University of Economics in Katowice

Journal of Economics and Management

ISSN 1732-1948 Vol. 31 (1) • 2018

Mangalani Peter Makananisa

Operational Research (OR)
Tax, Customs and Excise Institute (TCEI)
South African Revenue Service (SARS)
pmakananisa@sars.gov.za

Jean Luc Erero

ORCID ID: 0000-0002-7901-3108

Operational Research (OR)
Tax, Customs and Excise Institute (TCEI)
South African Revenue Service (SARS)
JErero@sars.gov.za

Predicting South African personal income tax – using Holt–Winters and SARIMA

DOI: 10.22367/jem.2018.31.02

Accepted by Editor Ewa Ziemba | Received: October 18, 2017 | Revised: December 13, 2017; January 15, 2018; January 23, 2018 | Accepted: January 26, 2018.

Abstract

Aim/purpose – Over estimation and under estimation of the Personal Income Tax (PIT) revenue results in an unstable economy and unreliable statistics in the public domain. This study aims to find a suitable SARIMA and Holt–Winters model that suits the sample monthly data for PIT well enough, from which a forecast can be generated.

Design/methodology/approach – This study uses the aspects of time series model (Holt–Winters and SARIMA) and regression models with SARIMA errors to simulate the structure which followed the historical actual realization of PIT. The quarterly data were obtained from quarter1, 2009 to quarter 1, 2017 for the purpose of modelling and forecasting. The data were divided into training (quarter 1, 1995 to quarter 1, 2014) and testing (quarter 2, 2014 to quarter 1, 2017) data sets. The forecast from quarter 2, 2017 to quarter 1, 2020 were also derived and aggregated to annual forecast.

Findings – Holt–Winters, SARIMA and Time Series Regression models fitted captured the movement of the historical PIT data with higher precession.

Research implications/limitations – The generated forecast is recommended to avoid several model revisions when locating the actual PIT realisation. However, monitoring of this model is crucial as the prediction power deteriorate in a long run.

Originality/value/contribution – The study recommends the use of these methods for forecasting future PIT payments because they are precise and unbiased when forecasting are made. This will assist the South African authorities in decision making for future PIT revenue.

Keywords: South African Revenue Service (SARS), Personal Income Tax (PIT), Autoregressive Integrated Moving Averages (SARMA), Holt–Winters (HW).
JEL Classification: H24, C15, E37.

1. Introduction

Income tax of interest to this study is a collectible tax from individuals referred to as Personal Income Tax (PIT). Theoretically PIT is expected to grow because earnings and prices normally increase every year, although it may be affected by economic shocks such as recession and drought. For this reason, state authorities with advanced technology will always analyse the internal data against third party data to find out if there was incorrect reporting of information which could lead to a smaller collection of income tax revenue.

PIT plays a significant role in the South African economy, as individuals generally receive most of their income as salary/wages, pension/retirement payments and investment income (interests and dividends). Some individuals may also have a business income which is taxable as a personal income, such as sole proprietors and partners (SARS & National Treasury, 2015).

The reporting or predicting of future revenue collection with some certainty is vital for planning and decision making because it will negatively impact the economy of a country if not done properly. The use of time series models to forecast the continuation of historical data patterns is becoming more popular in the financial world due to their accuracy and reliability. Moreover, this is also applicable to environments such as biology, tourism, and natural climatic phenomena (Hyndman, Makridakis, & Wheelwright, 1998).

In academic circles, there has been a substantial interest given to the application of models such as non-linear modelling methods for analysing tax revenue over times. Tsay (2002) pointed out that non-linear models were evolved to bilinear models by Granger & Andersen (1978), threshold autoregressive model by Tong & Lim (1980) and Markov switching model by Hamilton (1989). The particularity of Markov switching type models is that variables related to taxpayer behaviour are determined exogenously by the use of an unobservable state variable, while in the threshold model they are determined endogenously. Due to policy evaluation, an acknowledgement of the advantages of threshold models currently exists in comparison with the Markov switching type models. Consistent and comparative analysis of tax policy scenarios are made possible by threshold models mostly because of behavioural test used to explain the asymmetric outcome in observed data. As described by Tsay (2002), the non-linear models require ceiling space to attain a reliable linear estimate of the conditional mean equation as opposed to the traditional piecewise linear model. Furthermore, African Tax Institute (2012) emphasises on equipping revenue forecasting analyst with necessary mathematical and computer skills to master revenue forecasting techniques under the learning courses revenue forecast and tax analysis.

There is little national literature on modelling of tax revenue, specifically looking at the personal income tax of South Africa. Hence, this independent study will focus on modelling and forecasting the South African PIT using time series model. This will assist the South African authorities in deriving the forecast which are more precise and unbiased. Therefore, it will improve the decision making for future PIT revenue management.

The national literature includes the work of Van Heerden (2013) focusing on the PIT structure and its implication on individual tax revenue in South Africa. She used a micro-simulation tax model to determine what can be done to optimise individual taxes so that they can minimise the negative burden of taxation on the performance of the economy. The author found that lowering tax rates stimulates potential for improved levels of efficiency with tax burden.

Boonzaaier (2015) tested the possible tax revenue asymmetries relative to the business cycle via a smooth transition ARMA framework. Its model allowed for the estimation of two separate regimes, i.e. a low and high growth regime where the movement in tax revenue at all times should be governed by a weighted average of two different linear models. The findings indicated that tax revenue collections do react differently depending on whether it is in a high or a low growth phase impacting revenue forecasting process and calculation of cyclically adjusted budget balance. Conversely, the study did not include the out of sample forecast given the impact of the cyclical changes (higher or lower grow phases).

Over estimation and under estimation of tax revenue results in an unstable economy and unreliable statistics in the public domain. The burning question might be which types of model(s) can better predict/forecast variable(s) of interest. Every model has some advantages and disadvantages. Nonetheless, they depend on the use of the models because explanatory models are good for sensitivity analysis while time series are limited for sensitivity analysis. In fact, the later minimises biasness when forecast are made.

The aim of this paper is to find a suitable SARIMA and Holt–Winters model that suits the sample monthly data for PIT well enough, from which a forecast can be generated. The R-statistical software was used for modelling and forecasting purposes. In this respect, we attempted to forecast South Africa's PIT using the mentioned methods because they have not been explicitly applied in a South African context for tax forecasting and publication purposes. Section 2 presents the literature review. Section 3 described the PIT historical contribution to total tax and the methodology is presented in section 4. Section 5 shows the models application and results on PIT, and section 6 discusses model results and limitations. Conclusions and recommendations are provided in sections 7.

2. Literature review

Various literatures around the world support the use of time series models. Some of the literatures on the ARIMA/SARIMA and Holt–Winters time series models include the work of Jayesekara & Passty (2009). These authors used ARIMA models, including the dummy variables for seasonal adjustment, as a univariate benchmark model to forecast the net income tax revenue in Cincinnati City. The monthly data were obtained from the Cincinnati Income Tax Division (CITD) for the period January 1970 to April 2009, although the data used to carry out the estimation were reduced to start from January 1989 due to changes in tax rates. In order to reach the two final ARIMA models, the best fit models were selected based on the model with minimal Akaike Information Criterion (AIC), the minimum Root Mean Squared Error (RMSE) and the model with the highest R-Squared (R^2).

