

ECONOMIC BULLETIN ECONOMIC BULLETIN

Gauteng Province

**ECONOMIC
BULLETIN 2**

2024



GAUTENG
PROVINCIAL GOVERNMENT
REPUBLIC OF SOUTH AFRICA

GGT2030
GROWING GAUTENG TOGETHER

Establishing the assumptions of the cashflow model for forecasting Gauteng Provincial Government Own Revenue Collection

Abstract

This study delves into the intricacies of modelling the provincial own revenue estimates, with a specific focus on enhancing credibility and robustness within the context of the Gauteng Province. Building upon previous research methodologies, the study employs advanced statistical techniques, including Vector Autoregression (VAR) modelling, Cholesky impulse response analysis, and Granger causality testing. Through these methodologies, the study assesses the impact of various macroeconomic variables on revenue collection, thereby revealing the underlying drivers of provincial revenue streams.

Key findings reveal that factors such as the Household Disposable Income (HDI), Vehicle Sales (VS), Household Final Consumption Expenditure (HFCE), and Consumer Confidence Index (CCI) exert significant influence on specific revenue sources in both the short and long term. Notably, the study uncovers complex dynamics, including negative short-term relationships between total own revenue and HFCE, alongside delayed positive impacts in the long run.

Table of Contents

1. Introduction.....	3
2. Literature Review	3
3. Gauteng Government Revenue Trends	8
3.1. Correlation Analysis.....	9
3.2. Seasonality Analysis	9
4. Methodology.....	12
4.1. Variables	12
4.2. Vector Autoregressive Model (VAR).....	13
4.3. Choleskey Impulse response functions.....	14
4.4. Granger Causality Test.....	14
5. Results	15
5.1. Stationarity Test	16
5.2. Cointegration Test	17
5.3. The optimal lag selection	18
5.4. VECM Estimation Output.....	19
5.5. Cholesky impulse response	21
5.6. Granger Causality.....	23
5.7. Results Discussion	25
6. Conclusion.....	26
REFERENCES	28
Annexure A: VAR Model Stability Test.....	30
Annexure B: VECM Output	31
Annexure C: Scatter plots.....	33

1. Introduction

This economic bulletin aims to enhance the credibility of the selection of independent variables for the cashflow provincial own revenue forecasts for Gauteng. It assesses if the independent variables are suitable for inclusion in the model based on their relative correlations and causality with the main revenue streams.

The research focuses on the largest two revenue sources and tests the assumptions used to improve the credibility of the revenue forecasts. The provincial own revenue sources mainly comprise of fees collected for motor vehicle licenses, which make up about 74.8 per cent of total own revenue, gambling licenses (19.9 per cent)¹ and hospital patient fees (5.3 per cent) as per actual revenue collected in the 2022/23 financial year.

The forecast credibility of the model is dependent on the selection accuracy of the independent macroeconomic variables and their respective buoyancy ratios. Tax revenue scenarios can be simulated by changes to the forecasted independent macroeconomic variables and changes to buoyancy ratios. These are the two critical factors altering the forecasts. This study identifies the most suitable independent economic variables and uses the Vector Autoregressive (VAR) model to examine the effect of the selected macroeconomic variables.

Section 2 of the study provides the literature background on tax revenue forecasting techniques. Section 3 provides a correlation and seasonality analysis of provincial own revenue sources. This is followed by Sections 4 to 6, where the model description is provided, data preparations are outlined, and the estimation results of the models are discussed.

2. Literature Review

The Cash Flow forecast model estimates tax revenues by tax type, based on the actual collection of own revenues to date and the relationship that exists between categories of own revenue and the underlying macroeconomic independent variable. The own revenue category is the

¹ Gambling licenses are made up of 5.2 per cent for horse racing and 14.6 per cent for casino taxes.

dependent variable in the equation and the independent variables include, Household Final Consumption Expenditure (HFCE), Household Disposable Income (HDI), motor Vehicle Sales (VS), and the Consumer Confidence Index (CCI). The model equation below is applied to forecast the own revenue:

$$Y_t = c + \beta_i X_t$$

Where: Y_t is the dependent variable, namely the own revenue of Gauteng, c is a constant term, β_i is the coefficient in the model and X_t is an independent/explanatory variable (i.e., GDP). The β_i in the model measures the relationship between relative changes in actual own revenue collection to changes in the proxy economic tax base, i.e., GDP, which essentially, is the tax buoyancy ratio (Jenkins, Kuo and Shukla, 2000). The model calculates the current year buoyancy ratio as a scenario.

Forecasting of tax revenue forms the core basis for the public purse and comprehensive medium-term budget policy. Jenkins, et al. (2000) defines tax buoyancy as the response of tax revenue to changes in national income and to changes in the tax system, including changes in tax rates. The paper defines tax elasticity as the response to changes in national income alone. It asserts that by these definitions, tax buoyancy is the appropriate tool for estimating the effect of changes in tax policy and tax elasticity for forecasting tax revenue for budget purposes. Equation 1 below expresses tax buoyancy as follows:

Equation 1

$$E_{TY}^b = \frac{\Delta T}{\Delta Y} * \frac{Y}{T^b}$$

Where:

E_{TY}^b = Buoyancy of tax revenue income

T^b = Total tax revenue

ΔT^b = Change in total tax revenue

Y = Income

ΔY = Change in income/GDP

Tax elasticity as:

Equation 2

$$E_{rr} = \frac{\% \Delta T}{\% \Delta Y}$$

Where:

E_{rr} = *Elasticity of tax revenue to income or GDP*

ΔT = *Change in Revenue*

ΔY = *change in income/GDP*

It is stated that in both cases, disaggregation will usually result in more accurate predictions, though sufficient data must be present to make disaggregation feasible. For tax buoyancy, the example provided is disaggregating it into different types of taxes, such as income tax and sales tax. For tax elasticity, the example is forecasting different sectors of the economy separately, even forecasting individual large businesses separately if data allows.

The paper uses what it calls microsimulation models to predict the effects of tax policy changes on different groups of people, differentiating across such factors as income group or number of children per household. This type of modelling increases the data requirements, as data must exist to differentiate taxpayers along the desired differences while matching groups to the taxes they pay. Furthermore, sufficient data must exist for statistically viable sample sizes within each group. For example, if one wishes to differentiate taxpayers by 3 income groups, 4 groups by number of children and 10 groups by average yearly expenditure on transport, multiplying these groupings together produces 120 possible combinations. Each of the 120 possible combinations of these factors must individually yield enough households to generate enough observations for forecasting purposes. When possible, differentiated forecasts allow policy makers to assess the likely impact of policy changes more accurately upon selected groups of people, such as the poor and vulnerable.

The Division of the Budget (DOB) of the New York state (NY) government within the United States of America (US), released a paper detailing the methodology it follows when forecasting the revenue that NY will receive in future (Division of the Budget, New York, 2004). This begins with a forecast of the US economy because NY is both affected by the broader US economy and accounts for a large share of it. Focus is also placed on the finance and business services industry because it accounts for a large percentage of NY's Gross Domestic Product by Region (GDP-R) and employment. Thus, changes in that industry greatly impact NY residents' ability to pay taxes and make taxable purchases.

The DOB notes that economic realities are complex, and no model can fully capture all the intricacies of these processes, thus, there will always be uncertainties and unpredictability. As such, economic forecasts are often calculated with a confidence interval, a range of numbers within which the future outcome is likely to fall. When reporting to stakeholders, this is most often simplified to a point estimate, a single number around which the range extends. The mid-point of the confidence interval is most often selected to be the point estimate that is reported. The DOB, however, asserts that this ignores that the cost incurred with forecast error is asymmetric. In this case, overestimating future receipts is more troublesome than underestimating them. This is because departments finding themselves unable to fund the projects they have initiated, is more of a problem than the government finding itself with more funds than budgeted.

The DOB gave an important role to expectations in their model when forecasting the behaviour of consumers and businesses. Consumers are assumed to take a long-term view of their own income when making decisions, while the business sector is assumed to consider predictions of such factors as prices and interest rates. However, they are also assumed to make changes gradually only when expectations change, and this is due to the habits of households and the institutional inertia of firms. Low-income households are an exception and are assumed to base their decisions on current conditions because they typically do not have reserves to draw upon and their challenges often require more immediate solutions. Economic agents in the financial sector are another exception, as they are assumed to adjust immediately when they receive new information because this type of rapid response is typical in the finance industry.

