Forecasting Tax Revenue and its Volatility in Tanzania

Chimilila, Cyril[14]

Abstract
Forecasting tax revenue and its predictability is important for government budgeting and tax administration purposes. This study used monthly tax revenue data for a period of 182 months spanning January 2000 to February 2015. The study applied ARMA and combined forecast models, and GARCH models to forecast tax revenue and its volatility, respectively. Tax revenue was found to increase steady over the period, although with a persistent volatility which increases over time. The observed volatility was found to be associated with taxes from bases (income) which have high volatility. Based on various forecast accuracy evaluation criteria, the study recommends combined forecasts and GARCH(1,1) models for forecasting monthly revenue and its volatility, respectively. The study further recommends enhanced diversity of taxes through widening consumption tax base within the existing tax portfolio so as to enhance its contribution to revenue collection and reduce volatility.

Keywords: Tax Revenue Forecasting, Combined Forecasts, ARMA and GARCH

[14] The Author is a Lecturer at the Institute of Tax Administration. Correspondence address: P.O. Box 9321, Dar es Salaam; Email cyrilignas@yahoo.co.uk or cchimilila@tra.go.tz.

1.0 Introduction

Governments need financial resources to provide public goods and services to the citizens. In order to determine the level of provision of these goods governments need to ascertain the availability of financial resources, which includes tax and non-tax revenues. Thus within governments budgetary procedures frameworks, fiscal forecasting and monitoring techniques have emerged crucially important. Fiscal forecasts include revenue and expenditure forecast. This study focuses on the former.

In Tanzania, the Public Finance Act (2001) stipulates the government budget process which encompasses revenue forecasting. The task of revenue forecasting is done at the Ministry of Finance and the Tanzania Revenue Authority. However, revenue target estimates have sometime been imposed to TRA by the government when it perceives that the tax authority provides a lenient achievable target which may be surpassed when collection effort is increased; this requires consensus – consensus forecasting. Nonetheless, statistical techniques have been a mainstay of revenue forecasting.

Trends in government revenue collection in Tanzania indicate persistent growth in tax revenue. This steady growth may be explained by, among others, widening of tax base, changing tax rates, increasing collection efforts, and promoting voluntary compliance through simplified tax payment procedures (Mwakasindile, 2011; TRA, 2013). Despite the growth, actual revenue collections have seldom meet targets. For instance in the financial year 2013/2014 monthly revenue collection between July 2013 and June 2014, except for April and June 2014 all other months recorded collections below targets (TRA, 2014). Failure to meet these targets can be a result of higher targets driven by 'optimism' or an ambition of reaching some per cent of GDP target. Other reasons can be inefficient administration in tax collection or forecast accuracy when projected revenue collections are beyond potential realized collection. Failure to meet revenue collection targets creates stress to the tax administration as well as the government as it fails to meet some expenditure commitments. However, if proper forecasting techniques are employed, revenue collection and its dynamics can be well predicted and accounted for with reasonable accuracy.

In forecasting tax revenue apart from the judgemental methods, quantitative methods employing static and dynamic models are also used. Static models include GDP based models. The GDP based models, for example, uses tax elasticity and buoyancy they assumes linearity - tax revenue responds at a certain percent with changes in GDP. Despite several merits of dynamic models, they sometime fail to capture dynamics in tax revenue and may produce less accurate forecasts. In this paper dynamic time series models have been chosen because they are capable of using past information to predict dynamics in revenue collection; in particular the issue of volatility. This study therefore undertakes to forecast revenue using dynamic models (ARMA and combined forecasts) and volatility using GARCH models. Further, ARMA and GARCH models are ideal for high frequency data like those recorded monthly over a long period of time.

The objectives of this study are threefold. First, to model and forecast monthly tax revenue collection; second, to model monthly tax revenue volatility; and thirdly, to propose policy recommendations for stabilization of revenue collection volatility. This study used monthly tax revenue data spanning January 2000 to February 2015. The study used ARMA and combined forecast models, and GARCH models to forecast tax revenue and its volatility, respectively. Tax revenue was found to increase steady over the period, although with a persistent volatility which increases over time.

The remainder of this study is organised as follows. Section 2 reviews the literature on revenue forecasting. While section 3 sketches out the methodology, section 4 presents and discusses the results. Section 5 gives conclusion and recommendations.

2.0 Review of Literature on Revenue Forecasting

Various tax revenue forecasting techniques are available (see Jenkins et al., 2000). They can be categorized into two, quantitative and qualitative methods. Qualitative methods are also referred to as judgmental forecasts; they are based on human judgement. These methods are prone of bias and conservatism. However, they are common in situations where data are scarce or non-existent, or when historical data are no longer representative such as in the event of structural shifts. A special case of judgment forecasting is consensus forecasting[15]. This method is determined by institutional setup which comprises parties involved in setting revenue targets. Studies on forecasting accuracy and institutional arrangements (e.g. Voorhees, 2004) argue that, on one side consensus forecasting diminishes forecast bias and increases forecast accuracy as it takes the politics out of the revenue forecast accuracy. On the other side, they argue that time-consuming nature of consensus forecasting can increase lags between forecast preparation and forecast use, in turn potentially reducing accuracy. Thus errors in revenue forecasts are not solely caused by the fluctuations in the economy, but also are attributable to the institutional structures and the degree of consensus required for the forecast. Borrowing from Boyd and Dadayan (2014), “it is good practice to try to insulate forecasting from the political process and consensus forecasting can help achieve that.”

Quantitative methods include causal and extrapolation techniques. Causal methods involve simple and multiple regression of a dependent variable (tax revenue) and some other independent variables (e.g. income, imports, consumption). Causal models however are static and hence have low ability of capturing dynamics in data collected over a long period of time. Extrapolation techniques are commonly referred to as time series techniques. Extrapolation techniques make use of past data to predict the future. These methods are very accurate compared to the former making them more popular. The most widely used extrapolative techniques are: naïve models which assume the current situation is the same as previous; moving averages and exponential smoothing which use averages of the most recent data to calculate forecasts; trend line analysis which regresses a variable on some function of time – linear, quadratic, logarithmic, etc.; autoregressive models which regress a variable on its past values; and Box-Jenkins models. Box-Jenkins models are considered quite accurate approach to forecasting. By combining autoregressive and moving average processes (ARMA), Box-Jenkins provide more objective forecasts because they are able of revealing regularities in the data that would be overlooked by other methods.

Major challenges to revenue collection in developing countries like Tanzania are prevalence of discretionary changes in tax systems, coupled with inadequacy of data and limited forecasting skills. For example, Fjeldstad et al. (2014) reported the influence of political and

[15] Consensus forecasting is revenue projection developed in agreement through an official forecasting group representing both the executive and legislative branches (NASBO, 2008).

economic factors and underdeveloped administrative capacities as the major challenges for sound fiscal policy in Angola where they found no evidence of application of any sophisticated forecasting methods or adherence to economic growth projections by the revenue forecasters.

The above-mentioned challenges of date adequacy and skills form major impediments for using sophisticated time series models in revenue forecasting in developing countries; literature shows that time series models, like ARMA, are extensively used in developed countries than in developing countries due to the said challenges. Since this study focuses on monthly revenue forecasts it benefits from the advantage of a large sample size.

The accuracy of forecasts, as measured by the size of forecast error, is very important. Accurate revenue forecasts are a key element for the design and execution of sound fiscal policies. Large forecast errors can lead to substantial budget management problems. Auerbach (1995) distinguishes between three types of errors: policy errors, economic errors and technical (behavioural) errors. While forecast errors can never be entirely avoided, a model needs to perform better both in-sample and out-of-sample. However, most models perform better for in-sample forecasts than out-of-sample forecasts. It is therefore customary for forecasters to estimate several models and compare them in terms of forecast performance. Several measures are used to check forecast accuracy by measuring the size and distribution of the errors. The common used measures are Mean Absolute Error - MAE, Mean Absolute Percentage Error - MAPE, Mean Square Error - MSE, and Root Mean Square Error – RMSE (Johnston and DiNardo, 1997; Leal et al., 2007).

The question as to which forecasting model performs better is still unresolved. As Leal et al. (2007) noted, from a study of the available literature it is not clear which method fiscal and monetary authorities, international economic organisations, financial market analysts, rating agencies or research institutes should be adopting when preparing their forecasts. Because of differing situations between economies to which the models have to be applied, there is no single model that outperforms others universally.