The ARIMA models fitted predicted the Cincinnati net income tax well, captured the seasonality in the data throughout the sample used, and the within sample estimates were compared with the actuals for the period 2006 to 2009. Furthermore, the ARIMA model was considered to forecast the net income tax starting from January 2008 (a portion of the in-sample) to verify the effectiveness of the model. Moreover, data were converted to bi-monthly in order to construct a bi-monthly model. In this respect, Jayesekara & Passty (2009) recommended the use of ARIMA for the CITD for short-term forecasts of net income tax.

Similarly Chatagny & Soguel (2009) used the ARIMA model to estimate tax revenues (an amalgamation of PIT and Corporate Income Tax (CIT)) for all 28 cantons or districts in Switzerland. The main aim of the study was to prove that forecast bias can be reduced by using univariate time series models. Tax revenue data for the period 1944 to 2006, together with the observed official forecasts were obtained from the districts. In addition, the time series data were divided into two samples (1944 to 2006 and 1976 to 2006) due to some districts not having recorded historical data in certain years. To assess the ARIMA performance against the observed forecasts for the two sample periods, the mean percentage error was used to classify the over, under and zero error percentage on district. The results from the mean percentage error showed that the observed forecasts under estimated tax revenue in almost all cantons and ARIMA models had a Mean Percentage Error (MPE) close to zero in the two sample periods. The study concluded that bias from an observed forecast can still be reduced by using simple univariate models, ARIMA.

Pelinescu, Anton, Ionescu, & Tasca (2010) analysed the Romanian local budget with the aim of assisting officials to create efficient plans and manage local income and expenditure using a good strategic management tool. This arose as a result of the local authorities finding it difficult to predict future revenues to construct their annual budgets. Using the historical data from the first quarter of 2000 to the third quarter of 2010, the authors applied the Holt–Winters multiplicative and additive models to forecast total local revenue and own revenue of local authorities. The E-views software was used to build and run the Holt–Winters equations and to select a model that minimised the Root Mean Squared Error (RMSE). The study recommended the use of a Holt–Winters model as a tool for multi-annual budget forecasting because it is user-friendly and provides stable forecasts.

A similar study was conducted in Romanian by Brojba, Dumitru, & Belciug (2010) who used ARIMA models on monthly earning data for the period 2007 to 2008 (the economic crisis period) to model the total budget revenue. The ARIMA models captured the data movement during the economic crisis because the data contained or showed the trend and seasonality. The fitted values were close to the actuals and the study concluded that ARIMA models can be used to set targets and sound future developments. Nonetheless, the models have its limitations as the parameters are sensitive to sample selection, with the most accurate forecasts being for the short-term.

Forecasting future revenue with maximum precision is important for country's economy, as it leads to a better overall distribution of future budgets. From the above literature review it can be seen that time series models have proven to be useful methods for forecasting tax revenues. These types of methods are more precise for short-term forecasts (two to three years) because their precision declines over a longer period.

More knowledge on the variables of interest must be obtained when generating long-term forecasts because they are precise, and the model must be well defined, and some statistic must be considered. These statistics are such as root mean squared error, mean absolute percentage error, Akaike information criterion, and many others. The following section summaries the SARIMA and Holt–Winters models theories and assumptions.

3. PIT historical contribution to total tax

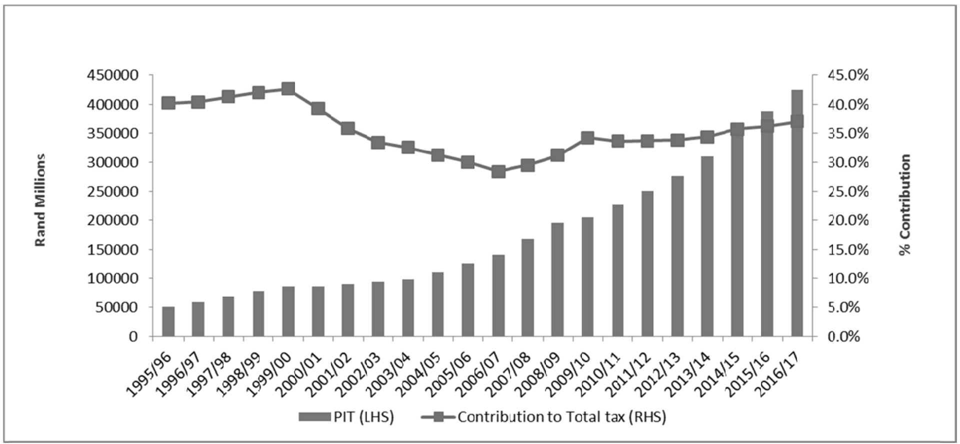
South African PIT is a tax levied on the taxable income (gross income less exemptions and allowable deductions) of individuals and trusts. It is determined for a specific year of assessment. Taxable capital gains form part of taxable income. Most individuals receive their income as salaries or wages, pension or annuity payments, and/or investment income (interest and dividends). Some individuals, such as sole proprietors and partners, may also have business income which is taxable as personal income.

This main source of revenue (PIT) is made up of sub sources such as employee's tax (or pay as you earn), provisional tax and assessed tax with employees tax contributing around 95% and the other two sub sources sharing the remaining 5% of PIT. The employee's tax is collectable by employer on behalf of employee, hence provisional tax payable by any person who derives income other than remunerations, an allowance or advance and assessed tax paid on final assessments of the tax return (SARS & National Treasury, 2017).

PIT remain the largest source of the South African revenue with its contribution to total tax of 40.2% in 1995/96 hence its relative contribution has declined to 37.2% in 2016/17 SARS fiscal year (Figure 1). This tax is levied per tax brackets from which the tax payer belongs to with regards to their generated income per specific period (SARS & National Treasury, 2017).

Figure 1 shows the contribution of this main source of revenue (PIT) to the total tax revenue for the period 1995/96 to 2016/17 fiscal years.

Figure 1. Personal income tax and its contribution to total tax



Source: SARS & National Treasury (2015-2017).

4. Research methodology

The Time Series Regression methodology is seen as the basis for modelling and forecasting the quantitative variable(s) of interest as their forecast are unbiased. Thus everything that is needed to project the continuation of the historical pattern is already included in the history of the variable. Here, more weight is given to the recent data in this study. The three popular Time Series Regression methods – Holt–Winters Model, SARIMA Model and the Time Series Regression with ARIMA error were used to predict South Africa's PIT revenues.

4.1. Holt–Winters model

Holt–Winters methods are an extension of simple exponential smoothing and Holt's trend corrected exponential smoothing method. Simple exponential smoothing is used to forecast series when there is no trend or seasonal pattern, while Holt's trend corrected exponential smoothing is applicable when a time series displays a changing level (mean) and the growth rate (slope) for the trend. However, work on state space models with a single source by Hyndman, Koehlor, Synder, & Grose (2002) provided a statistical framework for the exponential smoothing methods. Holt–Winters methods are designed for a time series that exhibits a linear trend and seasonality. These methods include the: additive Holt–Winters model and multiplicative Holt–Winters model. However, this study will concentrate on the additive Holt–Winters model.

Let $y_t = y_1, y_2, y_3,..., y_T$ be the time series of interest where $t = 1, 2, 3,..., T$. The additive Holt–Winters models is appropriate when a time series y_t has a linear trend with an additive seasonal pattern for which the level (mean), the growth rate and the seasonal pattern may be changing (Hyndman, Koehlor, Synder, & Grose, 2002). This model can be described as:

$$y_t = (l_t + tb_t) + s_t + e_t \tag{4.1}$$

Where l_t is the mean (level value), b_t the growth rate, and s_t is the fixed seasonal pattern at time t. The additive Holt–Winters method can be summarised as follows.

