International Journal of Thesis Projects and Dissertations (IJTPD)
Vol. 3, Issue 3, pp: (54-66), Month: July - September 2015, Available at: www.researchpublish.com

Volatility Modelling of Corporate Income Tax (CIT) in Rwanda

Lucie NIYIGENA[1], Dr Marcel Ndengo[2], Dr. Joseph K. Mung'atu[3]

[1]Jomo Kenyatta University of Agriculture and Technology, Applied Statistics, Kigali Campus, Rwanda
[2]Jomo Kenyatta University of Agriculture and Technology, Kigali Campus, Rwanda
[3]Jomo Kenyatta University of Agriculture and Technology, Department of Agriculture and Actuarial Science, Kenya

Abstract: **In order to test whether the assumption of homoscedasticity is valid or not for Rwanda Corporate Income tax, we estimated a correctly specified Autoregressive Integrated Moving Average (ARIMA) model of the underlying time series, this helped to remove the linear dependence in the series. To test for Autoregressive Conditional Heteroscedasticity (ARCH) effects the residuals of the mean equation from the ARIMA model have been used. It was verified through the Autocorrelation Function (ACF), Partial Autocorrelation Function (PACF) and Ljung-Box Q-statistics that the residuals of the mean equation did not show any significant serial correlation and that the model provided a good fit. The most appropriate specification model the volatility of Corporate Income tax collections in Rwanda was found to be a GARCH (1,1) model given the characteristics of the time series and the overall fit of the model. The autocorrelation functions of the residuals and squared residuals were also examined to confirm the model adequacy.**

Keywords: **Volatility, Corporate income tax, Autoregressive, and Heteroscedasticity.**

I. INTRODUCTION

Constant variance in the residuals, also known as homoscedasticity, is one of the assumptions in the classic normal linear regression model. This assumption does not always hold in practice, however, as volatility may be clustered around a certain level during certain periods and move towards another level during other periods. For example, Corporate Income Tax (CIT) collections in Rwanda maintain a non-linear relationship with calculated tax liability (Total amount of tax that an entity is legally obligated to pay to an authority as a result of the occurrence of a taxable event), which can be mainly attributed to the structure of the Income Tax Act as it relates to the impact of tax loss carry forwards and foreign currency gains/losses on corporate profitability as well as the role of special tax treatment for certain sectors.

This would have the consequence that accounting profits do not necessarily extend to tax liability. Given the impact of these factors, especially during periods of low economic activity, one would expect an increasing level of volatility in tax collected from companies during these periods, as corporate profits become an unreliable indicator for actual tax collected.

It could therefore be helpful in certain instances to model the conditional variance of corporate income tax (CIT) collections as opposed to only using the long-run (unconditional) variance, given that the variance does not remain constant over the measured period.

For many years in Rwanda as in other countries taxes plays the important role in economic development. They have different functions such as; to collect funds to be used in financing government works (financial function), promoting national economic development (economic function), and promoting the social welfare of the population (Social function). Regarding income taxes on corporations, nearly all countries assess them, but the provisions and rates differ widely. Since industrialized countries generally have larger corporate sectors than less-developed countries, corporation income taxes in developed countries tend to be greater in relation to national income and total government revenue, except in major mineral-producing areas of less-developed countries.

Variations in growth and cyclicality of revenues arise from two factors; one is the industry mix and performance of industries within a country. Another important factor is the composition of each country's tax portfolio. Countries choose from a variety of tax instruments when designing their tax structure, including general sales, selective sales, personal income, corporate income, license, property, and severance taxes (same as mining royalties in Rwanda). Each of these tax instruments responds differently to upturns and downturns in the economy. For example, a sales tax on food is fairly stable because people will buy food in good times and in bad. However, tax revenues from capital gains depend largely on the stock market's performance and thus can be volatile. The problem for Revenue Authorities trying to predict revenues is that stock market fluctuations and other cyclical events have a larger impact on incomes at the top, causing revenues from income taxes to vary widely from year to year. Therefore, this research intends to model Rwanda Corporate Income tax (CIT) by using Conditional volatility model (i.e. ARCH-GARCH model).

II. REVIEW OF PREVIOUS STUDIES ON THE SUBJECT OF STUDY

Autoregressive Conditional Heteroscedasticity (ARCH) and General Autoregressive Conditional Heteroscedasticity (GARCH) models have become very popular in that they enable the econometrician to estimate the variance of a time series at a particular point in time. Clearly, asset pricing models indicate that the risk premium will depend on the expected return and the variance of that return. The relevant measure is the risk over the holding period, not the unconditional risk. Similarly, a portfolio manager who uses value-at-risk (VaR) might be unwilling to hold a portfolio with a 5% chance of losing $1 million. The assessment of the risk should be determined using the conditional distribution of asset returns. To use Engle's example of the importance of using the conditional variance rather than the unconditional variance, consider the nature of the wage-bargaining process. Clearly, firms and unions need to forecast the inflation rate over the duration of the labor contract. Economic theory suggests that the terms of the wage contract will depend on the inflation forecasts and the uncertainty concerning the accuracy of these forecasts. Let $\mathbf{E_t\pi_{t+1}}$ denotes the conditional expected rate of inflation for $\mathbf{t+1}$ and let $\boldsymbol{\sigma^2_{\pi t}}$ denotes the conditional variance. If parties to the contract have rational expectations, the terms of the contract will depend on $\mathbf{E_t\pi_{t+1}}$ and $\boldsymbol{\sigma^2_{\pi t}}$ as opposed to the unconditional mean or the unconditional variance.

The rational expectations hypothesis asserts that agents do not waste useful information. In forecasting any time series, rational agents use the conditional distribution, rather than the unconditional distribution of the time series. Hence, any test of the wage bargaining model above that uses the historical variance of the inflation rate would be inconsistent with the notion that rational agents make use of all available information (i.e., conditional means and variances).

Engle's 2003 Nobel Prize (shared with Clive Granger) is a testament to the importance of ARCH models. Theoretical models using variance as a measure of risk (such as mean-variance analysis) can be tested using the conditional variance. As such, the growth in the use of ARCH/GARCH methods has been nothing short of impressive. In fact, there are so many types of models of conditional volatility that it is common practice to refer to the entire class of models as ARCH or GARCH models.

III. METHODOLOGY

According to Enders (2010), one can attempt to forecast the variance inherent in corporate income tax collections using a this model; $\mathbf{CIT_{t+1} = \varepsilon_{t+1}x_t}$, where $\mathrm{CIT_{t+1}}$ represents the variance of corporate income tax collections at time t+1, $\varepsilon_{t+1} \sim N(0,\delta^2)$ is a white noise process with variance σ^2 and x_t is an explanatory variable at time t. When x_t is considered to be a constant, the variance of CIT simplifies to a white noise process and is considered homoscedastic.