Cited in Leal et al. (2007), Bretschneider et al. (1989) compare the forecasting accuracy of different forecasting methods. On the basis of their results, they favour a combination of judgement and simple econometric equations, against time series and complex econometric models. Several other studies favour either time series or simple regression models. But there is no clear cut. Litterman and Supel (1983) provide some evidence to support the combining of different forecasting techniques. Although recommending a single forecast method features prominent in forecast literature, it is sometime possible to arrive at a more accurate forecast by using combination of several forecasts. Combined forecasts were proved to yield more accurate results in many cases. Variants of combination forecasts include simple average forecasts, weighted forecast and linear combination forecasts. The simple average is the most-widely used combining method (see Bunn, 1985). However, this method is challenged for not being able to incorporate dynamics in the forecasts. This study adopts a linear combination forecasts. This method was used because out-of-sample forecasts for ARMA were observed to be very optimistic compared to trends of in-sample forecasts. Further, liner combination produced better forecasts than ARMA or moving averages used singly.

Dynamic models are data intensive and also need to be updated from time to time; for dynamic models, the more the data better forecasts. Updating models is important because unforeseen event can change the whole calculation of forecasted value and thus forecasting of any series is a continuous process rather than a one-time calculation (Nandi et al., 2014). Literature however, has not established the frequency and timing of updating and thus it remained largely a matter of discretion. It is worth noting that there is no statistical evidence that frequent forecast updates lead to greater accuracy, and this is consistent with past research (e.g. Boyd and Dadayan, 2014), and as such this paper doesn't specify such kind of recommendation for models it estimates.

3.0 Methodology of the Study

This study mainly employs a time series approach using Box-Jenkins models to model and forecast monthly tax revenue. GARCH models are used to forecast volatility. The study uses monthly and annual total tax revenue collection data published by the Bank of Tanzania[16]. The monthly revenue data covers the period of 182 months from January 2000 to February 2015. This sample size is adequate for estimation of ARMA models; according to Garrett and Leatherman (2000) the generally accepted threshold for ARMA estimation is 50 data points.[17]

Data analysis undertaken involves descriptive analysis, stationarity tests, model fitting and forecasting. Descriptive analyses are used to explore internal properties of data. Analytical models used include ARMA, combined forecasts and GARCH models. This study followed standard procedure for estimation of ARMA models which has three steps: identification, estimation, and forecasting (Johnston and DiNardo, 1997). Identification involves checking for stationarity and determination of the order of the model. The best models are selected by using Akaike Information Criteria (AIC) and Bayes-Schwarz Information Criteria (BIC), forecast performance measures and other statistical criteria. AIC is an asymptotically model selection criterion. AIC provides a trade-off between goodness of fit and the complexity of model specification (Akaike, 1974). Root Mean Square Error (RMSE) was used to compare forecast performance of the models. In building the model it was assumed that no substantial discretionary changes occur in the out-of-sample forecast period.

3.1 Stationarity test

Before the model is estimated, stationarity tests were performed using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. If the time series is not stationary the t-distribution will have non-standard distributions and thus test results may be misleading. A serious problem is the possibility of finding spurious regressions, i.e. having significant regression results while the variables have no long-run relationship (Johnston and DiNardo, 1997). The ADF and Phillips-Perron tests were performed with level and differenced data. Both ADF and PP tests were used; Phillips-Perron (PP) test relaxes the assumption of homoskedasticity and thus tend to be more powerful than the Augmented Dickey-Fuller test in the presence of heteroskedasticity (Hamilton, 2006).

The ADF is specified as,

---

[16] Economic Statistics: Government Budgetary Operations (both monthly and fiscal year) available through http://www.bot.go.tz/Publications/PublicationsAndStatistics.asp

[17] For details on determination of minimum sample size for time series analysis see Dharan, B.G. (1985).

$$\Delta y_t = \alpha_0 + \gamma y_{t-1} + \sum_{i=2}^{\beta} \beta_i \Delta y_{t-i} + \varepsilon_t \qquad \varepsilon \sim IID\left(0, \sigma^2\right) \tag{1}$$

H_0: $\gamma = 0$ (nonstationary, i.e. unit root)
H_1: $\gamma < 0$ (stationary, i.e. no unit root)
where,
y is monthly tax revenue
ε is a white noise process
γ is the stationarity coefficient
α_0 and β_i are parameters to be estimated

3.2 Analytical Models

3.2.1 ARMA

For the purpose of forecasting tax revenue ARMA models were specified and estimated. ARMA is a mixed process of a time series that has an Autoregressive part which comprise lagged variables of the endogenous variable and a Moving Average part that contain current and past error terms. A general ARMA(p,q) is specified as,

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + ... + \alpha_p y_{t-p} + \varepsilon_t + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + ... + \beta_q y_{t-q} \tag{2}$$

In order to determine the order of the ARMA model, Autocorrelation (AC) and Partial Autocorrelation (PAC) graphs were plot and significant lags were determined. Cut-off points of the AC indicate order of Moving Average (MA) part and cut-off points of PAC indicate order of Autoregressive (AR) part of the model.

3.2.2 GARCH

In economic data time-varying volatility is more common than constant volatility, and accurate modelling of time-varying volatility is of great importance. The ARMA models are used to model the conditional expectation of a process given the past, but in an ARMA model the conditional variance given the past is constant. Since the conditional variance in monthly tax revenue is not constant due to observed clusters of high and low volatility, GARCH models were used to model separately volatility of monthly revenue collection. GARCH models treat heteroskedasticity as a non-linear variance to be modelled. The formulation of GARCH was provided by Tim Bollerslev (1986) as a generalization of the ARCH process developed by Robert Engel (1982).[18]

The GARCH models are expressed as,

GARCH (p,q): $$\sigma^2 = \omega + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \tag{3}$$

where
σ^2 is conditional variance (i.e. variance conditional to past observed variances)
ε_t^2 is the squared residual (i.e. innovation)
ω, α_i, β_i are parameters to be estimated
$\omega > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$ and $\Sigma(\alpha_i + \beta_i) < 0$ (for convergence)

---

[18] ARCH stands for Autoregressive Conditional Heteroskedasticity and GARCH is Generalized Autoregressive Conditional Heteroskedasticity.

The GARCH(p,q) is a mix of both ARCH(p) and GARCH(q) processes. Estimation of GARCH commences with testing for presence of ARCH effect. As iterated in Johnston and DiNardo (1997) the test of presence of ARCH has three important steps: regress y_t on regressors and obtain the residuals (ε_t); estimate an OLS specified as ε_t^2 on constant and ε_{t-i}^2 terms i.e. $\varepsilon_t^2 = \hat{\alpha}_0 + \hat{\alpha}_1\varepsilon_{t-1}^2 + \hat{\alpha}_2\varepsilon_{t-2}^2 + \hat{\alpha}_3\varepsilon_{t-3}^2 + ... + \hat{\alpha}_p\varepsilon_{t-p}^2 + u_t$; and test the joint significance of $\hat{\alpha}_1 ....\hat{\alpha}_p$. Significance of the estimated coefficients implies persistence of volatility. The test for ARCH effect was performed using Engle's langrage multiplier test.

4.0 Results and Discussion of the Results

4.1 Descriptive analysis

4.1.1 Summary Statistics

Summary statistics of tax collection are presented in Table 1 below. Results indicate that the average monthly tax collection is increasing persistently over the sample period. In the year 2000 average tax revenue collection was Tanzania Shillings (TZS) 63,745.3 million per monthly while in 2014 it was TZS 812,634 million. Variance in monthly revenue collection was observed to increase year after year from 8,115 in 2000 to 114,461 in 2014. Monthly variability within each year fairly compares across years, except for the year 2001 where coefficient of variation was 8.3 (lowest). The overall average figures however have indicated a high variability with a coefficient of variation of 77.7% which is higher than 50% (for moderate variability). The coefficients of skewness, except for the year 2000 indicate that tax revenue collections are slightly positive skewed. Also Kurtosis coefficients which are less than 3.0 indicate that the tax collection distributions are platykurtic (flat peaked). Generally, the tax collection distributions have more or less the same patterns across years.

