TAX REVENUE

PIDE Working Papers
No. 2023:12

Pakistan Institute of Development Economics

Federal Tax Revenue Forecasting of Pakistan: Alternative Approaches

Syeda Um Ul Baneen

Pakistan Institute of Development Economics
Islamabad, Pakistan

E-mail: publications@pide.org.pk
Website: http://www.pide.org.pk
Fax: +92-51-9248065

Designed, composed, and finished at the Publications Division, PIDE.

PIDE Working Papers
No. 2023:12

Federal Tax Revenue Forecasting of Pakistan: Alternative Approaches

Syeda Um Ul Baneen
Pakistan Institute of Development Economics, Islamabad.

PAKISTAN INSTITUTE OF DEVELOPMENT ECONOMICS
ISLAMABAD
2023

CONTENTS

List of Tables

ABSTRACT

This study is an appraisal of FBR's existing tax revenue forecasting technique, i.e., the buoyancy approach. It compares the effectiveness of the buoyancy approach and alternative forecasting methods. Using 1984-85 to 2018-19 annual data, predictions are made using different theoretical, statistical, and machine learning methods. The root mean square error (RMSE) suggests that the Least Absolute Shrinkage and Selection Operator (LASSO) and Elastic Net approaches provide better forecasts for total tax and federal excise duty. The Box Jenkins methodology provides the most accurate direct and sales tax estimates, while custom duties are best predicted through the Vector Autoregressive model. On the other hand, mean absolute error (MAE) recommends Marginal Tax Rate(MTR) to forecast all federal taxes of Pakistan. In comparison, the buoyancy approach predictions are not accurate for any of the federal taxes.

JEL Classifications: B26 B41 C13 H20

Keywords: Tax Revenue Forecasting, Buoyancy Approach, LASSO, Marginal Tax Rate

1. INTRODUCTION

Every year, there are forecasting errors in Pakistan's federal tax revenues. From 1970 to 2020, the federal government has significant forecast errors for all types of tax revenues. Using 1970-2014 data, two-thirds of the time, total federal taxes were overestimated by more than five percent (Qasim & Khalid, 2016). Errors in tax revenue forecasts directly affect budget formulation. The consequences of such errors are misleading tax revenue targets leading to wrong allocation of resources resulting in arbitrary cuts at the end of fiscal year such as a cut in development expenditures and/or an increase in the unanticipated debt burden. This, from policy perspectives, have long-run monetary and fiscal effects.

In Pakistan, the Federal Board of Revenue (FBR) is responsible for revenue collection at the federal government level. FBR produces annual forecasts of its tax streams before the budget formulation so that the expenditures can be planned at the federal government and through resource sharing (National Finance Commission-NFC) the provincial governments. Once the expenditures are set the deficits are estimated and financing options such as Bank borrowing are set. FBR uses the Buoyancy approach to forecast tax revenues(FBR, 2021). In this approach, we estimate the parameter by calculating the proportionate change in tax revenue with respect to the proportional change in the tax base. The estimates are then multiplied by nominal GDP growth targets for predicting the future estimates of tax revenues.

As reported earlier, forecasting errors were mainly on the higher side, resulting in fiscal mismanagement. A revenue prediction error results in over/underbooking of government expenditures, which, if called back, can result in wastage or under-provisioning. If not adjusted, it would result in extra financing, thus putting pressure on the monetary side as well. Further, given Pakistan's federating nature, any revenue prediction error would also carry over to the provinces through the revenue-sharing channel. A possible reason for these repeated errors could be using an inappropriate forecasting method over the years. Several alternative forecasting methods are available for better prediction of revenue streams. This study estimates some alternatives to buoyancy approach for federal tax revenue forecasts. These methods have been compared to suggest the most accurate predicting method for each federal tax head.

The paper provides literature review on the alternative methods efficacy, followed by methodology and data description. Then results are provided based on alternative methods estimations, and the conclusion is drawn at the end of the paper.