Suppose that the time series $y_t = y_1, y_2, y_3,..., y_T$ exhibits linear trend locally and has a seasonal pattern with constant (additive) seasonal variation and the level, growth rate and seasonal pattern may be changing. Then, the estimate l_t for the level, the estimate b_t for the growth rate and the estimate s_t for the seasonal factor of the time series in time period t are given by the smoothing equation:

$$l_t = \alpha(y_t - s_{t-1}) + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{4.2}$$

$$b_t = \beta(l_t - l_{t-1}) + (1-\beta)b_{t-1} \tag{4.3}$$

$$l_t = \alpha(y_t - s_{t-1}) + (1-\alpha)(l_{t-1} + b_{t-1}) \tag{4.4}$$

Where α, β and γ are smoothing constants between 0 and 1, l_{t-1} and b_{t-1} are estimates in time period $t-1$ for the level and growth rate respectively, and s_{t-1} is the estimate in $t-1$ for the seasonal factor. The application of this method requires that a time series have a linear trend with an additive seasonal pattern.

The Additive Holt-Winters requires initial values for l_t, b_t and s_t which are estimated as follows:

$$l_0 = \frac{1}{s}(y_1 + y_2 + \ldots + y_s), \tag{4.5}$$

$$b_0 = \frac{1}{s}\left[\frac{y_{s+1} - y_1}{s} + \frac{y_{s+2} - y_2}{s} + \ldots + \frac{y_{s+s} - y_s}{s}\right] \tag{4.6}$$

$$s_1 = y_1 - l_0, \tag{4.7}$$

$$s_2 = y_2 - l_0 \tag{4.8}$$

$$\text{and } s_s = y_s - l_0 \tag{4.9}$$

where smaller s is the number of seasons.

The forecast at time period T + $\wp$ will be given by:

$$y_{T-\wp} = l_T + \wp b_T + s_T, \tag{4.10}$$

Where l_T the smoothed estimate of the level is at time *T*, b_t is the smoothed estimate of the growth at time *T* and s_T is the smoothed estimate of the seasonal component at time *T*.

4.2. SARIMA model

The SARIMA model relates the current time series observation to its historical seasonal occurrence; this is called model fitting. This means that the series to be forecast is generated by a random process with a structure that can be described. The description is given in terms of the randomness of the process rather than the cause and effect used in regression models (Pindyck & Rubinfeld, 2010). For this type of model the data must be stationary around the mean and the variance.

Let $y_t = y_1, y_2, y_3, \ldots y_T$, be the time series of interest where *t* = 1, 2,..., *T*, and $w_t = \nabla_s^D \nabla^d y_t$ be the stationary variable derived from y_t, then the general SARIMA equation is represented by equation 4.11 (Shumway & Stoffer, 2006).

$$\Phi_P(B^s)\phi_p(B)w_t = \delta + \Theta_Q(B^s)\theta_q(B)\eta_t \quad (4.11)$$

Where $w_t = (1-B^s)^D(1-B^d)y_t = (1-B^d - B^{sD} + B^{sD+d})y_t$ are the product of seasonal differencing D and non-seasonal differencing d, s is the series seasonality which takes the value 4 for quarterly time series data and 12 for monthly time series data, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^p$ are the non-seasonal AR components of order p, $\Phi_P(B^s) = 1 - \Phi_1 B^s - \Phi_2 B^{2s} - \ldots - \phi_P B^{Ps}$ are the seasonal AR components of order P, $\theta_q(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q$ are the non-seasonal MA components of order q, $\Theta_Q(B^s) = 1 + \Theta_1 B^s + \Theta_2 B^{2s} + \ldots + \Theta_Q B^{Qs}$ are the seasonal MA components of order Q, δ is the constant term and η_t the disturbance or error term at time t, and B is a back shift operator with $B^i(w_t) = w_{t-i}$ and $B^j(\eta_t) = \eta_{t-j}$.

The Autocorrelation Function (ACF) which is represented by ($\widehat{\rho}_k$) at lag k provides a partial description of the process for modelling purposes (Pindyck, & Rubinfeld, 1998). This is useful for the selection of the autoregressive and moving averages components and is given by:

$$\widehat{\rho}_k = \frac{\gamma_k}{\gamma_0} = \frac{\sum_{t=1}^{T-K}(w_t - \overline{w})(w_{t+k} - \overline{w})}{\sum_{t=1}^{T}(w_t - \overline{w})^2} \quad (4.12)$$

A better way to check model adequacy is to analyse the residuals of the series obtained from the model. If the model is correctly specified and the parameters are reasonably close to the true value, the residuals should have nearly the properties of white noise. This means that they should behave more or less like independent, identically distributed normal variables, with zero mean and common variance. Hence, the residuals will be stationary in both the mean and variance. The most popular residual diagnostic to check for the models adequacy is the Ljung–Box (LB) statistic, which tests for the joint hypothesis that all the $\widehat{\rho}_k$ (autocorrelation) up to a certain lag are equal to zero. The LB test statistic in equation 4.13 follows a chi-square (χ^2) distribution with m degrees of freedom for larger sample n.

$$LB = n(n+2)\sum_{k=1}^{m}\left(\frac{\widehat{\rho}_k^2}{n-k}\right) \quad (4.13)$$

4.3. Time series regression with SARIMA

The regression model with SARIMA error is simply the expansion of a linear regression model. Since variables are recorded over time interval, the dependent and independent variable(s) should be stationary for the simplification of the model fitting. The linear regression will take the following form:

$$\nabla^d y_t = \beta_0 + \sum_{i=1}^{l} \beta_i (\nabla^d x_{it}) + e_t \qquad (4.14)$$

Where ∇^d is the stationary differencing of order d, y_t is the variable of interest at time t, x_{it} is the i^{th} explanatory variable at time t, at the error term at time t, β_0 and β_i represent the constant term and the coefficient of the i^{th} explanatory variable, respectively. The error term from the model fitted assume that there is autocorrelation within the current et and the previous errors $(e_{t-1}, e_{t-2}, e_{t-3}, ...)$ and the ARM A model can be fitted and be mathematically represented as follows (Hyndman, Makridakis, & Wheelwright, 1998):

$$e_t = \frac{\Theta_Q(B^s)\theta_q(B)}{\Phi_P(B^s)\phi_p(B)}\eta_t = \Phi^{-1}{}_P(B^s)\phi^{-1}{}_p(B)\Theta_Q(B^s)\theta_q(B)\eta_t \qquad (4.15)$$

When the assumption hold, the model is now called Time Series Regression with SARIMA error and can be represented as follows:

$$\nabla^d y_t = \beta_0 + \sum_{i=1}^{l} \beta_i (\nabla^d x_{it}) + \Phi^{-1}{}_P(B^s)\phi^{-1}{}_p(B)\Theta_Q(B^s)\theta_q(B)\eta_t \qquad (4.16)$$

The SARIMA fitted on the error tern e_t is naturally expected to further reduce the unexplained part of the variation in e_t to η_t, thus $\eta_t < e_t$ resting in more accurate fitted/predicted values.

