ISSN: 2755-0214

Journal of Economics & Management Research

Review Article

Open Access

Estimating Vat Refunds the Case of South Africa

Cathrine Thato Koloane * and Mangalani Peter Makananisa

[1]Independent researcher, University of Kwa-Zulu Natal, Durban, South Africa

[2]Independent researcher, University of South Africa, Pretoria, South Africa

ABSTRACT

This study intends to estimate VAT refund levels in South Africa in an ideal situation where there are well-equipped, incorruptible officials and a proper VAT system is in place. Understanding the dynamics behind the behaviour of VAT and its main drivers is crucial and could have a huge benefit to the country's economy with regards to closing the tax gap related to this tax type. Using the data from various sources (VAT refunds and some macroeconomic variables), a Vector Autoregression (VAR) model was used to estimate the level of VAT refunds in South Africa. The model estimates VAT refunds for the period 2021/22 to be R242.7 billion, while the VAT refunds forecast for the period 2022/23 and 2023/24 amounts to R254.6 billion and R267.3 billion, respectively. Furthermore, VAT refunds contribute on average 17.5% to the total tax for the forecast period of 2021/22-2023/24. The study also indicates that the growth in VAT refunds is influenced by the growth in domestic VAT collections, increasing employment rate and the growth in both agriculture and construction GDP. The estimated level of VAT refunds can serve as an important consideration in the national budgeting processes in South Africa. Adequate provisions can be made to enable proper planning and distributions to government departments. To our knowledge, this study is the first of its kind for South Africa. In summary, the South African tax authority should not deviate from the primary goal of building sound VAT systems based on improved voluntary compliance through effective systems of self-assessment.

***Corresponding author**

Cathrine Thato Koloane, Independent researcher, University of Kwa-Zulu Natal, Durban, South Africa. Tele phone No: +2782 456 4669, E-Mail: tkoloane09@gmail.com

Received: September 29, 2021; **Accepted:** October 04, 2021; **Published:** October 07, 2021

Keywords: VAT Refunds, Vector Autoregression, South Africa.

Introduction

Value-Added Tax (VAT) is an indirect tax on the consumption of goods and services in the economy. Revenue is raised for government by requiring certain businesses (i.e. vendors) to register and to charge VAT on the taxable supplies of goods and services. These vendors act as agents for government in collecting the VAT. VAT is charged at each stage of the production and distribution process and it is proportional to the price charged for the goods and services. It is intended to tax all consumption of goods and services except those that are exempt or zero-rated. VAT is payable on the supply of goods and services as well as on the importation of goods into the South Africa territory [1].

Tax revenue plays a crucial role in the South Africa's economy as it is a pillar of the country's development. VAT contributes between 25% and 26% to the total tax revenue. Together with the other two direct taxes, i.e. Personal income tax (PIT) and Corporate income tax (CIT), these taxes contribution around 80% to the overall tax revenue [2].

VAT refunds grew by 1.5% in 2019/20, which is quite modest compared to the 2018/19 growth which was mostly driven by the VAT rate hike that came into effect on 01 April 2018. The modest growth was driven by wholesale and retail trade, catering and Accommodation, construction as well as electricity, gas and water sectors. The VAT rate hike was done to balance the loss of revenue from the negative effect of the 2008/09 recession and to meet set revenue targets in order to properly allocate to essential government services [2].

VAT refunds contributed around 17% to the total tax collected over the past 11 years. This is a significant dent to the state funds and it would be prudent of the government to be aware of the magnitude of this amount and plan appropriately for it. The ratio of the VAT refunds to gross VAT increased from 31.6% in 2010/11 to 40.8% in 2020/21, averaging 39.2%. However, the ratio of VAT refunds to current GDP increased from 3.7% in 2010/11 to 4.6% in 2020/21, averaging 4.2% has announced that for each rand claimed as VAT refunds, 87 cents was declared and input tax claim of one rand and eighty-three cents was observed on average [2,3].

In an attempt to compensate the needy, the government of South Africa was obliged to provide necessary assistance by introducing a broad based VAT system with concessions restricted to a minimum. Consequently, the current VAT system in South Africa is more differentiated and has a narrower base than was originally intended. This has given precedence to more interested parties to request for preferential treatment for their goods and services [4].

To maintain a balanced financial flow there should be some reasonable level of refunds released to the public. This includes minimisation of misuse of government systems for personal gain (fraud) by both officials and the citizens of the country. Being able to release the refund that are due within an acceptable time and at acceptable levels becomes a crucial matter as it will ensure a balanced flow of cash for both individuals and the government.