2. LITERATURE REVIEW

There are a large number of methods implemented for tax revenue forecasting across the globe. Each economy forecast taxes considering its economic, social, and organisational indicators. Moreover, academics and researchers have tried many methods which provide better estimates than classical methods. They are divided as theoretical/orthodox methods, statistical methods and machine learning methods.

2.1. Theoretical/Orthodox Methods

Theoretical or orthodox methods of tax revenue forecasts refer to the methods built specifically for tax forecasting. These are practiced by most economies and also recommended by international economic and financial organisations. Jenkins et al. (2000) provided a brief guide on the importance and practice of tax revenue forecasting. They covered the buoyancy approach, the elasticity method, GDP-based methods and microsimulation techniques for predicting tax revenues. IMF suggests 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). Bayer (2013) described the basic methodology for predicting tax revenues in terms of the macroeconomic approach and from a microeconomic perspective. It also described methods for assessing the quality of predictions.

The theoretical methods are also adopted by some researchers in Pakistan too. Qasim & Khalid (2016) did an accuracy analysis of tax receipt forecasts of Pakistan. The paper identified that there was a significant difference between the budget estimates of Federal Revenue receipts and actual revenue receipts. They also found that revenue forecasts were as bad after 1990 reforms as before them. Firdawss & Karim (2018) applied the effective tax rate approach (ETR), marginal tax rate approach (MTR), regression approach and elasticity approach to forecasting the federal tax revenue of Morocco for the years 2017 and 2018. Gumbo & Dhliwayo(2018) used exponential smoothing, ETR, and elasticity approaches to predict Zimbabwe's VAT revenue for 2012 and 2013. The study found that Exponential Smoothing had relatively lower forecast errors. Some studies in the past also used IMF-recommended regression models. Reed (1983) used two regression models to forecast Indiana's tax revenues. The first model was estimated using two-stage least squares.The second model was an adaptation of the first model.

2.2. Statistical Methods

This section summarises a literature survey of Statistical tax forecasting techniques. They are categorised as univariate and multivariate methodologies. Univariate approaches involve only one variable, while multivariate approaches involve a number of variables.

Grizzle & Klay (1994) used the moving average method, the average change method, linear regression and the curve fitting method to forecast the tax revenue of 28 US states. The fitting method produced the lowest MAPE. 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 et al. (2015) established an ARIMA(0,1,0) model to forecast the Income Tax Revenue of the Philippines for the year 2014-2020. Forecasts revealed no significant difference between the actual values of Income Tax Revenue inspected through Paired T-Test. Makananisa (2015) used different time series models to forecast central taxes in the South African Revenue Service. The findings revealed that the SARIMA model accurately predicted personal income tax and value-added tax.

Streimikiene et al. (2018) forecasted the total tax revenue of Pakistan for the FY2017 using ARIMA and Vector Autoregressive (VAR) 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. Ofori et al. (2020) forecasted Ghana's value added tax revenue through two methods. The first method was ARIMA with Intervention, and the second method was Holt linear trend method. 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. The results were more accurate than official forecasts. Reed (1983) used the Box-Jenkins methodology to forecast the state tax revenue of Indiana. Nandi et al. (2014) used the Box Jenkins methodology to forecast the total tax revenue of Bangladesh.

Cirincione et al. (1999) scrutinised the impact of the time series estimation method selection, the data length, and the data frequency on forecasting accuracy. The study concluded that exponential smoothing models were the most accurate. Pavlík (2011) focused on forecasting income tax when the time series was short. EGARCH proved to be the best prediction method. Krol (2010) used the Bayesian Vector Autoregressive (BVAR) model to predict the tax revenue of California. The BVAR model outperformed VAR and random walk models. Yu et al. (2015) used an error correction mechanism to improve accuracy using a single forecasting model. 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. Greoning et al. (2019) used co-integration analysis to recognise factors that affect company tax revenue in Swaziland. Combined forecasting was proven to cause a 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 Autoregressive Moving Averages (ARIMA), the State Space exponential smoothing (ETS) model and Bayesian Vector Autoregressive (BVAR) model. Based on RMSE, the results proved the superiority of the ETS model.