From the year 2000 to 2017, the median tax on cigarettes among the various US states rose from 34 US cents to 157 cents (Pew Charitable Trusts, 2018). There are typically two motives for increasing this tax, improving public health by discouraging the smoking of cigarettes and collecting more money from those who continue despite the higher cost. The paper notes that these goals are contradictory because reducing cigarette consumption reduces government revenue from taxing their sale. This applies to all so-called sin taxes on price-sensitive products. It was found that tobacco use has fallen by a large enough degree that tax income from tobacco taxes has fallen despite higher per-unit taxes. Alcohol being a much more popular product has led to it being more resilient to price increases and total tax revenue from alcohol has increased over time instead of declining. The tax rate on alcohol sales has also increased more slowly due to its popularity, as the average voter would more greatly resent the government for increasing the tax on alcohol than the tax on tobacco. Recreational marijuana has been a fruitful source of tax revenue for the states that have legalised it, but the legal market is too new to draw strong conclusions about long-term revenue potential.

The paper suggests that sin taxes can be employed to discourage activities that are deemed harmful to the public good and to fund short-term government projects but should not be used to fund recurring expenses or otherwise relied upon. This is due to their volatility and the long-term tendency of sin tax revenue to fall as they succeed at discouraging the taxed activity. Even the mostly reliable alcohol tax has a degree of volatility due to changes in culture, both in terms of the amount consumed and which alcohols are considered fashionable. Division of the Budget, New York, 2004, found that these changes in alcohol culture were largely cyclical and could thus be approximately predicted by running regressions of the consumption of spirits, beer, and wine against their own lags.

H2 Gambling Capital (2020) asserts that higher taxes and increased regulation on gambling in Denmark decreased the quantity of online gambling being conducted in that country's online gambling websites. This is because consumers can easily use websites based in other countries for online gambling if higher taxes cause Denmark's sites to offer smaller prizes relative to buy-in. This was of concern not only because of reduced tax revenue for the Danish government. Many countries have fewer consumer protections in place for online gambling and the paper asserts that this market shift would thus put Danish gamblers at risk. The paper therefore argues

for a smaller gambling tax increase than the one proposed at the time, to reduce the size of the move from Danish online gambling to the websites of other countries. This phenomenon can be extrapolated to other sin taxes to an extent. For example, persons living near a border can potentially purchase alcohol or tobacco products from stores in the neighbouring country if domestic taxes push prices significantly higher than next door. This could potentially cause health risks if safety standards applied to these products are looser in the neighbouring country.

Gambling is a large source of sin tax revenue in the US states that allow it. However, that revenue has stagnated because opening new casinos provides only a short-term boost to tax revenue while the casino is still new. In the medium- to long-term, gamblers reduce their purchases of lottery tickets to fund their casino gambling. This results in casinos being largely revenue-neutral for state governments. In several states, taxing e-cigarettes at the same rate as tobacco products caused a chilling effect on the legal market due to the high tax burden on tobacco. This led to black markets being created as some consumers sought to evade these taxes while others were simply finding it challenging to legally access a product where many of the legal businesses had shut down due to the tax increase. This outcome was not predicted ahead of time because the e-cigarette market had existed for too short a time to generate enough data to make reliable forecasts.

3. Gauteng Government Revenue Trends

In this section, the correlation and seasonality analysis of the provincial own revenue sources is presented. Using these tools, we gain deeper insights into the variability of revenue collection and make more informed recommendations for policy decisions.

3.1. Correlation Analysis

Table 1 below is a correlation matrix of own revenue sources and the selected macroeconomic variables from April 2009 to March 2023.

Table 1: Correlation Matrix

	VS	MOTOR	HFCE	HDI	GEMP	CCI
VS	1.000000	0.784880	0.915326	0.908806	0.587016	-0.462505
MOTOR	0.784880	1.000000	0.845496	0.846879	0.481063	-0.464961
HFCE	0.915326	0.845496	1.000000	0.999463	0.530083	-0.568976
HDI	0.908806	0.846879	0.999463	1.000000	0.514726	-0.568866
GEMP	0.587016	0.481063	0.530083	0.514726	1.000000	-0.200136
CCI	-0.462505	-0.464961	-0.568976	-0.568866	-0.200136	1.000000

Source: Author's Computation based on Bloomberg & Quantec Research Information

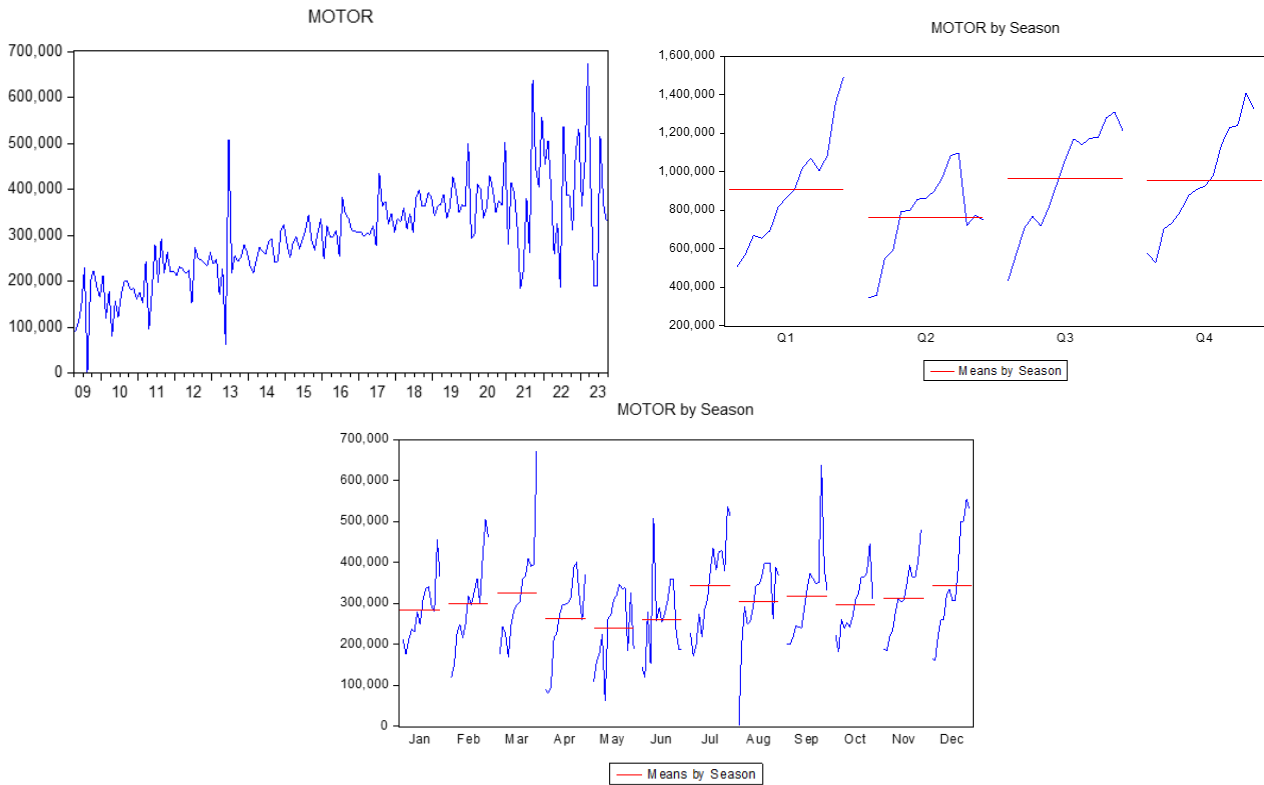
Motor vehicle licenses are highly positively correlated to HFCE, HDI and VS. In the case of GEMP, there is a weak positive correlation and a weak negative correlation with CCI.

3.2. Seasonality Analysis

The seasonal effects present in the own revenue sources are presented in the following section and a discussion of the observed trends, along with an explanation for the deviations. As shown below, there is evidence of seasonality in MVL and casino taxes revenue collection. MVL revenue collection starts low on average during the second quarter of the calendar year, which is the first quarter of government fiscal year (April to June).

MVL revenue collection are the lowest (trough) in May. MVL revenue collection increase to their peak in July and are the largest during the third quarter of the calendar year (2nd quarter in the fiscal year). They moderate during the final quarter of the calendar year and once again in the 1st calendar quarter (last quarter of the fiscal year).

Figure 1: Motor License Revenue Seasonality Trends – April 2009 to March 2023

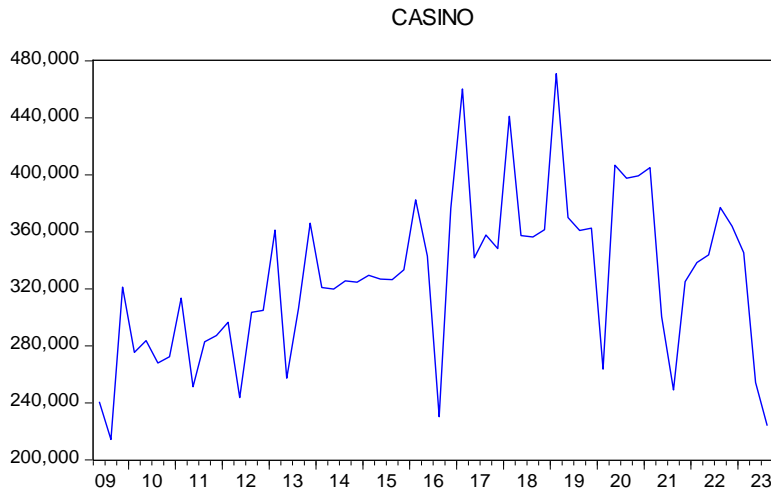


Source: GPT, 2024.

