.	0.203529
Bias Proportion	0.011812
Variance Proportion	0.247414
Covariance Proportion	0.740774
Theil U2 Coefficient	0.684554
Symmetric MAPE	35.72445

Direct Tax

Forecast: DLDTF	
Actual: DLDT	
Forecast sample: 1 35	
Adjusted sample: 10 35	
Included observations: 26	
Root Mean Squared Error	0.088913
Mean Absolute Error	0.067593
Mean Abs. Percent Error	280.4075
Theil Inequality Coef.	0.272816
Bias Proportion	0.003416
Variance Proportion	0.393228
Covariance Proportion	0.603356
Theil U2 Coefficient	0.163214
Symmetric MAPE	54.49412

Fderal Excise Duty

Forecast: DLFEDF	
Actual: DLFED	
Forecast sample: 1 35	
Adjusted sample: 10 35	
Included observations: 26	
Root Mean Squared Error	0.115672
Mean Absolute Error	0.101134
Mean Abs. Percent Error	148.7669
Theil Inequality Coef.	0.522669
Bias Proportion	0.001029
Variance Proportion	0.734684
Covariance Proportion	0.264287
Theil U2 Coefficient	0.596494
Symmetric MAPE	114.9589

Cutom Duties

Forecast: DLCUSTOMSF	
Actual: DLCUSTOMS	
Forecast sample: 1 35	
Adjusted sample: 12 35	
Included observations: 24	
Root Mean Squared Error	0.157679
Mean Absolute Error	0.124007
Mean Abs. Percent Error	664.6882
Theil Inequality Coef.	0.560650
Bias Proportion	0.001578
Variance Proportion	0.860137
Covariance Proportion	0.138285
Theil U2 Coefficient	0.178694
Symmetric MAPE	98.99915

Sales Tax

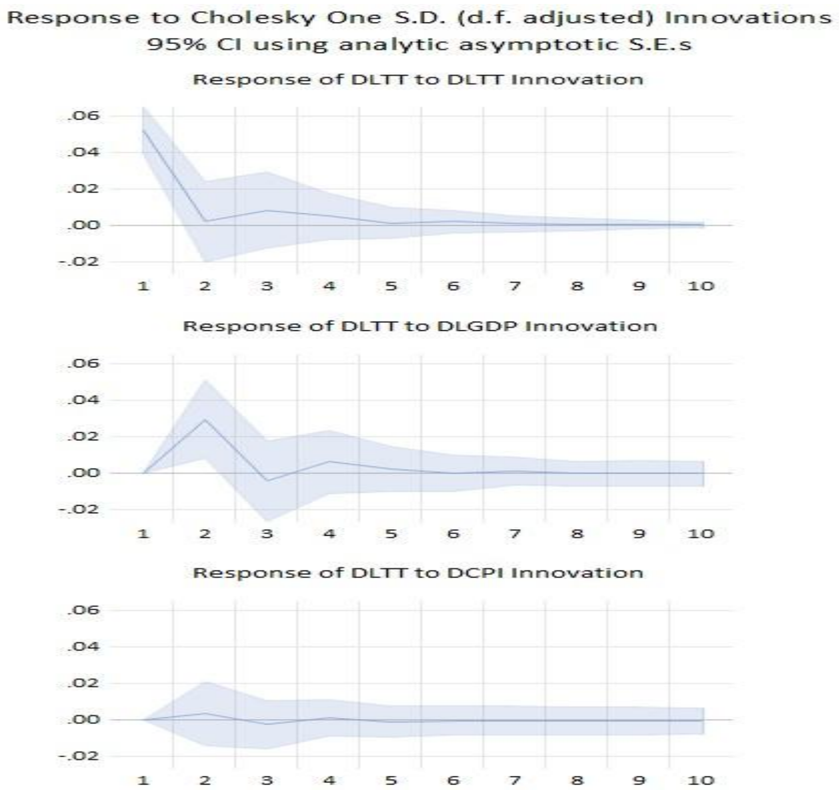
Forecast: DLSALESF	
Actual: DLSALES	
Forecast sample: 1 35	
Adjusted sample: 8 35	
Included observations: 28	
Root Mean Squared Error	0.088962
Mean Absolute Error	0.073241
Mean Abs. Percent Error	124.5185
Theil Inequality Coef.	0.227098
Bias Proportion	0.110534
Variance Proportion	0.224261
Covariance Proportion	0.665205
Theil U2 Coefficient	0.614135
Symmetric MAPE	51.11734

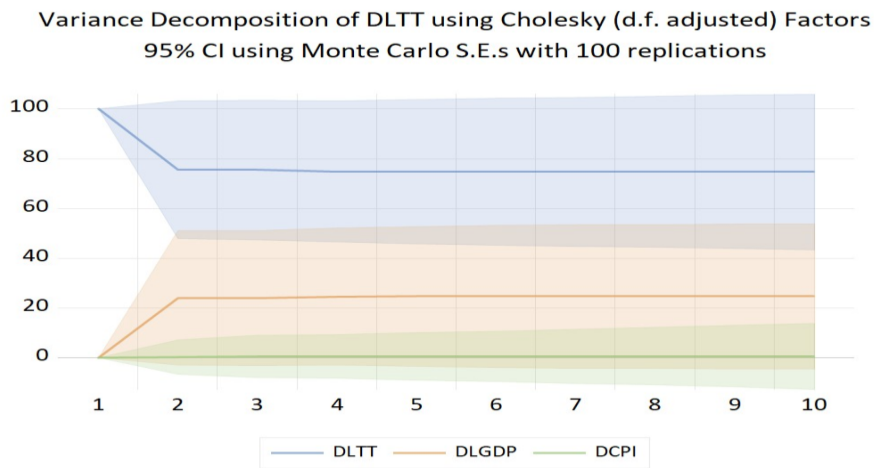
Appendix 2A.3

Sutructural Vector Autoregressive (SVAR) Model Forecasts

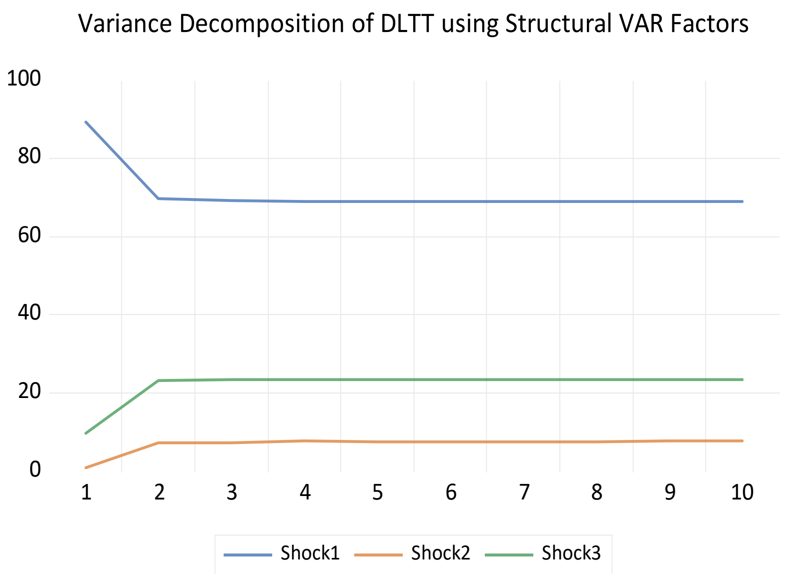
<u>Total Tax</u>

Short Term Cholesky





Long Term Blanchard-Quah



Forecast Errors

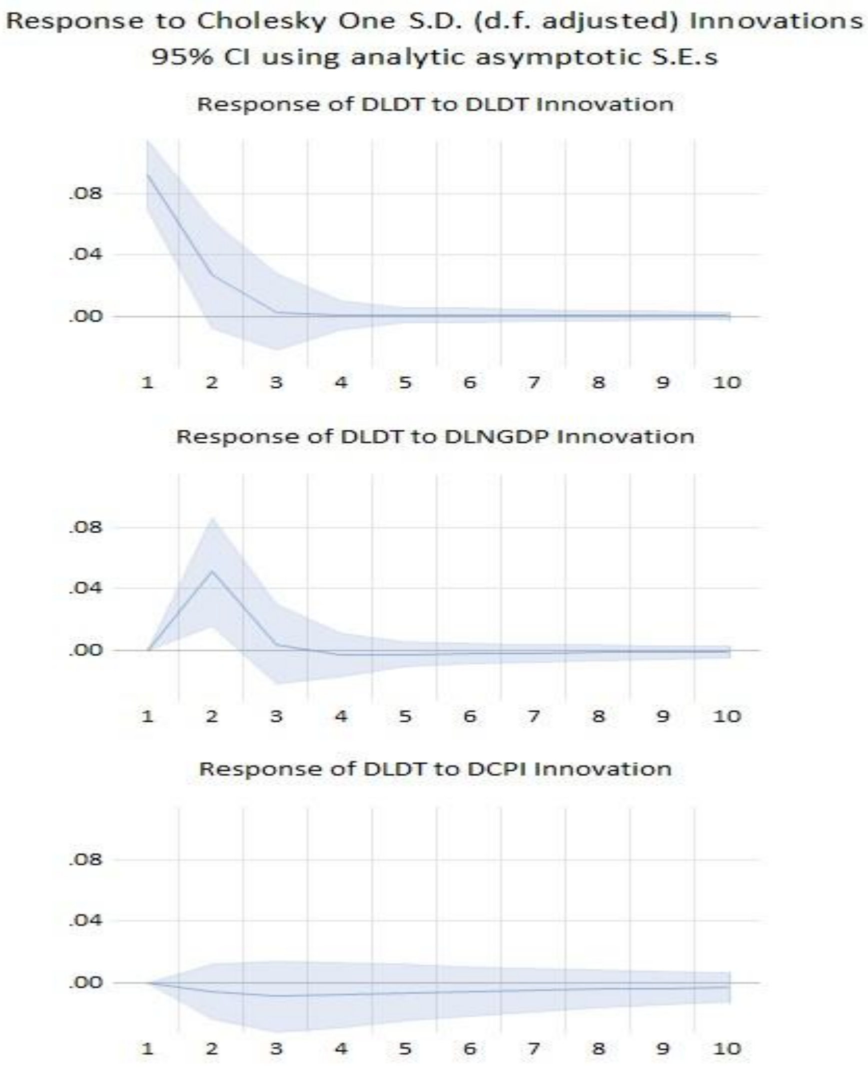
Forecast Evaluation
Sample: 1 34
Included observations: 34