This study uses a VAR model to estimate VAT refund levels for the fiscal year 2021/22 and also gives preliminary estimates for 2022/23-2023/24. The statistical software STATA version 15.0 was used to conduct the analysis. After an introduction, the paper presents the literature review in Section 2. It then describes the research methodology in section 3, model results and analysis in section 4, discussion in section 5 and finally offers a conclusion and recommendations in section 6.

Literature Review

Several literatures exist around the world on the application of regression on VAT revenue and most specifically on VAT refunds. This includes the work of who emphasise that there should be a balance between the VAT incentives and ballooning refunds [5]. In the study, the authors used transactional firm data from 2009 to 2012 to evaluate the effect of VAT incentives on refunds. The study uses linear regression method adopted from [6], however some model advancement or modifications were included. The findings suggest that the factors influencing the ballooning refunds bill in Zimbabwe are VAT fraud and zero rating goods. Consequently, the study emphasises administrative efficiency and closing the loopholes by reducing zero-rated goods.

Determine factors affecting VAT revenues in Bandung, part of Indonesia, with an attempt to solve the problem of increasing taxable entrepreneurs. Monthly data for the period 2012 to 2016 was used in building a regression model and a descriptive research method was also used. The model outcomes suggest that taxable employers partially influence VAT receipts and the refunds highly influence VAT receipts.

Analyses the factors affecting VAT revenue by assessing the relationship between VAT revenue and economic indicators at a macro level [7]. The economic indicators, which include Gross Domestic Product (GDP), GDP per capita, consumption expenditure, household consumption (HHC) and government expenditure, imports, exports and unemployment were used to explain the importance of VAT collection. Linear regression was constructed on panel data and it was observed that GDP, consumption (for both government and household), imports and exports influence VAT income positively. VAT income is mostly influenced by HHC than by government expenditure.

Attempted to improve the Russian VAT refunds model for exported goods by finding a possible alternative way of handling VAT refunds [8]. The study used the Russian and foreign experience comparative analysis method, primarily American and Chinese method. This study concluded that borrowing Chinese systems to regulate VAT refunds would be beneficial because it is suitable for export-orientated countries.

The relationship between VAT and the economy is a two-way stream. This implies that VAT also affects economic growth. Analyses the effect of VAT and tax on the Bangladesh economy [9]. The effect of VAT policy on Bangladesh is that fewer individuals and corporations share the burden of taxes. In Bangladesh there are other taxes (supplementary duties) incorporated within VAT, which brings complexity in VAT rules. The implementation of an effective VAT system was recommended to make tax workable in Bangladesh.

Analyses the relationship between VAT and economic growth in Nigeria [10]. The Engle and Granger counteraction technique was used on data spanning from 1994 to 2012. The study concluded that VAT has a positive effect on the Nigerian economy. Furthermore, for productivity enhancement, the study recommended that the state should put measures in place to increase VAT contribution.

Paper also attempted to estimate the reasonable level of VAT refunds using the countries' cross sectional or panel data. The generalised regression equation for panel data was obtained to estimate the level of VAT refunds for different countries, included were several countries from advanced, transitional and emerging economies. Some of the key findings were that refund levels vary from state to state, more than 50% of the countries on the survey do not provide late interest payments on refund, refunds abuse is only a component of VAT fraud and there are similarities in the nature of VAT refund abuse across the states. However, the VAT refund distribution might differ from country to country, as there are different dynamics that identify the countries' behaviour towards refunds, for example:

- Some countries are advanced to emerging economies
- Countries' laws and systems might differ, and
- Different tax rates are applied per country preference.

This implies that each and every country is unique and care should be taken when generalising country results. Empirically the use of panel data might distort the precision in estimating the level of reasonable refunds for a specific country. This study suggests observing the relationship that exist between a country's economic activities and VAT refunds rather than using cross sectional or panel data from different countries to estimate VAT refund levels.

Research Methodology

VAR Model

A Vector Autoregressive (VAR) model is a general framework used to describe interrelationship among stationary variables. The model is commonly used for forecasting systems of interrelated time series and for analysing the dynamic impact of random disturbances on the system of variables. A VAR model is different from univariate autoregressive models because it allows feedback to occur between the variables in the system of equations.

