ISSN (online) 2664-9535 (print) 2664-9527 https://atcr.kra.go.ke

How long is a long run? Tax Revenue Forecasting: A case study of Tanzania

Masoud Mohammed Albimana [1]
Issa Moh'd Hemedb[2]

[1] Institute of Tax Administration (ITA), Dar-es-Salaam, Tanzania
[2] Zanzibar University, Zanzibar, Tanzania

Received 18 March 2022
Accepted for publication 05 April 2022
Published 19 April 2022

Abstract

This paper intends to examine whether using long run sample size has more forecasting power than short run sample size. The sample size ranges from 1996 to 2016 and 2000 to 2015. Ordinary Least Square (OLS) method was used to forecast three components of tax revenues including total revenue (TR), Pay As You Earn (PAYE) and Value-added Tax (VAT). The results show that, both TR and PAYE forecasts are slightly better when using short run sample size. However, for VAT, forecasting power is slightly better when using long run sample period. This reveals that, in contrast to other fields, forecasting tax revenue using the short run sample size data could be more useful. We believe that, the long run period is subjective and field oriented. Also, the nature of the tax can have different implications in selection of sample size and data frequency.

Keywords: Forecasting, Tax Revenue, VAT, PAYE, Tanzania.

1. Background

Forecasting models using time series have frequently been used in financial markets and economic growth models, but it has not been well-articulated within tax revenue studies. The selection of sample size is a key component in forecasting and should carefully be conducted prior to analysis, but this aspect has been very often overlooked in tax studies.

Theoretical studies of time series insisted on using large sample size which would capture a “long run” period in order to ensure accuracy of estimation and forecasting (Hakkio and Rush, 1991; Hawley et al, 2019). However, Hakkio and Rush (1991) failed to have consensus to answer this question “How long is a ‘long run‘?’. They did add that, the length of the ‘long run’ may differ depending on the field. For some academic fields, the long run can be a decade while for others it can be a month (Hawley et al, 2019) . It might be true that in some fields, adding some independent variables increase additional observations on long-run hence shorter sample size might be acceptable.

In the empirical literature, researchers often face limitations of using relatively short span of data due to lack of longer span data. On one front, too long time series data leads to structural breaks that we saw in exchange rate system for African countries and trade openness, which all emerged effectively in the early 1980s. On the other, estimation using short time series are subjected to different claims from different studies (e.g.; DeCarlo& Tryon, 1993; Huitema & McKean, 1991, 1994; DeCarlo & Tryon, 1993; Solanas et al., 2010; Krone et al., 2017)

The minimum number of sample size recommended in time series forecasting differs. However, a considerable consensus ranges from 30 to 50 observations (Hakkio and Rush, 1991; Poole et al., 2002; McCleary et al., 1980; Warner, 1998). The general conclusion from some literature is that, the quality of estimation coefficients increases with an increasing number of sample size.

In turn, some researchers tend to choose higher frequency data for forecast purposes (Su Zhou, 2001; Hakkio and Rush, 1991; Lahiri and Mamingi, 1995; Choi and Chung, 1995; Ng, 1995). Su Zhou (2001) suggested that, using fixed sample size of 20 to 50 years, moving from low frequency to higher frequency data, may either double or even triple the power of the tests. The validity of this suggestion is still doubtful as it was based on small annual data (See examples; Bahmani-Oskooee, 1996; Masih and Masih, 1996; Taylor, 1995).

The empirical testing of previous discussions is still limited especially regarding tax and fiscal studies. Henceforth, the present paper is motivated to to compare forecasting result of a long run sample size of 24 years (88 quarterlies) with that of a short run sample size of 18 years (44 quarterlies) .This objective is motivated by debates presented by previous studies (Poole et al., 2002; McCleary et al., 1980; Warner, 1998; DeCarlo & Tryon, 1993; Huitema & McKean, 1991, 1994; De Carlo & Tryon, 1993; Solanas et al., 2010; Krone et al, 2017). The next sections of this paper includes, review of theoretical and empirical studies, methodology, empirical discussions and conclusion with recommendation. Table 1 provide snapshot contribution of VAT and PAYE to total Tax revenue which implies a significant importance. This implies that, forecasting of this tax categories can be very crucial in Tax Administration. VAT contributes around 29 percent while PAYE contributes around 12 to 17 percent.

See annex table 1 Contribution of VAT and PAYE to TRA Mainland (%)

2. Literature Review

2.1 *Discussion of a short and long run sample size.*

The theory of taxation states that tax revenue is collected by various means with respect to different types of tax. The literatures summarised several factors, among them three main tax revenues, total tax revenue (TR), Pay-As-You-Earn (PAYE) tax and Value-added Tax (VAT). Total Tax revenues, depend on nominal GDP (Klazer, 2013; Bayer, 2015). It implies that, as economic growth expands, it stimulates the growth of tax base such as house hold consumption, domestic investment and international trade. It observed that, economic development, international trade and income level of the country are also factors affecting Tax revenues. (Bird et al, 2008). Generally, taxation depend on dynamic factors including history and institutions of the country. (See, Víctor Mauricio Castañeda Rodríguez, 2018).