Assuming that x_t is not a constant, the assumption of homoscedasticity is not valid anymore, and the variance of $\mathrm{CIT_{t+1}}$ conditional on the value of independent variable x_t can be expressed as follows: $\mathbf{Var(CIT_{t+1}|x_t) = x_t^2\,\boldsymbol{\sigma}^2}$.

Based on this relation, the variance of $\mathrm{CIT_{t+1}}$ will vary directly with x_t^2. In practice, the above specification is modified and estimated in logarithmic form to obtain a linear regression equation of the following form: $\mathbf{\ln}(\boldsymbol{CIT_t}) = \boldsymbol{a_0} + \boldsymbol{a_1}\,\mathbf{\ln}(\boldsymbol{x_{t-1}}) + \boldsymbol{e_t}$.

A problem with this linear regression specification, apart from the fact that x_t affects the mean of $\ln(\mathrm{CIT_t})$, is that a specific event or factor is assumed to be the reason for the change in variance. In addition, since a linear regression equation was estimated, the assumption of constant variance needs to hold, which does not make this specification helpful for the purposes of modeling conditional volatility.

ARIMA Processes:

The ARMA process y_t with AR of order p and MA of order q (ARMA(p, q)) is illustrated as by the following equation; $y_t = \alpha_1 y_{1-1} + \cdots + \alpha_p y_{t-p} + u_t + m_t u_{t-1} + \cdots + m_q u_{t-q}$. The process is stable and stationary if α(z) ≠ 0 for |z| ≤1, and it is invertible if m(z) ≠ 0 for |z| ≤1. If the process is stable, it has a pure (possibly infinite order) MA representation from which the autocorrelations can be obtained. Conversely, if the process is invertible, it has a pure (infinite order) AR representation. For mixed processes with nontrivial AR and MA parts, the autocorrelations and partial autocorrelations both do not have a cutoff point but taper off to zero gradually. A stochastic process y_t is called an ARIMA(*p*,*d*,*q*) process ($y_t \sim$ ARIMA(*p*,*d*,*q*)) if it is I(*d*) and the *d* times differenced process has an ARMA(*p*,*q*) representation, that is $\Delta^d y_t \sim$ ARMA(*p*,*q*). For processes with distinct seasonality, so-called seasonal models are sometimes considered.

Autoregressive Conditional Heteroscedasticity (ARCH) Processes:

Robert Engle (1982) provides a way of modeling the mean and variance of a series at the same time. Continuing from the equation in the previous section (CIT collections in Rwanda), should the variance of the residuals be heteroscedastic in nature, these movements in the variance can be approximated by an ARMA (i.e. white noise) process. The conditional variance can for example then be represented as an AR(q) process using the squares of the estimated residuals, as shown in Enders (2014): $\hat{\varepsilon}_t^2 = \boldsymbol{\alpha_0 + \alpha_1 \hat{\varepsilon}_{t-1}^2 + \alpha_2 \hat{\varepsilon}_{t-2}^2 + \cdots + \alpha_q \hat{\varepsilon}_{t-q}^2 + v_t}$.

Where $\boldsymbol{v_t}$ follows a white noise process (i.e. a random process of random variables that are uncorrelated, have mean zero, and a finite variance). This type of equation is known as an autoregressive conditional heteroscedastic model or ARCH model. In the case where the estimated coefficients of $\boldsymbol{\alpha_1}$ up until $\boldsymbol{\alpha_n}$ are not significantly different from zero, the estimated variance will equal the constant $\boldsymbol{\alpha_0}$. The above equation can then be used to forecast the conditional variance. Two additional considerations arise from the above specification. First, maximum likelihood is the preferred estimation method when estimating the mean and variance simultaneously. Second, it is better to specify $\boldsymbol{v_t}$ in a multiplicative form as opposed to additive. As a result, the following set of higher-order ARCH (*q*) processes was considered by Engle (1982): $\boldsymbol{\varepsilon_t = v_t \sqrt{\alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2}}$, Where v_t mimics a white noise process such that its variance equals one, v_t and the elements of the error process ε_{t-i} are independent of each other. α_0, α_1, ... α_q are constants such that α_0>0 and 0 ≤α_1,…, α_q≤1, It can be shown that the ε_t process exhibits a zero mean and contains no autocorrelation. In addition, the unconditional mean and variance are not affected by the ε_t process, while the conditional mean of the ε_t process equals zero. It is the conditional variance of the ε_t process, however, that has important implications. The conditional variance of ε_t can be expressed as a function of ε_{t-1}^2 and a constant in the case of an ARCH (1) model, as shown in Enders (2014): $\boldsymbol{E[\varepsilon_t^2 | \varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2}$. It will therefore follow that a large value of ε_{t-1}^2 will result in a large conditional variance. It is necessary to restrict the coefficients to positive values, while for stability purposes it is necessary for α_1 to be further restricted to 0 ≤ α_1 ≤1.

To summarize the properties of ARCH models:

- The conditional and unconditional expectations of the error terms are zero;
- The $\boldsymbol{\varepsilon_t}$ set of elements are serially uncorrelated;
- The errors, however, are not independent and are related through their second moment;
- The conditional variance can be represented as an autoregressive process, which results in conditional heteroscedastic errors; and
- The conditional heteroscedasticity in the errors will ultimately result in the volatility series being heteroscedastic.

Generalized Autoregressive Conditional Heteroscedasticity(GARCH) Processes:

As an extension of Engle's paper, Tim Bollerslev published a paper titled Generalized Autoregressive Conditional Heteroscedasticity in 1986, which allowed the conditional variance to be represented by an ARMA process. It is now assumed that the error process can be represented as follows: $\boldsymbol{\varepsilon_t = \nu_t\sqrt{h_t}}$, Where the variance of ν_t equals one, and the term h_t can be represented as follows: $\mathbf{h_t = \alpha_0 + \sum_{i=1}^{q}\alpha_i\varepsilon_{t-i}^2 + \sum_{i=1}^{p}\beta_i h_{t-i}}$. First, similar to the case of ARCH models, the conditional and unconditional means of ε_t are zero given that the elements of the set ν_t follows a white noise process. Second, it can now be seen that the conditional variance of ε_t follows an ARMA process (as opposed to only an AR process in the case of ARCH models), which is represented by h_t. This type of model is known as a GARCH (p,q) model as it allows for both autoregressive and moving average components in describing the variance. It should be clear that any ARCH model can be represented by a GARCH model.