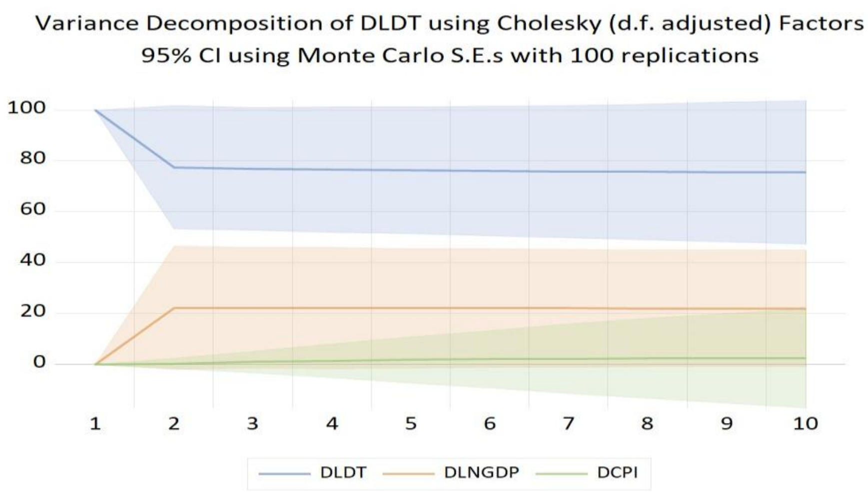
Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DCPI	34	1.837308	1.587079	81.62570	0.221140
DLGDP	34	0.036886	0.029824	26.96882	0.147039
DLTT	34	0.052382	0.044653	135.6055	0.205783

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

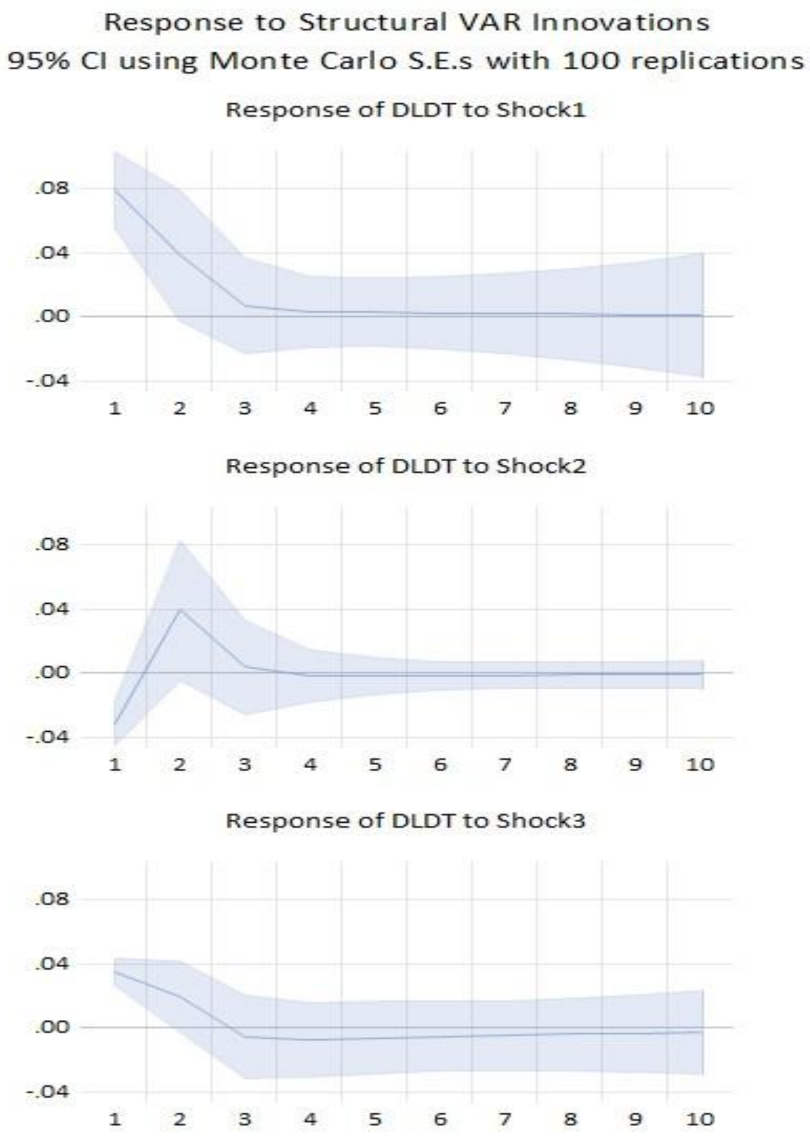
Direct Tax

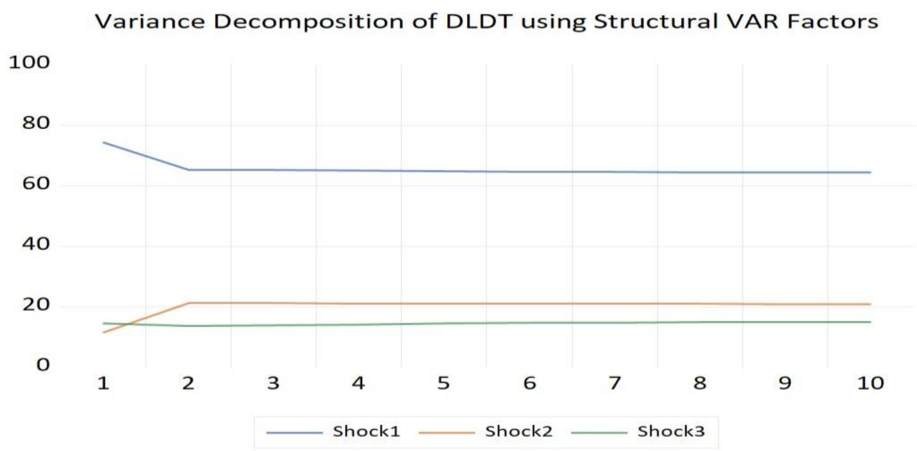
Short Term Cholesky





Long Term Blanchard-Quah





Forecast Errors

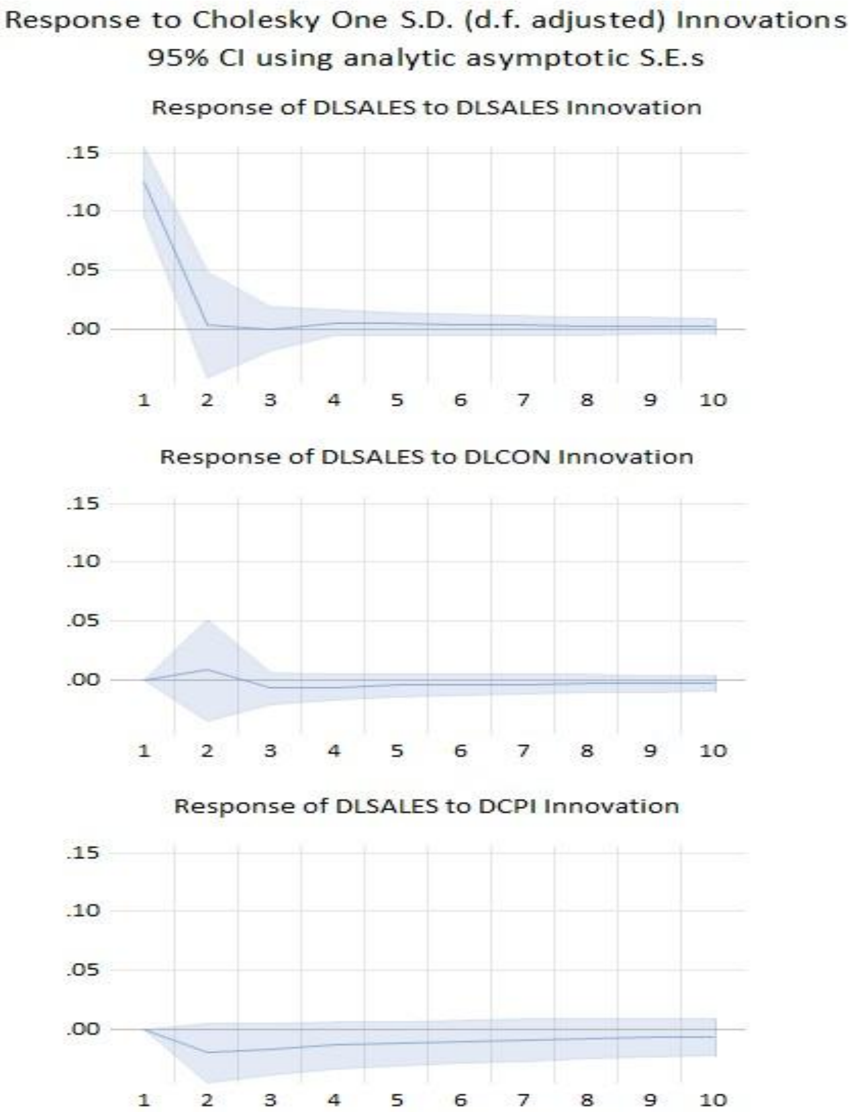
Forecast Evaluation
Sample: 1 34
Included observations: 34

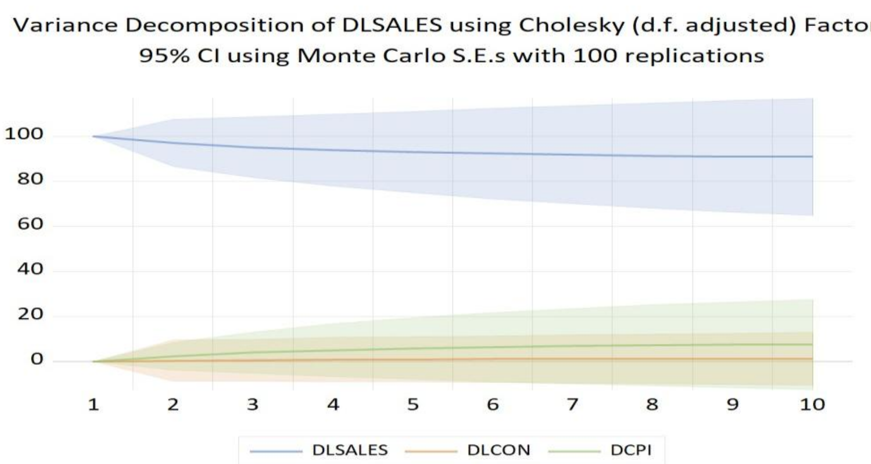
Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DCPI	34	2.105484	1.817591	96.90466	0.262876
DLDT	34	0.101356	0.073373	279.7958	0.307526
DLNGDP	34	0.048882	0.037250	29.47694	0.172454