There are few specific empirical tax related studies are described. For example,Bayer (2015) compared the performance of tax revenue forecasting using long run and short run sample size in Sweden. He found that, short run sample size performs slightly better compared to long run sample size using VAT and PAYE. Also, Streimikiene et al (2018) examine best revenue forecast between ARIMA, VAR and OLS in Pakistan. The results suggested that, among these models the A.R.I.M.A. model gives better-forecasted values. Gosolov (2022) explained that, forecasting of macro-economic forecast and tax revenue were any based on closely bases of each tax type. Forecast of tax revenue can be evaluated into two approaches on the one hand, as simply, unconditional of unlikely outcome. On other hand, can be conditional based on accuracy of the macro-economic variables. These two approaches depend on the purpose of the evaluation. Using current year estimates, estimate errors of Excise tax and VAT were found relatively low given their size. Overall errors during the estimation, were highly affected by crisis shock of 2008 in respective tax. Once the year of crisis 2008 and 2009 removed, the RMSE errors declined. (Roinn Airgeaadis, 2019).

Other studies were specific to relevant to the area of sample size. For example, Krone et al. (2017) examined the performance of AR (1) parameter for a time (T) ranging from 10 to 100 and found that as T increases, bias, including bias of the standard error, decreases. Jebb et al. (2015) argued that, the length of time series data can vary, but is generally not less than 20 observations to qualify as long run, and many models require at least 50 observations for accurate estimations (McCleary et al., 1980). At the very least, a time series should be long enough to capture the phenomena of interest. Su Zhou (2001) conducted a simulation study and showed that studies with sample sizes of annual data spanning 30 to 50 years or those using higher frequency data yield greater forecasting capability and less distortion.

More recently, Qin Linet al. (2019) conducted a simulation and empirical studies to investigate the accuracy of ARIMA forecasting under four different lengths of time series from 5 to 30 years of historical data. They observed that the ARIMA model with the shortest time series holds the lowest forecasting. Although empirically models with a smaller sample size may converge to a solution and generate parameter estimates, the estimates may contain bias, which may affect inferences in applied research (McNeish, Daniel M.; Stapleton, Laura M. (2016). Predictions based on limited records are unlikely to be as good as those based on a large number of samples (Hernandez et al., 2006). They are reliable for under or overestimations, irrespective of the formula used in the calculation (Springate, 2012).

Gavilanes (2020) performed simulations of Monte Carlo on 6, 10, 20 and 500 samples and used the regression techniques of OLS, bootstrapping and others leading to his dissuasion of using lower sample sizes in generating significant relationship in the regression (Hryniewicz and Kaczmarek, 2015). Generally, more observed indicators per factor could reduce its magnitude coefficients (Marsh et al., 1998)Many studies are increasingly using broader data sets these days. Armstrong (2011) too claimed that, the benefit of using long run data significantly enhances statistical power, but there are also data interpretation problems associated with this increased power that are less well known and addressed.

In other words, large sample size may also impose problems. The problems can affect all statistical procedures to some degree, such that those that use t-statistics or F-statistics tests to compare mean differences between groups, or those that use correlation and regression, are particularly vulnerable. Generally, the above discussion implies unresolved debate whether using long run or short run sample size calls for superior results. To make it worse, the evidence is particularly limited in the fiscal and tax studies.

3. Methodology

3.1 Description of the variables and Sample Size.

This paper utilizes annual data from 1996 to 2016 (20 years) to answer the our objective, is there a significant difference between short run sample size [2000 - 2016 (16 years)] and long run sample size [(1996-2016 (20 years)]. This objective will forecast tax revenue in each quarter of 2017. Three common taxes VAT (Value-added tax); PAYE (Pay As You Earn), Total tax revenue (TR) will be used as samples due their great contribution to total Tanzania's tax revenue handled by Tanzania Revenue Authority (TRA).

We selected tax base for each tax, VAT, PAYE and TR, based on previous studies as explained in the next section. Before long run estimation using OLS, all data were tested if they were stationary at first difference using Augmented Dickey Fuller (ADF) test, which is widely used in time series analysis. Then, we tested for cointegration using Johansen and Juselius (1991) multivariate cointegration test to ensure that our forecasting regression is not spurious.

We chose to calculate tax elasticity/buoyancy using Ordinary Least Square (OLS) method, as it is the best method compared to point estimate or average point estimates. OLS use regression techniques to minimize the errors between actual and the forecasted values. The best regression was selected based on several forecasting criteria such as root mean squared error (R.M.S.E), mean average error (M.A.E), mean absolute percentage error (M.A.P.E) and Theil's inequality coefficient (TH.I.C). In some cases, we used Adjusted R-squared and Akaike information criteria (AIC). At the end, we compared the forecast values with actual values to come up with forecast power.