To Build a volatility model for an asset return series consists of four steps, Ruey S. Tsay, (2005, p106):

1. Specify a mean equation by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.

2. Use the residuals of the mean equation to test for ARCH effects.

3. Specify the volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations.

4. Check the fitted model carefully and refine it if necessary.

For $\boldsymbol{\varepsilon_t = \nu_t - \mu t}$ be the residuals of the mean equation. The squared series $\boldsymbol{\varepsilon_t^2}$ is then used to check for conditional heteroscedasticity, which is also known as the *ARCH* effects. Two tests are available. The first test is to apply the usual Ljung–Box statistics *Q(m)* to the $\{\boldsymbol{\varepsilon_t^2}\}$ series; see McLeod and Li(1983). The null hypothesis is that the first *m* lags of ACF of the $\boldsymbol{\varepsilon_t^2}$ series are zero. The second test for conditional heteroscedasticity is the Lagrange multiplier test of Engle (1982).

IV. RESULTS

ARIMA model for corporate income tax:

In order to test whether the assumption of homoscedasticity is valid or not for Corporate Income tax, we firstly estimate a correctly specified ARIMA model of the underlying time series. The squared residuals of the obtained regression equation would then serve as the basis of whether the construction of an ARCH-GARCH model would be appropriate. We have time series data on CIT (Corporate Income Tax) collections in Million Rwandan francs (Rwf). Data are distributed quarterly from 1996Q1 to 2015Q1 with summary statistics: Mean of CIT = 10762.64, Standard deviation of CIT = 10224.37, Minimum of CIT returns = 453.43, and the Maximum of CIT returns = 42568.

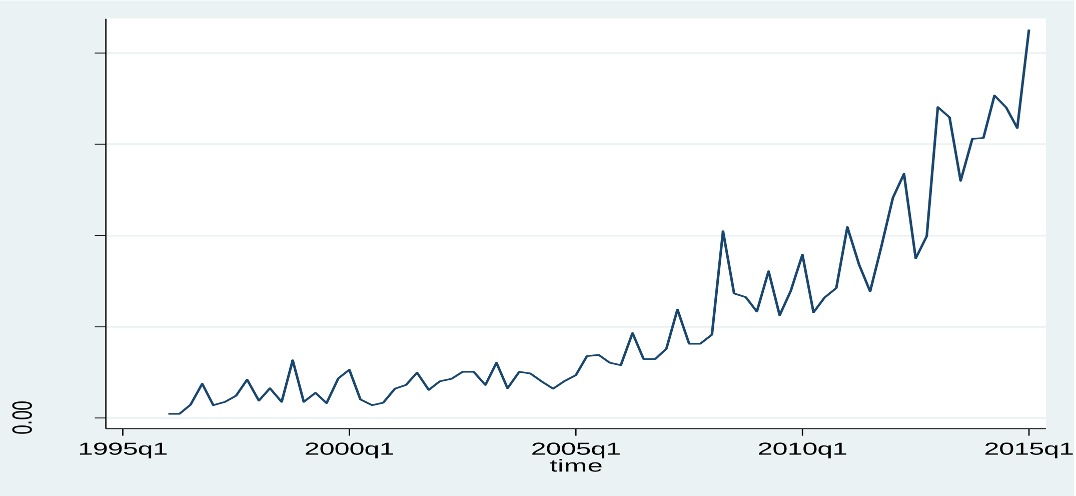

Figure 1. Quarterly Corporate Income Tax Collections (in Million Rwf)

The following observations can be made for Corporate Income Tax collections over the period

1996 Q1 - 2015 Q1, as shown in Figure 1:

i. Original CIT collections series is not stationary.

ii. The magnitude of seasonal fluctuations increased over the measured period, as indicated by the recurring spikes that increases over time.

iii. Plotted data appear to be related to an exponential function, natural logarithmic transformation can be used for smoothing,

iv. A certain upward trend can be observed, it is likely that the differenced transformation can provide more useful information, since the magnitude of corporate income tax collections in the later years is lower the trend implicit in the collections of the earlier years.

An initial attempt to address the non-constant variance and non-stationary of a time series, normally involves applying a logarithmic transformation to the time series. The logarithmic transformation of corporate income tax after adjustments is shown in Figure 2.

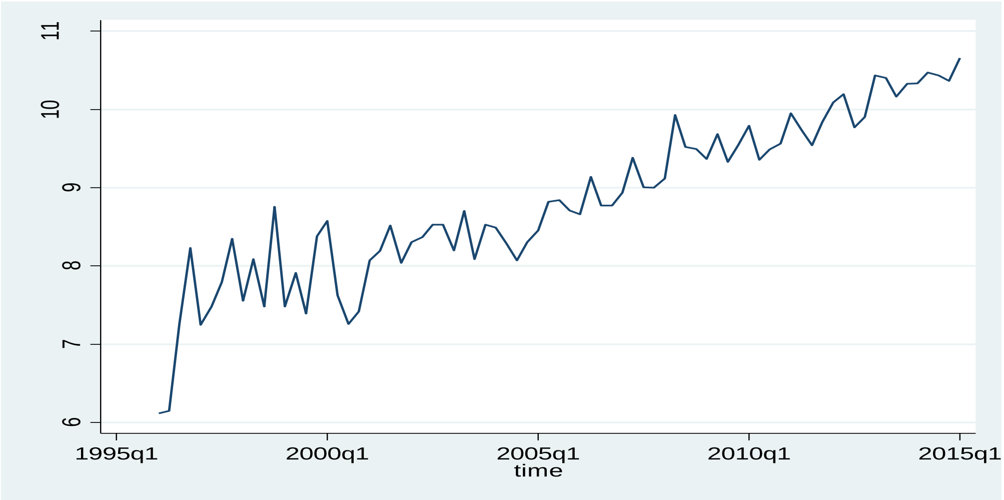

Figure 2. Logarithm of Corporate Income Tax Collections

Figure 2 shows that the logarithmic transformation, made the variance more uniform over time, also addressed to a certain extent the varying magnitude of the seasonal fluctuations as well as forcing the time series to resemble more closely a linear trend. It should be noted at this point that the underlying characteristics of the time series will not always conform perfectly to the required assumptions, and should be taken into consideration when using a model whose accuracy relies on the validity of these assumptions. Following the visual analysis, the next step would be to analyze the autocorrelation and partial autocorrelation function of the time series.

With the requirement of stationarity, the stationary process should have a mean and variance that do not change over time and the process does not have trends.