There are three types of VAR models namely, the reduced form, the recursive form and the structural VAR model. Recursive VAR models considers each variable to be a function of its own past and the past values of other variables in the model but also allows some variables to be functions of other concurrent variables. By imposing these short-run relationships, the recursive model allows for structural shocks.

Structural VAR (SVAR) models include restrictions that assist in identifying causal relationships beyond those that can be identified with reduced form or recursive models. These causal relationships can be used to model and forecast impacts of individual shocks, such as monetary policy decisions [11]. A SVAR model can be represented by equation (1).

$$\mathrm{BY_t} = \Gamma_0 + \Gamma_1\, \mathrm{Y_{t-1}} + \Gamma_2\, \mathrm{Y_{t-2}} + \cdots + \Gamma_p\, \mathrm{Y_{t-p}} + \varepsilon_t \qquad (1)$$

Where $\varepsilon_t \sim$ i.i.d. D $(0, \Sigma)$, ε_{1t}, ε_{2t}, ..., ε_{nt} are structural errors

From this model, a reduced VAR model with p lags can be expressed as equation (2) where each variable is a linear function of its own past values and past values of all the other variables in the system.

$$Y_t = \Phi_0 + \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_p Y_{t-p} + a_t \quad (2)$$

Where Y_t can represent three variables such that $Y_t = \begin{bmatrix} Z_t \\ X_t \\ W_t \end{bmatrix}$,

the at's are called reduced form errors, a linear combination of ε_t's, $E(a_t) = 0$, $E(a_t a_t') = \begin{Bmatrix} \Omega & t = \tau \\ 0 & t \neq \tau \end{Bmatrix}$, Φ_i are matrices. For the 1st lag matrix, $\Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix}$.

A typical equation of the system can be expressed as equation (3) below. Each equation has the same repressors [12].

$$Z_t = c_0 + \phi_{11}^{(1)} Z_{t-1} + \phi_{12}^{(1)} X_{t-1} + \phi_{13}^{(1)} W_{t-1} + \cdots + \phi_{11}^{(P)} Z_{t-p} + \phi_{12}^{(P)} X_{t-p} + \phi_{13}^{(P)} W_{t-p} + a_{1t} \quad (3)$$

A reduced VAR model with only two variables, y_t and x_t, each of which expressed as a linear function of p lags of itself, plus an error term and a constant can be expressed as equation (4) and (5).

$$y_t = \alpha + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{p} \phi_j x_{t-j} + \epsilon_{1t} \quad (4)$$

$$x_t = c + \sum_{i=1}^{p} \beta_i y_{t-i} + \sum_{j=1}^{p} \phi_j x_{t-j} + \epsilon_{2t} \quad (5)$$

Since the systems of equations contain p lags, this model is known as a VAR (p) model. If the time series are not stationary then the VAR framework needs to be modified to allow consistent estimation of the relationships among the series. If y_t and x_t are not stationary at level, but stationary after first differencing, i.e. I(1), then equation 4 and 5 above can be expressed as equation 6 and 7 below respectively.

$$\Delta y_t = \alpha + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \sum_{j=1}^{p} \phi_j \Delta x_{t-j} + \varepsilon_{1t} \quad (6)$$

$$\Delta x_t = c + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \sum_{j=1}^{p} \phi_j \Delta x_{t-j} + \varepsilon_{2t} \quad (7)$$

If y_t and x_t are I (1) and cointegrated, then the system of equations is modified to allow for the cointegrating relationship between the I (1) variables. Introducing the cointegrating relationship leads to a model known as the Vector Error Correction Model (VECM) [13].

Goodness of Fit Tests

The model can only be deemed reliable and appropriate to use if it satisfies model diagnostics tests such as the Jarque-Bera normality test and the Lagrange multiplier test. The section below will provide a theory on the two tests used in the study.

Jarque-Bera Normality Test

The Jarque–Bera (JB) test statistic is a goodness-of-fit test to test whether sample data have the skewness and kurtosis matching a normal distribution. The test statistic is always nonnegative. If it is far from zero, it signals that the data does not have a normal distribution.

The test statistic is defined as

$$JB = \frac{n}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \quad (8)$$

Where n is the degrees of freedom, S is the sample skewness and K is the sample kurtosis.

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{3/2}} \quad (9)$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{2}} \quad (10)$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of third and fourth central moments, respectively, $\bar{x}$ is the sample mean, and is the estimate of the second central moment, the variance [14].