3.2 Model Specification:

The theory of taxation states that tax revenue is collected by various means with respect to different types of tax. In our regression, we have three main tax revenues, total tax revenue (TR), Pay-As-You-Earn (PAYE) tax and Value-added Tax (VAT).

TR is expressed as a function of nominal GDP (Klazer, 2013; Bayer, 2015). It implies that, as economic growth expands, it stimulates the growth of tax base such as house hold consumption, domestic investment and international trade. Also, house hold final consumption (HFC) can act as substitute if there is a degree of distortion in GDP trend as it is one of the main determinants of nominal GDP (Bayer, 2015).

Second model specification is PAYE which is expressed as a function of salaries and wages (SW) and unemployment rate (UEM) (Bayer, 2015). The amount of wage and salary of employees can be linked directly to the amount of tax collected through PAYE. The relationship between PAYE and salaries and wages is expected to be positive while that of PAYE and unemployment rate is expected to be negative. The negative relationship is due to belief that, as unemployment increases, wages and salaries rolled out shrinks, in turn, causing PAYE collection to decline subsequently (IMF, 2005). In some cases, we can also use expansion of GDP as explanatory variable for PAYE since it has a positive correlation with profits gained by individual taxpayers which makes up the tax base for PAYE (IMF, 2005; Bayer, 2015).

The third is a VAT Model that uses two different explanatory variables namely total household final consumption (HFC) and nominal GDP. Theoretically, VAT is charged from purchases of goods and services. Thus, household final consumption is its best proxy (See, Jenkins et al, 2000; IMF,1985; Bayer, 2015). In some cases, we can also use expansion of GDP as explanatory variable for VAT due to its positive correlation with total consumption and VAT collection (IMF, 2005; Bayer, 2015).

3.3 Data Sources:

Table 4.1 reports the sources of data for each variable. The estimation of these data was done based on the objectives of this study and time scope. For all three models, the data range is from 1996 to 2016, whereby data were grouped into four conditions, long run, short run, low frequency and high frequency. Three main sources for the data collected include World Bank, International Labour Organization (ILO) and Tanzania Revenue Authority. The variables used in this study are Total Revenue (TR), Value-Added Tax (VAT), Pay-As-You-Earn tax (PAYE), Unemployment (UEM), Wage and Salaries (WAGE), Household Final Consumption Expenditure (HC) and Nominal GDP Per Capita (NGDP).

See annex Table 4. 1 Description of the Variables

5. Findings and discussions

4.2 Empirical Results and Discussions

4.2.1 Unit root and Cointegration Results

Before estimation of OLS, we have to verify whether our variables are free from unit root problem and has long run relationships. To determine whether the series has unit root problem, we applied the Augmented Dickey-Fuller (1979) tests, the results of all four sub-samples (short-term, long-term, higher frequency and lower frequency time series) indicated that the series become stationary at first difference (Table 1).

The cointegration results are presented in Table 2A and 2B for lower and higher frequency time series data respectively. After considering both trace statistics and Max-Eigen, the results confirm the existence of a long-run relationship. Generally, this concludes that we can estimate and forecast using OLS method, as expected.

See annex Table 1: Unit root test

See annex Table 2A and 2B: Long-term Vs Short-term; Cointegration Results

4.2.4.4 Forecasting Long Run and Short Run Sample Sizes

4.2.4.1 OLS regression results*:*

This study examines whether using long run time series (1996 to 2016) data is better than the short run time series data (2000-2016). As explained earlier, we use OLS to estimate and forecast all three components of tax namely TR, PAYE and VAT. The result OLS estimates for both spans are presented in Appendix 1 . Table 3 shows the result for forecasting criteria and Table 4 shows the forecast and actual values in both spans together with the difference between the two. Using diagnostic test, we used goodness of fit measurement (Adjusted R-squared), which revealed that, shorter time series models have high adjusted R-squared of between 97 to 98 percent compared to 93 to 95 percent for long run time series models. In fact, the highest coefficient of determination can be taken as the best model.

The findings presented in Appendix 1 show that household consumption (HC) has a significant negative effect on TR but a non-significant one on VAT. The nominal GDP (NGDP) has positive effects on TR and VAT with statistical significance, as expected. The impact of WAGE is also positive and significant, as expected, while Unemployment (UEM) is not significant. The results also align with Streimikiene (2018) who also found that indirect tax are the main determinants of tax revenue. Also, Eugene and Chineze (2016) Eugene and Chineze (2016) also demonstrated a presence of a positive linear relationship between tax base and tax revenues.

4.2.4.2 Results of Forecasting for Long run and Short Run Sample Sizes

4.2.4.2.1 Forecasting Evaluation Criteria

Before forecasting, we looked at the four evaluation criteria for forecasting (R.M.S.E., M.A.E., M.A.P.E and TH.I.C). The results for all four criteria are as shown in Table 3. However, the results obtained are ambiguous . However, the criteria do suggest that, the data for short run are slightly better at forecasting for TR and PAYE tax components and that the long run time series data are better for VAT tax component.