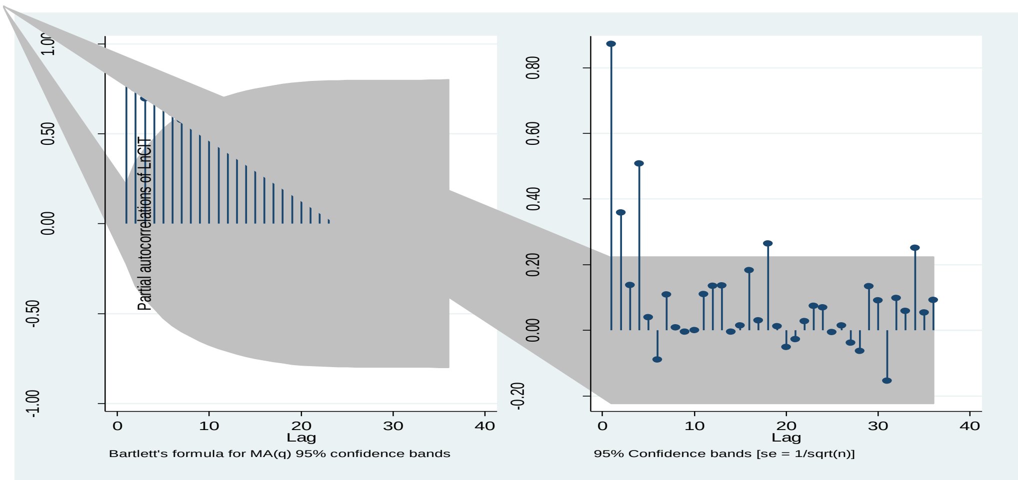

Figure 3. Natural Log of Corporate Income Tax, Correlogram

The majority of the autocorrelations are statistically significantly different from zero and ACF is a slow decay function, which is indicating non-stationarity, and PACF cuts off at lag 1 or 2.

Let test whether the natural logarithm of corporate income tax is trend stationary; if this proves to be false, the conclusion is made that the natural logarithm of corporate income tax will only become stationary after applying a differencing procedure.

Table 1. Augmented Dickey-Fuller Unit Root Test

dfuller trend regress lags(2)						
Augmented Dickey-Fuller test for unit root				Number of obs = 74		
----------		Interpolated Dickey-Fuller		---------		
	Test Statistic	1% Critical Value		5% Critical Value		10% Critical Value
Z(t)	-6.433	-4.097		-3.476		-3.166
MacKinnon approximate p-value for Z(t) = 0.000						
D.LnCIT	Coef.	Std. Err.	T	P>|t|	[95% Conf. Interval]	
LnCIT						
L1.	-1.146337	0.1781997	-6.43	0.00	-1.501835	-0.790838
LD.	0.1993836	0.1363017	1.46	0.15	-0.072531	0.471298
L2D.	0.1836546	0.1041824	1.76	0.08	-0.024184	0.391493
_trend	0.0469369	0.0076011	6.17	0.00	0.031773	0.062101
_cons	8.333503	1.281937	6.50	0.00	5.776108	10.890900

Based on the results shown in Table 1 of the Augmented Dickey-Fuller Unit Root Test, the null hypothesis that the de-trended corporate income tax collections contain a unit root can be rejected at all common significance levels. The line graph and correlogram of the de-trended corporate income tax collections are shown in Figure 4.

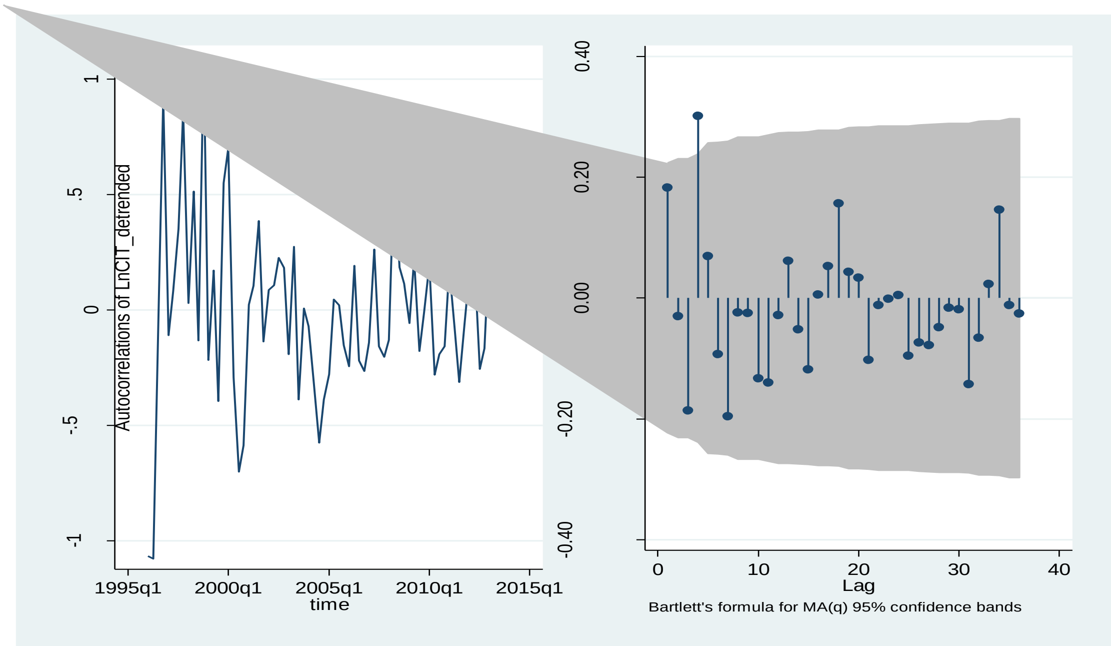

Figure 4. De-Trended Log of Corporate Income Tax Collections, Line and Correlogram

Still the line graph and the autocorrelation function correlogram show that the mean of de-trended corporate income tax collections continuously revert to the zero level, the time series has always a non-constant variance even after that the time series was de-trended. The most appropriate step would be to apply a difference transformation of the de-trended corporate income tax collections, which in the case of quarterly data is equivalent to taking the difference between the current quarter and the corresponding quarter in the previous year, and applying this procedure to the whole series. The resulting output is shown in Figure 5.