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

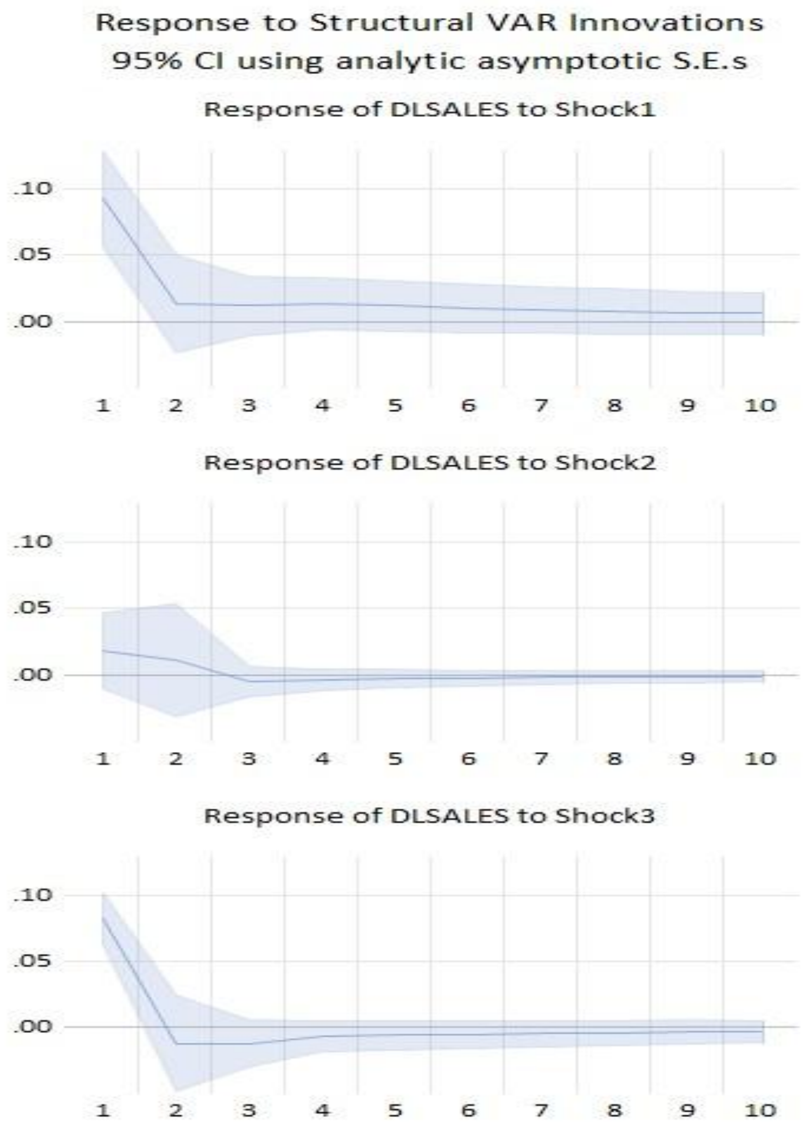
Sales Tax

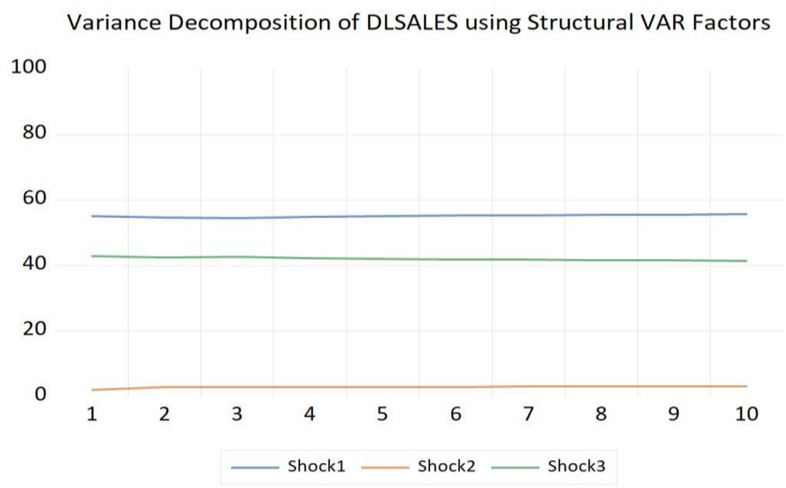
Short Term Cholesky





Long Term Blanchard-Quah





Forecast Errors

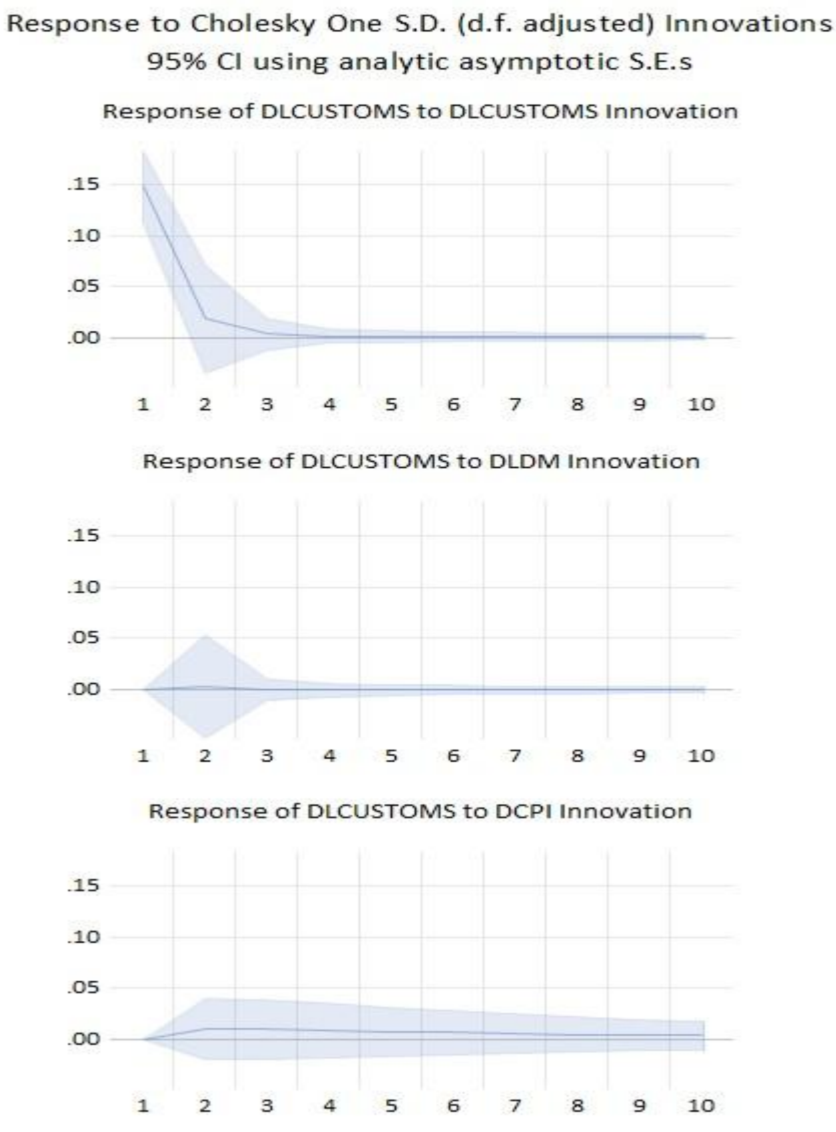
Forecast Evaluation
Sample: 1 34
Included observations: 34

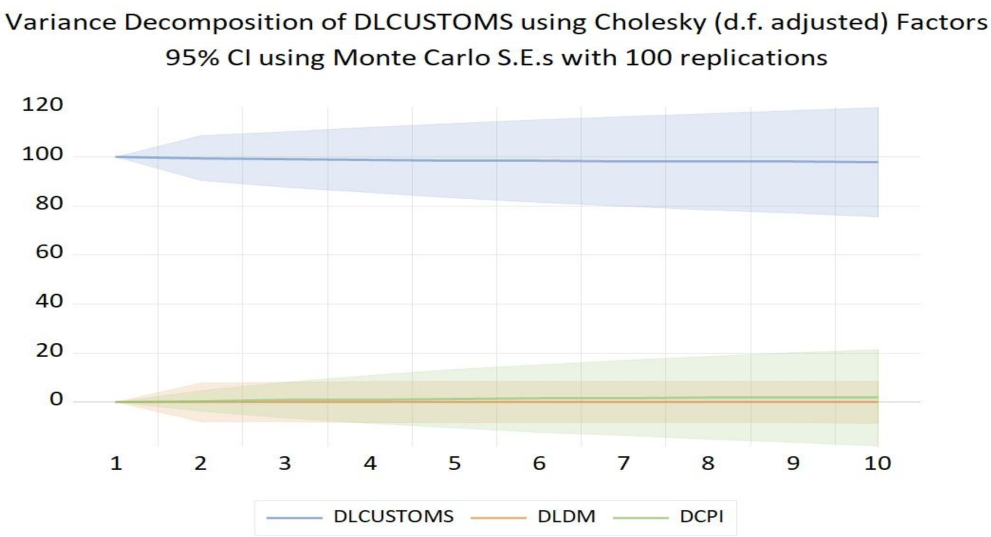
Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DCPI	34	2.122577	1.860454	105.8888	0.260603
DLCON	34	0.047133	0.037808	30.42340	0.170145
DLSALES	34	0.117898	0.087796	104.2536	0.314786