4.4. Some measure of accuracy

For a given time series, there may be several competing models for forecasting. According to Wei (2006), the models can be compared for goodness of forecasting using the criteria described below. Fit the models to the $t-1(0 < l \leq t)$ observations of the time series and use the fitted models to forecast the last l observed values of the series, we can calculate:

$$e_t(t-l+j) = y_{t-l+j} - \hat{y}_{t-l+j},\ j = 1,2,...,l \tag{4.17}$$

Where for $j = 1,2,...,l$, $\hat{y}_{t-l+j}$ is the forecast for y_{t-l+j} using any competing models; and compute:

$$PE = \frac{e_t(t-l+j)}{y_{t-l+j}} * 100\ , \tag{4.18}$$

$$MPE = \frac{1}{l}\sum_{j=1}^{l} \frac{e_t(t-l+j)}{y_{t-l+j}} * 100\ , \tag{4.19}$$

$$MAPE = \frac{1}{l}\sum_{j=1}^{l} \left| \frac{e_t(t-l+j)}{y_{t-l+j}} \right| * 100\ , \tag{4.20}$$

$$MSE = \frac{1}{l}\sum_{j=1}^{l} e_t^2(t-l+j)\ , \tag{4.21}$$

$$MAE = \frac{1}{l}\sum_{j=1}^{l} |\ e_t\ (t-l+j)\ | \tag{4.22}$$

The best model within the method for forecasting is the one with the smallest Mean Percentage Error (MPE), Mean Square Error (ME), Mean Absolute Error (MAE) or Mean Absolute Percentage Error (MAPE) depending on the criterion/criteria which one chooses to use. The other popular criteria used to select the best fitting model(s) is the R^2 which can be represented by equation 4.23, although R^2 is commonly used for explanatory models.

$$R^2 = \frac{\sum_{i=1}^{T} (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{T} (y_i - \bar{y})^2} \tag{4.23}$$

Two commonly used criteria for choosing the best model, according to Maindonald & Braun (2003) are the Akaike's Information Criterion (AIC) and the Bayesian Information Criterion (BIC).

The AIC is defined by the following equation.

$$AIC = -2\ln(Likelihood) + 2r \tag{4.24}$$

The AIC increases with the number of model parameters (r), and the best model has the smallest AIC. The BIC is an extension of the AIC and is given by:

$$BIC = -2\ln(Likelihood) + r\ln(T) \tag{4.25}$$

Where T is the number of observations in y_t. As with the AIC, the best model among competing models within the method has the smallest BIC.

5. Model application and results on PIT

The initial sample data from quarter 1, 1995 to quarter 1, 2017 in section 4 was divided into two samples, quarter 1, 1995 to quarter 1, 2014 and quarter 2, 2014 to quarter 1, 2017 for model training (fitting) and model testing (prediction) purposes, respectively. Moreover, the models out of sample forecast were generated for the period quarter 2, 2017 to quarter 1, 2020.

We can recall that the purpose of the study is to model and generate annual or Fiscal Year (FY) forecast for PIT. Normally, the FY span for SARS includes a fixed period starting from April to March for each year. This implies that the quarterly fitted/predicted values will then be aggregated to form FY fitted/predictions and be compared with the PIT actual for the corresponding FY which apply to the out of sample forecast. The following section summarises the additive Holt–Winters model fitted and the results thereof for personal income tax.

5.1. PIT additive Holt–Winters model

Unlike SARIMA models, Holt-Winters models look at a time series data of interest, separated into three components which are the level value, trend and seasonal. This model gives each component's weight on an interval of zero to one, which should be able to fit the model and forecast the future values. The PIT time series has a gradual increasing trend and an additive seasonality predictable over time. This implies that an additive Holt–Winters requires data to have a linear trend with an additive seasonal pattern suiting the PIT time series data. The Additive HW model is presented in Table 1.

Table 1. PIT additive Holt–Winters model coefficients

Smoothing Constant	Coefficient
alpha (α)	0.2667
beta (β)	0.0865
gamma (γ)	0.7333
Level Mean (β_0)	9.3401
Growth/Trend (β_1)	0.0305

Source: Authors' computation.

Table 1 represents the smoothing constants, level mean and growth/trend estimation for the additive Holt–Winters model in equation (5.1) fitted to PIT time series.

$$\ln(y_t) = (9.340 + 0.0305t) + s_t + e_t \tag{5.1}$$

Equation (5.1) is the PIT additive Holt–Winters model from Table 1. The estimate mean l_T for the level mean, the estimate b_T for growth/trend rate and the estimate s_T for the seasonal factor of the data in time period T are given by the smoothing equation:

$$\ln(y_t) = (9.340 + 0.0305t) + s_t + e_t \tag{5.2}$$

$$b_T = 0.0865(l_T - l_{T-1}) + (1 - 0.0865)b_{T-1} \tag{5.3}$$

$$S_T = 0.7333(y_T - l_T) + (1 - 0.7333)s_{T-1} \tag{5.4}$$

Where α = 0.168, β = 0.083 and γ = 0.479 are smoothing constants ranging between 0 and 1. The l_{T-1} and b_{T-1} are estimates in time period T − 1 for the level mean and growth rate, respectively, and S_{T-1} is the seasonal factor. More weight is given to the seasonal component of the model, verifying that indeed PIT is a seasonal time series data. The initial seasonal state for this quarterly data was computed as presented in Table 2.

Table 2. Initial values of the seasonal components

S_1	S_1	S_1	S_1
0.005	0.0299	−0.1547	0.1198

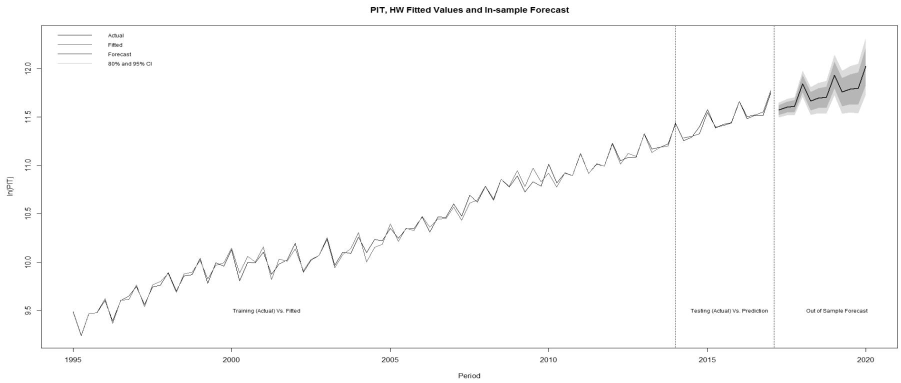
Source: Authors' computation.

5.2. PIT Holt–Winters model performance and forecast

This section focuses on the Holt–Winters model performance on the in-sample data (training and testing). The model good fit on the sample data enables the model to be used for forecasting purposes with some level of accuracy. Depending on the tolerance level or error range, the model can be regarded as good. The rule of thumb error preference is the 5% error rate; however, other may be comfortable with the maximum error of 10%.

The additive Holt–Winters model fitted on the training data (quarter 1, 1995 to quarter 1, 2014) indicates that the overall standard error was 0.0038, a good fit on the training data. The fitted values in Figure 2 follow or simulate the movement of the observed actual.

Figure 2. PIT Actual, Holt–Winters in-sample fitted and forecast



Source: SARS & National Treasury (2015, 2016 & 2017) and authors' computation.

The measure of accuracy on the training data is included in Table 3.