Table 1: Summary Statistics of Tax Revenue Collection from 2000 to 2015

Year	Obs.	Mean*	Std. Deviation	CV	Skewness	Kurtosis
2000	12	63,745.6	8,115.1	12.73045	-0.091963	1.951208
2001	12	72,457.1	6,021.5	8.31043	0.168992	1.777129
2002	12	85,313.6	9,716.8	11.38951	0.904976	2.892490
2003	12	101,165.3	14,327.8	14.16276	0.759446	2.706878
2004	12	123,649.3	16,099.3	13.02013	0.807492	2.814310
2005	12	145,183.6	21,058.3	14.50460	0.670063	2.059107
2006	12	183,306.7	34,137.6	18.62321	0.586587	2.370463
2007	12	248,322.0	44,732.5	18.01391	0.312372	2.066897
2008	12	313,960.2	55,364.3	17.63418	0.844934	2.615228
2009	12	348,356.5	53,355.2	15.31626	0.754329	2.638597
2010	12	399,054.6	71,602.3	17.94298	1.142051	3.760105
2011	12	489,334.3	90,132.4	18.41939	0.644466	2.204779
2012	12	599,247.0	119,280.9	19.90513	0.878178	2.795777
2013	12	696,340.3	120,485.4	17.30266	0.699803	2.165187
2014	12	821,632.3	114,461.1	13.93094	0.147744	1.638918
2015	2	721,697.5	37,858.5	5.24576	-	-
Overall	182	317,232.0	246,513.0	77.70748	0.911068	2.813154

* Measured in TZS million
Source: Own computations

4.1.2 Characteristics of Monthly Revenue Collection

Time series data have four components; trend, seasonality, cyclical variation and randomness. Results in Figure 1 show that monthly tax revenue collections exhibited a steady upward trend over the sample period. The increasing trend is explained by expanding tax base as income increase over time, increased collection efforts and changes in the tax system. The period from 2000 to 2005 (second term of third political reign) was marked by enhanced efforts and trade liberalization which boost revenue collection from foundations laid in the first term (1995 to 2000) in which major reorganization in the tax administration occurred including formation of an autonomous tax authority in 1996 as well as more economic liberalization. The period from 2005 to 2015 was during the fourth political reign which was characterized by upscaling efforts and further economic liberalization. Thus as a result of these efforts both income and tax revenue grew steady; GDP growth was sustained at an average of 7.0% per annum. The growth of monthly tax collection is however exponential implying that collection increases at an increasing rate. Seasonality in tax collection was observed with alternating peaks and troughs at consecutive time intervals. Overall monthly collection data is not stationary as indicated by an upward trend and also volatility which increases over time.

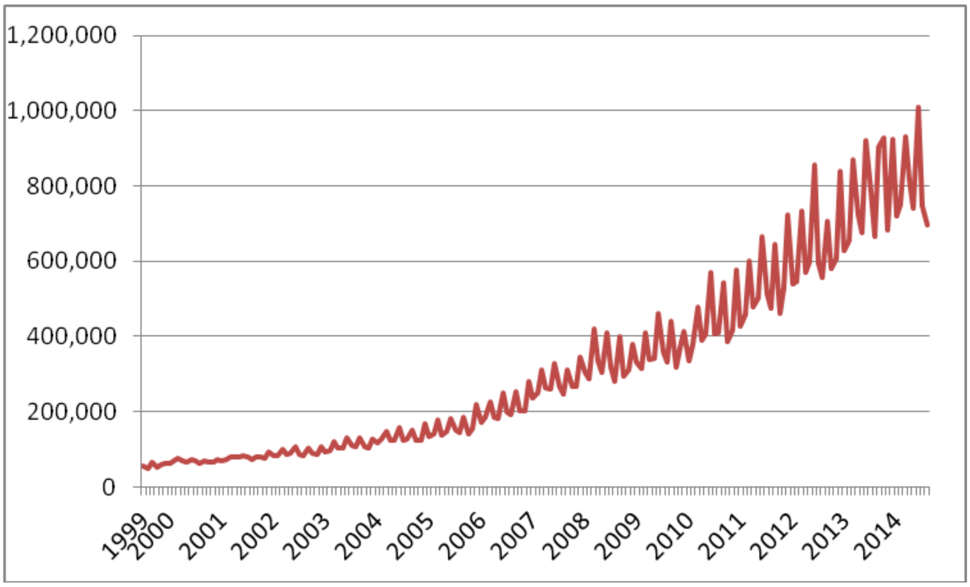


Figure 1: Monthly Tax Collection (TZS million)
Source: Government Budget Operations Reports

4.1.3 Growth and Composition of Annual Tax Collection

Tax revenue collection by tax type is presented in Table 2. Results in Table 2 show that tax collection increased steady from TZS 4.42 trillion in 2009/10 to TZS 9.36 trillion in 2013/14. This growth is attributed by increased efforts in collection and steady economic growth. Annual growth rates of total tax revenue were compared and it was found that there is alternating high and low growth rates in consecutive years, a pattern which can be well explained by business cycle of the Tanzania economy which is characterized by alternating increasing and falling GDP growth. So the effect of the business cycles is clear because under a progressive tax structure, revenue would automatically rise or fall with the increase or decrease, respectively, in income.

Further, taxes on imports and income taxes (especially corporate income and employment income taxes) are the major contributors. Taxes on consumption (VAT) contribute less than the two afore-mentioned taxes. Low contribution of VAT is largely attributed by challenges in its administration as well as the overall structure of the economy (large share of non-market output and low level of value addition).

Table 2: Annual Tax Collection (TZS million) 2009/10 to 2013/14

Source	2009/10	2010/11	2011/12	2012/13	2013/14
Tax revenue	4,427,834	5,295,589	6,480,478	7,729,986	9,364,943
Taxes on imports	1,916,612	2,283,257	2,555,536	2,915,215	3,535,758
Sales/VAT and excise on local goods	934,063	1,064,072	1,336,916	1,466,562	1,607,136
Income taxes	1,334,020	1,660,385	2,246,784	3,019,556	3,778,546
Other taxes	243,139	287,875	341,242	328,653	443,504

Source: Government Budgetary Operations Reports

4.1.4 Volatility in Monthly Tax Collection

Preliminary analysis of volatility (month-to-month growth fluctuations) was conducted using percentage changes in tax collection for each tax type. The results of percentage changes in monthly tax revenue over a period of March 2014 to February 2015 are depicted in Figure 2. Results in Figure 2 indicate that all taxes are volatile. However, total monthly tax revenue and incomes taxes are more volatile and they follow more or less the same pattern, albeit income tax has more volatility than the rest of taxes. Taxes on imports are relatively more stable. These results suggest that the observed volatility in total monthly tax collection is explained to a large extent by the fluctuations in income taxes. According to Cornia and Nelson (2010), macroeconomic conditions and tax structures jointly determine the growth and volatility of tax revenues. For this reason, tax policy makers need to consider the natural tendencies of their economies when formulating tax policy.