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

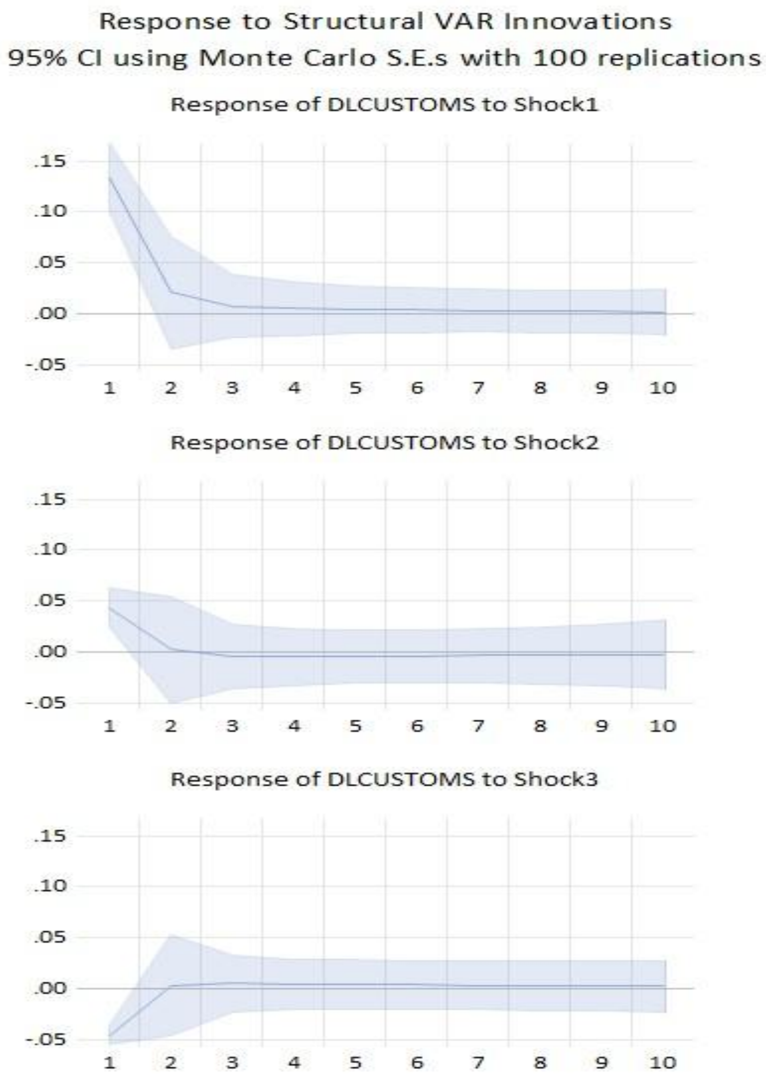
<u>Custom Duties</u>

Short Term Cholesky





Long Term Blanchard-Quah



Forecas Errors

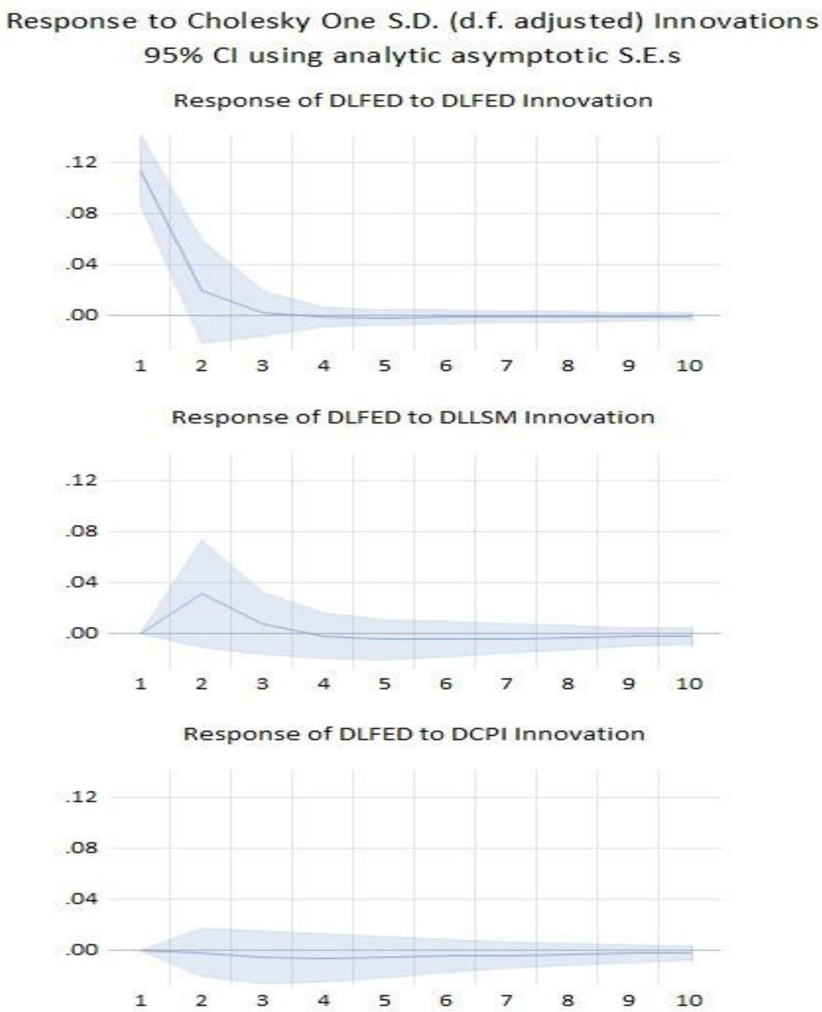
Forecast Evaluation
Sample: 1 34
Included observations: 34

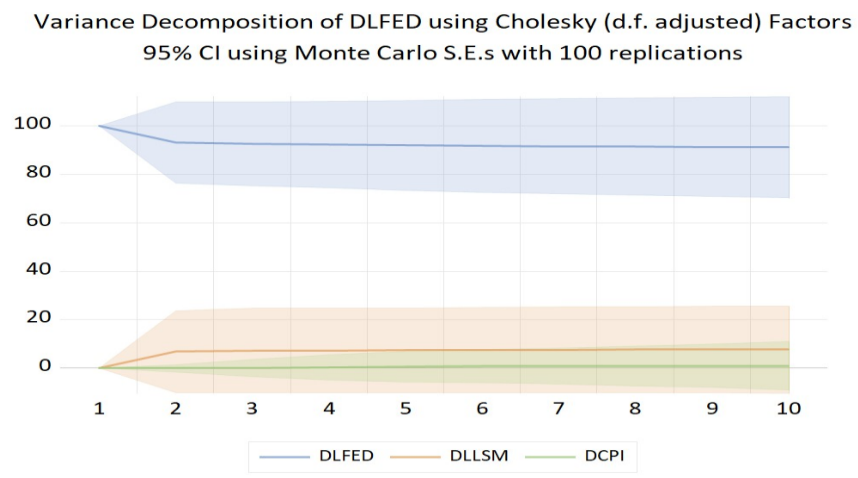
Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DCPI	34	2.114384	1.848266	103.6935	0.262004
DLCUSTOMS	34	0.141464	0.107707	641.7913	0.520648
DLDM	34	0.116723	0.094490	523.9644	0.368878