Table 3. Additive Holt–Winters measure of accuracy

	ME	RMSE	MAE	MPE	MASE
Training	−0.00047	0.03927	0.029054	0.28186	0.28411

Source: Authors' computation.

The additive Holt-Winters model was tested on the sample 2014 and quarter 2, 2014 to quarter 1, 2017. The predicted values follow or capture the movement in the actual values of ln (PIT), as shown in Figure 2. The predicted values were then aggregated to be in line with SARS fiscal year and be compared with the actual for the testing sample period as shown in Table 4.

Table 4. PIT testing data and additive Holt–Winters predicted values in rand millions

FY	Testing Data	Predicted Values	PE
2014/15	352,950	346,565	1.8%
2015/16	388,101	387,370	0.2%
2016/17	424,545	433,001	−2.0%

Source: SARS & National Treasury (2016) and authors' computation.

The percentage error on the three years 2014/15 to 2016/17 (12 quarters) in Table 4 clearly shows the model performance with an error of less than 5% on aggregate which is another indication that this model can be used for forecasting future PIT values.

The model out of sample forecast were generated for the period quarter 2, 2017 to quarter 1, 2020 with the 80% and 95% confidence intervals as shown in Figure 1. Table 5 shows the out of sample three years quarterly forecast.

Table 5. Cumulative PIT and Holt–Winters forecast in rand millions

Quarter	2017/18	2018/19	2019/20
Q01	106,049	116,265	127,465
Q02	215,309	236,050	258,789
Q03	325,510	356,867	391,245
Q04	464,299	509,026	558,062

Source: Authors' computation.

The quarterly forecast amounts to R464,3 billion, R509,0 billion and R558,1 billion for fiscal year 2017/18, 2018/19 and 2019/20, respectively. This represents an average growth rate between 9% and 10% for the three years forecast interval. It is a consistent growth rate when compared to the historical growth from 2010/11 fiscal year. The consistence growth due to that PIT is dependent on wages and salaries which normally increase once in a year.

5.3. PIT SARIMA model

As discussed in subsection 3.2, in order to fit the SARIMA model the data need to be stationary on the mean with a constant variance. The PIT original series was identified not to be stationary as it changes over time. The normal and seasonal difference was done on the PIT time series presented in section 4. The SARIMA model fitted on PIT using the maximum likelihood method was $SARIMA\ (0,1,1)(0,1,0)_4$ and can be mathematically represented as follows:

$$w_t = (1 - \theta_1 B)\eta_t \tag{5.5}$$

Where $w_t = (1-B)(1-B^4)y_t$, $y_t = PIT$, θ_1 is the first moving average (MA(1)) coefficient, B is a back shift operator with B is a back shift operator with $B^i(w_t) = w_{t-i}$ and $B^j(\eta_t) = \eta_{t-j}$, and η_t being an error term at time t. However, equation 5.5 can also be represented as equation 5.6.

$$y_t = y_{t-1} + y_{t-4} + y_{t-5} + \eta_t + \theta_1 \eta_{t-1} \tag{5.6}$$

Table 6 presents the maximum likelihood parameter estimation for the SARMA model from equation (5.5) fitted to the transformed PIT time series.

Table 6. PIT SARIMA model parameter estimated

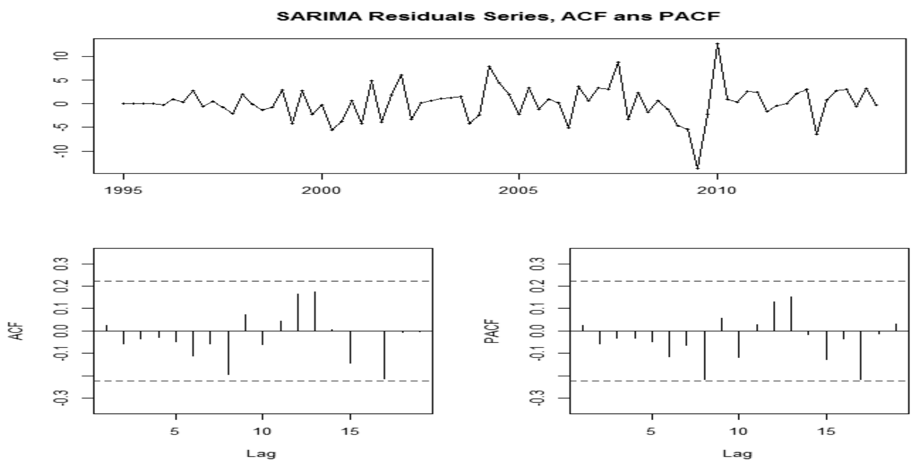
	η_{t-1}
θ_1	−0.00047
s.e	0.1263
t-ratio	−3.3512
p-value	0.0008

Source: Authors' computation.

The first moving averages (MA(1) or η_t) fitted on w_t PIT transformed data with the coefficient $\theta_1 = -0.4233$, t-ratio of −3.3512 and the standard error of 0.1263 was found to have a p-value significant at less than 1% level of significance (0.0008).

The SARIMA model fitted will be used to forecast future values if the model residual are white noise. For this reason this study examined the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for a white noise residual as illustrated in Figure 3.

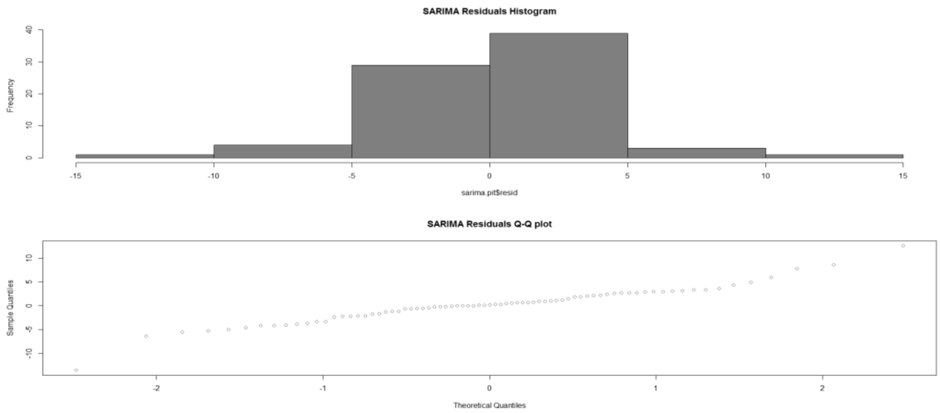
Figure 3. PIT SARIMA model residuals plot



Source: Authors' computation.

Figure 3 clearly shows that the model has the residuals which are highly independent from one another. The residuals ACF and PACF SARIMA are within the 95% CI boundaries, where T is the number of series observations. Based on the residual's 95% confidence intervals, the residuals are assumed to be not far from the zero line, that is, they come from a well-defined model. The model, the histogram and the q-q plot in Figure 4 further confirm the independence of the residuals, as the two plots showed the distribution of the residual overtime, with the most of the values centered around zero mean, though there was some skewness.

Figure 4. PIT SARIMA residuals histogram and Q-Q plot



Source: Authors' computation.