Revenue from MVL is the main source of revenue and shows a higher collection than other sources over the period under review.

Further analysis was conducted on the seasonality patterns in casino own revenue collections. Casino taxes are the second largest revenue generator for GPG. The figure below plots the collected revenue over time, quarterly, from the second quarter of 2009 to the third of 2023, in calendar years. To gain better understanding of the underlying trends present in the data and improve forecast capacity, seasonality analysis of the data was conducted using the EViews software.

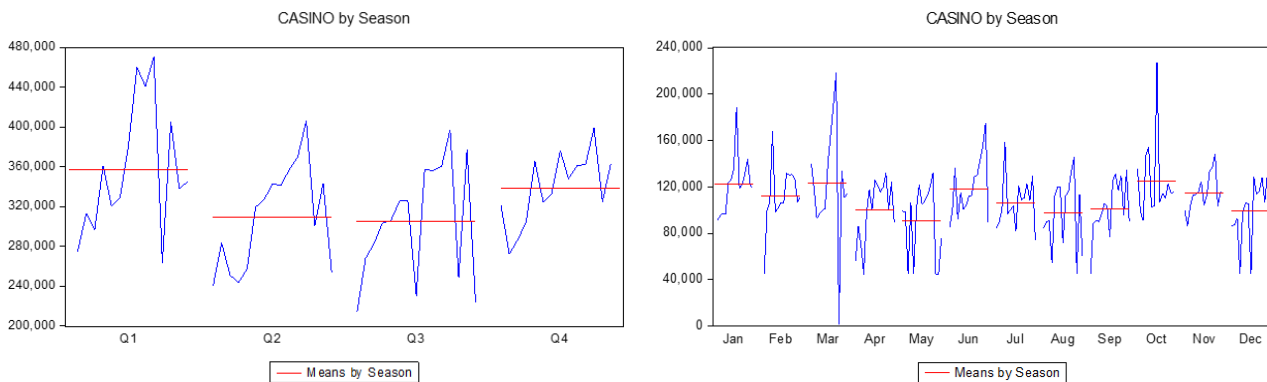
Figure 2: Casino Revenue Seasonality Trends – April 2009 to March 2023



Source: GPT, 2024.

As shown below, there is evidence of seasonality in casino revenue collection within each fiscal year. Collection starts low on average during the second quarter of the calendar year April to June. Generally, casino collections retreat to their lowest (trough) during the third quarter of the calendar year (second quarter of the fiscal year). They accelerate during the fourth quarter of the calendar year (third quarter of the fiscal year) and rise to their peak in the first quarter of the calendar year (last quarter of the fiscal year).

Figure 3: Casino Revenue Seasonality Trend by Season – April 2009 to March 2023



Source: GPT, 2024.

Note: Fiscal Year refers to April – March annually, begins start of calendar year Q2 and ends after Q1

The conclusion is that casino revenue starts off low, dips and then accelerates as the "financial" year progresses. For the casino revenue collection, it should be noted that the seasonality observed is attributed to the festive season, this is to say during the period of school holidays,

the entity normally realises an increase in terms of casino taxes. The festive season has a significant impact on the casino taxes.

4. Methodology

This study uses the Vector Autoregressive (VAR) model to examine the effect that the selected macroeconomic variables (household final consumption expenditure, household disposable income, vehicle sales, and consumer confidence index) have on the Gauteng Provincial Government (GPG) revenue sources of casino taxes, motor vehicles license taxes, and the total provincial own receipts. The data employed in the model was sourced from GPG In-Year Monitoring financial revenue database, Bloomberg, and Quantec Easydata. The frequency of the data is quarterly, spanning from 2008Q2 to 2023Q3. All the variables were transformed into a logarithm scale to improve the stability of the data, especially the revenue data.

4.1. Variables

The variables used in this study are listed below in Table 2. The parentheses define how the variables in the model estimation will be expressed.

Table 2: Variables

Dependent Variables	Description
Casino Revenue (Casino)	Nominal quantitative data was recorded and received as revenue by GPG from casinos registered under the Gauteng Gambling Board establishments.
Motor Licenses Revenue (Motor)	Nominal quantitative data was recorded and received as revenue by GPG from registered vehicles in Gauteng.

Total Own_Receipts Revenue (Own_Receipts)	Nominal quantitative data was recorded and received as revenue by GPG, which is the total revenue from different revenue sources due to GPG.
Independent Variables	Expected Sign
Household Disposable Income (HDI)	Nominal quantitative data. HDI is expected to show a positive sign in the estimation.
Household Final Consumption Expenditure (HFCE)	Nominal quantitative data. HFCE is expected to show a positive sign in the estimation.
Consumer Confidence Index (CCI)	Quantitative data is measured as an index. CCI is expected to show a positive sign in the estimation.
Total Vehicle Sales (VS)	Nominal quantitative data. VS is expected to show a positive sign in the estimation.

4.2. Vector Autoregressive Model (VAR)

Christopher Sims (1980) created the Vector Autoregression (VAR) model in 1980. The VAR model evaluates econometric models that represent linear equations of each variable. VAR allows for the observation of how a variable is influenced by its own and other variables' past values (Sims, 1980). Instead of just looking at how a variable is affected by different variables simultaneously, it also shows how the variable was affected by those same variables in the past. This multi-directional relationship between variables and their lagged variables enables the observation of variable behaviours without imposing any functional relationship among them (Christiano, 2012). Hence, Atolagbe & Abiodun (2021) were able to study the impact of macroeconomic variables they had selected in their study to assess the impact of the variables on the tax revenue in Nigeria using a VAR model. Therefore, this study will model the three revenue sources collected by the GPG and the selected macroeconomic variables using a VAR model, as expressed in Equation 3 below.

Equation 3: VAR model

$$\begin{bmatrix} casino_t \\ motor_t \\ own_receipts_t \\ HDI_t \\ HFCE_t \\ CCI_t \\ VS_t \end{bmatrix} = c + \sum_{j=1}^p A_j + \begin{bmatrix} casino_{t-j} \\ motor_{t-j} \\ own_receipts_{t-j} \\ HDI_{t-j} \\ HFCE_{t-j} \\ CCI_{t-j} \\ VS_{t-j} \end{bmatrix} + \varepsilon_t$$

Where:

- Y_t is a vector of all variables at time t , including casino, motor, own_receipts, HDI, HFCE, CCI, and VS.
- c is a vector of constants.
- A_1, A_2, \dots, A_p are coefficient matrices.
- ε_t is a vector of error terms.

4.3. Cholesky Impulse response functions

To enhance the VAR model, this study will incorporate impulse response functions (IRF) to mitigate information loss resulting from the inconsistency inherent in VAR forecasting, which can be challenging with its large multivariate systems (Enders, 2014). Additionally, the use of IRF will aid in understanding the gradual adjustments of the revenue sources following a shock in HDI, HFCE, CCI and VS.

4.4. Granger Causality Test

The Granger causality test established by Granger (1969) examines the causal relationship between two time-series variables. The Granger-causality test will assist in determining the causal relationship between the GPG revenue sources and the macroeconomic variables. The Granger Causality test economic approach is outlined below in Equation 4 and Equation 5.

Equation 4

$$Y_t = \sum_{k=1}^n \phi_k X_{t-k} + \sum_{k=1}^t \alpha_k Y_{t-k} + \epsilon_t$$

Equation 5

$$X_t = \sum_{k=1}^p \beta_k X_{t-k} + \sum_{a=1}^n \theta_a Y_{t-a} + \mu_t$$

Where ϵ_t and μ_t are uncorrelated white noise processes, ϕ_i , α_i , β_i , and θ_i are coefficients in the model, and m and n are the numbers of lags.

In this case, Y_t represents the response variable, and X_t is the explanatory variable

5. Results

The results section consists of a stationarity test section, which covers the stationarity results determined using the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. Furthermore, the optimal lag selection was used to establish the optimal lag length of the model. Roots of the characteristic polynomial are used to test the estimated VAR model for stability. The IRF and Granger Causality tests from the VAR estimation were used to observe the response of dependent variables and causality direction, respectively.

5.1. Stationarity Test

This sub-section tests for stationarity using the ADF and PP tests. The variables are considered stationary when their respective p-values from ADF or PP tests are less than 10%.