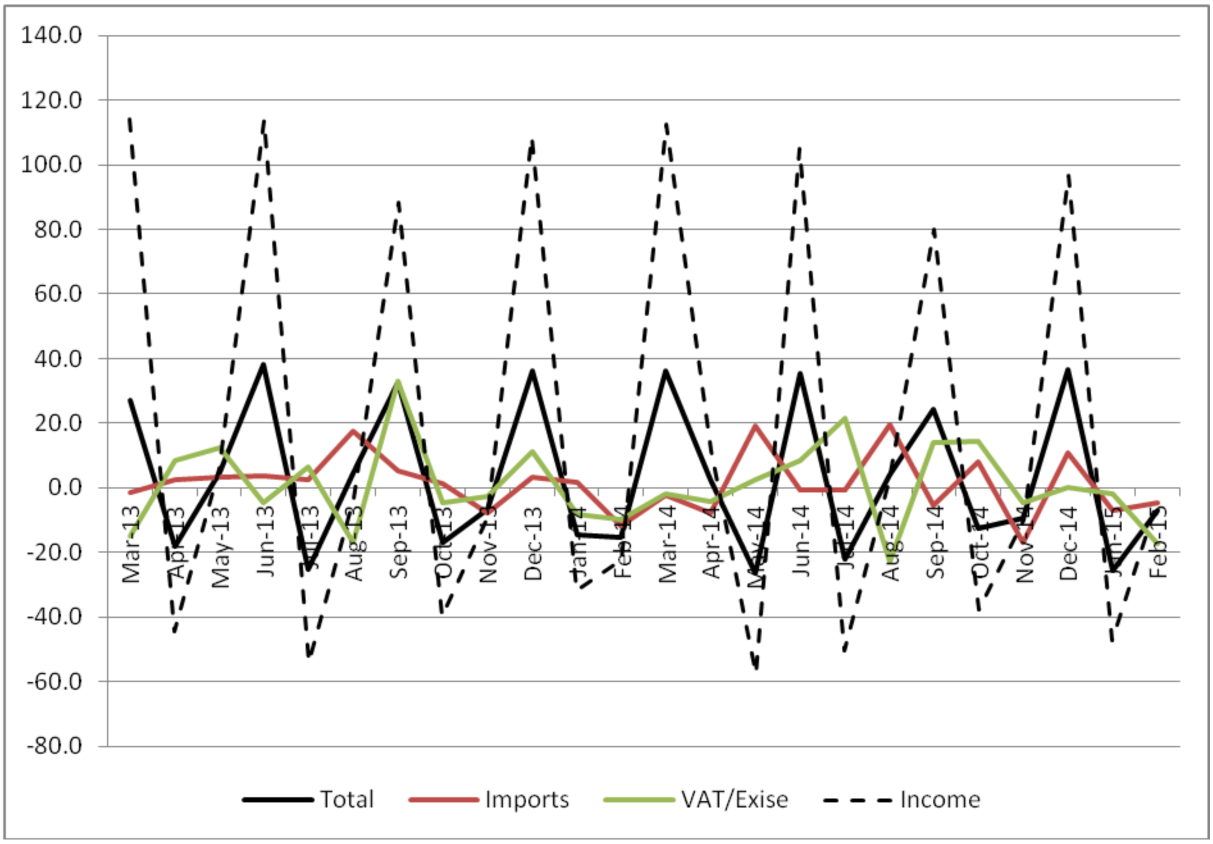
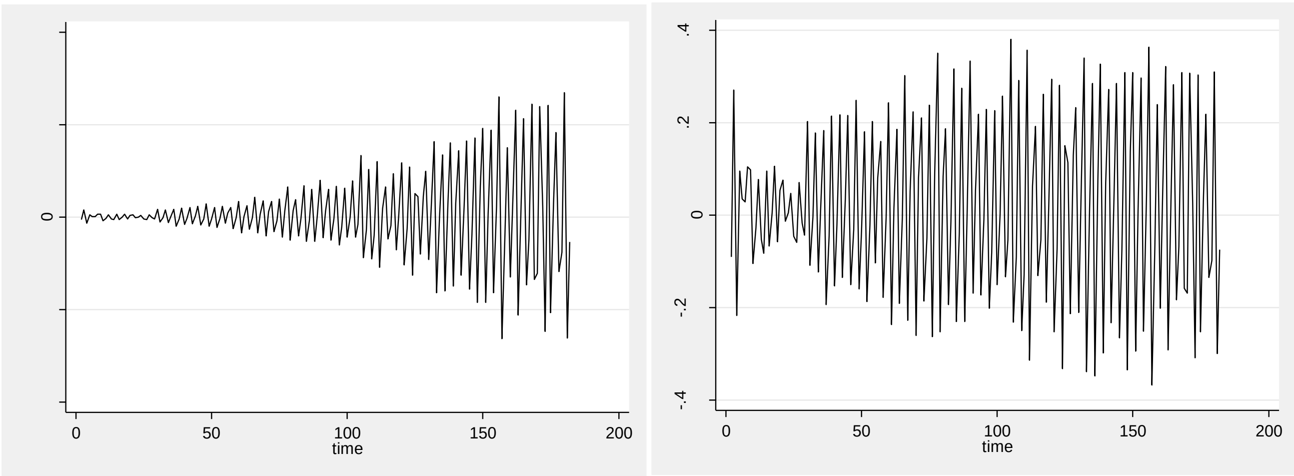


Figure 2: Monthly Percentage Changes in Tax Collection
Source: Government Budgetary Operations Reports

4.2 Test of Stationarity
Before a variable with time series data is used for modelling it is recommended to test for stationarity so as to avoid estimation of spurious relations. To test for stationarity, first tax collection trends were observed (Figure 1). The tax collection trend depicted in Figure 1 suggests that the time series is nonstationary because the series is trending. When first difference series were plot, they were found to be covariance nonstationary (Figure 3 left panel).

Figure 3: Trends of Difference (left) and Log difference (right) of Tax Collections



Source: Researcher’s computations

This preliminary test of stationarity suggests logarithm transformation to supress the variances. The plot of first difference of log values is observed to be stationary (Figure 3 right panel). However, statistical tests had to confirm these results. Results of statistical tests (Unit root tests) are presented in Table 3 below.

Table 3: ADF and PP Unit Root Tests

ADF test in Level			**PP test in Level**		
Calculated Z = -2.452	Critical values		Calculated Z = -0.941	Critical values	
	1%	-3.483		1%	-3.483
	5%	-2.885		5%	-2.885
	10%	-2.575		10%	-2.575

ADF test (After log & difference)			**PP test (After log & difference)**		
Calculated Z = -48.467	Critical values		Calculated Z = -50.05	Critical values	
	1%	-3.484		1%	-3.483
	5%	-2.885		5%	-2.885
	10%	-2.575		10%	-2.575

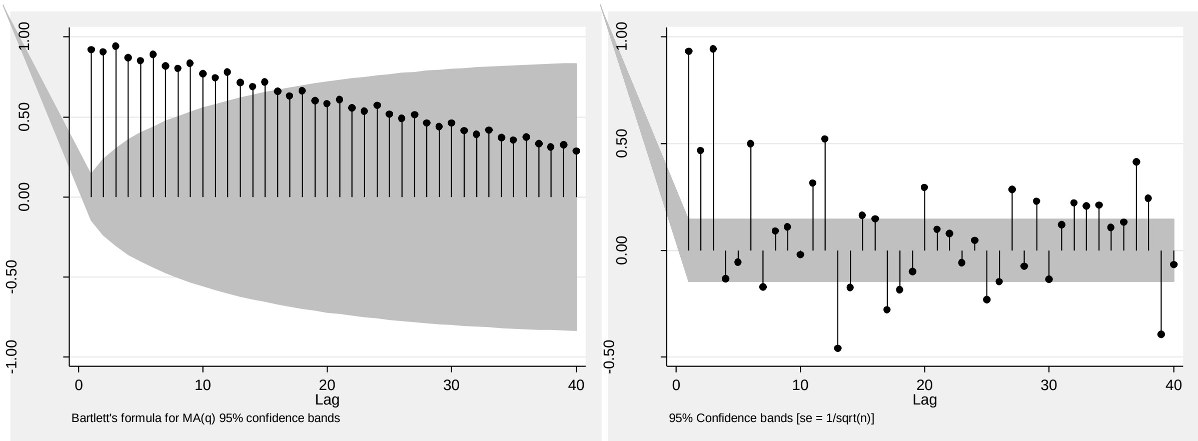
Results of Unit root test presented in Table 3 show that the ADF and PP test rejected stationarity of original data (in level) as the absolute values of calculated Z-statistic are less than the tabulated values. The ADF of first difference accepted stationarity with trend while

PP rejects trend and both tests reject presence of drift. The presence of exponentially growing covariance necessitated logarithm transformation in order to smooth the fluctuations existing in the data. The calculated statistics of differenced log values of tax collection for both ADF and PP are higher than the tabulated values at 1, 5 and 10 percent significant levels. Therefore a null hypothesis of stationarity was accepted; log tax revenue time series is integrated of order one i.e. I(1).

4.3 Identification of ARMA models

After identified the correct differencing order required to make the time series stationary, the next step is to find an appropriate ARMA form. The identification of the ARMA form is done by using a traditional Box-Jenkins procedure that uses plots of correlogram, autocorrelation functions (ACF) and partial autocorrelation functions (PACF). The correlogram of original data is presented in Appendix 1. The Correlogram, ACF and PACF graphs indicate that several autocorrelation and partial autocorrelation coefficients are significant ($p<0.05$). For ACF and PACF this is indicated by points outside the shaded area at different lag values. These results suggest a mixed ARMA model to be estimated. The cut-off points of ACF and PACF in Figure 4 indicate significant MA and AR terms, respectively that can be used to estimate the model.

Figure 4: Autocorrelation and Partial Autocorrelation Functions of Tax



Based on significant lags, several models were estimated. In the vein of a parsimonious model first three lags were chosen because they have high correlation coefficients. Further, the choice of these lags is based on administration practice of cascading annual revenue targets into monthly and quarterly targets which informs that farther past collections may have a small influence to the current monthly tax collection targets and volatility. Thirteen different ARMA models were identified and estimated (see Appendix 2).