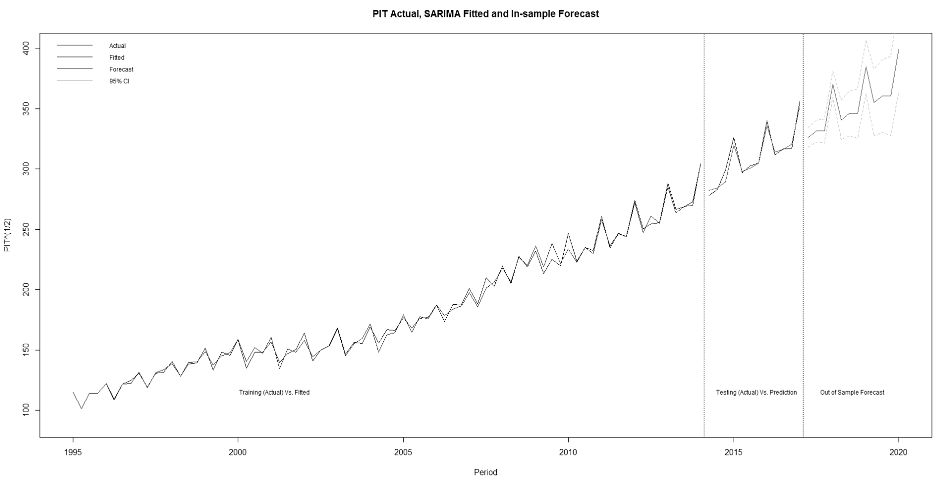
The more general statistics to verify the residual white noise would be the Ljung–Box (LB) test defined by equation 4.13 in subsection 4.2, which is calculated on the sampled residuals of the model fitted. The Ljung–Box test with a chi-squared of 21.868 from 20 degrees of freedom gave a p-value of 0.3477 was obtained from the SARIMA model. This shows that the residuals are independent or uncorrelated and assumed to be emanating from a well specified model and for this reason the model will be used for forecasting the continuation of the PIT historical patterns.

5.4. PIT SARIMA model performance and forecast

This section applied the SARIMA model to analyse the performance of the model training and testing data. Figure 5 present the in-sample fitted values and the model out of sample forecast from the SARIMA fitted model.

The SARIMA model fitted on the training data quarter 1, 1995 to quarter 1, 2014 performed exceptionally well with the fitted values mimicking the actual (Figure 5). The measures of accuracy on the training data are shown in Table 7.

Figure 5. PIT actual, SARIMA in-sample fitted and forecast



Source: SARS & National Treasury (2015-2017) and authors' computation.

Table 7. SARIMA measure of accuracy

	ME	RMSE	MAE	MPE	MASE
Training	164.6617	1505.243	996.2337	0.2616165	2.815479

Source: Authors' computation.

The SARIMA model was tested on the sample 2014 and quarter 2, 2014 to quarter 1, 2017. The predicted values follow or capture the movement in the actual values with minimal error observed as shown in Figure 4. The predicted values were then aggregated to be in line with SARS fiscal year and be compared with the actual for the testing sample period as shown in Table 8.

Table 8. PIT testing data and SARIMA predicted values in rand millions

FY	Testing Data	Predicted Values	PE
2014/15	352,950	345,890	2.0%
2015/16	388,101	381,850	1.6%
2016/17	424,545	417,809	1.6%

Source: SARS & National Treasury (2016) and authors' computation.

The percentage error on the three years 2014/15 to 2016/17 (12 quarters) are shown in Table 8. It clearly shows the model performance with an error of less than 5%. The models out of sample forecast were generated for the period quarter 2, 2017 to quarter 1, 2020 with the 95% confidence intervals as shown in Figure 5. Table 9 shows the out of sample three years quarterly forecast.

Table 9. Cumulative PIT and SARIMA forecast in rand millions

Quarter	2017/18	2018/19	2019/20
Q01	106,561	116,072	125,582
Q02	216,489	235,510	254,532
Q03	326,522	355,055	383,587
Q04	462,588	500,631	538,674

Source: SARS & National Treasury (2016) and authors' computation.

The quarterly forecast amounts to R462,7 billion, R500,6 billion and R538,1 billion for fiscal year 2017/18, 2018/19 and 2019/20, respectively. The annual forecasts from the SARIMA model in Table 9 do not differ much with the Holt–Winters forecast in Table 5, signifying that the actual realization will be around the generated forecast.

5.5. Time series regression model with SARIMA errors

The regression model with SARIMA error was defined as an expansion of a linear regression model in subsection 4.3. This model relates dependent variable to time and the effect of the explanatory variable(s) included in the model. This study fit the time series regression (TS-regression) which relates PIT to Total Compensation of Employees ($Comp_t$). The mathematical representation of this model is shown in equation 5.9.

$$\nabla^d y_t = \beta_1 \nabla^d Comp_t + e_t \tag{5.7}$$

$$e_t = (1 - B^4)^{-1}(1 + \theta_1 B)\eta_t \tag{5.8}$$

$$\nabla^d y_t = \beta_1 \nabla^d Comp_t + (1 - B^4)^{-1}(1 + \theta_1 B)\eta_t \tag{5.9}$$

where ∇^d is normal differencing of the data, in this case d = 1 and the rest of the notation already defined in the current and previous sections. The $\text{SARIMA}(0,0,1)(0,1,1)_4$ was fitted to the error term in equation 5.9 which reduces the model error e_t to η_t. The model's parameter estimates, standard error, t- ratio and p-value are shown in Table 10.

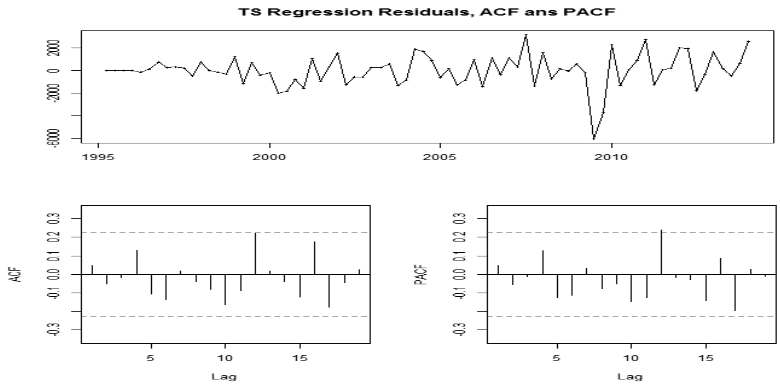
Table 10. TS-regression model parameter estimated

	θ_1	β_1
Coef	−0.719	0.236
s.e	0.120	0.053
t-ratio	−6.011	4.456
p-value	1.85e-09	8.35e-06

Source: Authors' computation.

The two estimated parameters are significant at less than 1% level of signific-ance. The model will be used for forecasting, thus the study need to examine the normality of the residuals. If the residuals are normal the assumption will be that the fitted values are also normal, hence the model can be used for forecasting purposes. Figure 6 shows the residuals from the model in equation 5.9.

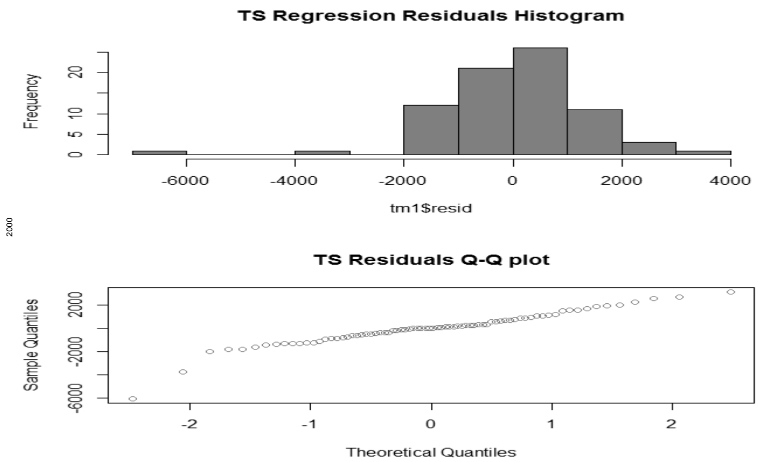
Figure 6. Residuals from TS-regression model



Source: Authors' computation.