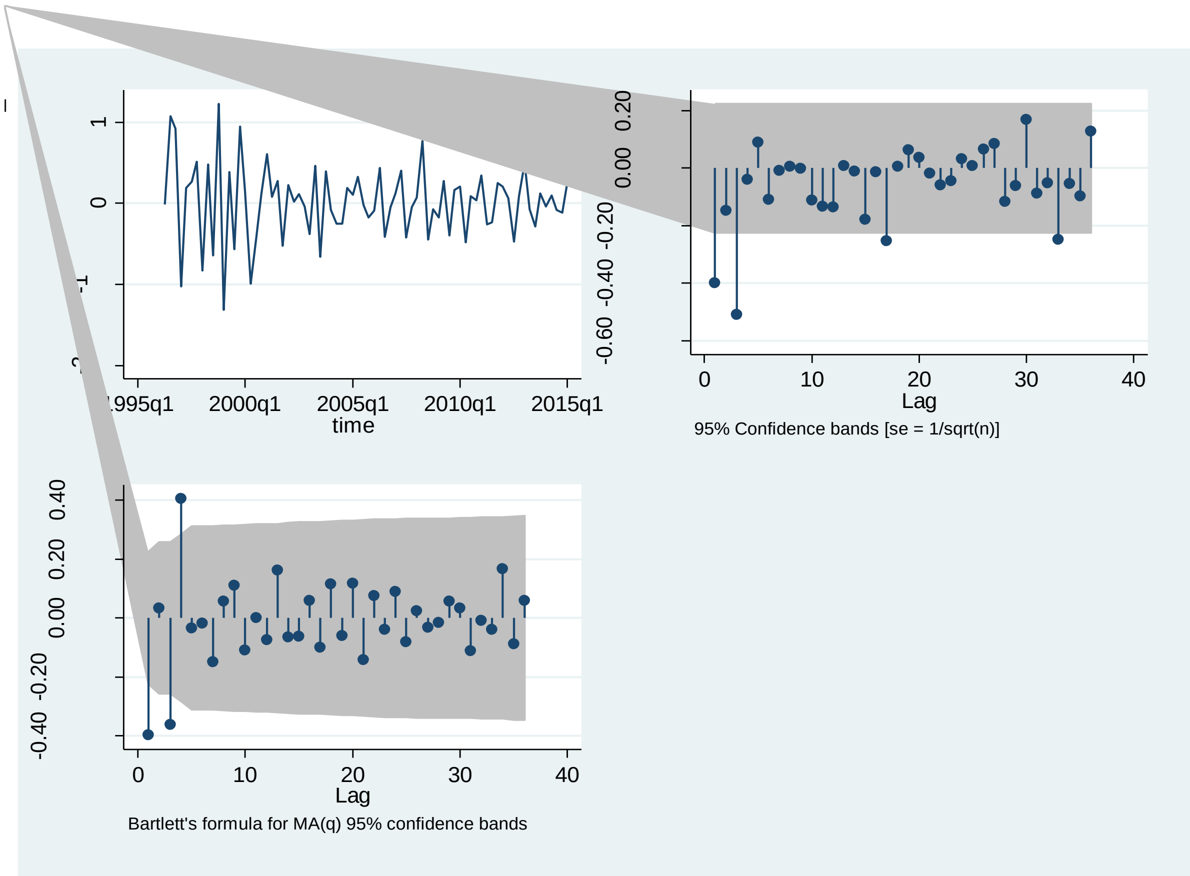

Figure 5. De-Trended and Differenced Log of CIT, Line Graph and Correlogram

Figure 5, which contains the combination of the line graph and the correlogram ACF and PACF of de-trended and differenced natural logarithm of corporate income tax, the line graph shows that now logarithm of corporate income tax appears to follow a white-noise process with no visible recurring seasonal pattern or trend. Looking at the correlogram, however, there is some evidence of non-constant variance in the series, since there are some spikes which are going out of critical region. This remaining non-constancy will be addressed in the specification of the model in the form of seasonal autoregressive, integrated and moving average terms. By applying an Augmented Dickey-Fuller Unit Root Test with no intercept or trend, the null hypothesis that the seasonally differenced and de-trended time series has a unit root can be rejected at the 5% significance level, thereby formally confirming stationarity.

Model Identification:

Since the described outcome is inconsistent with the given theoretical patterns and does not provide a clear model structure, it would make sense to estimate a set of alternative models and then selecting the model that performs best in terms of information criteria as well as a set of diagnostic tests. The set of alternative models are shown in Table 2. To enable comparability, all the equations were estimated over the period Q2 1996 to Q1 2015. The values in parentheses are the p-values of each estimated coefficient. Ljung-Box Q-statistics are also shown to test the null hypothesis that all the lags up until the selected lag length are collectively not significantly different from zero. The model to estimate is: $\mathbf{cit_t = \alpha_0 + \sum_1^p \alpha_i\, cit_{t-i} + \sum_0^q \beta_i\, \mathcal{E}_{t-i}}$, where cit_t represents the seasonally differenced and de-trended logarithm of corporate income tax, the β parameter is used to represent ordinary moving average processes and the α parameter is used for ordinary autoregressive terms.

International Journal of Thesis Projects and Dissertations (IJTPD)
Vol. 3, Issue 3, pp: (54-66), Month: July - September 2015, Available at: www.researchpublish.com

Table 2. Estimated ARIMA models

	1. ARIMA (1,0,0)	2. ARIMA (0,0,1)	3. ARIMA (1,0,1)	4. ARIMA (1,1,1)	5. ARIMA (1,1,3)	6. ARIMA (1,1,4)	7. ARIMA (3,1,1)	8. ARIMA (3,1,4)
Const	8.695835 (0.000)*	8.81522 (0.000)*	8.586621 (0.001)*	0.0434139 (0.000)*	0.04346 (0.000)*	0.054433 (0.03)*	0.0498652 (0.002)*	0.0481982 (0.000)*
L1.ar	0.93882 (0.000)	-	0.9941718 (0.000)*	0.2228497 (0.039)*	-0.67887 (0.000)*	-0.135868 (0.50)	-0.547345 (0.000)*	-0.6005166 (0.001)*
L2.ar	-	0.601266 (0.000)*					-0.50994 (0.000)*	-0.65992 (0.000)*
L3.ar							-0. 637332 (0.000)*	-0.6228793 (0.000)*
L1.ma			-.5442567 (0.000)*	-0.99999	0.08625 (0.711)	-0.391023 (0.036)*	-0. 052565 (0.762)	-0.0035198 (0.986)
L2.ma					-0.55607	-0.111466 (0.440)		0.1393855 (0.399)
L3.ma					-0.53017 (0.001)*	-0.377216 (0.004)*		-0.2087813 (0.259)
L4.ma						0. 514271 (0.000)*		0.0645309 (0.735)
AIC	105.2453	180.8785	92.33013	71.09018	65.03532	69.37207	62.66624	66.90471
BIC	112.2767	187.9099	101.7053	78.08238	76.68899	85.6872	76.65064	87.88131
Q(4)	0.000	0.01	0.02	0.000	0.91	0.01	0.98	0.102
Q(8)	0.000	0.00	0.037	0.047	0.53	0.061	0.57	0.000
Q(16)	0.01	0,04	0.000	0.000	0.81	0.87	0.80	0.01