4.4 ARMA Models Estimation

With the objective of arriving at a balance between capturing as much dynamics as possible and having a parsimonious model, different model selection criteria were used (Akaike Information Criteria, Bayes-Schwarz Information Criteria, Wald statistic, Log likelihood, and Durbin-Watson test) and forecast performance criteria (root mean square error - RMSE, and normality test). Comparison of evaluation criteria for the competing thirteen models is presented in Appendix 2. It can be seen from Appendix 2 that models (5), (7), (11) and (12) provides superior results compared to the rest of models. These models have higher R-squared, lower MAE and RMSE. Except for model (11), the AIC and BIC of these models fairly compares. Further, they have normally distributed error terms and fairly weak to no autocorrelation. Further, all parameter estimates of the AR and MA terms are significant at p<0.01, while the rest of the models have their MA terms insignificant at p<0.05. Model (5) is also more parsimonious than the other 3 models. Therefore model (5) is proposed for forecasting monthly tax revenue. The model estimation results are presented in Table 4 below. Results indicate that AR(3), MA(1) and MA(3) terms influence current monthly change in tax revenue collection significantly (at p < 0.01). The AR(3) term was found to influence current changes in revenue positively while MA terms have negative effects. The constant tern is also highly significant (p < 0.01) indicating the presence of a very small drift of about 0.0146 implying that in the long run the change in monthly revenue collection (in logarithm) would converge to an equilibrium level of 0.0146.

Table 4: Estimation Results of Model (5)

				Number of obs	=	181
				Wald chi2 (3)	=	3528.6
Log likelihood = 235.4623				Prob > chi2	=	0.0000

Dif_logtax	**Coef.**	**OPG Std. Err**	**z**	**P > |z|**	**(95% conf. Interval)**	
Constant	0.0146212	0.0024542	5.96	0.000	0.009811	0.0194315
ARMA						
ar						
L3.	0.9740063	0.0167705	58.08	0.000	0.9411367	1.006876
ma						
L1.	-0.4099128	0.1424941	-2.88	0.004	-0.689196	-0.1306296
L3.	-0.5911633	0.0934816	-6.32	0.000	-0.7743838	-0.4079427
/sigma	0.0649464	0.0048506	13.39	0.000	0.0554394	0.0744534

Thus the estimated ARMA model of monthly change in tax collection can be written as,

$$\Delta y_t = \underset{(0.0025)\ *}{0.0146212} + \underset{(0.0168)\ *}{0.9740063}\Delta y_{t-3} - \underset{(0.1424)\ *}{0.4099128}\varepsilon_{t-1} - \underset{(0.0935)\ *}{0.5911633}\varepsilon_{t-3} \quad (5)$$

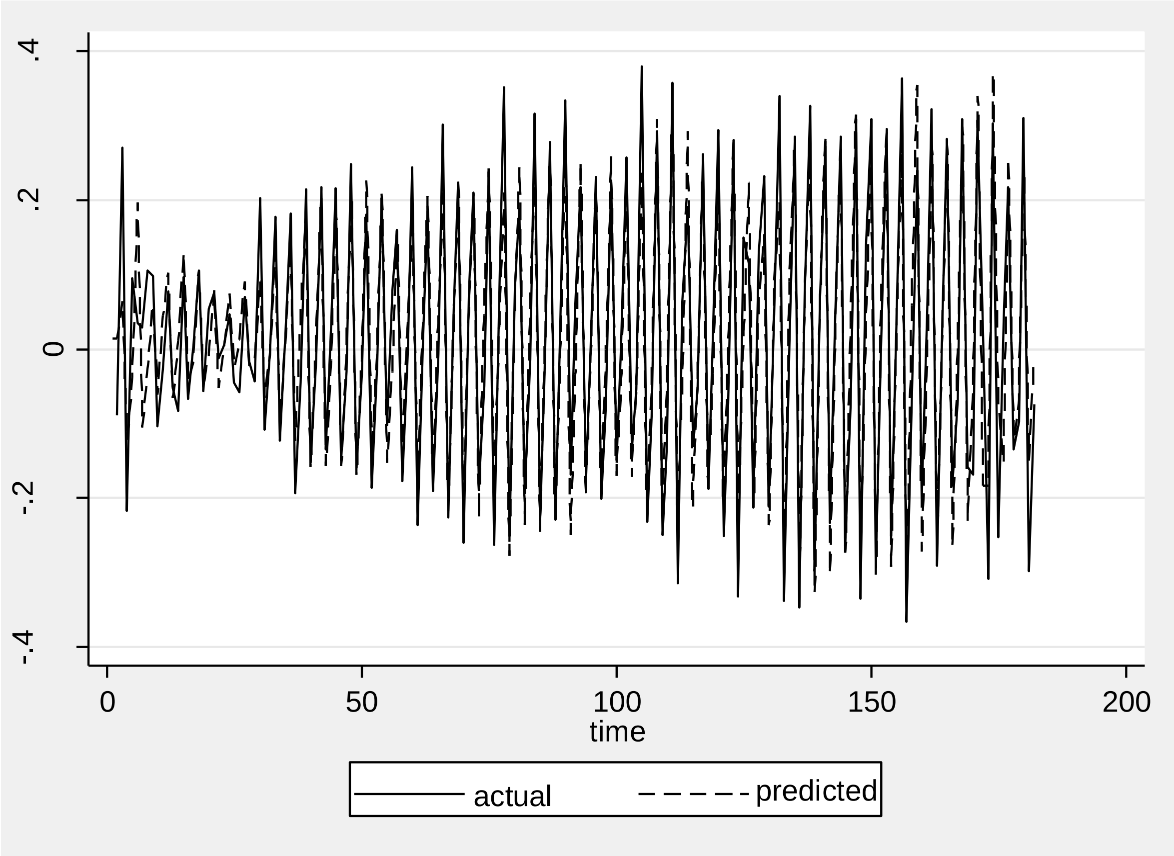
where y is the first difference of log tax revenue

* Significant at p<0.01, standard errors in parentheses.

4.4.1 Post-estimation Model Evaluation (in-sample forecast)

Post-estimation model evaluation was performed by examining the distribution of the error terms using correlograms of errors and squared errors (Appendices 3 and 4) and plots of actual and fitted data. It can be seen from these correlograms that the squared errors are very small ruling out autocorrelation. According to Johnston and DiNardo (1997), absence of autocorrelation leads to unbiased and consistent estimators. Graphical presentation (Figure 5) shows that the model close-fitting the data as the forecasts and actual trends move very close in the same pattern at almost all data points. Hence the estimated model provide good fit of the data and may be used for forecasting monthly tax revenue.

Figure 5: Actual and Predicted Revenue (log differenced)



4.4.2 Out of Sample Forecast

The estimated ARMA model is in first difference of logarithm. In order to perform out-of-sample forecasting the ARMA model was generalized for a future t period forecast y_t (in logarithm). This model (equation 6a) is used to generate 12-months out of sample forecasts.

The forecast value y_t is expressed as,

$$y_t = y_{t-1} + \Delta y_t, \quad \forall t \geq 1 \tag{6a}$$

Thus forecast at any future time t is generalized by the equation,

$$y_t = y_0 + \sum_{t=1}^{T} \Delta y_t, \quad \forall t \geq 1 \tag{6b}$$

Then ARMA forecasts (equation 6b) were compared with the 3-, 6- and 12-months smoothed centred moving average series (S). The in-sample forecast performance of the models may provide a clue on accuracy and reliability of these forecasts. The computed in-sample RMSE for ARMA, S(3), S(6) and S(12) are 0.0675, 0.3098, 0.1868 and 0.1465, respectively. These results indicate that ARMA model performs better than the centred moving average forecasts. The series S(3) and S(6) were not consider for further analysis due to their low forecast performance.

Out-of-sample forecasts presented in Figure 6 produce interesting results; for shorter periods up to 5 months the ARMA and S(12) series are closer than they are at farther periods (5-months and beyond). However, as ARMA and S(12) started to diverge each other from period 7 onwards a faster loss of forecast power of either of these models after period 6 was suspected (it is a common phenomenon in forecasting that the farther forecast the higher the forecast errors). ARMA forecasts were observed to increase faster than S(12) with a gradient higher than that of the actual series.

In view of widening out of out-of-sample forecasts this paper could neither recommend the very optimistic forecasts of ARMA and could be the underestimated forecasts of S(12). Instead a balanced forecast was arrived at by using a combination of forecasts method; naïve ARMA forecasts were not used because they would be much correlated with ARMA forecasts. First, simple average combination were performed followed by linear combination forecasts of ARMA and S(12), and then compared their forecast performances. It was found that linear combination method[19] provides better forecasts (with lowest RMSE of 0.0629) compared to other forecasts described above. The obtained combination weights are 0.9595 and 0.0378 for ARMA and S(12), respectively with a constant of 0.0354 (see Table 5). Although this linear combination is no much improvement in terms of in-sample forecast accuracy compared to ARMA, it is recommended because of its moderately optimistic out-of-sample forecasts.