2.3. Machine Learning Methods

This section reviews machine learning methods used for forecasting, including tax revenues and other variables. Ticano, et al. (2017) predicted income tax revenues of Brazil. This work introduced a hybrid model based on Genetic Algorithms (GAs) and Neural Networks (NNs) for a multi-step forecast of tax revenue collection. The results were more accurate than the outcome that the Federal Revenue of Brazil's Secretariat estimated with the indicators method. 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. Jang (2019) presented an auxiliary tax prediction system that was based on an artificial neural networks. The system can help experts to predict tax revenues efficiently.

Waciko & Ismail (2020) used shrinkage methods to forecast and compare the GDP of India and Indonesia. Under the MSE criteria, Elasic Net performed better than LASSO and Ridge Regression. Nakroji & Aminu (2022) used different machine learning (ML) techniques for inflation forecasting of Nigeria. They found ridge regression and artificial neural networks to be the best methods. Chung et al. (2022) did revenue forecasting by local governments. Their findings revealed that traditional statistical forecasting methods outperformed ML algorithms overall. But one of ML algorithms, k-nearest neighbors (KNN) algorithm, was more effective in predicting property tax revenue. Noor, et al. (2022) tested the forecast ability of feed-forward neural network (FFNN), random forest, and linear regression for tax revenue forecasts of Malaysia. After conducting several experiments, it was found that FFNN achieved the highest accuracy.

The literature survey showed that many tax revenue studies addressing the method issues are available. Most of this research is for advanced economies. An inclusive study of Pakistan's federal tax revenue forecasting errors was needed. There should be research that appraises the methodological issues in the existing federal tax revenue forecasting mechanism. The current system needs better alternatives too. This study will fill this gap.

3. DATA AND METHODOLOGY

To analyze the different forecasting methods of federal tax revenues, the methodology is divided into three parts theoretical or orthodox methods, statistical methods, and machine earning methods. This exercise will be repeated for a different type of federal taxes. Then the accuracy of these methods is judged using root mean square error and mean absolute error. A detailed description of these methods is provided below:

3.1. Theoretical or Orthodox Methods

Theoretical tax revenue forecasting methods involve some fiscal or revenue theory or logic in their estimation. The current study uses the effective tax rate, marginal tax rate, and Buoyancy approaches as theoretical forecasting methods. Following are brief reviews of these methods.

3.1.1. *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 utilise 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 suggests a tax base replacement to examine 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:

$$Effective\ tax\ rate = \frac{Tax\ revenue}{Proxy\ tax\ Base} \quad \dots \quad \dots \quad \dots \quad \dots \quad (1)$$

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

$$Tax(f) = Taxbase(f) * (ETR) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (2)$$

Once we determine that the ETR is constant, we can forecast tax revenue by multiplying the forecasted tax base by 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).

3.1.2. *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:

$$Marginal\ tax\ rate = \frac{\Delta Tax\ revenue}{\Delta Proxy\ tax\ base} \quad \dots \quad \dots \quad \dots \quad \dots \quad (3)$$

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

$$Tax(f) = MTR * Taxbase(f) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (4)$$

This study estimates marginal tax rate and forecasts for all taxes using Equations 3 and 4.

3.1.3. 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:

$$Buoyancy = \frac{\Delta T / T}{\Delta GDP / GDP} \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (5)$$

And tax revenue forecast through Buoyancy:

$$Tax_f = GDP\ grwoth * Tax\ buoyancy \quad \dots \quad \dots \quad \dots \quad \dots \quad (6)$$

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). In this study, the method calculates the buoyancies of all tax heads. The buoyancies are then be multiplied by nominal GDP growth targets to get the forecasts.

3.2. 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. In this study, a univariate regression technique, i.e., Box-Jenkins methodology, and a multivariate approach, i.e., the Vector-Autoregressive method, are used. In this connection, the literature identifies that multilateral institutions such as the IMF sometimes also use 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.