** Indicates statistical insignificance at 5% level*

Given that both the ACF and PACF show significant spikes at selected lags, it was decided to estimate both autoregressive-based as well as moving average-based models. The lags initially chosen to be included in the models were based on the ACF and PACF patterns of the differenced and de-trended CIT. The majority of the estimated models included a combination of lags 1, 3 and 4. To select a model to use, we look at the significance of the coefficients and also at the lower Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), but usually, there are a few models that perform similarly. After analyzing all models in table 5, we found that model 7 is the only autoregressive based model that performed comparatively well. No evidence of residual autocorrelation was found, although the information criteria ruled out Model 7 as the best model among the set of alternatives. It provides an alternative moving average representation, and incidentally was also the best model in terms of the information criteria, as well as the Ljung-Box Q statistics. Model 7 only contains four variables: one ordinary moving average term to account for the serial correlation at the third and three auto-regression terms at third lag, all of them they impact negatively to the corporate income tax collection at time t.

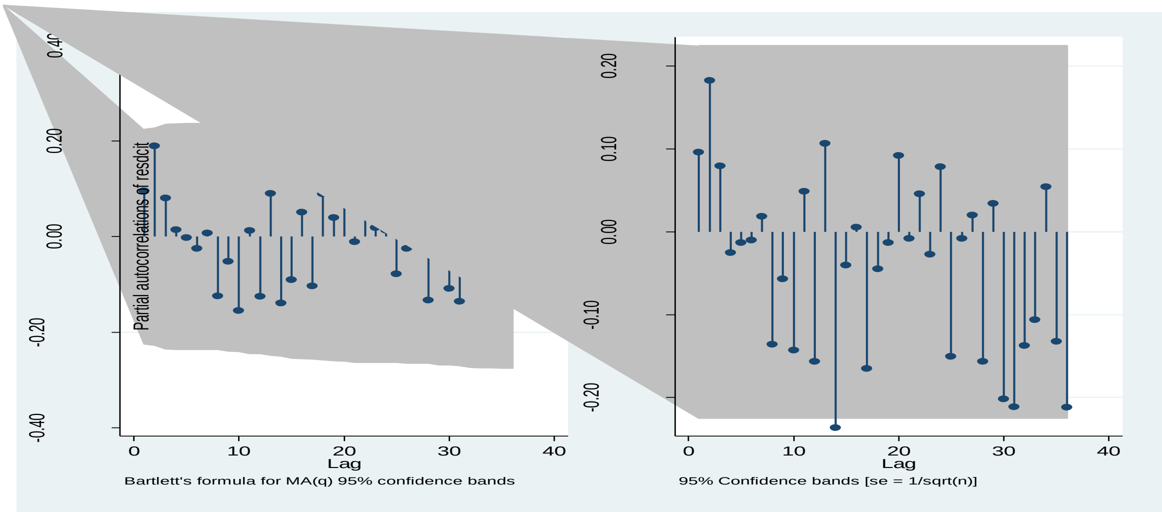

Figure 6. Residuals of Model 7, Correlogram

The correlogram of residuals of Model 7, shows that overall, no significant autocorrelation remains as the null hypothesis of no significant autocorrelation cannot be rejected at conventional significance levels. Only one individual lag exhibited statistically significant serial correlation, at the sixteenth lag, but it was decided to not explicitly account for this autocorrelation in the model for reasons of keeping the model as parsimonious as possible as well as no evident reason why corporate income tax collections in the current period would be correlated to collections that occurred sixteen quarters ago.

Diagnostic Checks for Model Adequacy:

Based on the probability of Skewness/Kurtosis tests for normality, the null hypothesis that the residuals of Model 7 are normally distributed can be rejected at the 5% level of significance. The residuals of Model 7 are following a right skewed distribution as the Skwness is equal to 0.815. While the kurtosis is equal to 5.174 > 3, which indicates that the residuals have a leptokurtic distribution, which has a more peaked shape than the normal bell curve.

To check for parameter stability a particular procedure performed is to estimate the model recursively. The recursive estimation period was only commenced from 2000Q4, given that the autoregressive term is present in the model. The recursive estimation results for the three coefficients are shown in Figures 8 respectively, along with their 95% confidence bands.

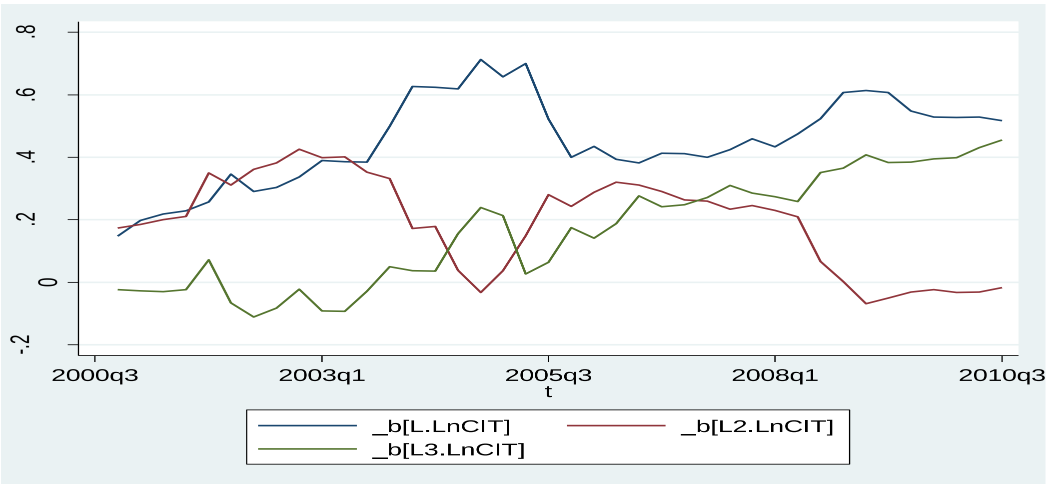

Figure 7. Recursive Estimation of AR(3) coefficients

It can be deduced that the coefficients have been relatively volatile over time. For instance, the AR(1) coefficient fluctuated within a range of 0.2 and 0.8 over the measure period, while the AR(2) coefficient predominantly stayed within the range of 0 and 4. The autoregressive term appears to have a structural breakdown at 2005q3 time, moving from an autoregressive term value of approximately 0.7 to roughly 0.4. In addition, at 2009q1 time, there appears to be a significant shift in the values of all the model coefficients. Most of these apparent shifts in the coefficients can to some extent be explained by extreme observations in the de-trended and differenced logarithm of CIT collections. If one looks at Figure 7, there does not appear to be any permanent shift in the level of the series. However, there appears to be periods in which the relative volatility is more pronounced than in other periods. The assumption can therefore be made that the heteroscedastic nature of the residuals of Model 7 may to some extent contribute to the instability of the parameters, and may not be the best representation of the underlying data process. It would therefore make sense to examine the autocorrelation function and partial autocorrelation function of the squared residuals to test for the presence of ARCH errors.