Under the normal distribution, the JB test statistic asymptotically has a χ^2 distribution with two degrees of freedom. The null hypothesis is a joint hypothesis of the skewness being zero and the excess kurtosis being zero. Samples from a normal distribution have an expected skewness of 0 and an expected excess kurtosis of 0 (which is the same as a kurtosis of 3). As the definition of JB shows, any deviation from this increases the JB statistic.

When using the statistic in multiple regression analysis, the correct estimate can be represented by equation (11) below.

$$JB = \frac{n-k}{6}\left(S^2 + \frac{1}{4}(K-3)^2\right) \quad (11)$$

Where n is the number of observations and k is the number of repressors when examining residuals to an equation [15].

Lagrange Multiplier Test

A Lagrange multiplier (LM) test is used to test for autocorrelation in the residuals of VAR models. The formula for the LM test statistic at lag j is represented by equation (12).

$$LM_S = (T - d - .5)\ln\left(\frac{|\hat{\Sigma}|}{|\tilde{\Sigma}_S|}\right) \quad (12)$$

Where T is the number of observations, $\hat{\Sigma}$ is the maximum likelihood estimate of Σ, the variance- covariance matrix of the disturbances from the VAR, $\tilde{\Sigma}_S$ is the maximum likelihood estimate of Σ from the augmented VAR and d is the number of coefficients estimated in the augmented VAR. The augmented VAR is derived as follows:

If there are K equations in the VAR, e_t can be defined as K x 1 vector of residuals. After creating the K new variables e1, e2, eK containing the residuals from the K equations, one can augment the original VAR with lags of these K new variables. For each lag s, an augmented regression is formed in which the new residual variables are lagged s times [16].

R-Squared

R^2 is a goodness of fit measure for linear regression models. This statistic indicates the percentage of the variance in the dependent variable that is explained by the independent variables collectively. R^2 ranges from a scale of 0 to 1 where 0 indicates that the model does not explain any of the variation in the dependent variable around its mean and 1 indicating that the model explains all the variation in the response variable around its mean. The mathematical formula for R^2 is as shown below.

$$R^2 = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2} \quad (13)$$

Advantages of VAR Models

- The model is systematic and flexible enough to capture complex real-world behaviour
- As discovered by it outperforms a lot of univariate techniques with better forecasting performance [17].
- It is able to capture the interrelationships between variables in a specified system of equations [11].

The VAR model will be utilised in this study to estimate VAT refund levels in South Africa.

Model Results and Analysis

Data

Time series was obtained for the period 2002 to 2020 for the following variables: domestic VAT collections, imports VAT, government expenditure, household consumer expenditure (HHCE), shares, mining GDP, finance GDP, GDP per capita, agriculture GDP, construction GDP, wholesale GDP, employment and exports. However due to the non-stationary of the variables and significance of variables, only domestic VAT collections, employment, construction GDP and agriculture GDP were included in the final VAR model. The dummy variable (dum1) for quarter 1 was also included. Excluding the derived dummy variable, the other variables mentioned in this section were sourced from Tax Statistics, South African Reserve Bank and Statistics South Africa (open sources).

VAR Model

VAR model is only constructed if the variables are integrated of order 1. All the variables were tested for stationary and Table 4.1 shows the results from the Augmented Dicky-Fuller (AD-F) test. The variables show unit roots at level (0) and become stationary after first differencing since the calculated t-statistics from the AD-F equations are larger in absolute terms than the critical values at 5% level of significance. The mathematical natural logarithm (ln) transformation was applied to all variables to render the variables as normal as possible.

Table 4.1: AD-F unit root test at level and first difference

Unit root	Statistic	Refunds	Domestic VAT	Construction GDP	Employment	Agriculture GDP
I (0)	Test Statistic	-1.915	-1.91	0.453	-2.16	-1.714
	5% critical value	-3.481	-3.482	-3.477	-3.479	-3.478
I (1)	Test Statistic	-4.967	-3.718	-7.548	-5.336	-4.856
	5% critical value	-3.482	-3.483	-3.478	-3.48	-3.479

Source: Author's computation

The VAT refunds equation with all the retained variables stationary at first level is presented in equation 13.

$$\Delta \ln(ref_t) = c + \emptyset dum1_t + \sum_{i=1}^{4}(\theta_i \Delta \ln(ref_{t-i}) + \gamma_i \Delta \ln(domvat_{t-i}) + \delta_i \Delta \ln(emp_{t-i}) + \vartheta_i \Delta \ln(constGDP_{t-i}) + \alpha_i \Delta \ln(agricGDP_{t-i})) \quad (13)$$

where c= -2.23 represent the constant term, $\emptyset$ =0.13 represents the coefficient of the dummy variable, ref represents the refunds, domvat represents domestic vat, constGDP represents construction GDP, emp represents employment and agricGDP represents agriculture GDP. The variable coefficients up to lag four are presented in Table 4.2.