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

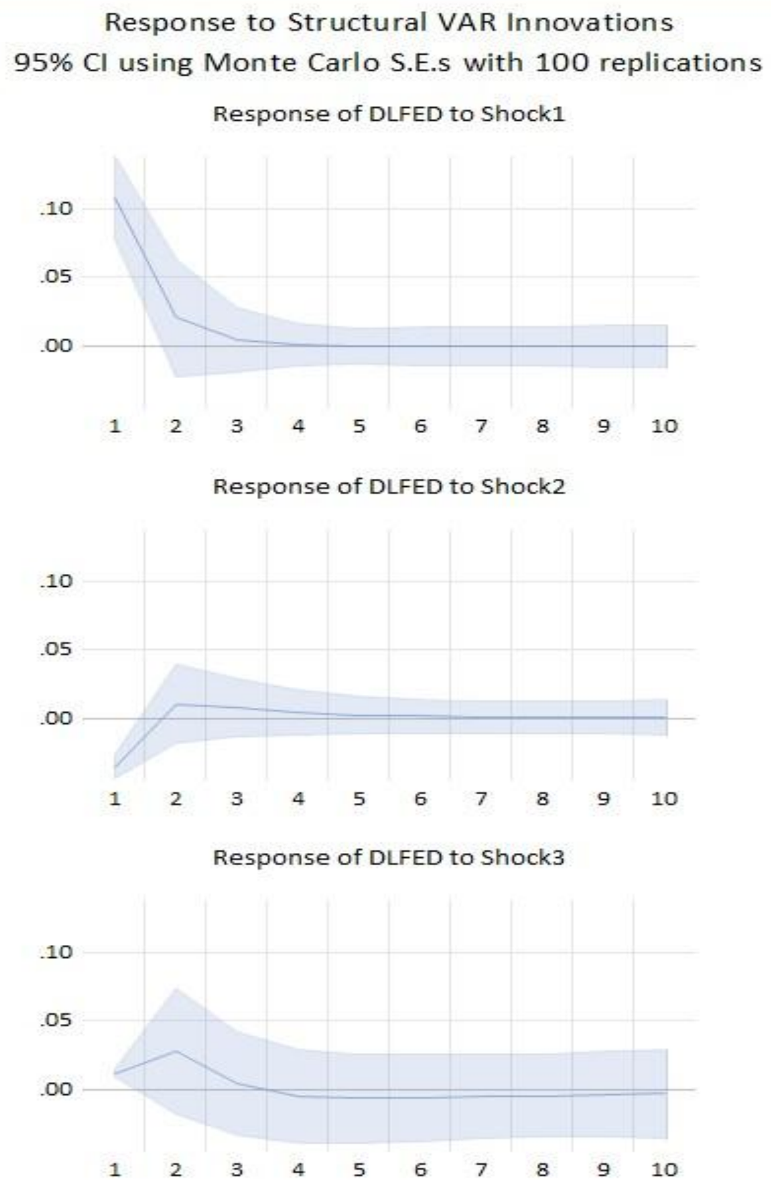
FED

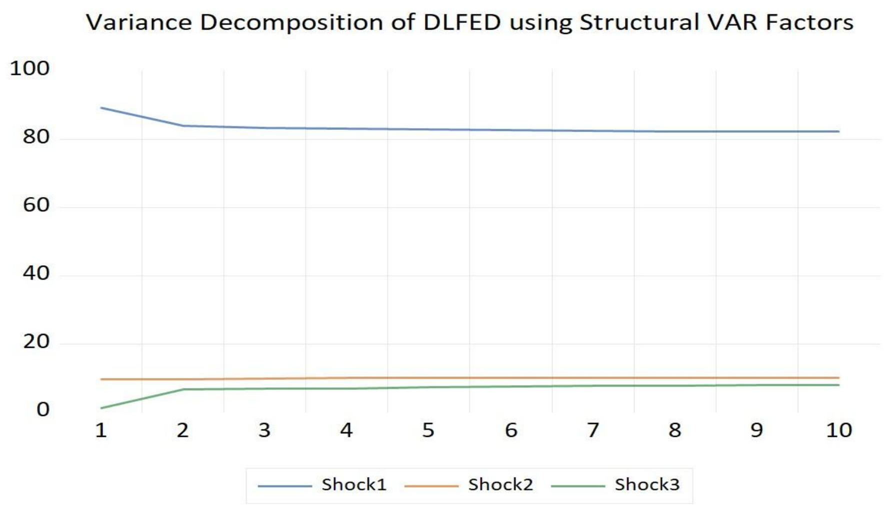
Short Term Cholesky





Long Term Blanchard-Quah





Forecast Errors

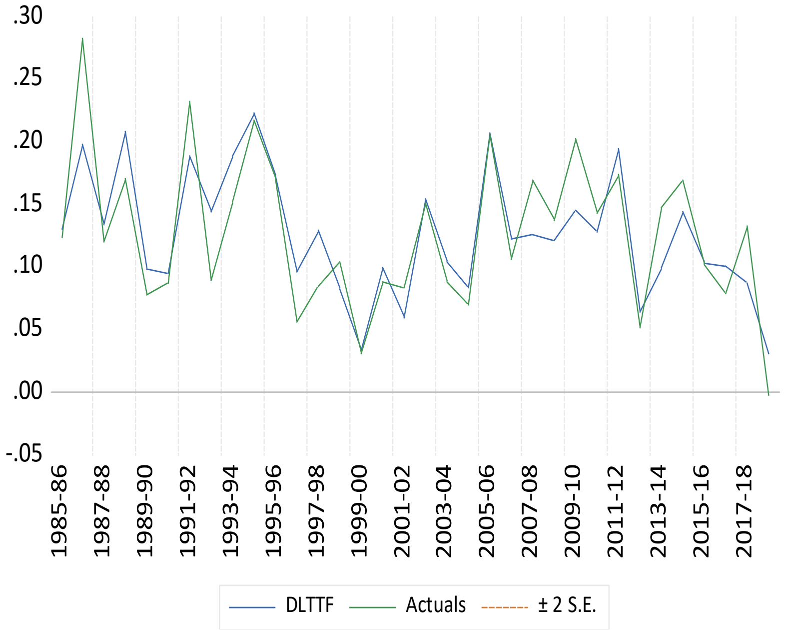
Forecast Evaluation
Sample: 1 34
Included observations: 34

Variable	Inc. obs.	RMSE	MAE	MAPE	Theil
DCPI	34	2.182367	1.900748	100.8134	0.272664
DLFED	34	0.111685	0.093700	140.1306	0.506974
DLLSM	34	0.064273	0.049881	162.4934	0.234096

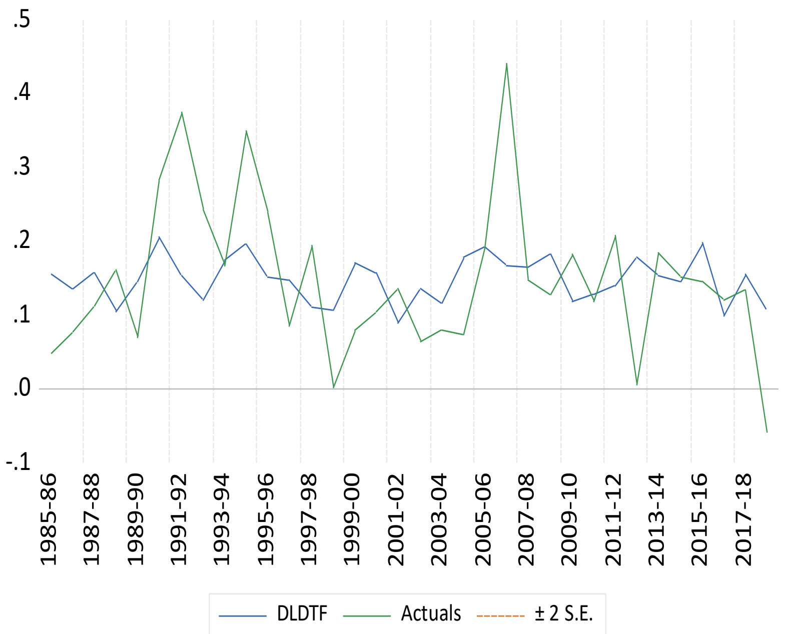
RMSE: Root Mean Square Error
MAE: Mean Absolute Error
MAPE: Mean Absolute Percentage Error
Theil: Theil inequality coefficient

Appendix 2A.4

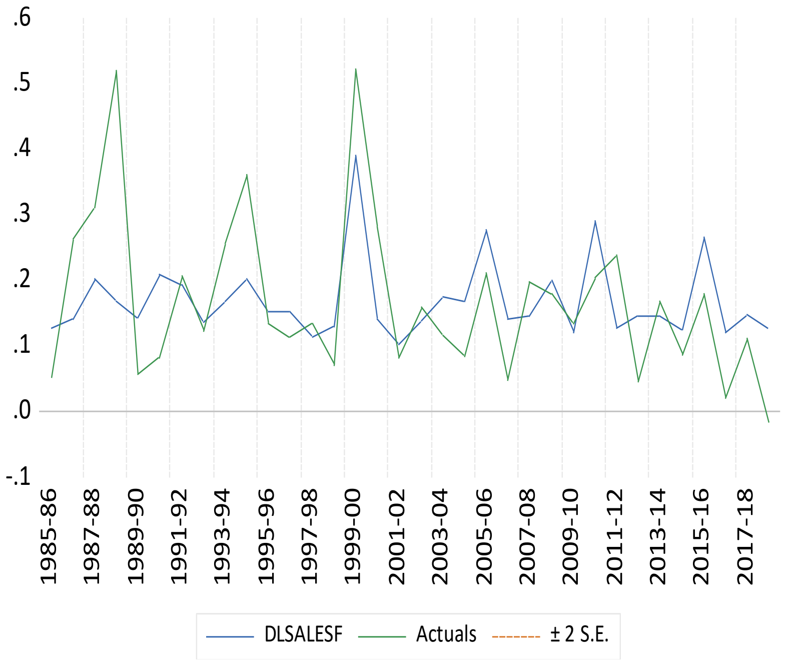
Elastic Net Forecasts



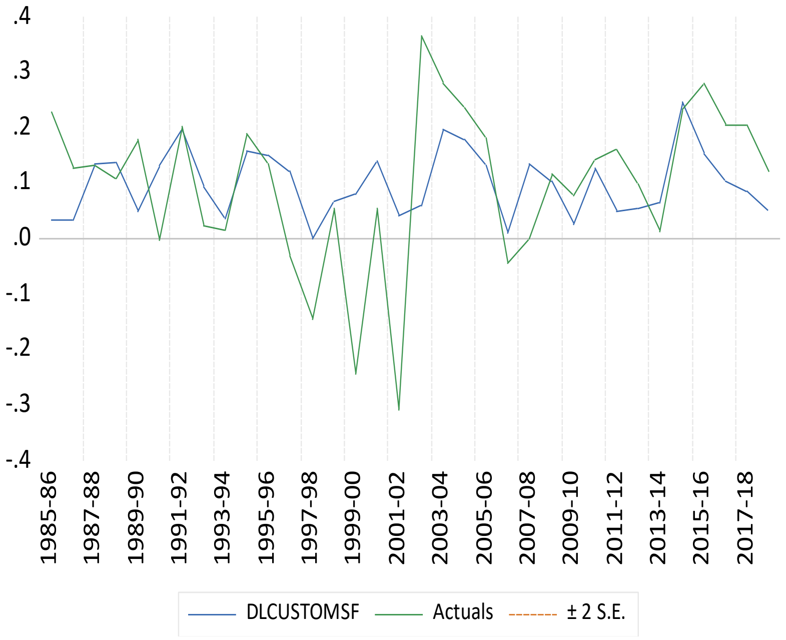
Forecast: DLTTF	
Actual: DLTT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.031472
Mean Absolute Error	0.024876
Mean Abs. Percent Error	45.14149
Theil Inequality Coef.	0.114454
Bias Proportion	0.000000
Variance Proportion	0.125081
Covariance Proportion	0.874919
Theil U2 Coefficient	0.409550
Symmetric MAPE	25.22539



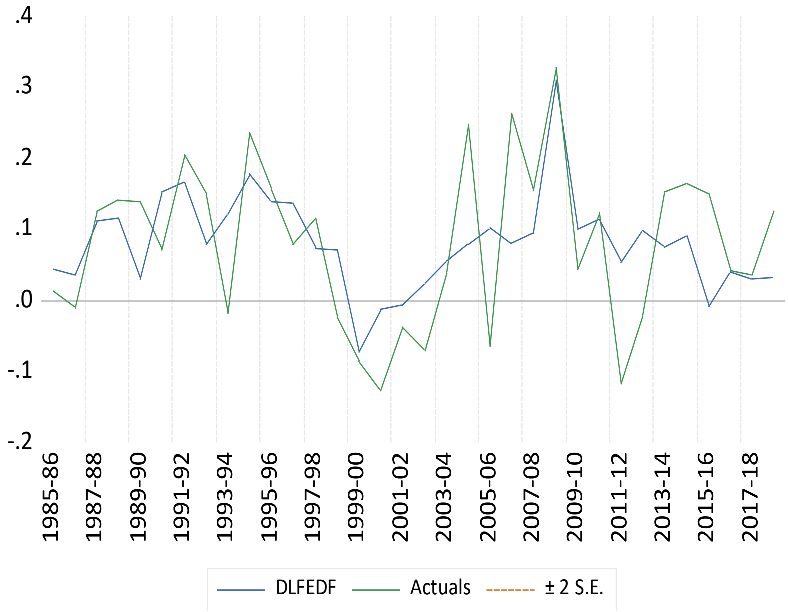
Forecast: DLDTF	
Actual: DLDT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.098301
Mean Absolute Error	0.077233
Mean Abs. Percent Error	261.7760
Theil Inequality Coef.	0.295954
Bias Proportion	0.000000
Variance Proportion	0.544849
Covariance Proportion	0.455151
Theil U2 Coefficient	0.748926
Symmetric MAPE	57.04637



Forecast: DLSALESF	
Actual: DLSALES	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.102228
Mean Absolute Error	0.079564
Mean Abs. Percent Error	92.22594
Theil Inequality Coef.	0.263773
Bias Proportion	0.000000
Variance Proportion	0.388165
Covariance Proportion	0.611835
Theil U2 Coefficient	0.494718
Symmetric MAPE	51.89763