The Ljung-Box Q-statistics for the squared residuals indicates that the null hypothesis of no autocorrelation can be rejected at the 5% level of significance. This informally suggests that overall there are significant heteroscedasticity present in the series. To test for the presence of ARCH errors more formally, this test is also known as the ARCH test.

ARCH-GARCH model for corporate income tax:

Table 3. ARCH Heteroscedasticity Test

Source	SS	df	MS		Number of obs = 76	
					F(1, 74) =	7.96
Model	.517840254	1	.517840254		Prob > F =	0.0062
Residual	4.74929869	73	.065058886		R-squared=	0.0983
					Adj.R-squared =	0.0860
Total	5.26713894	74	.071177553		Root MSE =	0.25507
ehat2	Coef.	Std.Err.	t	P>|t|	[95% Conf. Interval]	
ehat2 L1.	.3134316	.1110959	2.82	0.006	.0920177	.5348455
_cons	.0881481	.0327086	2.69	0.009	.0229599	.1533362

H_0: no ARCH effects vs. H1: ARCH(p) disturbance

The F-statistic p-value shows the significance at the 5% level that the null hypothesis of the model coefficients being jointly equal to zero can be rejected, which implies that there are some type of GARCH effect in the residuals given that only an individual lagged value was found to be significant. The proper order of the GARCH process will be obtained by modeling the time series and the conditional variance at the same time, using maximum likelihood. First, a low order GARCH(p,q) process will serve as the base, after which further refinements will be made if necessary. Recall that the seasonally differenced and de-trended logarithm of corporate income tax collections was best represented by the following ARIMA specification: $\mathbf{cit_t = \alpha_0 + \beta_1 \mathcal{E}_{t-1} + \beta_2 \mathcal{E}_{t-2} + \zeta_1 \mathcal{E}_t + \mathcal{E}_t}$. This equation will be estimated simultaneously along with the equation for the variance.

Table 4. GARCH (1,1) Model

Sample:	1996q2-2015q1			Number of obs =	76	
Distribution:	t			Wald chi2(.) =		
Log likelihood =	-37.87319			Prob > chi2 =		
D.LnCIT_detrended		OPG				
	Coef.	Std.Err.	z	P>|z|	[95%	Conf. Interval]
LnCIT_detrended						
_cons	0.013091	0.0388458	0.34	0.036	-0.06305	0.0892274
ARCH						
arch						
L1.	0.1533595	0.115922	1.32	0.018	-0.07384	0.3805625
garch						
L1.	0.7793728	0.1615559	4.82	0.000	0.462729	1.096017
_cons	0.0077416	0.0120089	0.64	0.519	-0.0158	0.0312787

The coefficient of the mean model is statistically significant at the 5% level. The error distribution of the residuals was assumed to follow a student's *t* distribution as opposed to the Gaussian (normal) distribution, given the kurtosis characteristics of the original ARIMA model. The coefficients in the variance equation are all statistically significant at the 5% level, except for the intercept term. Overall the GARCH (1,1) model appears to be a good fit. Compared to the ARIMA model, the obtained sum of squared residuals is fairly similar, but the information criteria seem to suggest that the GARCH model is superior. In addition, the mean absolute percentage error also shows some improvement.

Diagnostic Checks for Model Adequacy:

The residuals of the GARCH (1, 1) model show that the model provides a good fit, with significant autocorrelation in the residuals only remaining at the third lag. It can also be seen that the autocorrelation in the squared residuals at the fourth lag is significant at the 5% level, while no evidence of autocorrelation can be found in the remaining squared residuals.

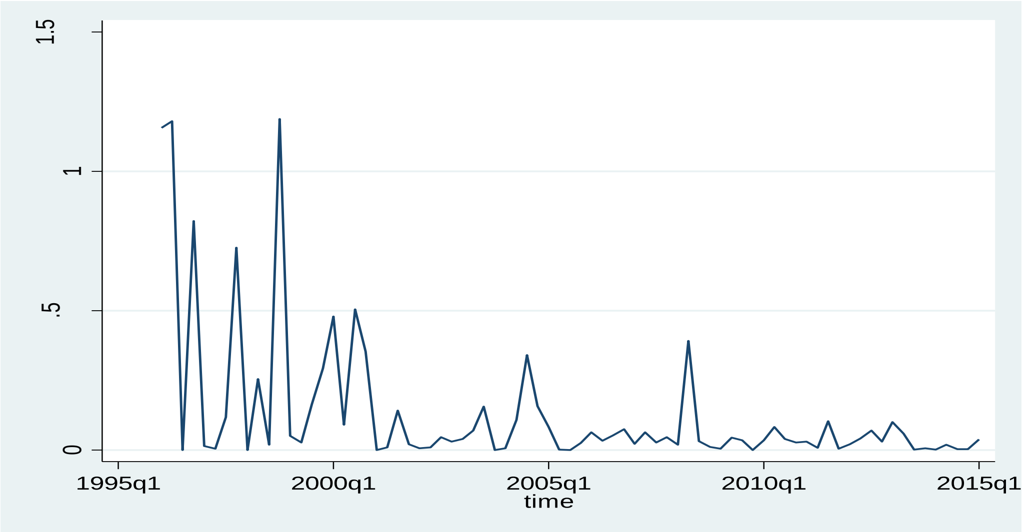

Figure 8. Residuals of GARCH (1, 1) Model, Line Graph

Compared to the residuals of ARIMA Model 7, the residuals of the GARCH (1,1) model are clearly more homoscedastic in nature. The overall distribution of the remaining residuals are more closely approximates normality, the degree of skewness are closer to zero, while the amount of excess kurtosis reduces to less than two (from twenty initially). However, the Jarque-Bera test statistic continues to reject the null hypothesis of normality in the residuals, which to a large extent can be explained by the inherent nature of quarterly corporate income tax collections. This will imply that the normality assumption will not be appropriate, and a more leptokurtic distribution function should be chosen.