See annex Table 3: Revenue Categories Forecasting Error (Long-term Vs Short-term Samples)

4.2.4.2.2 Forecasting Results

Similar to forecasting criteria, TR and PAYE models perform slightly better in short run than long run period. The difference is between 0.252 (short run) and 0.339 (long run) for TR. Other values are as indicated in Table 4. In real value term, for the short run period in 2016, the forecasted value of TR is 22.852, equivalent to TZS 7608.5 billion while the actual value is 22.6 percent, equivalent to TZS 6,532.54. The difference was 0.252 percent, equivalent to TZS 1,075.96 billion. The forecasted and actual values of real data are in Appendix 2 while their graphs are in Appendix 3 (TR), Appendix 4 (PAYE) and Appendix 5 (VAT).

Our results support previous literature which believe that, using large sample size is not necessary better in forecasting, and that it may be associated with poor interpretation and faces more vulnerable problems (Armstrong, 2011). Jebb et al. (2015) proposed that sample size should be at least 20 years. Rose and Yellen (1989) used only 25 years in their estimation. We also support Hakkio and Rush (1991), who suggested that, the long run of sample size is subjective to a field.

In contrast, we believe that VAT model performs better if the sample size is longer. For data in 2017, the difference between forecasted and actual values are 0.086 for long run period and 0.164 for short run period. These results concur with previous literature including Springate (2011), McNeish et al. (2016) and Gavilanes (2019) who agreed that long run sample size makes better forecast results.

See annex Table 4: To compare the actual and forecasted values

6. Summary and Conclusion

Generally, for the first objective, after considering the four evaluation criteria (R.M.S.E., M.A.E., M.A.P.E and TH.I.C), we found that shorter span models are better and more accurate in their forecasts, with the exception of VAT model which showed otherwise. It is between 0.252 and 0.339 percent for TR during 2016, a considerable difference when converted to real values (in TZS) as shown in Appendix 2. For PAYE the difference was 0.313 and 0.262 for long run data and short run data respectively.

Therefore, we suggest that forecasting tax revenue be done shorter time period give better and more accurate forecasts. This is because shorter period is less influenced by exogenous factors, thus better and more accurate forecast (Bayer, 2015). However, for VAT tax forecast, a longer time span and higher frequency data make for better and more accurate forecast. This, we support the conclusion by Jebb et al. (2015) and Rose and Yellen (1989) who suggested that sample size could be at least 20 to 25 years as minimum sample size. In addition to that, we support Hakkio and Rush (1991) in the sense that sample size is subjective to a field. The nature of the tax can have implication in selection of sample size and data frequency, as being seen in the case of VAT.

Having said all that, our experiment had only tested two time periods (17 vs 21 years). This limits our result to just comparisons between two conditions. Future studies should divide the time series into more variation to see the extent shorter samples give the most accuracy. It would be useful to know the how short is short and what is considered short period.

7. References

[1] Armstrong, J. S. (2011). Illusions in Regression Analysis (No. 81663). University Library of Munich, Germany.

[2] Bahmani-Oskooee, Mohsen. (1996). Decline of the Iranian rial during the postrevelutionaryp eriod: A productivitya p-proach. Journal of Developing Areas, 30:477-92.

[3] Bayer, O. (2015). Relevance of Input Data Time Series for Tax Revenue Forecasting. Procedia Economics and Finance, 25: 518-529.

[4] Bird, R., Martinez-Vazquez, J., & Torgler, B. (2008). Tax effort in developing countries and high income countries: The impact of corruption, voice and accountability. Economic Analysis and Policy, 38(1), 55–71.

[5] Brady, Gordon L.; Magazzino, Cosimo (2017). The Sustainability of Italian Public Debt and Deficit. International Advances in Economic Research, 23(1), 9–20. doi:10.1007/s11294-016-9623-7.

[6] Choi, I., Chung, B.S., (1995). Sampling frequency and the power of tests for a unit root: a simulation study. Economics Letters

[7] Danninger, M. S., & Kyobe, M. A. J. (2005). Revenue Forecasting—How is it done? Results from a Survey of Low-Income Countries (No. 2005/024). International Monetary Fund.

[8] DeCarlo, L. T., & Tryon, W. W. (1993). Estimating and testing autocorrelation with small samples: A comparison of the c-statistic to a modified estimator. Behaviour Research and Therapy, 31, 781–788. https://doi.org/10.1016/0005-7967(93)90009-J

[9] Eugene, N., and Chineze, E. A. (2015). Productivity of the Nigerian tax system (1994–2013). International Journal of Business Administration, 6, 30–40. doi:10.5430/ijba.v6n4p30

[10] Gavilanes, J. M. R. (2020). Low sample size and regression: A Monte Carlo approach. Journal of Applied Economic Sciences (JAES), 15(67), 22-44.