Table 4.2: Model lagged coefficients

Lag(t-i)	θ_i	γ_i	δ_i	ϑ_i	α_i
1	0.107	0.789	-0.961	0.443	0.179
2	-0.038	-0.432	1.883	-0.232	-0.031
3	0.304	1.453	-0.068	0.201	0.034
4	-0.030	-1.262	-0.757	-0.376	-0.010

Source: Author's computation, i=1, 2, 3, 4

Table 4.3 below shows the overall explanatory power of the system of equations. Our model of interest, ln(ref), has an R2 = 0.974, indicating that 97.4% of the variation of refunds is being explained by the chosen explanatory variables.

Table 4.3: R-squared of the system of equations

Equation	Parms	R^2	χ^2	Prob
lnref (*)	22	0.974	2 662.354	0.00
lndomvat	22	0.994	12 371.780	0.00
lnempl	22	0.969	2 282.134	0.00
lnConstrGDP	22	0.991	7 825.129	0.00
lnAgricGDP	22	0.956	1 563.457	0.00

Source: Author's computation, (*) – model test of interest, Parms - Parameters

Furthermore, the model results indicates that domestic VAT has a positive effect on VAT refunds at 5% level of significance at lag 1 and 2, employment at lag 2 and agriculture GDP at lag 1. However, construction GDP shows a positive effect on VAT refunds at lag 1 only at 10.3% significance level (Detailed model output is in the Appendix).

Model Diagnostics Tests

There are several test statistic criteria used to test for serial correlation in residuals, however, the study used the Lagrange multiplier to test for the null hypothesis that there is no autocorrelation at lag order. Table 4.4 below shows the results of the test and at 25 degrees of freedom (df), there was no serial correlation at lags 2 to 4 as p-value > 0.05 (5%).

Table 4.4: Lagrange-Multiplier Test

Lag	χ^2	df	Prob
1	39.163	25	0.035
2	37.735	25	0.050
3	27.900	25	0.313
4	15.412	25	0.931

Source: Author's computation

A well-fitting model will yield residuals that are normally distributed. Table 4.5 below shows the results of the Jarque-Bera test, where the null hypothesis states that the residuals are normally distributed. Of interest is the ln(ref) model which shows a p-value of 0.501 > 0.05 which means that the null hypothesis cannot be rejected. Therefore, the residuals are normally distributed.

Table 4.5: Jarque-Bera Test

Lag	χ^2	df	Prob
ln(ref) (*)	1.382	2	0.501
ln(domvat)	240.352	2	0.000
ln(empl)	5.200	2	0.074
lnConstrGDP	7.220	2	0.027
lnAgricGDP	0.564	2	0.754
All	254.718	10	0.000

Source: Author's computation, (*) - model test of interest

Estimation of the Vat Refunds

The diagnostic checks above indicates that the residuals are normally distributed with no serial correlation and the VAR model explains over 97% of the variation in VAT refunds, therefore the fitted VAR model can be used to forecast VAT refunds. In order to further test the robustness of the fitted model the refunds were divided into training data (2010/11 – 2019/20) and test data (2020/21).

Table 4.6 shows the aggregated in-sample actual vs fitted VAT refunds and the performance or percentage error for the period 2010/11 to 2020/21. Though there was some under performance in the latter part of the test data, the model performs better on the training data for 2020/21 fiscal year with an error margin aggregating to 1.36%.

Table 4.6: In-sample actual and fitted refunds in rand million

Fiscal Year	Refunds	Refunds Est.	%Error
2010/11	103 647	122 551	-15.43%
2011/12	131 009	133 763	-2.06%
2012/13	138 820	144 902	-4.20%
2013/14	156 879	155 933	0.61%
2014/15	162 138	166 848	-2.82%
2015/16	167 055	177 644	-5.96%
2016/17	181 574	188 338	-3.59%
2017/18	191 071	198 976	-3.97%
2018/19	229 152	209 623	9.32%
2019/20	232 516	220 370	5.51%
2020/21	228 193	231 335	-1.36%

Source: Tax Statistics (2011 – 2020) and Author's computation

Table 4.7 shows that the estimated VAT refunds levels for the period 2021/22 amounts to R241.9bn and the estimates for 2022/23 and 2023/24 amounts to R252.9bn and R263.8bn, respectively.