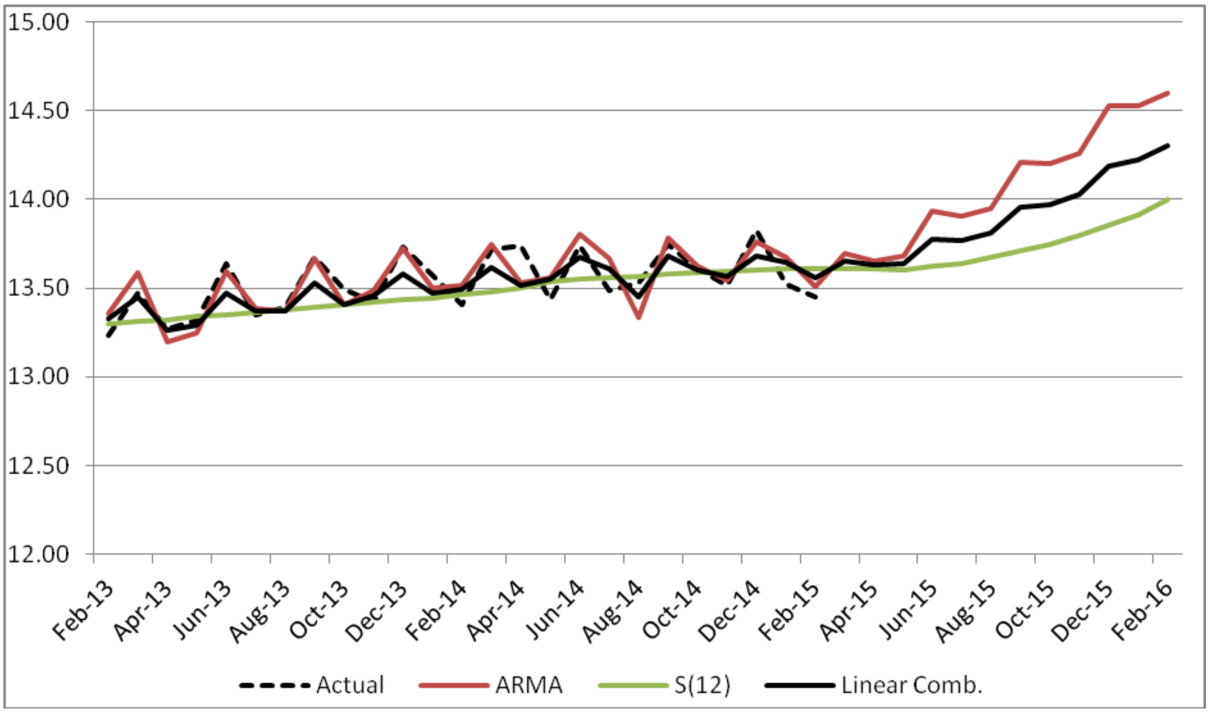
Table 5: Linear Combination of ARMA and S(12)

Number of obs	=	171	R-squared	=	0.9938
F (2, 168)	=	13566.34	Adj. R-squared	=	0.9938
Prob > F	=	0.0000	Root MSE	=	0.0629

Log_tax	**Coef.**	**OPG Std. Err**	**t**	**P > |t|**	**(95% conf. Interval)**	
ARMA	0.9595	0.0353	27.15	0.000	0.889796	1.029367
S(12)	0.0378	0.0350	1.080	0.282	-0.031422	0.107144
Constant	0.0354	0.0754	0.470	0.639	-0.113361	0.184212

[19] Combining weights were determined by using Ordinary Least Squares (OLS). For details see Granger & Ramanathan (1984), Aksu & Gunter (1992) and Gunter (1992). They recommend the use of OLS combination forecasts with the weights restricted to sum to unity.

Figure 6: Out of Sample 12-Months Forecasts (Log Tax Revenue)



Source: Own computations

4.5 GARCH Models Estimation of Revenue Volatility

To test for ARCH effect graphical observation and Engel's Lagrange multiplier (LM) test were used. In Figure 3 above volatility clustering are observed which indicates the presence of ARCH effect. The results of LM test (Table 6) confirm the presence of ARCH effect as indicated by the significance of Chi-square statistics of the first five lags at $p<0.01$. Therefore, the test rejected the null hypothesis that the errors are not autoregressive conditional heteroskedastic. These results warranty estimation of the GARCH models using these lags.

Table 6: Engel's LM Test for Autoregressive Conditional Heteroskedasticity (ARCH)

Lags (p)	Chi2	df	P > chi2
1	72.805	1	0.0000
2	87.194	2	0.0000
3	154.353	3	0.0000
4	153.496	4	0.0000
5	152.702	5	0.0000

H_0: no ARCH effects, H_1: ARCH(p) disturbance

Since the Engle's LM test failed to reject a null hypothesis of no ARCH, the immediately question is which order of GARCH to be used? Empirical studies (e.g. Bera and Higgins, 1993; Javed and Mantalos, 2013) agrees on the performance of standard GARCH(1,1) model rather than attempting to determine the ‘appropriate’ lag values. They claim that first lag of conditional variance captures a large share of information on volatility clustering in the data. However, this assumption can be misguided if the real data has high order variance structure. In this later case a true specification of volatility structure for the candidate data is required. Limiting to a maximum of two lags, six different GARCH models were estimated and compared using AIC, BIC and RMSE to determine the best model. These criteria have been used in other volatility studies also (e.g. Alberg et al., 2008).

Results in Table 7 shows that all models fairly compares in terms of RMSE. However, these models differ considerably in AIC and BIC. Model (4) was found to balance these criteria. It has the lowest AIC, BIC and RMSE. By having a relatively lower AIC and BIC it implies that model (4) minimized information loss. A next candidate is model (1). When estimation results of models (1) and (4) were compared, model 1 was found to have relatively lower standard errors of estimated parameters. Lower standard errors imply that results are more reliable. Model (4) has all parameter estimates of GARCH, except for a constant, insignificant at p<0.05 while model (1) has a significant GARCH(1) term. Therefore model (1) which is GARCH(1,1) is selected for modelling and forecasting volatility of monthly tax revenue.

Table 7: Comparison of GARCH Models

Model	Specification	AIC	BIC	RMSE
(1)	ARCH(1)GARCH(1)	-91.67375	-78.87976	0.0114402
(2)	ARCH(1)GARCH(2)	-93.07970	-80.28571	0.0125081
(3)	ARCH(2)GARCH(1)	-93.41996	-80.62597	0.0112022
(4)	ARCH(2)GARCH(2)	-74.14166	-61.34767	0.0101905
(5)	ARCH(1,2)GARCH(1)	-91.80689	-75.81440	0.0114687
(6)	ARCH(2)GARCH(1,2)	-98.75352	-82.76103	0.0115867

Estimation results of model (1) are presented in Table 8. Results in Table 8 show that GARCH(1) term is statistically significant (p<0.01). Further, results indicate that first lag of the squared variance positively contribute to the current disturbance (p<0.01). Likewise, the first lags of the squared error terms have positive effect on the current disturbance, although it is insignificant at p<0.05. The coefficients are all less than unity implying that in the long run disturbance converges to a constant. However, the sum of coefficients of the model is marginally different from unit (0.9966) which implies that the process converges very slowly and thus volatility would persists for a long period. ARCH term has a relatively very low coefficients indicating that in the long run volatility is driven by GARCH term.

Table 8: Estimation of GARCH Model

	Number of obs	=	181
Distribution: Gaussian	Wald chi2 (.)	=	.
Log likelihood = 49.83688	Prob > chi2	=	.

Dif_logtax		Coef.	OPG Std. Err	z	P > |z|	(95% conf. Interval)	
dif_logtax							
	_cons	0.0188315	0.0129501	1.45	0.146	-0. 0065502	0. 0442132
ARCH							
	arch						
	L1.	0. 1756534	0. 1410118	1.25	0.213	-0. 1007247	0. 4520316
	garch						
	L1.	0. 8207645	0. 1473613	5.57	0.000	0.5319416	1.1095870
	_cons	0. 0005361	0. 0009652	0.56	0. 579	-0. 0013557	0. 0024279

The estimated mean equation and GARCH model are presented in equations (8) and (9), respectively as,

$$\Delta y_t = 0.0188315 + \varepsilon_t \quad (8)$$

$$\underset{(0.00129)}{\sigma_t^2 = 0.0005361} + \underset{(0.14101)}{0.1756534\varepsilon_{t-1}^2} + \underset{(0.0.00096\ )*}{0.82076455\sigma_{t-1}^2} \quad (9)$$

* Significant at p<0.01, standard errors in parentheses.