[11] Golosov, M. (2002). Tax revenue forecasts in IMF-supported programs. Working papers, (WP/02/236).

[12] Hakkio, Craig S., and Mark Rush. (1991). Cointegration:H ow shorti s the long-run? Journal of InternationalM oney and Finance 10:571-81.

[13] Hawley, S., Ali, M. S., Berencsi, K., Judge, A., & Prieto-Alhambra, D. (2019). Sample size and power considerations for ordinary least squares interrupted time series analysis: a simulation study. Clinical epidemiology, 11, 197.

[14] Hooker, M.A., 1993. Testing for cointegration: power versus frequency of observation. Economics Letters 41, 359–362.

[15] Huitema, B. E., & McKean, J. W. (1991). Autocorrelation estimation and inference with small samples. Psychological Bulletin, 110, 291–304. https://doi.org/10.1037/0033-2909.110.2.291

[16] Hernandez, P. A., Graham, C. H., Master, L. L., & Albert, D. L. (2006). The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography, 29(5), 773-785.

[17] Hryniewicz, O., & Kaczmarek, K. (2015). Forecasting short time series with the bayesian autoregression and the soft computing prior information. In Strengthening Links Between Data Analysis and Soft Computing (pp. 79-86). Springer, Cham.

[18] IMF (1985). Financial Policy Workshop: Chapter 9 Workshop 7. The case of Kenya. https://doi.org/10.5089/9780939934003.071

[19] IMF (2014): Revenue Fundamentals, Fiscal Forecasting, and the Effective Tax Rate Approach. Fiscal Analysis and Forecasting Workshop Bangkok, Thailand, June 16 – 27, 2014. Retieved from https://www.imf.org/external/region/tlm/rr/pdf/aug2.pdf.

[20] Jebb, A. T., Tay, L., Wang, W., & Huang, Q. (2015). Time series analysis for psychological research: Examining and forecasting change. Frontiers in Psychology, 6, 1–24.

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

[22] Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economics Dynamics and Control 12, 231–254.

[23] Klazar, S. (2003): Efektivnost predikcí daňových příjmů v ČR. Praha, University of Economics, Prague, 2003. Doctoral thesis.

[24] Krone, T., Albers, C. J., & Timmerman, M. E. (2017). A comparative simulation study of AR(1) estimators in short time series. Quality & Quantity, 51:1–21. https://doi.org/10.1007/s11135-015-0290-1

[25] Lahiri, K., Mamingi, N., (1995). Testing for cointegration: power versus frequency of observation — another view. Economics Letters 18:381–386

[26] Masih, Abul M., and Rumi Masih. (1996). Empirical tests to discern the dynamic causal chain in macroeconomic activity: New evidence fromT hailand and Malaysia based on a multivariatec ointegration/vectorer ror-correction modeling approach. Journal of Policy Modeling 18:531-60.

[27] Marsh, H. W., Hau, K. T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. Multivariate behavioral research, 33(2), 181-220.

[28] McCleary, R., Hay, R. A., Meidinger, E. E., & McDowall, D. (1980). Applied time series analysis for the social sciences. SAGE.

[29] Perron, P., (1989). Testing for a random walk: a simulation experiment of power when the sampling interval is varied. In: Raj, B. (Ed.), Advances in Econometrics and Modelling, Kluwer, London

[30] PERRON, PIERRE (1989). Testing for a Random Walk: A Simulation Experiment of Power When the Sampling Interval Is Varied,' in B. Raj, ed., Advances in Econome/rics and Modeling, Kluwer Academic Publisher.

[31] Poole, M. S., McPhee, R. D., & Canary, D. J. (2002). Hypothesis testing and modeling perspectives on inquiry. In M. L. Knapp & J. A. Daly (Eds.), Handbook of interpersonal communication (3rd ed., pp. 23–72). SAGE

[32] Qin, L., Shanks, K., Phillips, G. A., & Bernard, D (2019). The Impact of Lengths of Time Series on the Accuracy of the of Dahlberg's formula. The European Journal of Orthodontics, 34(2), 158-163.

[33] Solanas, A., Manolov, R., & Sierra, V. (2010). Lag-one autocorrelation in short series: Estimation and hypotheses testing. Psicológica, 31, 357-381.https://www.uv.es/psicologica/articulos2.10/9SOLANAS.pdf

[34] Streimikiene, D., Rizwan Raheem, A., Vveinhardt, J., Pervaiz Ghauri, S., & Zahid, S. (2018). Forecasting tax revenues using time series techniques–a case of Pakistan. Economic research-Ekonomska istraživanja, 31(1), 722-754.

[35] Taylor, Mark P. (1995). Modeling the demand for U.K. broad money. The Review of Economics and Statistics, 75:112-7.