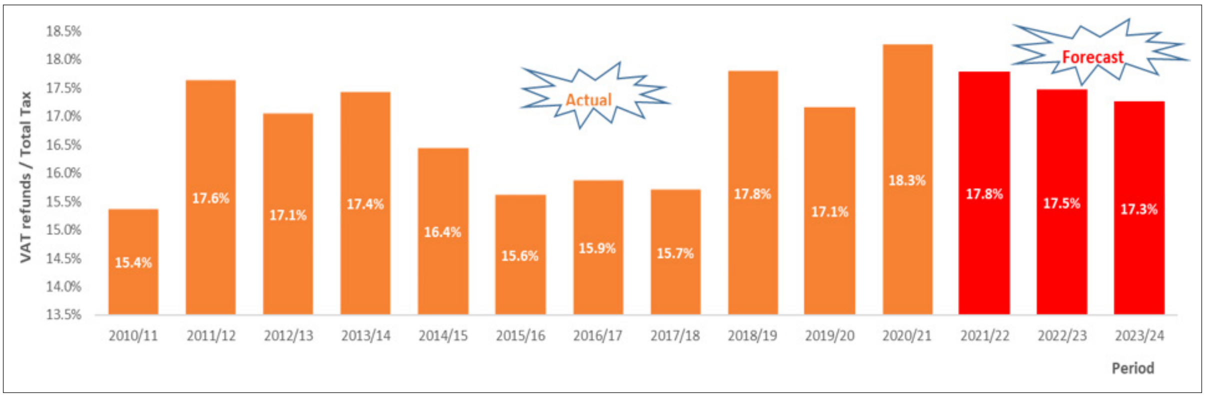
Table 4.7: VAT refunds forecast in rand million

Fiscal Year	Refunds
2021/2022	242 671
2022/2023	254 573
2023/2024	267 287

Source: SARS and Author's computation

The medium term estimates revenue collections to be approximately R1 365.1 trillion, while the forecast indicated a total amount of R1 457.6 trillion and R1 548.5 trillion for the period 2021/22 and 2023/24, respectively. Figure 4.2 below shows the ratio of VAT refunds to total tax for the period 2010/11 to 2020/21 (actual) and the forecast for three fiscal years 2021/22 to 2023/24 [4].

The actual realization of this ratio for the period 2010/11 to 2020/21 increases from 15.4% in 2010/11 to 18.3% in 2020/21. However, the forecast from the VAR mode1 shows a decrease of 17.8% in 2021/22 to a further slight decrease of 17.5% in 2022/23 and 17.3% 2023/24 in relation to the overall medium term estimate. The decrease of 17.8% in 2021/22 is similar to the projected decrease shown in the 2021/22 Printed Estimates from the South African National Treasury where the VAT refunds forecast are R241.2billion.



Source: Author's computation

Discussion

The VAR model is an econometric model which links current occurrences with the lagged values of the same variable and lagged values of other endogenous variables. The advantage of the time series VAR model used in this study is its power to test the economic interdependency or relationship between variables. The designed model examines the relationship between some macro-economic variables including gross domestic product at a sectorial level and VAT refunds. Compared to countries panel data model, this model claims superiority as it is country specific and accurately predicts VAT refunds over a certain time interval.

The fitted model showed that VAT refunds are positively influenced by domestic VAT, employment and agriculture GDP at 5% level of significance. Construction GDP is only significant at 10.3% level of significance. This means that any positive movement in this variable will result in increased VAT refunds. South Africa's economy has been stagnant in recent years, with growth rates of 0.79% and 0.15% recorded in 2018 and 2019 respectively, and heavily contracted in 2020 due to the COVID-19 pandemic showing a decline of 6.96%. With deterioration economic activities, jobs are lost and many businesses close down. This will not bode well for the growth in VAT refunds. If the current situation continues, VAT refunds will surely decrease.

The larger the domestic VAT collections, the higher the expectation of VAT refunds to be released to the vendors. For the past five year the registered total vendors increased from 742 388 in 2016/17 to 831 821 in 2019/20. The active vendors were 432 072 (58%) in 2016/17 and 449 597 in 2019/20 or 54% of total vendors for the same period [2]. However, the total refunds were R289.1 billion in 2016/17 increasing to R346.8 billion in 2019/20.