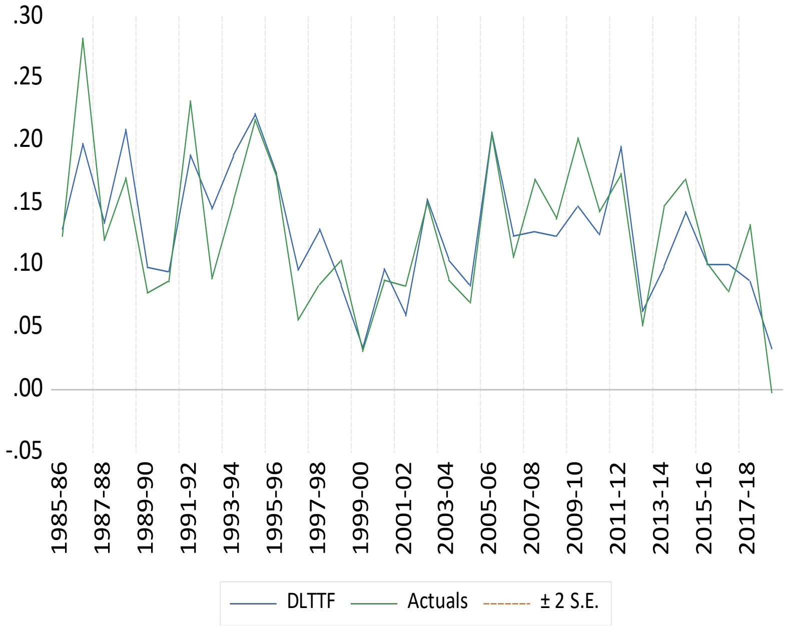
Forecast: DLCUSTOMSF	
Actual: DLCUSTOMS	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.128287
Mean Absolute Error	0.093922
Mean Abs. Percent Error	750.9610
Theil Inequality Coef.	0.445226
Bias Proportion	0.000000
Variance Proportion	0.411634
Covariance Proportion	0.588366
Theil U2 Coefficient	0.116440
Symmetric MAPE	90.86923



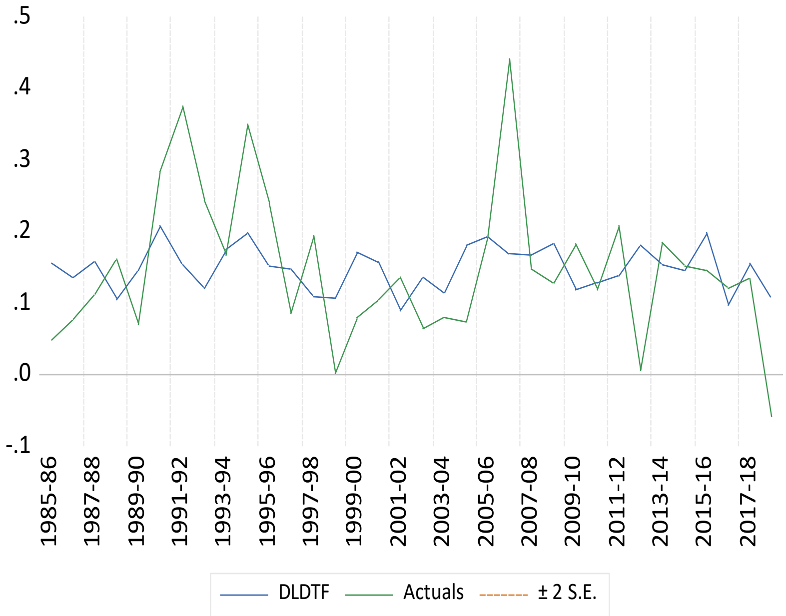
Forecast: DLFEDF	
Actual: DLFED	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.089866
Mean Absolute Error	0.072244
Mean Abs. Percent Error	123.9544
Theil Inequality Coef.	0.371179
Bias Proportion	0.000000
Variance Proportion	0.252505
Covariance Proportion	0.747495
Theil U2 Coefficient	0.401252
Symmetric MAPE	92.02326

Appendix 2A.5

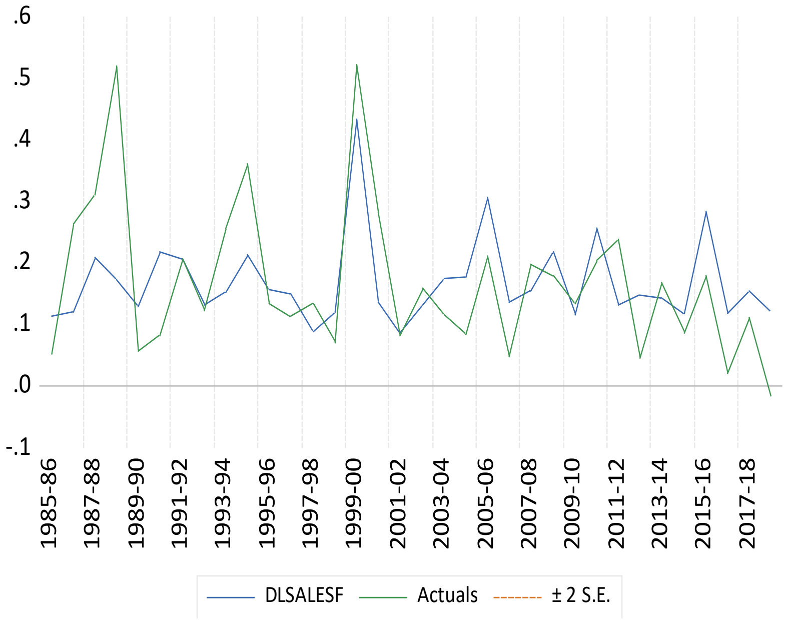
LASSO Forecasts



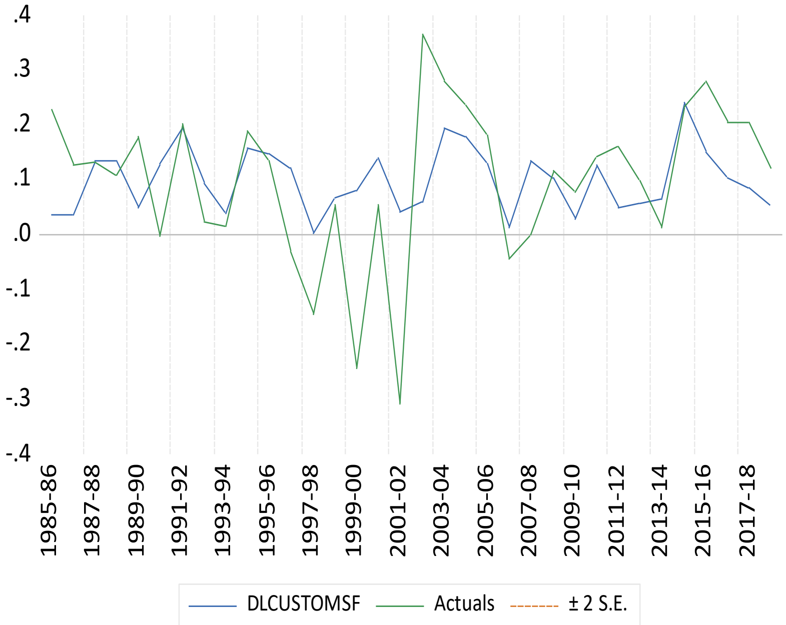
Forecast: DLTTF	
Actual: DLTT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.031535
Mean Absolute Error	0.024879
Mean Abs. Percent Error	46.39931
Theil Inequality Coef.	0.114694
Bias Proportion	0.000000
Variance Proportion	0.126290
Covariance Proportion	0.873710
Theil U2 Coefficient	0.407916
Symmetric MAPE	25.14545



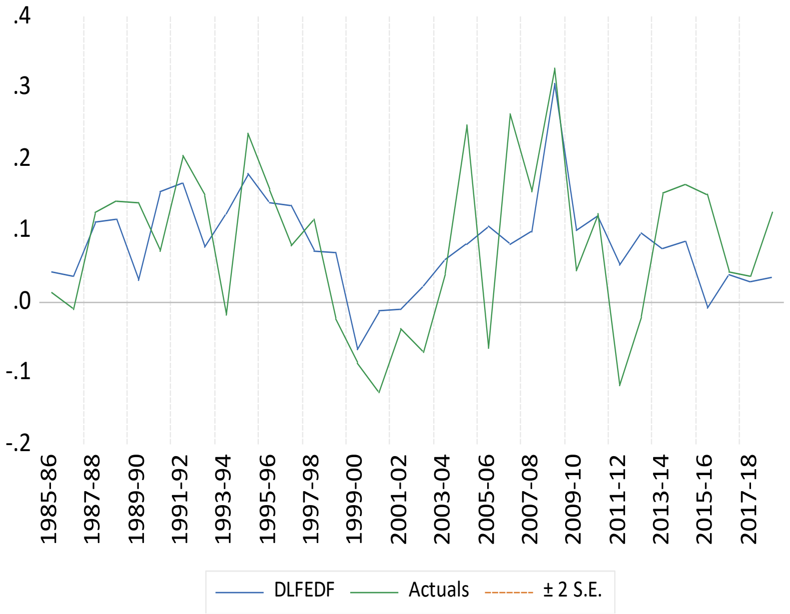
Forecast: DLDTF	
Actual: DLDT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.098299
Mean Absolute Error	0.077328
Mean Abs. Percent Error	261.3548
Theil Inequality Coef.	0.295862
Bias Proportion	0.000000
Variance Proportion	0.537697
Covariance Proportion	0.462303
Theil U2 Coefficient	0.744227
Symmetric MAPE	57.16472



Forecast: DLSALESF	
Actual: DLSALES	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.101071
Mean Absolute Error	0.078662
Mean Abs. Percent Error	89.79759
Theil Inequality Coef.	0.258520
Bias Proportion	0.000000
Variance Proportion	0.285991
Covariance Proportion	0.714009
Theil U2 Coefficient	0.518231
Symmetric MAPE	51.80909



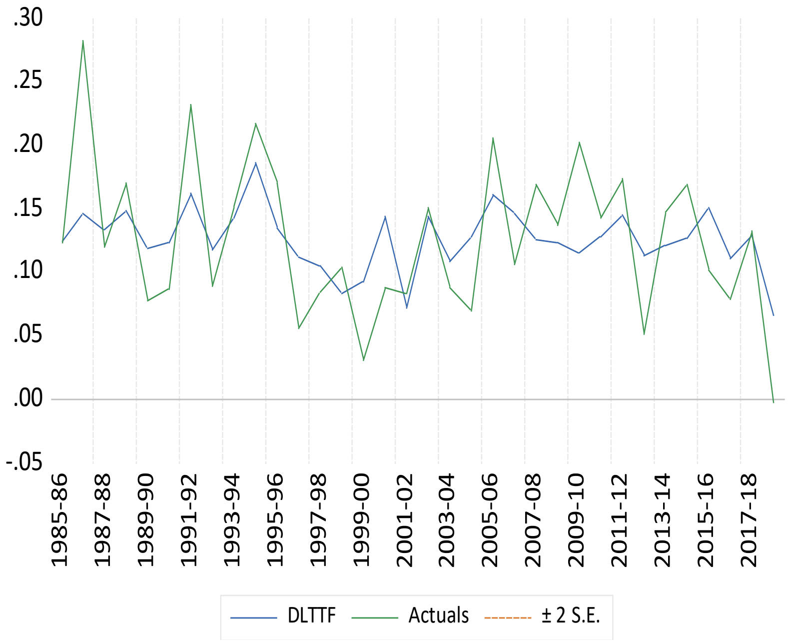
Forecast: DLCUSTOMSF	
Actual: DLCUSTOMS	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.128300
Mean Absolute Error	0.093917
Mean Abs. Percent Error	752.0282
Theil Inequality Coef.	0.446547
Bias Proportion	0.000000
Variance Proportion	0.428079
Covariance Proportion	0.571921
Theil U2 Coefficient	0.116590
Symmetric MAPE	90.46867