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

3.2.2. *Vector Autoregressive Method (VAR)*

Christopher Sims proposed Vector Autoregressive 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 all variables involved in the system. The study follows Yu et al. (2015) and Greoning et al. (2019) to adopt VAR or VECM for different tax heads.

3.3. Machine Learning Methods

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

3.3.1. *LASSO*

LASSO is the acronym for Least Absolute Shrinkage and Selection Operator. It is a statistical formula to regularise data models and feature choices. It is utilised in regression methods for more accurate predictions. It uses the shrinkage technique. Shrinkage is the point where values are shrunk to the mean. This regression is appropriate for models with high multicollinearity levels or when you want to mechanise some parts of model selection, like variable choice/parameter elimination (Chan-Lau, 2017). The mathematical equation of this regression is given as

$$L = \sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j) + \lambda \sum_{j=1}^{p} |\beta_j| \quad \dots \quad \dots \quad \dots \quad \dots \quad (7)$$

Here, λ shows 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 minimising the variance.

3.3.2. *Ridge Regression*

Ridge regression is also one of the machine learning shrinkage methods. Hoerl & 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 minimises it close to zero. It is given as

$$R = \sum_{i=1}^{n}(y_i - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p}\beta_j^2 \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad (8)$$

The comparison of Equations 7 and 8 shows that LASSO regression takes the absolute magnitude of the coefficients, while ridge regression takes the square. For comparison purposes, we have used ridge regression using all potential regressors to get the tax revenue forecasts.

3.3.3. *Elastic Net*

The Elastic Net Regression method is a convex combination of the Ridge Regression Penalty and the LASSO Penalty. Zou and Hastie developed it in 2005. As the name suggests, it is flexible in adopting the λ value. Elastic Net overcomes the weaknesses of LASSO and ridge regressions as it involves both square and absolute magnitudes of the coefficients. This study follows Waciko & Ismail (2020) in using Elastic Net to forecast tax revenues.

3.4. Evaluation Criteria

These methods' performances are judged based on root mean square error (RMSE) and mean absolute error (MAE).

Table 1

Errors Description and Formulas

Description	Formula
Root Mean Square Error The root mean square error (RMSE) has been used as a standard statistical metric to measure model performance in different social sciences, including Economics.	$RMSE = \sqrt{\frac{\sum(Y_t^p - Y_t^A)^2}{n}}$ Here $Y_t^A = Actual\ Data$ and $Y_t^p = Predicted\ Data$
Mean Absolute Error(MAE) The mean absolute error (MAE) is another useful measure widely used in model evaluations.	$MAE = \frac{\sum \lvert Y_t^p - Y_t^A \rvert}{n}$

For this study, the following variables have been used to assess the accuracy of these tax revenue forecasting methods for Pakistan.

Table 2

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	Yearly federal excise duty revenue collection by FBR (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-Agricultural GDP (cfc)	Non-agricultural GDP measured at current factor cost (million Rs.)
Large Scale Manufacturing	Value addition of Large Scale manufacturing in GDP(million Rs.)
Dutiable Imports	Imports on which duty is leviable (million Rs.)
Total Consumption	Total Consumption expenditure GDP (million Rs.)
Nominal GDP growth Target	Nominal GDP growth estimate in percentage
Consumer Price Index	Consumer Price index taking base price of 2010
Total Population	Total population of Pakistan
Exchange Rate	PKR to the US. Dollar Exchange rate
Labour	Total Employed Labour

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. It is extracted from the official website of FBR, different issues of the Economic Survey of Pakistan, budget documents, annual plans of the Ministry of Planning, Development and Special Initiatives, and the World Bank WDI data set.

4. RESULTS AND DISCUSSION

As discussed in the methodology section, all types of methodologies are applied to different types of federal taxes. i.e., total tax, Direct tax, sales tax, custom duties, and federal excise duty. The root mean square error and mean absolute error for all eight forecasting methods are discussed in this section. In this section, ways are mentioned in the form of abbreviations.