The construction sector in South Africa was affected by the lockdown regulations, hence the value added decreased by 20% in quarter 3 of 2020. The outlook remains subdued with respect to infrastructure investment. However, this sector anticipates growth through government's initiatives for capital spending [18]. The construction refunds increased from R5.7billion in 2016/17 to R6.9billion in 2019/20 fiscal year [2].

The agriculture, forestry and fishing sector recorded an annual gross value added of 11.3% for the first half of calendar year 2020 compared to prior year same period. The stronger growth was due to production recoveries in livestock, grain and horticulture [19]. The strong international exports demand is expected to continue supporting growth in production as the hard lock down is fading away. This sector's refunds increased from R12.7 billion in 2016/17 to R16.6 billion in 2019/20 fiscal year [2].

The fitted VAR model estimated VAT refunds for 2021/22, 2022/23 and 2023/24 to be R242.7 billion, R254.6 billion and R267.3 billion respectively. This forecast represents a slight decrease from 17.8% to 17.3% ratio of VAT refunds to total tax revenue forecast for 2021/22 and 2023/24 respectively.

The limitation of this model is that it assumes the continuation of some historical patterns to the future, accepting some system faults to also continue into the near future. The current model is a data end receiver model and can adjust the forecast for the rectified system as more data becomes available. However, this model is frequently applied on the aggregated data (to capture data patterns) rather than cash flow or transactional data.

The estimated level of VAT refunds can serve as an important consideration in the national budgeting processes in South Africa. Adequate provisions can be made to enable proper planning and distributions to government departments. However, more focus on tax compliance is also another dimension that the South African government can venture into to increase revenue collection, minimising tax avoidance and increasing the tax base [20, 21].

Conclusion and Recommendation

The objective of this research was to estimate VAT refunds in South Africa using the VAR model. This multivariate time series model was able to capture the relationship between macroeconomic variables (domestic VAT, employment, construction and agriculture GDP) and VAT refunds.

Like many other models, the reliability of the fitted VAR model was derived by observing its performance on the in-sample data. Though there was some under performance in the latter part of the test data, the model performs better, reducing the error margin to 1.36% for fiscal year 2020/21. Over 97% of the variation of refunds is being explained by the chosen endogenous variables.

Furthermore, two diagnostic tests were carried out to check the fitness and reliability of the model i.e. the Lagrange-multiplier for serial correlation in residuals and Jarque-Bera test for white noise residuals. The constructed model passed the diagnostic test hence permitting it to be used in forecasting the continuation of the historical VAT refunds.

The model results indicate that domestic VAT has a positive effect on VAT refunds at 5% level of significance at lag 1 and 2, employment at lag 2 and agriculture GDP at lag 1. However, construction GDP shows a positive effect on VAT refunds at lag 1 only at 10.3% significance level. The model estimates VAT refunds for the period 2021/22 to be R242.7 billion, while the estimates for the period 2022/23 and 2023/24 amounts to R254.6 billion and R267.3 billion, respectively. On average, VAT refunds make up 17.5% of the total tax for the forecast period of 2021/22-2023/24.

High VAT refunds can either be an indication of a growing economy or a high indication of fraud. The South African tax authority should have sound VAT systems, able to detect any unusual VAT activities, based on improved voluntary compliance through effective systems of self-assessment. Thus, beyond the current study, we recommend that compliance studies be undertaken to better understand tax administration issues (including refunds), thereby increasing the tax base.