[36] Taylor, Mark P. (1995). Modeling the demand for U.K. broad money. The Review of Economics and Statistics, 75:112-7.

[37] Warner, R. M. (1998). Spectral analysis of time-series data. Guilford Press.

[38] Víctor Mauricio Castañeda Rodríguez (2018). Tax determinants revisited. An unbalanced data panel analysis, Journal of Applied Economics, 21:1, 1-24, DOI:10.1080/15140326.2018.1526867.

[39] 42)Zhou, S. (2001). The power of cointegration tests versus data frequency and time spans. Southern Economic Journal, 906-921.

ANNEX

Table 1: Contribution of VAT and PAYE to TRA Mainland (%)

Tax	2015/2016	2016/2017	2017/2018	2018/2019	2019/2020	2020/2021
P.A.Y.E.	**17.0**	**16.0**	**15.4**	**15.4**	**14.3**	**12.3**
Aggregate VAT	**26.9**	**27.9**	**29.4**	**30.5**	**29.0**	**29.2**
VAT Domestic Products	3.5	3.2	3.7	3.8	3.5	3.6
VAT Domestic Services	10.4	11.9	12.2	12.2	12.0	11.1
VAT on Imports	13.0	12.7	13.5	14.4	13.6	14.5

Source: TRA (2022)

Table 4. 1 Description of the Variables

Variable	Description	Measurements	Data source
TR	Total Tax Revenue includes all tax collected by TRA	Total value of tax revenues collected per year expressed in Local currency (TZS) but changed into percentage form.	TRA Website (2021)
VAT	Value-Added Tax is a tax charged at 18 percent from the difference of sales and purchases.	Total VAT expressed in local currency (TZS) but changed into percentage form.	TRA Website (2021)
PAYE	Pay-As-You-Earn is a tax charged from employee's monthly salary and/or wages.	Total PAYE values expressed in local currency (TZS) but changed into percentage form.	TRA Website (2021)
UEM	Unemployment refers to the share of the labour force that is without work but available for and seeking employment.	Unemployment, total (% of total labour force) (modelled ILO estimate)	International Labour Organization
WAGE	Wage and salaried workers (employees) are those workers who hold the type of jobs defined as "paid employment jobs,"	Wage and Salaried workers, total (% of total employment)	World Bank national accounts data, and OECD National Accounts data files.
HC	Household final consumption expenditure (formerly private consumption) is the market value of all goods and services, including durable products purchased by households.	Household final consumption expenditure (constant 2010 US$) but changed into local currency (TZS) for the respective year.	World Bank national accounts data, and OECD National Accounts data files.
NGDP	GDP at purchaser's prices is the sum of gross value-added by all resident producers in the economy plus any product taxes and minus any subsidies not included in the value of the products.	GDP in local currency	World Bank national accounts data, and OECD National Accounts data files.

Variable	Long-Term	Short-Term
	Level Form (ADF Test)	
TR	-1.19884	-1.333248
VAT	-0.4506	-0.928571
PAYE	-0.9848	0.839267
NGDP	-2.1881	-0.171538
WAGE	0.5582	-1.079307
HC	-1.20601	-1.956834
UEM	0.9248	-1.412337
	First Difference (ADF Test)	
TR	-5.1157***	-4.671942***
VAT	-12.7100***	-3.310962**
PAYE	-3.9528***	-3.357858
NGDP	-3.6118***	-3.236243**
WAGE	-3.1562**	-4.167693***
HC	-5.3948***	-3.357858**
UEM	-3.8434**	-3.370349**

Table 1: Unit root test

The tests are performed on the log-levels of the variables. ADF, to the Augmented Dickey-Fuller test. When it is required, the lag length is chosen according to the Akaike information criterion ***p < 0.01, ** p < 0.05, *p < 0.10

Table 2A and 2B: Long-term Vs Short-term; Cointegration Results

A. Short-run period

	TR	VAT	PAYE	
Hypothesized No. CE(s)	**Trace Statistic**	**Trace Statistic**	**Trace Statistic**	**Critical Value**
None *	67.7055***	57.2675***	42.2262***	29.7971
At most 1	11.1609	9.1511	11.5413	15.4947
At most 2	0.54358	0.6018	3.3676	3.8411
Hypothesized No. CE(s)	**Max-Eigen Statistic**	**Max-Eigen Statistic**	**Max-Eigen Statistic**	**Critical Value**
None *	56.5446***	48.1164***	30.6848***	21.1316
At most 1	10.6173	8.5493	6.1737	14.2646
At most 2	0.5435	0.6018	3.3676	3.8415

B. Long-run period

	TR	VAT	PAYE	
Hypothesized No. CE(s)	**Trace Statistic**	**Trace Statistic**	**Trace Statistic**	**Critical Value**
None *	31.0170**	36.8225***	59.6014***	29.7971
At most 1	8.0779	10.3910	26.9143***	15.4947
At most 2	1.7227	0.6740	0.0300	3.8411