Tax structure has effects on volatility of tax revenue[20]. As indicated in preceding explanations, income taxes are culprit of the volatile revenue situation. The presence of the observed long term volatility can be well explained by the permanent income hypothesis. As the permanent income hypothesis predict, consumption is more stable than incomes because of the options to smooth consumption. Thus widening consumption tax base would reduce the volatility in tax collection. These results are supported by the findings of other studies which can be separated into two; those which suggest substitution of more progressive taxes with a less progressive tax in order to reduce revenue volatility and improve revenue forecasting (e.g. Thompson and Gates, 2007; Boyd and Dadayan, 2014), and those which emphasize on diversification of taxes, than substitution of taxes, as a means to reduce revenue volatility (e.g. Crain, 2003). The fact that income taxes in Tanzania are more progressive than consumption taxes, based on our findings and these recommendations of previous studies this paper argue that widening the base of consumption taxes may help achieve enhanced revenue collection and reduction in tax revenue volatility in Tanzania.

[20] Tax Structure and volatility. http://oregoneconomicanalysis.com/2015/02/27/tax-structure-and-volatility/

5.0 Conclusion and Recommendations

5.1 Conclusion

The study found that monthly tax revenue persistently increased over the period owing to expansion in incomes, discretionary changes in administration and tax rates. It was further found that monthly tax revenue has high volatility which grows over time. High volatility is linked to reliance on income tax which was observed to have high volatility that follows the same pattern as the total monthly revenue volatility.

Based on different econometric procedures and criteria several models were estimated and compared, ultimately the best models were suggested for forecasting monthly revenue and its volatility; a linear combination of ARMA and S(12) for forecasting monthly tax revenue and GARCH(1,1) for volatility.

5.2 Recommendations

This study recommends linear combination of ARMA and S(12) for forecasting monthly tax revenue and GARCH(1,1) for forecasting volatility. In order to have better forecast accuracy this study recommends proper timing of forecasts for shorter time lags between production and usage of forecasts; the farther forecasts the higher forecast error. However, as afore-mentioned, this paper couldn't specify exact timing for updating of the models; forecasters' need to observe trends that may affect forecast accuracy significantly.

The policy implication from this study is the need to enhance diversity in taxes within the existing tax portfolio so as to reduce volatility. This recommendation corroborates other studies such as Crain (2003). Low volatility implies more predictability of revenues for budgeting and tax administration purposes. It is however worth mentioning that even the best-designed tax portfolio would not eliminate volatility in tax revenue growth.

As income taxes comprise a large share of taxes and are very volatile than consumption taxes, enhancement in collection of consumption taxes such as value added tax (VAT) is also recommended. Enhancing consumption taxes collection will increase tax revenue and reduced volatility hence improve forecast; consumption is more stable and their taxes are difficult to evade. To achieve this, expansion of the tax base by increasing value addition of economic activities. Further, reduction of the informal sector and improvements in enforcement to bring more people in the tax net is important.

6.0 References

Akaike, H. (1974). A New Look at the Statistical Model Identification. IEEE Transactions on Automatic Control 19(6): 716-723.

Aksu, C., & Gunter, S. (1992). An Empirical Analysis of the Accuracy of SA, OLS, ERLS and NRLS combination forecasts. International Journal of Forecasting 8:27–43.

Alberga D., Shalita H., and Rami Y. (2008). Estimating stock market volatility using asymmetric GARCH models. Applied Financial Economics 18: 1201–1208.

Auerbach, A.J. (1995). Tax Projections and the Budget: Lessons from the 1980's. American Economic Review 85: 165-169.

Bera, A.K. and Higgins, M. L. (1993). ARCH Models: Properties, Estimation and Testing, Journal of Economic Surveys 4:305-362.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics 31(3) 307–27.

Boyd, D. J. and Dadayan, L. (2014). Revenue Forecasting: State Tax Revenue Forecasting Accuracy. Rockefeller Institute of Government Technical Report.

Bretschneider, S.I., W.L. Gorr, G. Grizzle, and E. Klay (1989). Political and organizational influences on the accuracy of forecasting state government revenues. International Journal of Forecasting 5:307-319.

Bunn, D.W. (1985). Statistical Efficiency in the Linear Combination of Forecasts. International Journal of Forecasting 1:151–163.

Cornia, G. C. and Nelson, R. D. (2010). State Tax Revenue Growth and Volatility. The Federal Reserve Bank of St. Louis Regional Economic Development 6(1): 23-58.

Crain, William Mark (2003). Reliability of Revenues from Alternative Tax Instruments. In Volatile States: Institutions, Policy, and the Performance of American State Economies. Ann Arbor, MI: University of Michigan Press, pp. 73-81.

Dharan, B.G. (1985). A Priori Sample Size Evaluation and Information Matrix Computation for Time Series Models. J. Statist. Comput. Simul. 21:171-177.

Engle, R. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica 50(4):987–1007.

Fjeldstad, O-H, Jensen, S.K., and F.M. Paulo (2014). Poor Revenue Forecasting: A Major Challenge for Sound Fiscal Policy in Angola. Angola Brief Vol. 4. No. 1, February 2014. http://www.cmi.no/publications/file/5051-poor-revenue-forecasting.pdf

Garrett, T.A. and Leatherman, J.C. (2000). An Introduction to State and Local Public Finance. http://rri.wvu.edu/WebBook/Garrett/contents.htm

Granger, C.W.J. (1989). Combining Forecasts — Twenty Years Later. Journal of Forecasting 8: 167–173.

Granger, C.W.J., and Ramanathan, R. (1984). Improved Methods of Combining Forecasts. Journal of Forecasting 3:197–204.

Hamilton, L.C. (2006). Statistics with STATA: Updated for version 9. Thomson Higher Learning, Belmont, CA.

Javed, F. and Mantalos, P. (2013). GARCH-Type Models and Performance of Information Criteria. Communication in Statistics. Simulation and Computation 42(8):1917-1933.

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

Johnston, J., and DiNardo, J. (1997), Econometric Methods, Fourth Edition, New York: McGraw-Hill.

Leal, T., Perez, J. J., Tujula, M. and V. Jean-Pierre (2007). Fiscal Forecasting Lessons from the Literature and Challenges. European Central Bank WPS No. 843, December 2007. http://ssrn.com/abstract_id=1054881.

Litterman, R.B., and T.M. Supel (1983). Using Vector Autoregressions to Measure the Uncertainty in Minnesota's Revenue Forecasts. Federal Reserve Bank of Minneapolis Quarterly Review (spring): 10-22.

Mwakasindile, M. (2011). TRA Simplifies Tax Payment with M-Pesa. The Revenues (Quarterly Staff Magazine of the Tanzania Revenue Authority) 4 (27): April –June 2011.

Nandi, B.K., Chaudhury, M. and Hasan, G. Q. (2014). Univariate Time Series Forecasting: A Study of Monthly Tax Revenue of Bangladesh. EWUCRT Working Paper No 9. East West University, Dhaka, Bangladesh.

NASBO - National Association of State Budget Officers (2008). Budget Processes in the States. Washington, DC.

Tanzania Revenue Authority - TRA (2013). Fourth Corporate Plan (2013/2014 – 2017/2018). TRA, January 2013.

Tanzania Revenue Authority - TRA (2014). Annual Report 2013/2014.

Thompson, F. and Gates, B. L. (2007). Betting on the Future with a Cloudy Crystal Ball? How Financial Theory Can Improve Revenue Forecasting and Budgets in the States. Public Administration Review 67(5): 825-36.

Voorhees, W. R. (2004). More Is Better: Consensual Forecasting and State Revenue Forecast Error. International Journal of Public Administration 27(8-9):651-71.