Figure 6 presents the model residuals that are white noise, with their ACF and PACF falling within their 95% confidence interval or boundaries. The p-value of 0.283 was obtained from the Ljung–Box test for normal distributed residuals. The residual white noise was also shown by the histogram and the Q-Q plot in Figure 7.

Figure 7. Residuals histogram and Q-Q plot SARIMA TS-regression



Source: Authors' computation.

However, to proceed with out of sample forecast for PIT, the explanatory variable forecast need to be obtained first. The $SARIMA(2,1,2)(1,1,1)_4$ was fitted to the square root 1 transformed Compensation of Employees (Comp 2) and the out of sample forecast were generated. The Compensation SARIMA model is represented mathematical in equation 5.10.

$$w_t = \frac{(1+\theta_1 B+\theta_2 B^2)(1+\Theta B^s)}{(1-\phi_1 B-\phi_2 B^2)(1-\Phi B^s)}\varepsilon_t \qquad (5.10)$$

Where d and D are normal and seasonal differencing respectively, s = 4 for quarterly data used, thus $w_t = \nabla^d \nabla_s^D Comp_t^{\frac{1}{2}} = (1-B)(1-B^4)Comp_t^{\frac{1}{2}}$ and ε_t is the error term at time t. The model's parameter estimates, standard error, t-ratio and p-values are shown in Table 11.

Table 11. TS-regression model parameters estimated

	ϕ_1	ϕ_2	θ_1	θ_2	Φ_1	Θ_2
Coef	1.538	−0.863	− 1.608	0.780	− 0.676	0.834
s.e	0.092	0.073	0.133	0.128	0.193	0.148
t-ratio	16.760	− 11.872	− 12.115	6.093	− 3.506	5.651
p-value	4.82e-63	1.66e-32	8.83e-34	1.11e-09	4.55e-04	1.59e-08

Source: Authors' computation.

The estimated parameters are significant at less than 1% level of significance. The model will be used for forecasting. Subsection 4.6 presents the Time Series Regression with SARIMA error performance on the in-sample data and also shows the out of sample forecast.

5.6. Time series regression model performance and forecast

The time series regression with SARIMA errors built on the PIT training data quarter 1, 1995 to quarter 1, 2014 derived the fitted values that follows the actual realisation of PIT (Figure 6). The measure of determination R^2 was computed to be around 0.883 for the model fitted; this implies that 88% of the variation on the transformed PIT was explained by this model. Some measures of accuracy on the training data set are shown in Table 12.

Table 12. Measure of accuracy from TS-regression

	ME	RMSE	MAE	MPE	MASE	MASE
Training	0.0238	1375.01	962.7847	10.8045	59.6951	0.8088

Source: Authors' computation.

Moving out of the training data sample the model was tested on the sample 2014 and quarter 2, 2014 to quarter 1, 2017. The predicted values were then aggregated to form SARS fiscal years forecast and were compared with the actual for the testing sample period as shown in Table 13.

Table 13. PIT testing data and TS-regression predicted values in rand millions

FY	Testing data	Predicted values	Percentage Error
2014/15	352,950	343,200	2.8%
2015/16	388,101	374,672	3.5%
2016/17	424,545	407,153	4.1%

Source: SARS & National Treasury (2016) and authors' computation.

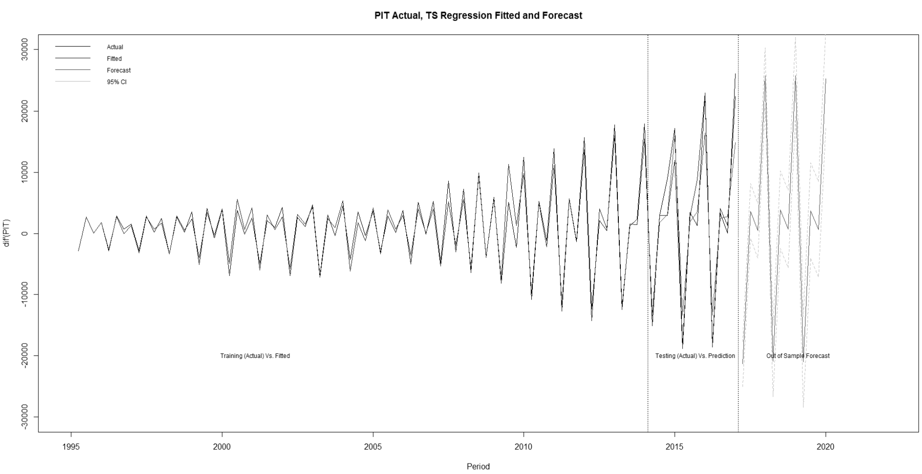
The percentage error on the three years 2014/15 to 2016/17 or 12 quarters in Table 13 clearly show the model performing with an error of less than 5% though the error was leaning towards 5% error as compared to the other methods above. The models out of sample forecast were generated for the period quarter 2, 2017 to quarter 1, 2020 with the 95% confidence intervals as shown in Figure 6. Table 14 shows the out of sample three years quarterly forecast.

Table 14. Cumulative PIT actual and TS-regression forecast in rand millions

Quarter	2017/18	2018/19	2019/20
Q01	93,604	102,635	111,924
Q02	190,869	209,127	227,588
Q03	288,580	316,347	343,910
Q04	412,067	449,262	485,476

Source: Authors' computation.

Figure 8. PIT actual, TS-regression in-sample fitted and forecast



Source: SARS & National Treasury (2015-2017) and authors' computation.

The quarterly forecast amounts to R412,1 billion, R449,3 billion and R485,5 billion for fiscal year 2017/18, 2018/19 and 2019/20, respectively. The annual forecast from the TS-regression model are little bit lower than Holt−Winters and SARIMA forecast in Tables 5 and 9, respectively. Figure 8 presents the in-sample fitted values and the model out of sample forecast from the SARIMA fitted model.

6. Discussion of results and limitations

Accurate forecasting of these taxes could further assist with predicting future collections, as well as detecting false reporting of information by individuals. However, in any economy there will always be some form of tax evasion and avoidance from individuals, which could be a result of a lack of tax knowledge and/or system manipulation.

This paper presented the modelling of the South African personal income tax in rand millions using the time series models Holt–Winters, SARIMA and TS-regression with SARIMA errors. The ultimate task was to introduce and to show the power of prediction when using these models if predicting the fiscal years outcomes. Though every model has its own advantages and disadvantages, the time series are seen as the base of forecasting (preferably for short term).

The SARIMA and HW model assumes all the explanatory variables are incorporated in the historical movements or patterns hence the TS regression with SARIMA Errors allows the inclusion of the explanatory variable(s) and SARIMA model for the un-explained portion of the dependent variable. However, there are other methods which can also be exploited for forecasting such as the GARCH model, the VAR models and so forth.

Table 15 compares the forecast from the three methods for the three years (2017/18, 2018/19 and 2019/20) with PIT as a variable of interest and the average (AVG) forecast from those models.