Hypothesized No. CE(s)	Max-Eigen Statistic	Max-Eigen Statistic	Max-Eigen Statistic	Critical Value
None *	22.9391**	26.4315***	32.6871***	21.1316
At most 1	6.3552	9.7169	26.8843***	14.2646
At most 2	1.7227	0.6741	0.030018	3.8415

* , **, *** denotes rejection of the hypothesis at the 0.1, 0.05 and 0.01 levelss of significance

**MacKinnon-Haug-Michelis (1999) p-values

Table 3: Revenue Categories Forecasting Error (Long-term Vs Short-term Samples)

Tax form	Observed Freq.	R.M.S.E	M.A.E	M.A.P.E	TH.I.C
T R	Long-Term.	0.4078	0.4027	1.772%	0.0088
	Short-Term.	0.1825	0.1804	0.7972%	0.0059
PAYE	Long-Term.	0.3934	0.3865	1.861%	0.0093
	Short-Term.	0.2468	0.2363	1.1377%	0.0109
VAT	Long-Term.	0.0899	0.0898	0.433%	0.0022
	Short-Term.	0.0993	0.0990	0.4769%	0.0024

Table 4: To compare the actual and forecasted values

Duration	Long-term	Short-term
Year	**2016**	**2016**
Actual Value (LNTR)	22.6	22.6
Forecasted Value (LNTRF)	22.939	22.852
Difference	**0.339**	**0.252**
Year	2017	2017
Actual Value (LNTR)	22.646	22.646
Forecasted Value (LNTRF)	23.113	22.96
Difference	**0.467**	**0.314**
Year	**2016**	**2016**
Actual Value (LNPAYE)	20.769	20.769
Forecasted Value (LNPAYEF)	21.082	21.031
Difference	**0.313**	**0.262**
Year	2017	2017
Actual Value (LNPAYE)	20.774	20.774
Forecasted Value (LNPAYEF)	21.234	21.186
Difference	**0.46**	**0.412**
Year	2016	2016
Actual Value (LNVAT)	20.715	20.715
Forecasted Value (LNVATF)	20.809	20.894
Difference	**0.094**	**0.179**
Year	2017	2017
Actual Value (LNVAT)	20.808	20.808
Forecasted Value (LNVATF)	20.894	20.972
Difference	**0.086**	**0.164**

Source: Author's Calculation

Appendix 1

OLS regression results for long-run and short run sample size

Table: OLS regression results for long-run

Dependent variable	**TR**	**PAYE**	**VAT**
Independent variable	**long-run**	**long-run**	**long-run**
Constant	-7.2912**	11.498***	-2.5982
	[3.3857]	[0.7547]	[1.8784]
HC	-2.2821**		-0.3899
	[1.0564]		[0.5861]
NGDP	3.1835***		1.1085**
	[0.9598]		[0.5325]
LNUEM		-0.3290	
		[0.3284]	
LNWAGE		3.6893***	
		[0.2177]	
Adjust R^2	0.936865	0.959204	0.956128
F-statistic	134.5527	212.6118	197.1412
Akaike info criterion	-0.213684	-0.569393	-0.96462
Breusch LM Test	4.351460	4.156069	2.717999
Hetero Test	4.334938	4.834098	2.465888

Source: Author's Calculation

OLS regression results for short-run sample size

Dependent Variable	**TR**	**PAYE**	**VAT**
Independent Var.	**short-term**	**short-term**	**short-term**
Constant	-5.6913***	10.764*** [0.4514]	-3.7908***
	[1.1146]		[1.1648]
LNHHFCE	-0.6204		-0.9969**
	[0.3946]		[0.4124]
LNGDPCUR	1.4945***		1.7449***
	[0.3697]		[0.3863]
LNUEM		-0.3181*	
		[0.1771]	
LNWAGE		3.9839***	
		[0.1394]	
Adjust R^2	0.984938	0.986821	0.979425
F-statistic	491.4352	562.5962	358.0267
Akaike info criterion	-2.028952	-1.640997	-1.940791
Breusch LM Test	5.382108	3.915611	5.290881
Hetero Test	2.318965	6.320043	3.467014