Table 3: ADF and PP Stationarity Test Results

Variables	H0: non-stationarity at levels		H0: non-stationarity at first differences	
	ADF test	P.P. test	ADF test	P.P. test
LNCasino	-5.221859*** (0.0003)	-5.163759*** (0.0004)	-6.622396*** (0.0000)	-21.59839*** (0.0001)
LNMotor	-1.718666 (0.7303)	-6.159827*** (0.0000)	-15.69025*** (0.0000)	-29.50673*** (0.0001)
LNOwn_receipts	-2.283209 (0.4358)	-6.215788*** (0.0000)	-5.991143*** (0.0000)	-44.40741*** (0.0001)
LNCCI	-5.673980*** (0.0001)	-5.673980*** (0.0001)	-8.351449*** (0.0000)	-21.57559*** (0.0001)
LNHFCE	-3.615930** (0.0366)	-3.487805** (0.0497)	-11.33369*** (0.0000)	-15.44848*** (0.0000)
LNHDI	-3.963925** (0.0151)	-3.885358** (0.0186)	-11.81343*** (0.0000)	-20.26578*** (0.0001)
LNVS	-4.470103*** (0.0036)	-4.422344*** (0.0042)	-6.880738*** (0.0000)	-20.17049*** (0.0001)

Source: Author's Computation based on Bloomberg & Quantec Research Information. *10%, ** 5%, 1%*** level of Significance;()=P-Value

At level, all variables are stationary except for two variables, LNMotor and LNOwn_receipts. At first difference, all the variables are stationary at a 1 per cent confidence level. Since there are non-stationary variables at level while others are stationary, the econometric approach is to test for cointegration amongst the variables (Johansen, 1991).

5.2. Cointegration Test

Table 43 shows the Johansen cointegration test, which was calculated using the VAR model estimation.

Table 4: Johansen Cointegration Test Results

Unrestricted Cointegration Rank Test (Trace)				
Hypothesised No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.693106	155.3521	125.6154	0.0002
At most 1	0.423976	85.65830	95.75366	0.2019
At most 2	0.323617	53.11350	69.81889	0.5002
At most 3	0.177312	30.04472	47.85613	0.7167
At most 4	0.135268	18.52922	29.79707	0.5272
At most 5	0.096851	9.954427	15.49471	0.2842
At most 6 *	0.064666	3.944254	3.841466	0.0470

Source: Author's Computation based on Bloomberg & Quantec Research Information.

Trace test indicates 1 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

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

According to the cointegration test, the VAR model has at least one equation with a pair of cointegrated variables. Cointegration implies a long-term relationship between variables, suggesting they move together in the long run despite short-term fluctuations.

Therefore, as per Johansen's (1991) recommendation on how to treat a VAR model with cointegrated variables, an error correction model will be used to mitigate the potential of running a spurious regression. In the case of a VAR model, the error correction model will be the Vector Error Correction Model (VECM).

5.3. The optimal lag selection

A VAR lag order selection criterion will be used to estimate a VAR with an appropriate and optimal lag length, and the results are represented in Table 4.

Table 5: VAR Order Selection

VAR Lag Order Selection Criteria
 Endogenous variables: D_LNCasino D_LNMotor D_LNOwn_receipts D_LNCCI D_LNHDI D_LNHCFE
 D_LNVS
 Sample: 2008Q2 2023Q3
 Included observations: 56

Lag	LogL	LR	FPE	AIC	SC	HQ
1	543.2238	NA	5.14e-17	-17.65085	-15.87867*	-16.96378*
2	603.3903	90.24978	3.69e-17	-18.04965	-14.50529	-16.67551
3	677.6890	92.87342*	1.81e-17*	-18.95318	-13.63663	-16.89197
4	725.3869	47.69781	2.89e-17	-18.90667	-11.81794	-16.15839
5	805.1027	59.78690	2.20e-17	-20.00367*	-11.14275	-16.56831

Source: Author's Computation based on Bloomberg & Quantec Research Information.

*Denotes the lag order chosen by the criterion. Sequential modified LR test statistic (each test at a 5% level); Final prediction Error (FPE); AIC, SC, and HQ stand for the Akaike, Schwarz, and Hannan-Quinn information criteria, respectively.

Furthermore, focusing on the three optimal lag lengths as per the serial correlation test, results in Table 6 show no evidence of serial correlation on lag lengths of 1,3 and 4. Due to the P-values being greater than 5 per cent, the null hypothesis of no serial correlation is not rejected.

Table 6: VAR Residual Serial Correlation LM Tests

Sample: 2008Q2 2023Q3

Included observations: 56

Null hypothesis: No serial correlation at lag h

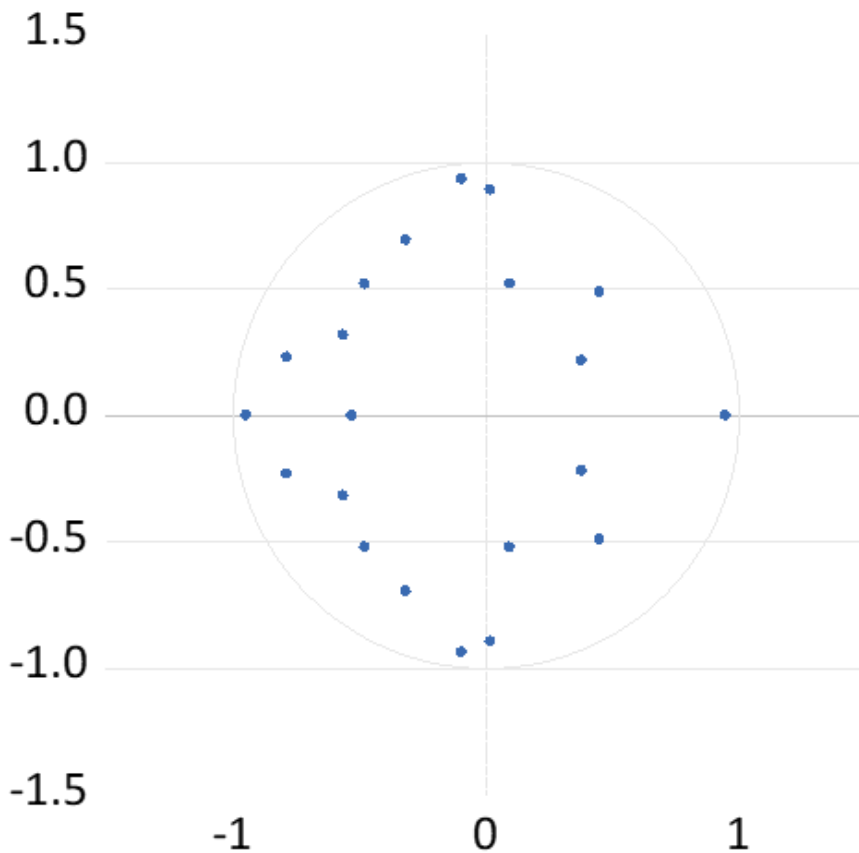
Lag	LRE* stat	Probability.
1	42.81276	0.7338
2	72.22328	0.0199
3	49.53311	0.4692
4	42.58124	0.7421
5	66.86370	0.0514

Source: Author's Computation based on Bloomberg & Quantec Research Information.

Since there are three lag lengths with no serial correlation, lag length three was chosen as an optimal lag length to estimate the VECM models. Firstly, the selection is based on the fact that the optimal lag selection criterion in Table 5 had two significant selection criteria: the LR and FPE. Secondly, from the estimation of the two models tested for stability, it was found that they are stable at an estimation of lag length 3. The Roots of Characteristic Polynomial test gave stability accession, showing that all the roots are inside the unit circle, as shown in Annexure A:

VAR	Model	Stability	Test
-----	-------	-----------	------

Inverse Roots of AR Characteristic Polynomial



. Furthermore, Stock & Watson (2001) also used the same rationale by assessing the optimal lag length and testing for the VAR model stability.

5.4. VECM Estimation Output

Table 6 Represents the VECM model estimation, where the first columns represent lagged explanatory variables, and the first row of the headings represents the response variables. The number inside the square brackets represents the t-statistics of the estimated coefficient for each variable in the model.