Table 15. Forecast comparisons in rand millions

FY	FY	Holt–Winters	SARIMA	TS-Regression	AVG Forecast
2017/18	2014/15	464 299	462 588	412 067	446 318
2018/19	2015/16	509 026	500 631	449 262	486 307
2019/20	2016/17	558 062	538 674	485 476	527 404

Source: Authors' computation.

The results for the PIT forecast from the Holt–Winters and SARIMA model converges and the TS-regression forecast are a bit lower than the forecast from the two models. This also confirms the assumption that pure time series are the bases for forecasting. The location of the actual realisation is assumed to be around the forecast from the three models with some acceptance error margin; hence the average forecast could be used.

As indicated earlier, these methods are good for short term forecast (i.e., three to five years forecast) as they lose prediction power when forecast period stretch forward. The effect of structural shock and other unknown phenomenon could results in a model forecast that are far away from the actual, therefore careful follow up of these models is recommended. The good forecasting practice will be to monitor the actual realisation as it shows up and revise the forecast if necessary.

7. Conclusions and recommendations

The Holt–Winters model, SARIMA model and TS-regression with SARIMA error were used to model and forecast personal income tax for the South African economy from which it was observed that the models captured the movement of these taxes with higher precision on the in-sample data used. This is shown by the computed in-sample measure of accuracy such as Mean Error (ME), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Percentage Error (PE), Mean Percentage Error (MPE), Mean Absolute Percentage Error (MAPE), Mean Absolute Squared Error (MASE), Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) on the training sample data spanning quarter 1, 1995 to quarter 1, 2014.

The sample data from quarter 2, 2014 to the quarter 1, 2017 were used to test the model strength by comparing the predicted values to the actual PIT collection for the same period. Using the sample with recent actual PIT collection up to quarter 1, 2017, the out of sample forecasts for the period quarter 2, 2017 to quarter 1, 2020 were generated from the fitted models. The quarterly forecast was then aggregated to form SARS fiscal year forecast (SARS fiscal year span from April to March).

The study recommends the use of these models when forecasting future PIT payments, because they are precise and unbiased in forecasting tax revenue with minimal error for short term, preferably the Holt–Winters and SARIMA models. The results of this study will assist the South African authorities with decision making for future revenue management, resources allocations and distribution of revenue collected to the government departments with accuracy. These types of models could also be used as back up or for a measure of correctness of the economic models, as they do not introduce bias when predictions are done.

The techniques portrayed in this paper are properly documented for the development of other models to be used for various purposes. Scope for further research on analysing the Value Added Tax (VAT), Corporate Income Tax (CIT) and Total tax revenue will depend greatly on techniques used in this study.

References

African Tax Institute. (2012). *Annual report for 2011*. South Africa: University of Pretoria.

Boonzaaier, W. (2015). *The Asymmetric Behavior of Tax Revenue over the Business Cycle*. ESSA Biannual Conference of the Economic Society of South Africa, 30 August – 1 September 2017, Grahamstown South Africa.

Brojba, L. C., Dumitru, C. G., & Belciug, A. V. (2010). On the use of some optimal strategies of fiscal administration during economic crisis, Romania. *Romanian Journal of Economic Forecasting*, *12*, 151-164.

Chatagny, F., & Soguel, N. C. (2009). *Tax revenue forecasting in the Swiss cantons: A time series analysis*. Chavannes-près-Renens, Switzerland: IDHEAP.

Jayesekara, N., & Passty, B. W. (2009). *City of Cincinnati income tax revenue: A forecast based on economic data and causal inference*. Cincinnati, Ohio: University of Cincinnati.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*, 357-384.

Hyndman, R. J., Makridakis, S., & Wheelwright, S. C. (1998). *Forecasting methods and application: Third edition*. New York: John Wiley & Sons.

Hyndman, R. J., Koehlor, A. B., Synder, R. D., & Grose, S. (2002). *A* state space framework for automatic forecasting using exponential smoothing methods. *International Journal of Forecasting*, *18*, 439-454.

Maindonald, J., & Braun, J. (2003). *Data analysis and graphics using R: An example based approach*. Cambridge: Cambridge University.

SARS, & National Treasury. (2015). *South African tax statistics* (national publication). Pretoria, South Africa: Author.

SARS, & National Treasury. (2016). *South African tax statistics* (national publication), Pretoria, South Africa: Author.

SARS, & National Treasury. (2017). *South African tax statistics* (national publication). Pretoria, South Africa: Author.

Shumway, R. H., & Stoffer, D. S. (2006). *Time Series Analysis and its applications with R examples* (3rd ed.). USA: Springer.

Tong, H., & Lim, K. S. (1980). Threshold autoregression, limit cycles and cyclical data. *Journal of the Royal Statistical Society*, *42*, 245-292.

Tsay, R. S. (2002). *Analysis of financial time series* (1st ed.). New York: John Wiley & Sons.

Van Heerden, Y. (2013). *Personal income tax reform to secure the South African revenue base using a micro-simulation tax model*. PhD thesis, Pretoria: University of Pretoria. Retrieved from https://repository.up.ac.za/handle/2263/40264

Wei, W. S. (2006). *Time series analysis: Univariate and multivariate.* Boston: Pearson.

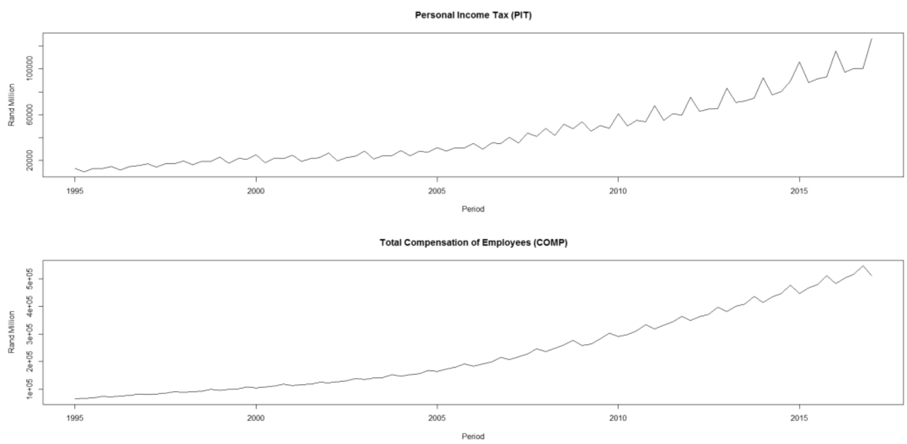
Pelinescu, E., Anton, L. V., Ionescu R., & Tasca R. (2010). The analysis of local budgets and their importance in the fight against the economic crisis effects. *Romanian Journal of Economic Forecasting*, Supplement, 17-32.

Pindyck, R. S., & Rubinfeld, D. L. (1998). *Econometric models and economic forecasts* (4th ed.). Boston: Irwin – McGraw-Hill.

Statistics South Africa (STATS SA) (2016). Retrieved from http://www.statssa.gov.za/?page_id=1854&PPN=P0441&SCH=6639

Appendix

PIT actual and total compensation of employees (COMP)



Source: SARS & National Treasury (2015-2017) and Statistics South Africa (2016).

Appendix is the South African quarterly actual PIT data and total compensation of employees (COMP) for the period Quarter 1, 1995 to Quarter 1, 2017.