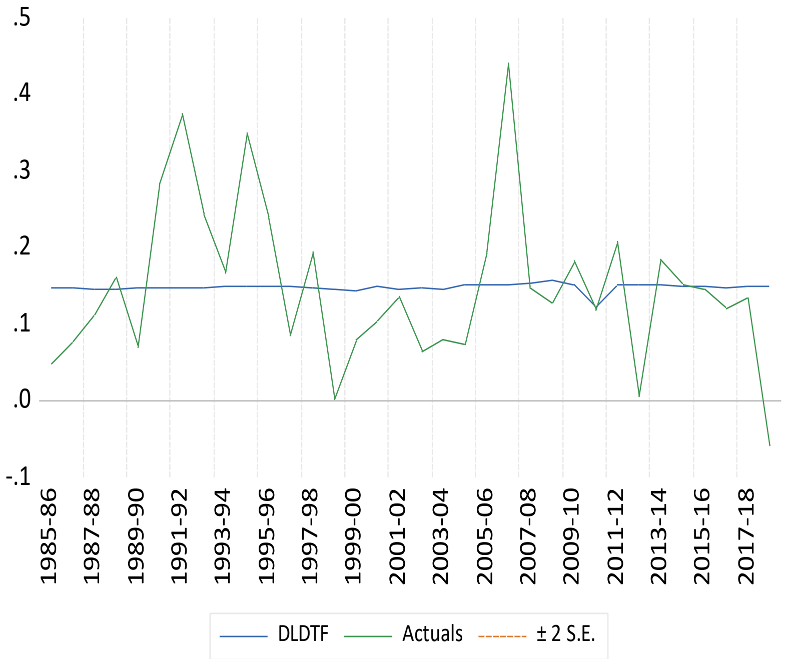
Forecast: DLFEDF	
Actual: DLFED	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.089895
Mean Absolute Error	0.072365
Mean Abs. Percent Error	123.5245
Theil Inequality Coef.	0.371644
Bias Proportion	0.000000
Variance Proportion	0.256271
Covariance Proportion	0.743729
Theil U2 Coefficient	0.399374
Symmetric MAPE	91.97676

Appendix 2A.6

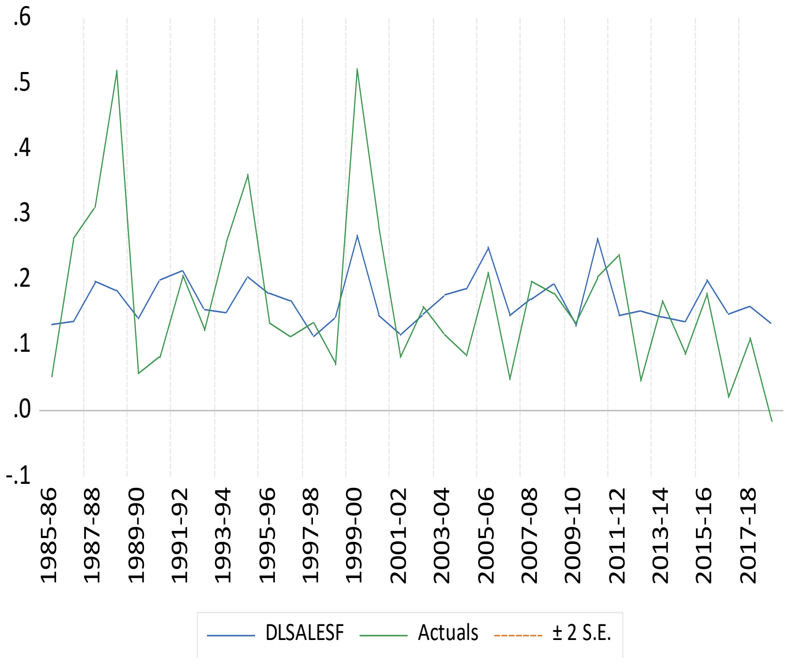
Ridge Regression Forecasts



Forecast: DLTTF	
Actual: DLTT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.046538
Mean Absolute Error	0.038033
Mean Abs. Percent Error	87.09327
Theil Inequality Coef.	0.173540
Bias Proportion	0.000000
Variance Proportion	0.580982
Covariance Proportion	0.419018
Theil U2 Coefficient	0.617631
Symmetric MAPE	35.64557

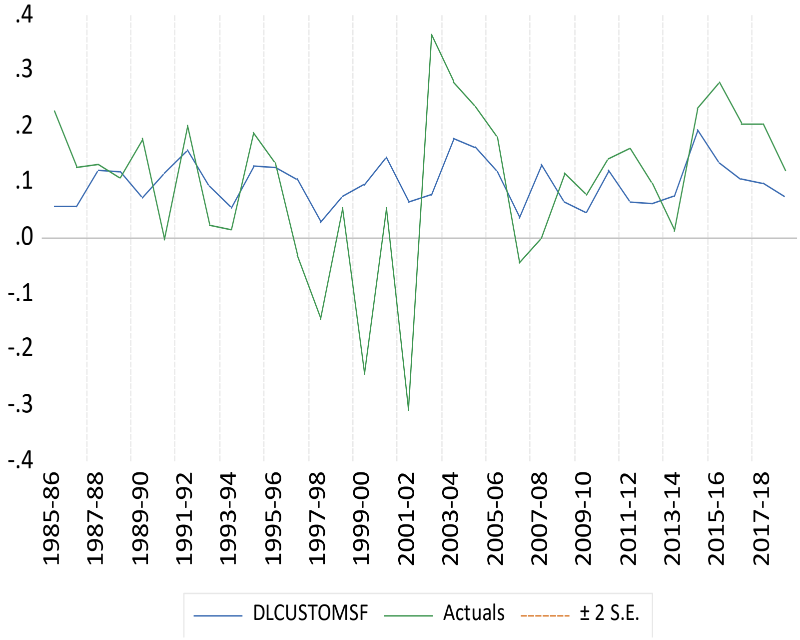


Forecast: DLDTF	
Actual: DLDT	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.102687
Mean Absolute Error	0.075331
Mean Abs. Percent Error	284.6556
Theil Inequality Coef.	0.311994
Bias Proportion	0.000000
Variance Proportion	0.909649
Covariance Proportion	0.090351
Theil U2 Coefficient	0.531386
Symmetric MAPE	53.86793



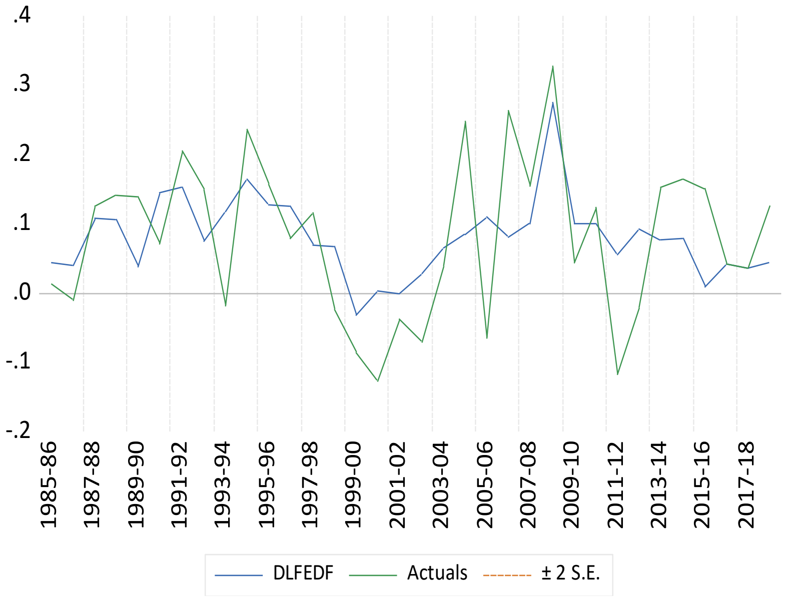
Forecast: DLSALESF
Actual: DLSALES
Forecast sample: 1 35
Adjusted sample: 2 35
Included observations: 34

Root Mean Squared Error	0.107579
Mean Absolute Error	0.082428
Mean Abs. Percent Error	99.01814
Theil Inequality Coef.	0.281671
Bias Proportion	0.000000
Variance Proportion	0.610857
Covariance Proportion	0.389143
Theil U2 Coefficient	0.581174
Symmetric MAPE	53.11383



Forecast: DLCUSTOMSF
Actual: DLCUSTOMS
Forecast sample: 1 35
Adjusted sample: 2 35
Included observations: 34

Root Mean Squared Error	0.130027
Mean Absolute Error	0.097196
Mean Abs. Percent Error	737.1242
Theil Inequality Coef.	0.464256
Bias Proportion	0.000000
Variance Proportion	0.595512
Covariance Proportion	0.404488
Theil U2 Coefficient	0.403747
Symmetric MAPE	88.85350



Forecast: DLFEDF	
Actual: DLFED	
Forecast sample: 1 35	
Adjusted sample: 2 35	
Included observations: 34	
Root Mean Squared Error	0.090694
Mean Absolute Error	0.075704
Mean Abs. Percent Error	126.2001
Theil Inequality Coef.	0.384919
Bias Proportion	0.000000
Variance Proportion	0.378893
Covariance Proportion	0.621107
Theil U2 Coefficient	0.427656
Symmetric MAPE	96.56471