Table 7: VECM Model Significant Variables Summary

	D(D_LNCasino)	D(D_LNMotor)	D(D_LNOwn_receipts)
D(D_LNCasino(-1))	—	—	[2.65710]
D(D_LNMotor(-1))	[-2.08338]	[-5.67961]	—
D(D_LNMotor(-2))	—	[-5.24830]	[-2.86265]
D(D_LNMotor(-3))	[0.48588]	[-4.86956]	[-3.82624]
D(D_LNOwn_receipts(-1))	[3.10364]	—	—
D(D_LNOwn_receipts(-2))	—	[2.83396]	—
D(D_LNOwn_receipts(-3))	—	[2.84764]	[2.04043]
D(D_LNHDI(-1))	[2.43552]	—	[2.28000]
D(D_LNHFCE(-1))	[-3.73388]	—	[-2.73471]
D(D_LNHFCE(-2))	[-2.57135]	—	—
D(D_LNHFCE(-3))	[-2.69408]	—	—
D(D_LNVS(-1))	[3.62973]	—	[2.51768]
D(D_LNVS(-2))	[3.87852]	—	[2.18180]
D(D_LNVS(-3))	[2.94127]	[-2.35262]	

Source: Author's Computation based on Bloomberg & Quantec Research Information.

Note= (-1) a variable with one lag, (-2) a variable with two lags, (-3) a variable with three lags.

This study focuses mainly on observing how the selected macroeconomic variables impact revenue variables, namely D_LNCasino, D_LNMotor and D_LNOwn_receipts. Moreover, variables in Table 7 were deduced from the VECM, and only significant variables were considered based on a t-statistic of 1.962 or greater. D(D_LNHDI(-1)) has a significant positive relationship with D_LNCasino and D_LNOwn_receipts. D(D_LNHFCE(-1)) has a negative relationship with D_LNCasino and D_LNOwn_receipts. D(D_LNHFCE(-2)) and D(D_LNHFCE(-3)) only have a negative relationship with D_LNCasino. D(D_LNVS(-1)) and D(D_LNVS(-2)) each have a significant positive relationship with D_LNCasino and D_LNOwn_receipts. D(D_LNVS(-3)) has a positive relationship with D_LNCasino and a negative relationship with D_Inmotor.

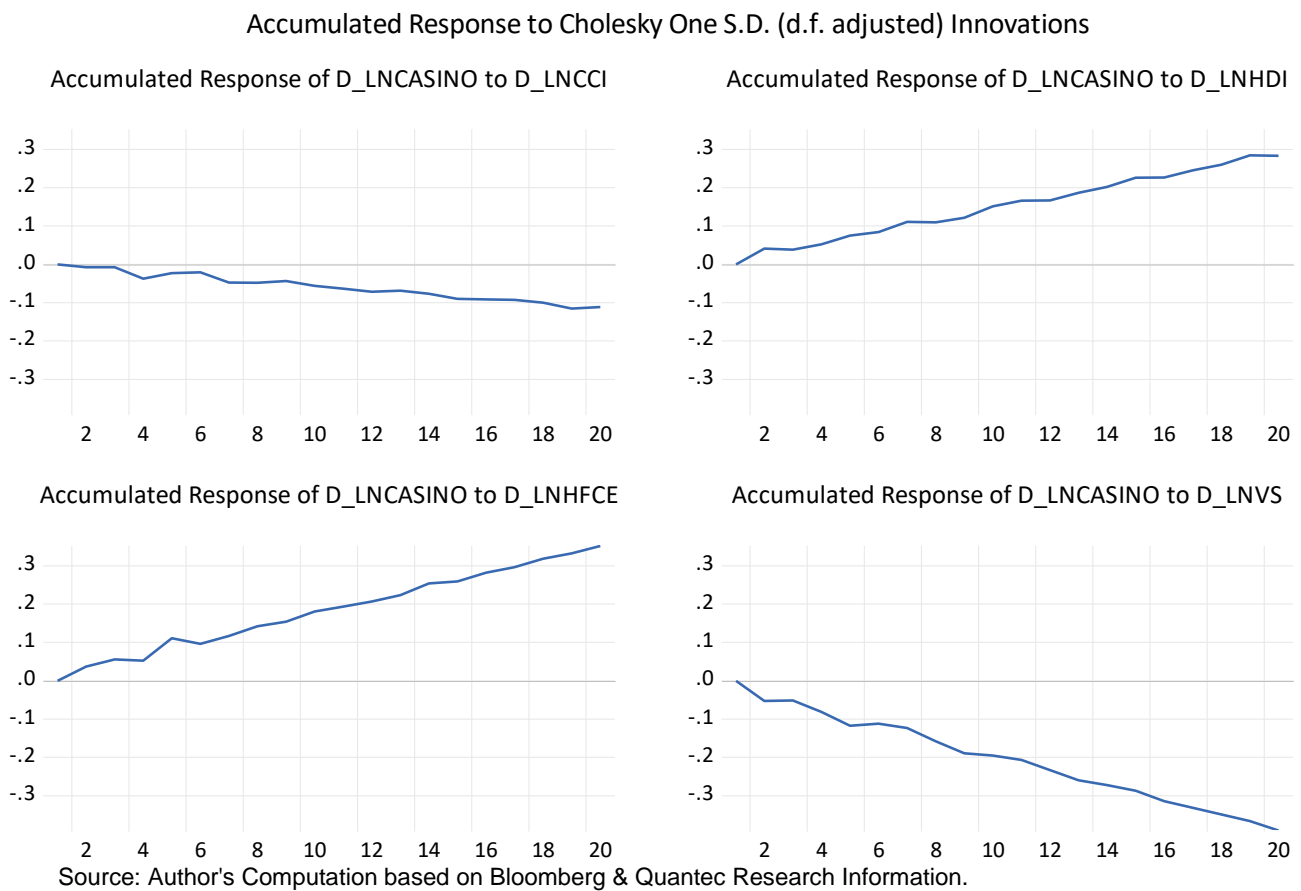
A t-statistic of 1.96 represents a 5 per cent level of significance.

5.5. Cholesky impulse response

This section shows the impulse response of the revenue streams, namely D_LNCasino, D_LNMotor and D_LNOwn_receipts, to a standard deviation respective increase shock in D_LNCCI, D_LNHDI, D_LNHFCE, and D_LNVS over 20 quarters.

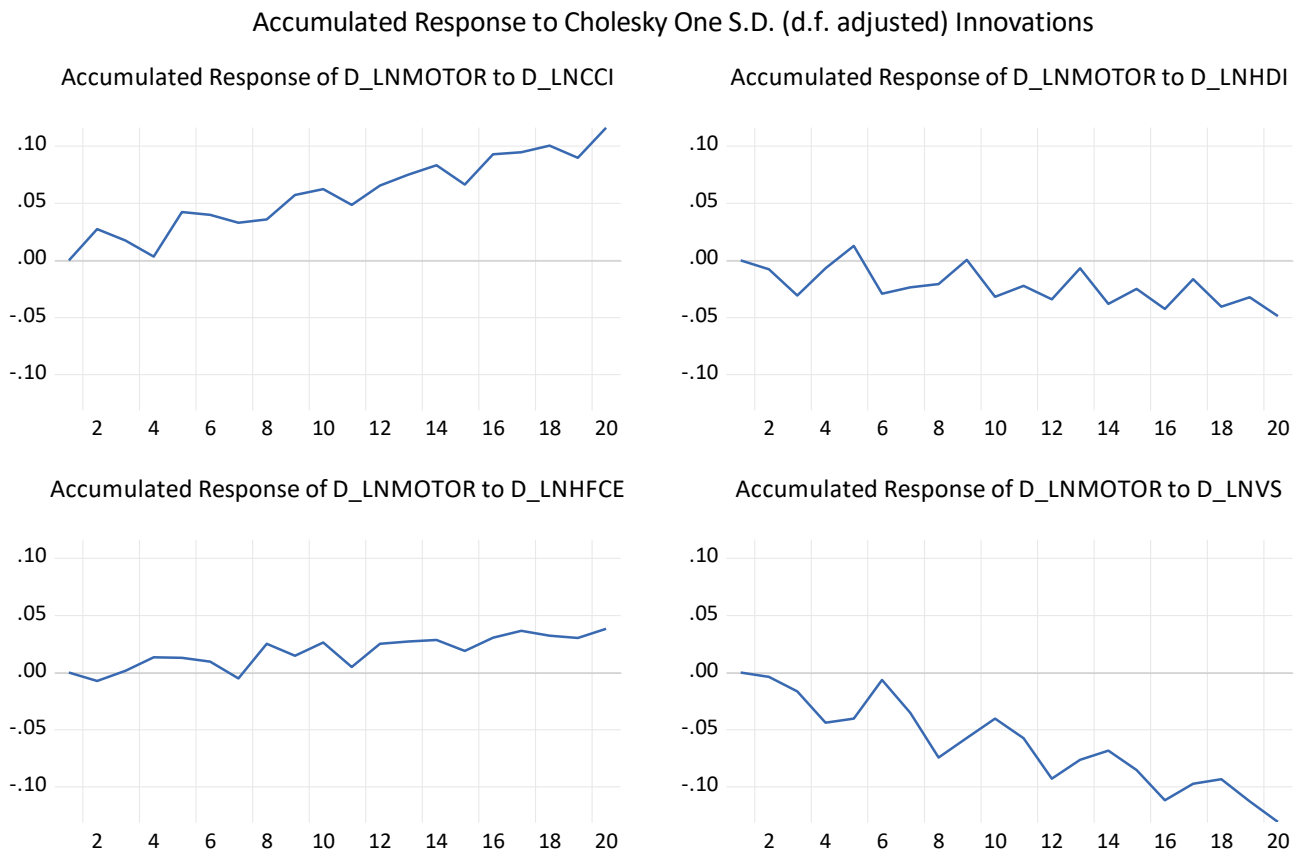
Figure 4 shows the response of the revenue sources given an increased shock of D_LNCCI, D_LNHDI, D_LNHFCE, and D_LNVS

Figure 4: The Response of D_LNCasino on Shock from Selected Variables



On the one hand, the overall response of the D_LNCasino is positive, given the shocks from the D_LNHDI and D_LNHFCE. On the other hand, the overall response of the D_LNCasino is negative, given the shocks from the D_LNCCI and D_LNVS.