7.0 Appendices

Appendix 1: Correlogram of Original Tax Revenue

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	0.9196	0.9318	156.45	0.0000		
2	0.9053	0.4665	308.92	0.0000		
3	0.9419	0.9437	474.89	0.0000		
4	0.8699	-0.1322	617.27	0.0000		
5	0.8504	-0.0541	754.09	0.0000		
6	0.8899	0.4997	904.76	0.0000		
7	0.8177	-0.1723	1032.7	0.0000		
8	0.8030	0.0900	1156.8	0.0000		
9	0.8370	0.1112	1292.4	0.0000		
10	0.7718	-0.0228	1408.4	0.0000		
11	0.7445	0.3147	1516.9	0.0000		
12	0.7809	0.5199	1637.1	0.0000		
13	0.7151	-0.4621	1738.4	0.0000		
14	0.6879	-0.1735	1832.7	0.0000		
15	0.7202	0.1652	1936.7	0.0000		
16	0.6592	0.1475	2024.4	0.0000		
17	0.6313	-0.2811	2105.3	0.0000		
18	0.6628	-0.1836	2195	0.0000		
19	0.6032	-0.0993	2269.7	0.0000		
20	0.5807	0.2943	2339.5	0.0000		
21	0.6104	0.0963	2416.9	0.0000		
22	0.5575	0.0801	2482	0.0000		
23	0.5369	-0.0572	2542.7	0.0000		
24	0.5693	0.0460	2611.4	0.0000		
25	0.5156	-0.2318	2668.1	0.0000		
26	0.4917	-0.1484	2720	0.0000		
27	0.5140	0.2852	2777.1	0.0000		
28	0.4616	-0.0754	2823.4	0.0000		
29	0.4397	0.2296	2865.7	0.0000		
30	0.4624	-0.1360	2912.8	0.0000		
31	0.4142	0.1214	2950.9	0.0000		

Appendix 2: Model Identification and Forecast Performance Evaluation

Criteria	(1) AR(3) MA(1)	(2) AR(3) MA(2)	(3) AR(3) MA(3)	(4) AR(3) MA(1,2)	(5) AR(3) MA(1,3)	(6) AR(3) MA(2,3)	(7) AR(3) MA(1,2,3)	(8) AR(1,3) MA(1)	(9) AR(1,3) MA(2)	(10) AR(1,3) MA(3)	(11) AR(1,3) MA(1,2)	(12) AR(1,3) MA(1,3)	(13) AR(1,3) MA(2,3)
AIC	-424.37	-367.74	-423.02	-423.47	**-460.92**	-421.07	**-459.51**	-411.75	-371.45	-421.63	**-422.53**	**-459.15**	-419.76
BIC	-405.18	-354.95	-410.22	-407.48	**-444.93**	-405.08	**-440.31**	-395.75	-355.46	-405.64	**-403.34**	**-439.96**	-400.57
Wald statistic	966.51	926.95	33195	464.08	**3528.6**	34705	**3502.2**	380.88	918.93	34887	**478.57**	**3387.0**	37702
P-value	0.0000	0.0000	0.0000	0.0000	**0.0000**	0.0000	**0.0000**	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	0.0000
Log likelihood	218.19	187.87	215.51	216.73	**235.46**	215.53	**235.75**	210.87	190.72	215.82	**217.27**	**235.58**	215.88
Durbin-Watson	2.0878	2.6641	2.9408	1.9390	**2.4028**	2.9493	**2.3391**	1.5068	2.6131	2.9142	**1.9195**	**2.3801**	2.9276
MAE	0.0572	0.0677	0.0578	0.0577	**0.0512**	0.0577	**0.0513**	0.0596	0.0667	0.0575	**0.0575**	**0.0513**	0.0574
RMSE	0.0736	0.0889	0.0777	0.0742	**0.0675**	0.0775	**0.0674**	0.0765	0.0862	0.0767	**0.0739**	**0.0675**	0.0764
Normality (χ^2)	2.1800	9.3400	8.3800	3.3700	**5.6300**	8.0500	**5.9900**	2.6000	4.3400	7.4500	**2.5200**	**5.3500**	7.1600
P-value	0.3360	0.0094	0.0151	0.1859	**0.0599**	0.0179	**0.0500**	0.2725	0.1144	0.0242	**0.2831**	**0.0689**	0.0279
R-squared	0.8564	0.7909	0.8403	0.8546	**0.8793**	0.8411	**0.8796**	0.8450	0.8033	0.8445	**0.8552**	**0.8795**	0.8456
Adj. R-squared	0.8556	0.7897	0.8394	0.8537	**0.8786**	0.8402	**0.8790**	0.8441	0.8022	0.8436	**0.8543**	**0.8788**	0.8448
F-statistic	1067.69	677.03	941.62	1051.7	**1303.8**	947.30	**1308.2**	975.88	731.07	971.93	**1056.8**	**1305.9**	980.56
P-value	0.0000	0.0000	0.0000	0.0000	**0.0000**	0.0000	**0.0000**	0.0000	0.0000	0.0000	**0.0000**	**0.0000**	0.0000

Appendix 3: Correlogram of Residuals

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	-0.2103	-0.2112	8.1383	0.0043		
2	0.0544	0.0143	8.6865	0.0130		
3	-0.2056	-0.2111	16.552	0.0009		
4	0.0775	0.0069	17.676	0.0014		
5	-0.0787	-0.0830	18.84	0.0021		
6	-0.0669	-0.1382	19.687	0.0031		
7	-0.0264	-0.0637	19.821	0.0060		
8	-0.0573	-0.1396	20.449	0.0088		
9	-0.1649	-0.3360	25.683	0.0023		
10	0.1100	-0.0634	28.027	0.0018		
11	0.0497	-0.0362	28.508	0.0027		
12	0.2587	0.2482	41.627	0.0000		
13	0.0122	0.1448	41.656	0.0001		
14	0.0040	0.0304	41.659	0.0001		
15	-0.0055	0.1634	41.665	0.0003		
16	-0.0270	0.0626	41.812	0.0004		
17	-0.0023	0.0412	41.813	0.0007		
18	-0.1083	-0.0294	44.197	0.0005		
19	-0.1004	-0.1049	46.257	0.0005		
20	-0.0657	-0.0562	47.146	0.0006		
21	-0.0918	-0.1097	48.891	0.0005		
22	0.0790	-0.0789	50.19	0.0006		
23	0.1074	0.0599	52.61	0.0004		
24	0.1175	0.0117	55.523	0.0003		
25	0.0429	-0.0313	55.914	0.0004		
26	-0.0681	-0.1715	56.906	0.0004		
27	0.1103	0.0523	59.523	0.0003		
28	-0.0545	0.0105	60.165	0.0004		
29	-0.0195	-0.0685	60.248	0.0006		
30	-0.1070	-0.0561	62.759	0.0004		
31	-0.0635	-0.0062	63.648	0.0005		

Appendix 4: Correlogram of Squares of Residuals

LAG	AC	PAC	Q	Prob>Q	-1 0 1 [Autocorrelation]	-1 0 1 [Partial Autocor]
1	0.3034	0.3035	16.942	0.0000		
2	0.1729	0.0995	22.471	0.0000		
3	0.2678	0.2215	35.819	0.0000		
4	0.1852	0.0600	42.236	0.0000		
5	0.0856	-0.0058	43.616	0.0000		
6	0.0667	-0.0041	44.458	0.0000		
7	-0.0195	-0.1151	44.53	0.0000		
8	0.0363	0.0627	44.782	0.0000		
9	0.0456	0.0527	45.182	0.0000		
10	0.0036	-0.0074	45.185	0.0000		
11	-0.0092	-0.0616	45.201	0.0000		
12	0.0162	0.0301	45.253	0.0000		
13	-0.0040	-0.0210	45.256	0.0000		
14	0.0674	0.1465	46.157	0.0000		
15	-0.0194	-0.1020	46.232	0.0000		
16	-0.0034	-0.0434	46.235	0.0001		
17	0.0475	0.0371	46.69	0.0001		
18	0.0037	-0.0323	46.693	0.0002		
19	-0.0440	-0.0813	47.089	0.0003		
20	0.0013	-0.0136	47.089	0.0006		
21	-0.0432	-0.1010	47.475	0.0008		
22	0.0047	0.0134	47.479	0.0013		
23	0.0440	0.1373	47.885	0.0017		
24	0.0509	0.0784	48.431	0.0022		
25	-0.0481	-0.1475	48.921	0.0029		
26	0.0376	0.0766	49.223	0.0039		
27	0.1213	0.2462	52.389	0.0024		
28	-0.0557	-0.2539	53.06	0.0029		
29	-0.0509	-0.0444	53.625	0.0036		
30	-0.0161	-0.0326	53.682	0.0050		
31	-0.0563	-0.1259	54.382	0.0058		