Appendix 2A.7

Effective Tax Rate Forecast

Log Dif Taxes

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86	0.04903	0.052918	0.013432	0.227558	0.123132
1986-87	0.077578	0.26277	-0.00998	0.126503	0.2817
1987-88	0.113464	0.31055	0.124581	0.132029	0.119362
1988-89	0.161843	0.519594	0.141218	0.108666	0.170708
1989-90	0.070849	0.05782	0.139515	0.178984	0.077361
1990-91	0.283524	0.082179	0.071236	-0.00287	0.086545
1991-92	0.374306	0.207059	0.203888	0.201874	0.232384
1992-93	0.242459	0.122776	0.147896	0.022457	0.089287
1993-94	0.167169	0.256071	-0.01866	0.015926	0.151901
1994-95	0.348784	0.36064	0.235632	0.189611	0.217316
1995-96	0.238237	0.135008	0.156721	0.135365	0.169624
1996-97	0.084595	0.113592	0.078856	-0.03216	0.055476
1997-98	0.193243	0.134985	0.114593	-0.14469	0.084403
1998-99	0.002845	0.071482	-0.02348	0.054313	0.103964
1999-00	0.080052	0.522122	-0.08565	-0.2444	0.030774
2000-01	0.105597	0.282544	-0.12636	0.054448	0.088282
2001-02	0.135399	0.081238	-0.03766	-0.30771	0.083514
2002-03	0.063336	0.158351	-0.06993	0.364325	0.150692
2003-04	0.07917	0.116122	0.040551	0.279627	0.087791
2004-05	0.072862	0.08469	0.247142	0.236825	0.069084
2005-06	0.194884	0.211766	-0.0646	0.181853	0.205085
2006-07	0.441006	0.048332	0.263413	-0.04497	0.106246
2007-08	0.148235	0.198763	0.155226	-0.00075	0.169006
2008-09	0.127512	0.179731	0.328092	0.115616	0.137347
2009-10	0.182934	0.133665	0.043229	0.076945	0.202052
2010-11	0.117714	0.204253	0.124197	0.142684	0.142253
2011-12	0.207661	0.239683	-0.11735	0.159903	0.174115
2012-13	0.005201	0.04569	-0.02121	0.098918	0.051492
2013-14	0.183689	0.167724	0.152155	0.013897	0.147692
2014-15	0.151989	0.087773	0.164599	0.232025	0.169012
2015-16	0.146474	0.180038	0.1504	0.278526	0.101602
2016-17	0.119754	0.020214	0.041064	0.203746	0.078864
2017-18	0.134543	0.11122	0.036137	0.204212	0.132168
2018-19	-0.06108	-0.01772	0.126293	0.11947	-0.00398

Log Dif Forecasts

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86					
1986-87					
1987-88	0.104486	0.314597	0.021434	0.28358	0.321603
1988-89	0.014479	0.356821	-0.05027	0.134392	-0.03427
1989-90	0.174542	0.51082	0.224938	0.046564	0.167143
1990-91	0.092262	0.104599	0.127089	0.245269	0.104259
1991-92	0.285999	0.073922	0.067886	0.134062	0.084669
1992-93	0.361058	0.163524	0.146152	-0.0188	0.206088
1993-94	0.258303	0.156316	0.2202	-0.06342	0.122773
1994-95	0.189664	0.280864	0.032748	0.198722	0.17563
1995-96	0.319341	0.335188	0.128115	0.21712	0.1873
1996-97	0.237088	0.14787	0.176954	0.088626	0.16661
1997-98	0.06942	0.07536	0.050005	0.066649	0.03288
1998-99	0.192043	0.129319	0.087022	-0.37762	0.070983
1999-00	-0.05488	0.026322	-0.05289	0.486777	0.054444
2000-01	-0.05208	0.348011	-0.00375	-0.42854	-0.12698
2001-02	0.083017	0.257687	-0.29396	-0.22166	0.243138
2002-03	0.154275	0.100483	0.049088	-0.20182	0.107122
2003-04	0.289182	0.37388	0.090392	0.443628	0.061224
2004-05	0.140308	0.152391	0.110078	0.130368	0.127216
2005-06	-0.089	0.106727	0.23418	0.278767	-0.05783
2006-07	0.234647	0.098634	-0.05006	0.193692	0.316336
2007-08	0.493179	0.116094	0.290894	0.26293	0.167875
2008-09	0.202839	0.280808	0.032094	0.079185	0.228274
2009-10	0.038884	0.077411	0.335226	0.141175	0.051065
2010-11	0.27432	0.222879	0.206147	0.251297	0.286202
2011-12	0.074266	0.131827	-0.00402	0.159723	0.055964
2012-13	0.133379	0.250871	-0.27971	0.190951	0.081425
2013-14	0.014592	0.038615	0.031852	0.129075	0.054729
2014-15	0.138524	0.125664	-0.06067	-0.05642	0.106274
2015-16	0.152825	0.079021	0.267632	-0.1836	0.164313
2016-17	0.052062	0.077725	0.212555	0.203058	0.12143
2017-18	0.120344	0.013871	0.056618	0.233373	0.075714
2018-19	0.150688	0.128052	0.035765	0.1736	0.140087

Errors

	Direct Tax	Sales Tax	FED	Customs	Total
RMSE	0.149383	0.160262	0.167768	0.248582	0.118094
MAPE	406.8221	177.4638	231.2742	1558.905	234.5774
MAE	0.128868	0.117634	0.142806	0.174639	0.095618

Appendix 2A.8

Marginal Tax Rate Forecasts

Log Dif Taxes

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86	0.04903	0.052918	0.013432	0.227558	0.123132
1986-87	0.077578	0.26277	-0.00998	0.126503	0.2817
1987-88	0.113464	0.31055	0.124581	0.132029	0.119362
1988-89	0.161843	0.519594	0.141218	0.108666	0.170708
1989-90	0.070849	0.05782	0.139515	0.178984	0.077361
1990-91	0.283524	0.082179	0.071236	-0.00287	0.086545
1991-92	0.374306	0.207059	0.203888	0.201874	0.232384
1992-93	0.242459	0.122776	0.147896	0.022457	0.089287
1993-94	0.167169	0.256071	-0.01866	0.015926	0.151901
1994-95	0.348784	0.36064	0.235632	0.189611	0.217316
1995-96	0.238237	0.135008	0.156721	0.135365	0.169624
1996-97	0.084595	0.113592	0.078856	-0.03216	0.055476
1997-98	0.193243	0.134985	0.114593	-0.14469	0.084403
1998-99	0.002845	0.071482	-0.02348	0.054313	0.103964
1999-00	0.080052	0.522122	-0.08565	-0.2444	0.030774
2000-01	0.105597	0.282544	-0.12636	0.054448	0.088282
2001-02	0.135399	0.081238	-0.03766	-0.30771	0.083514
2002-03	0.063336	0.158351	-0.06993	0.364325	0.150692
2003-04	0.07917	0.116122	0.040551	0.279627	0.087791
2004-05	0.072862	0.08469	0.247142	0.236825	0.069084
2005-06	0.194884	0.211766	-0.0646	0.181853	0.205085
2006-07	0.441006	0.048332	0.263413	-0.04497	0.106246
2007-08	0.148235	0.198763	0.155226	-0.00075	0.169006
2008-09	0.127512	0.179731	0.328092	0.115616	0.137347
2009-10	0.182934	0.133665	0.043229	0.076945	0.202052
2010-11	0.117714	0.204253	0.124197	0.142684	0.142253
2011-12	0.207661	0.239683	-0.11735	0.159903	0.174115
2012-13	0.005201	0.04569	-0.02121	0.098918	0.051492
2013-14	0.183689	0.167724	0.152155	0.013897	0.147692
2014-15	0.151989	0.087773	0.164599	0.232025	0.169012
2015-16	0.146474	0.180038	0.1504	0.278526	0.101602
2016-17	0.119754	0.020214	0.041064	0.203746	0.078864
2017-18	0.134543	0.11122	0.036137	0.204212	0.132168
2018-19	-0.06108	-0.01772	0.126293	0.11947	-0.00398

Log Dif Forecasts

FY	Direct Tax	Sales Tax	FED	Customs	Total
1984-85					
1985-86					
1986-87					
1987-88	0.435002	1.608714	-0.48767	-0.47844	0.998384
1988-89	0.118464	0.092124	2.088246	-2.23492	-1.21637
1989-90	0.792528	1.14576	0.860635	-0.09496	0.669459
1990-91	-0.72202	-1.67475	-0.28432	1.834362	-0.46705
1991-92	1.204276	-0.05782	-0.74354	-4.96169	-0.26765
1992-93	0.677976	1.078602	1.305456	3.662249	1.147235
1993-94	0.135648	0.144346	0.426086	-1.33865	-0.32276
1994-95	-0.35217	0.528182	-2.67254	4.149561	0.277479
1995-96	0.886076	0.51716	2.71533	-2.24543	0.432944
1996-97	0.0043	-0.41738	0.049313	-0.20995	0.135281
1997-98	-0.71821	-0.1434	-0.60925	-1.00218	-1.0258
1998-99	1.188774	0.600854	0.563701	1.975099	0.738648
1999-00	-4.19707	-0.5299	-1.25955	-0.56069	0.348149
2000-01	2.323587	1.16105	1.117207	0.534786	-0.92688
2001-02	1.766537	0.689636	-0.47999	-2.5165	0.899684
2002-03	0.228958	-0.53444	-0.11122	3.40365	0.591892
2003-04	-0.68865	0.539079	0.086434	-0.20926	0.153672
2004-05	-0.05695	-0.58384	-1.15925	-1.53447	-0.84853
2005-06	-0.27121	-0.16789	1.939997	0.235518	-0.33131
2006-07	0.661502	0.495511	-0.97754	0.420824	1.322061
2007-08	1.938325	-0.64707	1.701233	0.403235	-0.227
2008-09	-1.04136	1.450379	-1.00507	-5.66592	0.369541
2009-10	-0.16365	-0.40397	2.65371	6.786617	-0.24975
2010-11	0.971335	0.497704	-2.75943	2.481861	1.016549
2011-12	-0.74469	0.012667	0.440693	-2.99171	-0.74114
2012-13	1.147844	1.141744	0.714425	1.213811	1.045392
2013-14	-3.47557	-1.69254	-1.33448	0.115215	-1.22134
2014-15	3.536325	1.309562	1.270485	-2.77352	1.058921
2015-16	0.286168	-0.24097	2.713075	0.947056	0.574429
2016-17	-0.6992	0.084396	1.040683	0.803561	0.145635
2017-18	0.542805	-1.38664	-4.14933	0.358053	-0.6841
2018-19	0.152892	1.633915	-0.50718	0.319992	0.501288

Errors

	Direct Tax	Sales Tax	FED	Customs	Total
RMSE	1.434219	0.896368	1.60906	2.500503	0.75183
MAPE	2703.884	907.5076	1611.262	5715.665	1018.213
MAE	0.025377	0.019174	0.033039	0.046134	0.016748