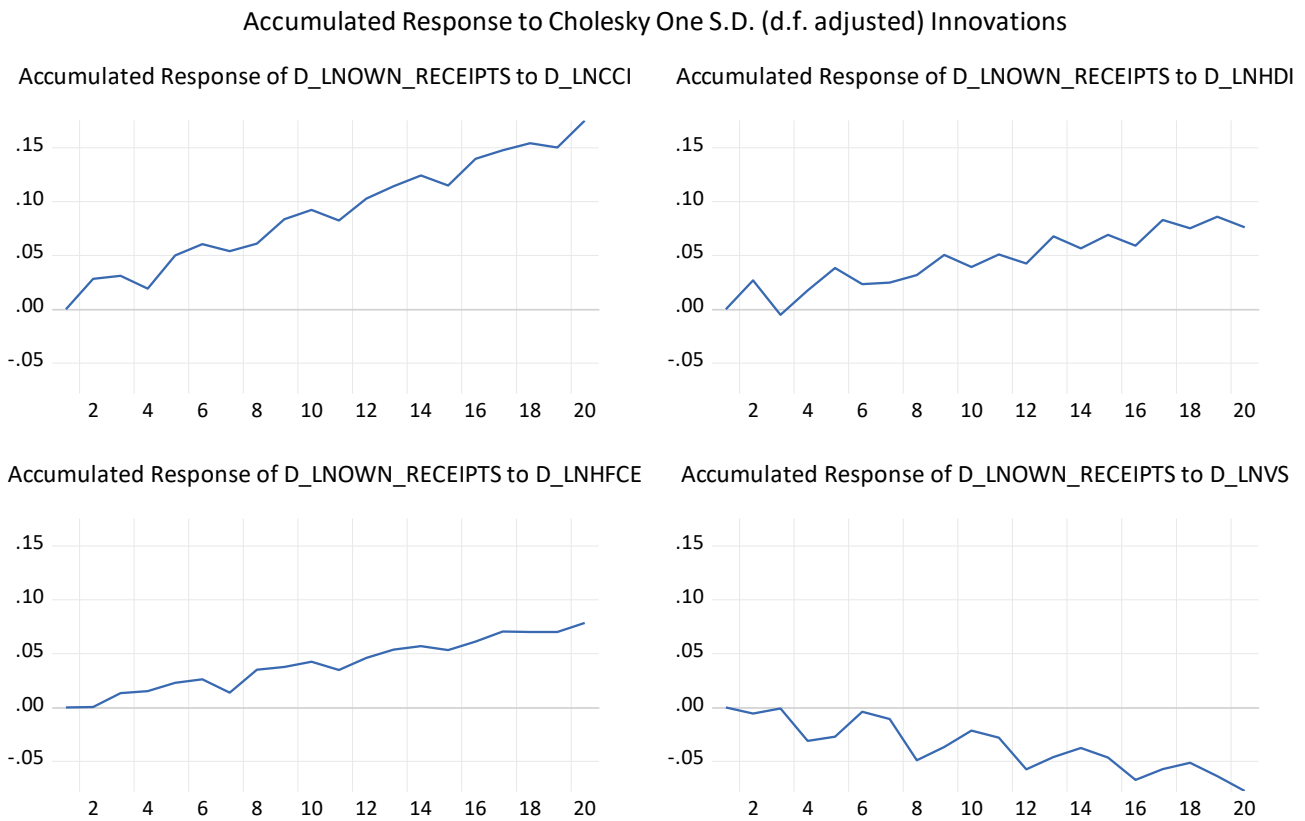
Figure 5: The Response of D_LNMotor on Shock from Selected Variables



Source: Author's Computation based on Bloomberg & Quantec Research Information.

The accumulated response of the D_LNMotor, given the respective increase shocks from the D_LNCCI, D_LNHFCE, can be observed in Figure 5, and it responds positively. D_LNVS and D_LNHI cause an overall negative response from the D_LNMotor. However, in the fifth quarter, the initial shock of LNVS increases the LNMotor.

Figure 6: The Response of D_LNOwn_receipts on Shock from Selected Variables



Source: Author's Computation based on Bloomberg & Quantec Research Information.

Figure 6 shows the accumulated response of the D_LNOwn_receipts is positive, given the respective increase shocks from the D_LNCCI, D_LNHDI, and D_LNHFCE. However, the LNVS initial shocks have an overall decreasing effect on D_LOwn_receipts.

5.6. Granger Causality

This section shows the significant Granger causality relationship of variables from the estimated VECM model.

The tables below show the Granger causality relationship of the dependent variable given in each respective table with their paired causal independent variables. The significance of the relationship is provided by the significant Chi-square test statistic p-values, which have a significance level of less than 5%.

Table 8: Significant Granger Causality Relationship [Dependant Variable: D(D_LNCASINO)]

Granger Causality Direction	Chi-sq (P-value)
D(D_LNMotor) Granger causes D(D_LNCasino)	15.45548 (0.0015)
D(D_LNOwn_receipts) Granger causes D(D_LNCasino)	19.64818 (0.0001)
D(D_LNHDI) Granger causes D(D_LNCasino)	7.999456 (0.0460)
D(D_LNHFCE) Granger causes D(D_LNCasino)	16.85827 (0.0008)
D(D_LNVS) Granger causes D(D_LNCasino)	16.09739 (0.0011)

Source: Author's Computation based on Bloomberg & Quantec Research Information.

D_LCasino is a Granger caused by the other two selected revenue sources and three selected macroeconomic variables as per Table 8. The macroeconomic selected variables that Granger causes the D_LNCasino are D_LNHDI, D_LNHFCE and D_LNVS.

Table 9: Significant Granger Causality Relationship [Dependant Variable: D(D_LNMOTOR)]

Granger Causality Direction	Chi-sq (P-value)
D(D_LNCasino) Granger causes D(D_LNMotor)	15.73197 (0.0013)
D(D_LNOwn_receipts) Granger causes D(D_LNMotor)	9.057847 (0.0285)
D(D_LNVS) Granger causes D(D_LNMotor)	12.58265 (0.0056)

Source: Author's Computation based on Bloomberg & Quantec Research Information.

D_LNMotor is a Granger caused by the other two selected revenue sources and only one selected macroeconomic variable, as per Table 9. Among this study's selected macroeconomic variables, D_LNVS is the only one that Granger causes D_LNMotor.

Table 10: Significant Granger Causality Relationship [Dependant Variable: D(D_LNOwn_receipts)]

Granger Causality Direction	Chi-sq (P-value)
D(D_LNCasino) Granger causes D(D_LNOwn_receipts)	15.73197 (0.0013)
D(D_LNMotor) Granger causes D(D_LNOwn_receipts)	9.057847 (0.0285)
D(D_LNHFCE) Granger causes D(D_LNOwn_receipts)	16.85827 (0.0008)
D(D_LNVS) Granger causes D(D_LNOwn_receipts)	12.58265 (0.0056)

Source: Author's Computation based on Bloomberg & Quantec Research Information.

D_LNOwn_receipts is a Granger caused by the other two selected revenue sources and two selected macroeconomic variables as per Table 10. The selected macroeconomic variables that Granger causes the D_LNOwn_receipts are D_LNHFCE and D_LNVS.

5.7. Results Discussion

The above results from the VECM estimation, Cholesky impulse response, and Granger causality show that the provincial own revenue sources, namely the casino tax, motor licenses tax, and the total own revenue, are impacted by the macroeconomic variables that were used in this study.