4.1. Total Tax

In Table 3, 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. The results deny Streimikiene, Ahmed, V Veinhardt, & Pervaiz (2018). They found out that Box Jenkins provided better predictions for FY17 for Pakistan.

Table 3

Statistical Errors for Total Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.108	0.093
Box Jenkins	0.05	0.036
Vector Autoregressive Model	0.063	0.051
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

4.2. Direct Tax

According to Table 4, 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.

Table 4

Statistical Errors for Direct Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.15	0.09
Box Jenkins	0.089	0.067
Vector Autoregressive Model	0.102	0.074
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. Sales Tax

Table 5 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.

Table 5

Statistical Errors for Sales Tax Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.196	0.159
Box Jenkins	0.088	0.073
Vector Autoregressive Model	0.122	0.092
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.

4.4. Custom Duties

According to table 6, the custom duties of Pakistan are best forecasted through Vector Autoregressive model with root mean square error (RMSE) of 0.114. 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.

Table 6

Statistical Errors for Custom Duties Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.204	0.142
Box Jenkins	0.157	0.124
Vector Autoregressive Model	0.114	0.091
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	2.500	0.046

Source: Authors' calculations

4.5. Federal Excise Duty (FED)

According to table 7, 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. According to these errors, LASSO, Elastic Net, and MTR are more appropriate to predict federal excise duty (FED) than any other method.

Table 7

Statistical Errors for FED Forecasting Methods

Estimation Methodology	RMSE	MAE
Buoyancy Approach	0.121	0.090
Box Jenkins	0.115	0.101
Vector Autoregressive Model	0.109	0.094
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.

5. CONCLUSION

Revenue forecasting is important because the government needs an accurate value for fiscal policy. Any forecast error would lead to mismanagement in fiscal policy operations. It will lead to government expense cuts if revenue forecasts are biased upwards.

If the expenditures are not cut, it will have an unanticipated increase in debt raising for deficit financing. For countries where there are federating setups and revenues are collected at the center and then distributed at the government's lower tiers, the revenue forecast errors further exacerbate the mismanagement in fiscal policy operations as it now runs in all governments.

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 and FED. Box Jenkins predicts accurately while forecasting sales tax and direct tax. Customs duties are better predicted through VAR. 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.

6. LIMITATIONS AND FUTURE DIRECTIONS

In this study, the statistical losses of forecasting methodologies are calculated in terms of MAE and RMSE. Future studies can also calculate the statistical accuracy in terms of biasedness and comprehension of the method. Other statistical accuracies like asymmetrical statistical loss function can also be calculated. The study has not focused on the political and administrative side of tax revenue forecasting. Future studies can incorporate these aspects.

REFERENCES

Bayer, O. (2013). Research of Estimates of Tax Revenue: An Overview. *European Financial and Accounting Journal*, *8*(3-4), 59-73.

Barnard, J. R., & Dent, W. T. (1979). *State Tax Revenue-New Methods of Forecasting*.

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., & Van de Sande, B. (1999). Municipal government revenue forecasting: issues of method and data. *Public Budgeting & Finance*, *19*(1), 26-46.

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

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.

Greoning, E., Zivanomoyo, J., & Tsaurai, K. (2019). Determinants of Company Tax Revenue in Swaziland (1990-2015). *Acta Universitatis Danubius: Oeconomica*, *15*(5), 7-37.

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.

Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. *Technometrics*, *12*(1), 69-82.

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

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

Molapo, M. A. (2018). Employing Bayesian Vector Auto-Regression (BVAR) Method as an Alternative Technique for Forecasting Tax Revenue in 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.

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.

Ofori, M. S., Fumey, A., & Nketiah-Amponsah, E. (2020). Forecasting Value Added Tax Revenue in Ghana. *Journal of Economics and Financial Analysis*, 63-99.

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

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

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

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

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)* (pp. 1-4). IEEE.

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

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

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*

Pakistan Institute of Development Economics
Post Box No. 1091, Islamabad, Pakistan

www.pide.org.pk