Source: Author's Calculation

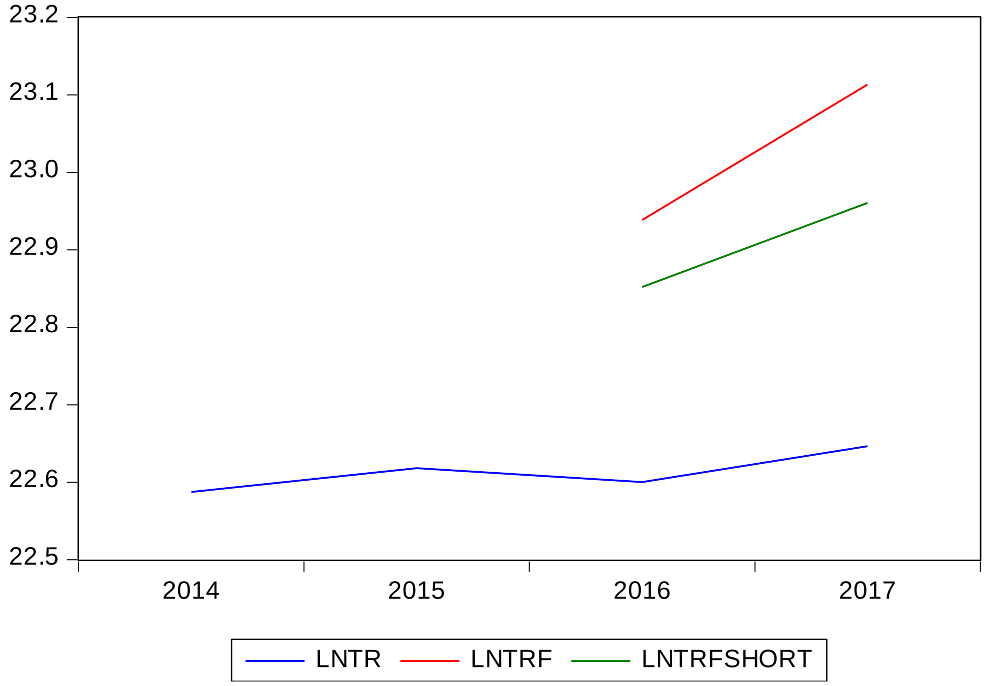
Appendix 2

Actual and forecasted values (in TZS billion) of tax revenue (TR) using OLS method.

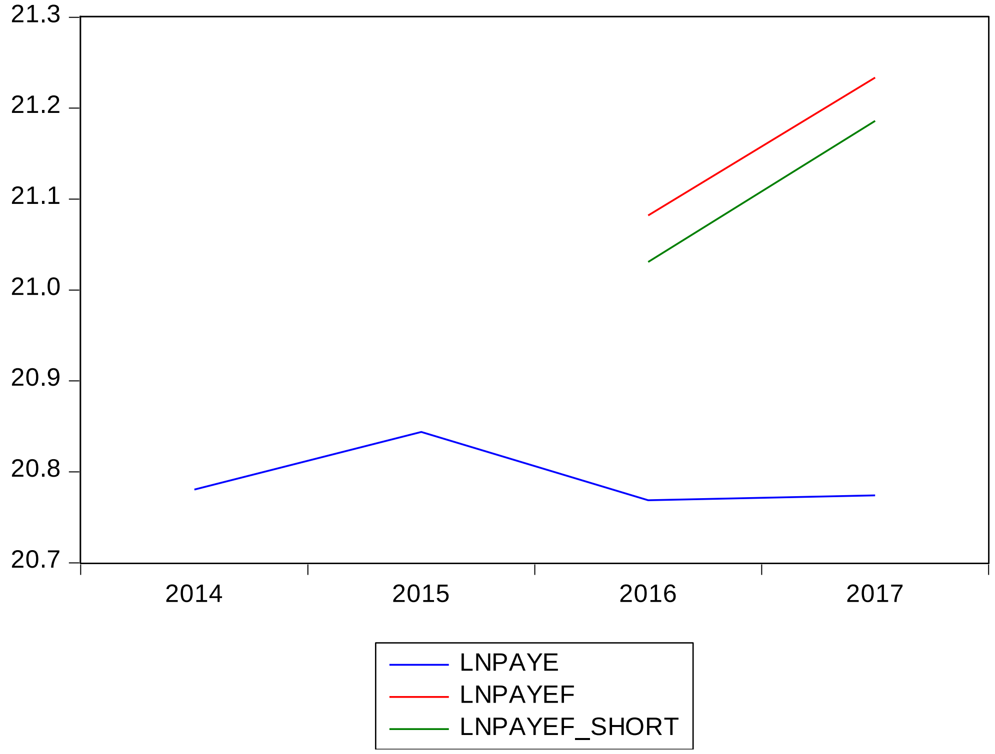
		2016			2017		
		Total Revenue	PAYE	Value-added	Total Revenue	PAYE	Value-added
Long-term	Forecast	9,163.43	1,089.41	1,431.31	10,914.33	1,185.63	1,665.95
	Actual	6,532.54	991.46	1,046.40	6,841.85	1,088.59	1,051.97
Short-term	Forecast	7,608.50	1,234.05	1,087.07	8,425.85	1,430.95	1,210.30
	Actual	6,532.54	991.46	1,046.40	6,841.85	1,088.59	1,051.97

Source: Author's calculation.

Appendix 3: Graphical Presentation of Total Revenue (TR) for long-run and short-run sample



Appendix D: Graphical Presentation of PAYE for long-run and short-run sample



Appendix 4:
Graphical Presentation of VAT for long-run and short-run sample