Appendix

Table 4.8: Detailed Model Results

lnref	Coef.	Std. Err.	z	P>|z|	95% Conf.	
					LL	UL
lnref						
L1.	0.107	0.111	0.960	0.335	-0.111	0.325
L2.	-0.038	0.116	-0.320	0.747	-0.266	0.191
L3.	0.304	0.111	2.750	0.006	0.087	0.521
L4.	-0.030	0.114	-0.260	0.795	-0.253	0.194
lndomvat						
L1.	0.789	0.376	2.100	0.036	0.051	1.527
L2.	-0.432	0.409	-1.060	0.291	-1.233	0.370
L3.	1.453	0.542	2.680	0.007	0.392	2.515
L4.	-1.262	0.445	-2.840	0.005	-2.135	-0.390
lnempl						
L1.	-0.961	0.683	-1.410	0.159	-2.300	0.377
L2.	1.883	0.864	2.180	0.029	0.190	3.577
L3.	-0.068	1.174	-0.060	0.954	-2.369	2.233
L4.	-0.757	0.975	-0.780	0.438	-2.667	1.154
lnConstrGDP						
L1.	0.443	0.272	1.630	0.103	-0.090	0.976
L2.	-0.232	0.316	-0.730	0.463	-0.851	0.388
L3.	0.201	0.357	0.560	0.573	-0.499	0.901
L4.	-0.376	0.232	-1.620	0.106	-0.831	0.079
lnAgricGDP						
L1.	0.179	0.071	2.520	0.012	0.040	0.319
L2.	-0.031	0.066	-0.480	0.634	-0.161	0.098
L3.	0.034	0.067	0.510	0.612	-0.097	0.165
L4.	-0.010	0.066	-0.150	0.883	-0.138	0.119
Dum1	0.125	0.079	1.590	0.111	-0.029	0.279
Constant	-2.234	3.997	-0.560	0.576	-10.068	5.601

Source: Authors computation, L - Lag, LL - Lower limit, UL – Upper limit, Coef – Coefficient

References

1. SARS (2015) Legal Counsel: Value-Added Tax', Guide for vendors. Available from URL https://www.sars.gov.za/legal-counsel/legal-counsel-publications/find-a-guide/value-added-tax-vat.
2. Tax Statistics (2011-2020) Available from URL https://www.sars.gov.za/about/sas-tax-and-customs-system/tax-statistics.
3. StatsSA (2021) Available from URL https://www.statssa.gov.za/?s=gdp&sitem= publications.
4. National Treasury (2020) Available from URL https://www.treasury.gov.za/mtbps/ FULLMTBPS.
5. Madzivanyika E, Kadenge PG (2015) VAT incentives and their effects on VAT refunds: the case of Zimbabwe (2009-2012). Public and Municipal Finance 4: 37-44.
6. Harrison G, Krelove R (2005) VAT Refunds: A Review of Country Experience, IMF Working Paper (WP/05/218).
7. Bikas E, Andurskaite E (2013) Factors affecting Value added tax revenue. 1st Annual International Interdisciplinary Conference. AIIC p- 24-26.
8. Zhuravleva IA (2021) Way to improve the Russian model of VAT refunds on goods exports. European proceedings of social and behavioural sciences (EpSBS).

9. Abdullah MN, Parvez K (2012) Effect of VAT and tax on economy: An analysis in the context of Bangladesh. Research journal of finance and accounting 3: 64-70.
10. Chidbu EE, Ali PI (2014) Econometric analysis of the Impact of Value added tax on economic growth in Nigeria. European journal of business and management 6: 32-36.
11. Aptech (2021) Available from URL https://www.aptech.com/blog/introduction-to-the-fundamentals-of-vector-autoregressive-models/.
12. Bauer (2021) Available from URL https://www.bauer.uh.edu/rsusmel/phd/ec p-2-6.
13. Kennedy P (2003) A guide to econometrics (5th ed.). United Kingdom: MPG Books.
14. Bowman KO, Shenton LR (1975) Omnibus test contours for departures from normality based on b1 and b2. Biometrika 62: 243-250.
15. Hall RE, Lilien DM (1995) EViews User Guide. Irvine: Quantitative Micro Software.
16. Johansen S (1995) Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.
17. Katara H (2021) Available from URL https://www.wallstreetmojo.com/r-squared-formula/.
18. Funke M (1990) Assessing the forecasting accuracy of monthly vector autoregressive models: The case of five OECD countries. International Journal of Forecasting 6: 363-378.
19. Budget review (2021) Available from URL https://www.google.co.za/url?sa=t&rct= j&q=&esrc=s&source=web&cd=&cad=rja&uact=8&ved=2ahUKEwi2oIzvqazzAhVHasAKHWjaDt4QFnoECBkQAQ&url=http%3A%2F%2Fwww.treasury.gov.za%2Fdocuments%2FNational%2520Budget%2F2021%2Fspeech%2Fspeech.pdf&usg=AOvVaw3M_Dx9-UQDYF2n7qXC009k.
20. IMF (2005) VAT refund; a review of country Experience. Working paper 05/218. Available from URL https://www.imf.org/en/Publications/WP/Issues/2016/12/31/ VAT-Refunds-A-Review-of-Country-Experience-18646.
21. Puspitawati L, Hartono WR (2019) The Effect of macro-economic variables on Value added tax (VAT). Advances in Economics, Business and Management Research 112: 214 -217.