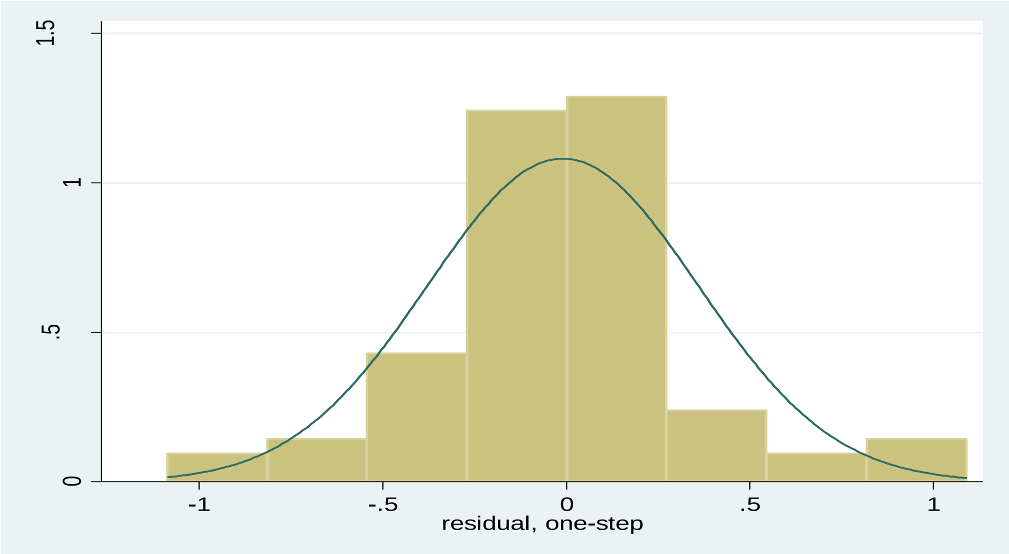

Figure 9. Residuals of GARCH (1, 1) Model, histogram

When comparing the results of the GARCH model and the ARIMA, the sum of squared residuals of the GARCH model show a minor improvement over that of the ARIMA model, and the information criteria seems to suggest that the GARCH model is superior. In addition, the mean absolute percentage error also shows improvement over the ARIMA model.

V. CONCLUSION

It has been observed in practice that standard ARIMA model of CIT collections does not appear to fulfill the requirement of constant variance, although these models are constructed using the assumption of homoscedasticity. It is hypothesized that CIT collections maintain a non-linear relationship with actual tax liability in a given fiscal year, which can be mainly attributed to timing issues, the impact of tax loss carry forwards and foreign currency gains/losses as well as the role of special tax treatment for certain sectors.

This non-linear relationship may give rise to situations during which company profits are rising but tax liability remains relatively stagnant or decreasing, depending on the assessed position of individual companies. This may lead to more heterogeneous payments and increased uncertainty or volatility in total payments received as a result. The overall impact of the above described effect, which is difficult to quantify, may give rise to periods of volatility that significantly deviates from the long-run variance, which implies that the variance in CIT collections are likely to be non-constant over time.

An ARIMA model was firstly specified to model the mean of the time series, since the validity of the conditional variance model is dependent on the mean equation being correctly specified. The next step was to test for ARCH errors through regressing the squared residuals of the mean equation on lagged values of itself and then testing the null hypothesis that the lagged squared residuals do not explain movements in the current period squared residuals. The most appropriate specification was found to be a GARCH (1,1) model given the characteristics of the time series and the overall fit of the model. The ACF and PACF of the residuals and squared residuals were examined to confirm the model adequacy. The results confirmed the suspicion that CIT collections do not exhibit constant variance over time, which needs to be taken into consideration when creating forward looking estimates of CIT collections.

REFERENCES

[1] WALTER Enders, (2014) Applied Econometric Time Series, 4th Edition.

[2] Ruey S. Tsay, (2005) Analysis of Financial Time Series 2nd Edition.

[3] Damodat N. Gujarati and Dawn C. Porter,(2008) Basic Econometrics, Fifth Edition.

[4] Robert F. Engle, (Jul., 1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica, Vol. 50, No. 4.

[5] Helmut Lutkepohl, Markus Kratzig (2004), Applied Time Series Econometrics.

[6] Tomasz WÓJTOWICZ and Henryk GURGUL, (2009), Long memory of volatility measures in time series.

[7] Tim BOLLERSLEV, (1986), Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 31 (1986) 307-327. North-Holland.

[8] McLeod, A. I. and W. K. Li (1983). Diagnostic checking ARMA time series models using squared residual autocorrelations.

[9] National treasury, Republic of South Africa, (2013/14), Economic Drivers of Tax Revenue.

[10] Xiong-Fei Zhuang and Lai-Wan Chan, (2004), Volatility Forecasts in Financial Time Series with HMM-GARCH Models.

[11] Peter Bloomfield and Ben Kedem Symposium, (July 31, 2009 Series), Modeling the Variance of a Time Series.

[12] Mei-Yuan Chen, (Feb, 25, 2013), Time Series Analysis: Conditional Volatility Models.

[13] Andrew C. Harvey, (1990), The Econometric Analysis of Time Series, 2nd Edition.

[14] Andersen, T.G., Bollerslev, T., Christoffersen, P.F., and Diebold, F.X. (2006), "Volatility and Correlation Forecasting".

[15] Alizadeh, S., Brandt, M.W., Diebold, F.X. (2002), "Range-based estimation of stochastic volatility models".

[16] Andersen, T.G. (1992). "Volatility". Working Paper, Finance Department, Kellogg School, Northwestern University.

[17] Andersen, T.G. (1994). "Stochastic autoregressive volatility: A framework for volatility modeling".

[18] Barucci, E., Reno, R. (2002). "On measuring volatility and the GARCH forecasting performance". Journal of International Financial Markets, Institutions and Money 12.

[19] Bates, D.S. (1996). "Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options". Review of Financial Studies 9.

[20] Battle, C., Barquin, J. (2004). "Fuel prices scenario generation based on a multivariate GARCH model for risk analysis in a wholesale electricity market". International Journal of Electrical Power and Energy Systems 26.

[21] Bauwens, L., Laurent, S. (2005). "A new class of multivariate skew densities with application to GARCH models". Journal of Business and Economic Statistics.

[22] Diebold, F.X., Lopez, J. (1995). "Modeling volatility dynamics". In: Hoover, K. (Ed.), Macroeconometrics: Developments, Tensions and Prospects. Kluwer Academic, Boston.