As per VECM estimation, the casino revenue has a positive short-term relationship with HDI and VS but a negative short-term relationship with HFCE. However, the impulse response of HFCE shows a long-term positive relationship with casino revenue. Therefore, HFCE positively impacts casino revenue despite evidence of negative short-term dynamics. The Granger causality also confirmed that HFCE causes casino revenue. In addition, VS impacts casino revenue positively in the short-run, given that VS has Granger causes casino revenue, and the VECM shows a positive relationship between these variables. Moreover, HDI impacts casino revenue positively in the short and long run, and HDI Granger causes casino revenue.

Furthermore, the impulse response shows that the initial shock from HFCE and CCI causes the motor licence revenue to increase over the short and long run. In addition, there is Granger causality of motor license revenue from VS, and VECM confirmed a negative short-run relationship.

According to the VECM, the total own revenue has a positive short-run relationship with HDI and VS and a negative short-run relationship with HFCE. However, the impulse response shows that the total own revenue increases due to HFCE, and the increase has a lag of about two quarters delayed, and the increase spans over the long run. Therefore, the negative short-run dynamics shown by the VECM could be capturing a delayed increase of the total own revenue emanating from a positive shock in HFCE.

Therefore, HFCE positively affects total own revenue with a lag. Moreover, impulse response confirms that HDI positively affects the total revenue in the short-run and long-run. On the contrary, with VS, there is evidence that it impacts the total own revenue only in the short-run. CCI impact on the total own revenue shows a long-term increasing trend. Furthermore, the Granger causality results reveal that the total own revenue sources are Granger and are caused by VS and HFCE.

6. Conclusion

This study assessed the assumptions and potential economic drivers used in the modeling process to develop provincial own revenue estimates that will enhance the credibility and robustness of provincial own revenue estimates. Such tools will be important in providing the Gauteng Provincial Treasury (GPT) with critical insights with respect to the credibility of revenue estimates.

This study is a follow up to the previous economic bulletin which introduced the concept of modeling provincial own revenue using the ARIMA and cashflow models to forecast own revenue collection. It assesses the various macroeconomic independent variables that may affect revenue collection, to understand the underlying drivers of own revenue collection and determine variables that will help to improve credibility of the models.

The study employed the VAR model and extended the estimation analysis by using the Cholesky impulse response and Granger causality to examine the effect of the selected macroeconomic variables on the provincial revenue sources. The study established that the HDI and VS drive the growth in the casino revenue in the short-run, and HDI and HFCE in the long-run. In light of the motor license revenue, HFCE and CCI causes the motor license revenue to increase over the short and long run.

Moreover, there are negative short-term dynamics between total own revenue and HFCE. Still, the impulse response analysis suggests a delayed positive impact of HFCE on total own revenue over the long run. Additionally, HDI and CCI also have positive impacts on total own revenue.

REFERENCES

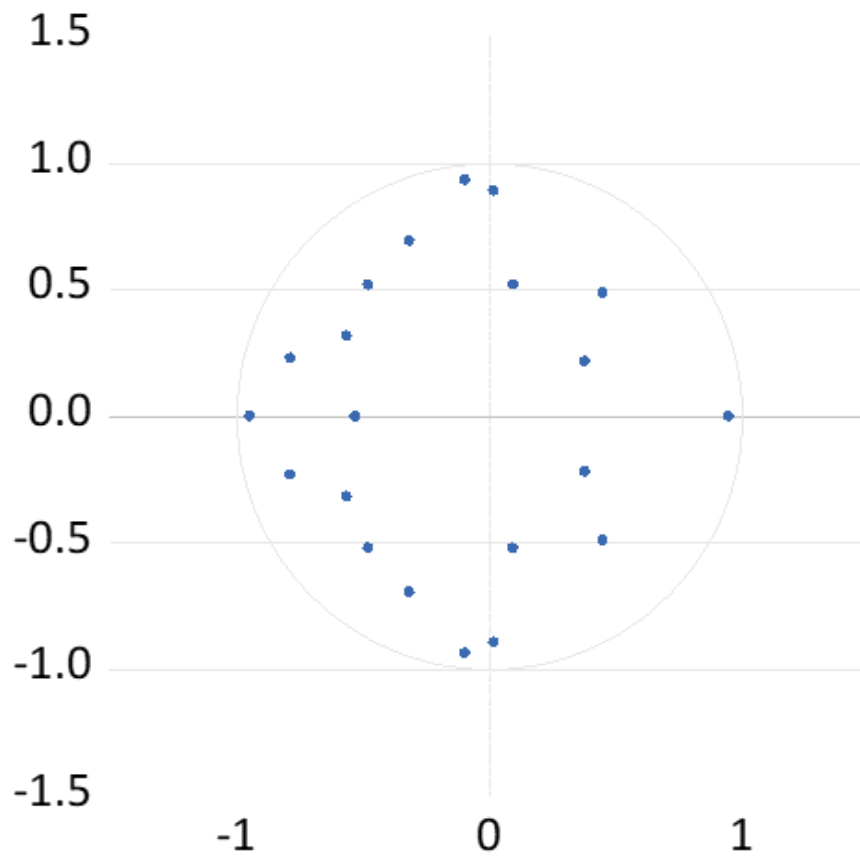
- Atolagbe, S. & Abiodun, A.A. (2021). Impact of some Macroeconomic Variables on Tax Revenue in Nigeria. *Nigerian Journal of Basic and Applied Sciences*, 29(2): 17–29. Available From: <https://www.ajol.info/index.php/njbas/article/view/221102> (Accessed 17 March 2024).
- Christiano, L.J. (2012). Christopher A. Sims and Vector Autoregressions. *Scandinavian Journal of Economics*, 114(4): 1082–1104.
- Division of the Budget, New York. (2004). "Revenue Estimating Methodology." New York State Government.
- Enders, W. (2014). Enders W. (2014) *Applied Econometric Time Series*. Wiley. Available From: https://books.google.com/books/about/Applied_Econometric_Time_Series.html?id=udXoAEACAAJ (Accessed 27 July 2023).
- Granger, C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37(3): 424. Available From: <https://www.jstor.org/stable/1912791> (Accessed 18 March 2024).
- Gujarati, DN and Porter, D.C. (2009). *Basic Econometrics*. 5th Edition, McGraw Hill Inc., New York.
- H2 Gambling Capital. (2020). *Denmark Online Gambling Market – Impact Analysis*. Danish Online Gambling Association.
- International Monetary Fund. (2005). *IMF Working Paper: Revenue Forecasting – How is it Done? Results from a Survey of Low-Income Countries*. Washington, DC, USA.
- Jenkins, G. et al. (2000). "Tax Analysis and Revenue Forecasting – Issues and Techniques." Harvard Institute for International Development, Harvard University.
- Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, 59(6): 1551–1580. Available From: <https://www.jstor.org/stable/2938278> (Accessed 6 March 2024).
- Pew Charitable Trusts. (2018). "Are Sin Taxes Healthy for State Budgets?." Rockefeller Institute of Government.

Sims, CA (1980). Macroeconomics and Reality. *Econometrica*, 48(1): 48. Available From: <https://www.jstor.org/stable/1912017> (Accessed 25 February 2024).

Stock, J.H. & Watson, M.W. (2001). Vector Autoregressions. *Journal of Economic Perspectives*, 15(4):101–115. Available From: <https://www.aeaweb.org/articles?id=10.1257/jep.15.4.101> (Accessed 28 February 2023).

Annexure A: VAR Model Stability Test

Inverse Roots of AR Characteristic Polynomial



Annexure B: VECM Output

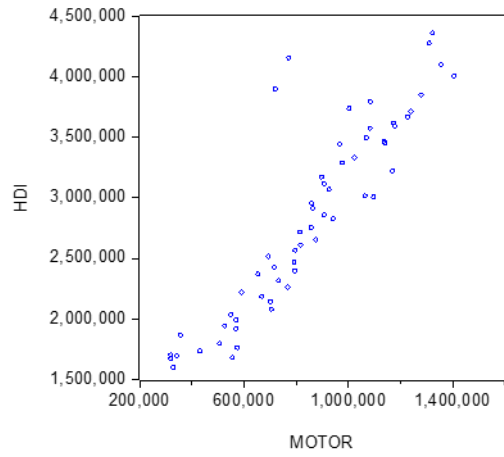
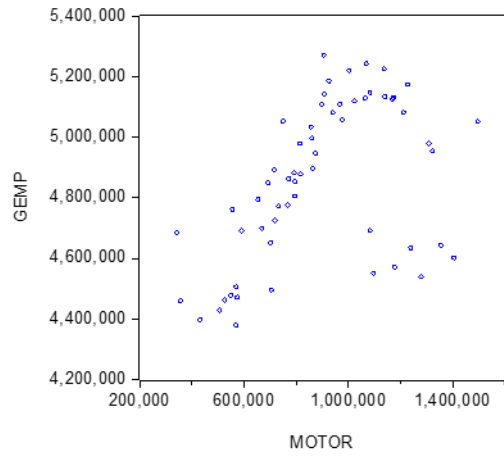
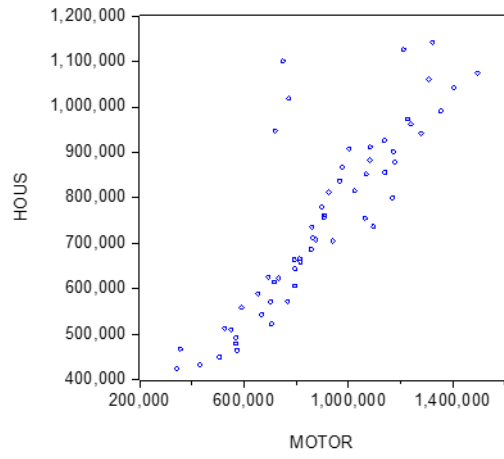
Vector Error Correction Estimates
 Date: 03/15/24 Time: 10:24
 Sample (adjusted): 2009Q3 2023Q3
 Included observations: 57 after adjustments
 Standard errors in () & t-statistics in []

Cointegrating Eq:	CointEq1
D_LNCASINO(-1)	1.000000
D_LNMOTOR(-1)	-1.982197 (0.91851) [-2.15806]
D_LNOWN_RECEIPTS(-1)	4.627371 (0.95620) [4.83935]
D_LNCCI(-1)	0.021939 (0.12692) [0.17285]
D_LNHDI(-1)	14.47715 (8.81504) [1.64232]
D_LNHFCE(-1)	-30.59308 (7.04277) [-4.34390]
D_LNVS(-1)	4.606260 (0.95797) [4.80835]
C	0.166241

Error Correction:	D(D_LNCASI)	D(D_LNMOT)	D(D_LNOWN)	D(D_LNCCI)	D(D_LNHDI)	D(D_LNHFCE)	D(D_LNVS)
CointEq1	-1.028444 (0.23049) [-4.46196]	-0.155528 (0.15244) [-1.02028]	-0.273923 (0.11263) [-2.43200]	1.646372 (0.75801) [2.17198]	0.109397 (0.04457) [2.45453]	0.139060 (0.04497) [3.09238]	0.287258 (0.12719) [2.25847]
D(D_LNCASINO(-1))	-0.110838 (0.25729) [-0.43079]	0.131360 (0.17016) [0.77197]	0.334075 (0.12573) [2.65710]	-0.171498 (0.84614) [-0.20268]	-0.006589 (0.04975) [-0.13244]	-0.036640 (0.05020) [-0.72992]	0.060987 (0.14198) [0.42955]
D(D_LNCASINO(-2))	0.060833 (0.26681) [0.22800]	-0.194942 (0.17646) [-1.10474]	0.008644 (0.13038) [0.06630]	-0.216584 (0.87745) [-0.24683]	0.030299 (0.05159) [0.58726]	0.007258 (0.05205) [0.13943]	0.127156 (0.14723) [0.86363]
D(D_LNCASINO(-3))	-0.054097 (0.18256) [-0.29632]	0.053412 (0.12074) [0.44238]	0.031184 (0.08921) [0.34955]	0.070836 (0.60038) [0.11799]	0.037512 (0.03530) [1.06261]	0.030086 (0.03562) [0.84469]	0.140201 (0.10074) [1.39167]
D(D_LNMOTOR(-1))	-0.988391 (0.47442) [-2.08338]	-1.782030 (0.31376) [-5.67961]	-0.463195 (0.23183) [-1.99799]	3.808947 (1.56019) [2.44133]	0.318015 (0.09174) [3.46659]	0.371399 (0.09256) [4.01261]	1.194659 (0.26180) [4.56331]
D(D_LNMOTOR(-2))	0.318159 (0.54095) [0.58814]	-1.877653 (0.35776) [-5.24830]	-0.756728 (0.26434) [-2.86265]	1.376434 (1.77901) [0.77371]	0.189528 (0.10460) [1.81188]	0.229563 (0.10554) [2.17514]	0.836696 (0.29851) [2.80287]
D(D_LNMOTOR(-3))	0.196787 (0.40501) [0.48588]	-1.304353 (0.26786) [-4.86956]	-0.757271 (0.19792) [-3.82624]	-0.476404 (1.33195) [-0.35768]	-0.005638 (0.07832) [-0.07199]	0.010421 (0.07902) [0.13189]	0.027358 (0.22350) [0.12241]
D(D_LNOWN_RECEIPT)	2.829012 (0.91152) [3.10364]	1.197115 (0.60284) [1.98580]	0.088799 (0.44542) [0.19936]	-8.333100 (2.99765) [-2.77988]	-0.602526 (0.17626) [-3.41844]	-0.711149 (0.17783) [-3.99893]	-2.170949 (0.50300) [-4.31601]
D(D_LNOWN_RECEIPT)	0.642054 (0.72654) [0.88372]	1.361716 (0.48050) [2.83396]	0.388576 (0.35503) [1.09448]	-4.121553 (2.38932) [-1.72499]	-0.334138 (0.14049) [-2.37840]	-0.393830 (0.14175) [-2.77843]	-1.301630 (0.40092) [-3.24659]
D(D_LNOWN_RECEIPT)	0.217220 (0.43156)	0.812757 (0.28541)	0.430301 (0.21089)	-0.853175 (1.41925)	-0.070779 (0.08345)	-0.093207 (0.08420)	-0.318764 (0.23815)

	[0.50334]	[2.84764]	[2.04043]	[-0.60115]	[-0.84816]	[-1.10701]	[-1.33852]
D(D_LNCCI(-1))	-0.068073 (0.09073) [-0.75027]	0.079253 (0.06001) [1.32074]	-0.003698 (0.04434) [-0.08340]	-0.701251 (0.29839) [-2.35014]	0.009503 (0.01754) [0.54163]	0.004321 (0.01770) [0.24410]	0.016704 (0.05007) [0.33363]
D(D_LNCCI(-2))	-0.140962 (0.11417) [-1.23463]	0.135470 (0.07551) [1.79409]	0.063065 (0.05579) [1.13035]	-0.427212 (0.37547) [-1.13779]	0.012046 (0.02208) [0.54562]	0.004263 (0.02227) [0.19136]	-0.003023 (0.06300) [-0.04798]
D(D_LNCCI(-3))	-0.145678 (0.10023) [-1.45337]	0.007006 (0.06629) [0.10569]	0.007925 (0.04898) [0.16179]	0.176797 (0.32964) [0.53634]	0.044413 (0.01938) [2.29146]	0.040048 (0.01956) [2.04789]	0.161758 (0.05531) [2.92444]
D(D_LNHDI(-1))	12.17486 (4.99887) [2.43552]	3.461460 (3.30603) [1.04701]	5.569504 (2.44276) [2.28000]	-36.34852 (16.4395) [-2.21105]	-3.598916 (0.96662) [-3.72320]	-2.709245 (0.97527) [-2.77795]	-5.887584 (2.75851) [-2.13433]
D(D_LNHDI(-2))	9.288853 (6.21905) [1.49361]	1.782785 (4.11301) [0.43345]	2.808554 (3.03902) [0.92416]	-25.32999 (20.4523) [-1.23849]	-2.857847 (1.20256) [-2.37646]	-2.177437 (1.21333) [-1.79460]	-3.257622 (3.43184) [-0.94923]
D(D_LNHDI(-3))	9.176654 (4.73229) [1.93916]	0.083690 (3.12973) [0.02674]	1.767651 (2.31250) [0.76439]	-18.62293 (15.5628) [-1.19663]	-1.675861 (0.91507) [-1.83140]	-1.483914 (0.92326) [-1.60725]	-3.299568 (2.61141) [-1.26352]
D(D_LNHFC(-1))	-22.83604 (6.11590) [-3.73388]	-6.164063 (4.04479) [-1.52395]	-8.173003 (2.98862) [-2.73471]	35.98027 (20.1130) [1.78890]	2.662031 (1.18262) [2.25097]	2.178440 (1.19320) [1.82571]	4.646858 (3.37492) [1.37688]
D(D_LNHFC(-2))	-15.89510 (6.18162) [-2.57135]	-4.629162 (4.08825) [-1.13231]	-5.456265 (3.02073) [-1.80627]	18.49022 (20.3292) [0.90954]	1.907267 (1.19533) [1.59561]	1.572242 (1.20602) [1.30366]	1.921766 (3.41119) [0.56337]
D(D_LNHFC(-3))	-12.07097 (4.48056) [-2.69408]	0.955870 (2.96325) [0.32258]	-0.737517 (2.18949) [-0.33684]	17.86344 (14.7350) [1.21232]	1.157038 (0.86640) [1.33546]	1.121161 (0.87415) [1.28257]	1.754828 (2.47250) [0.70974]
D(D_LNVS(-1))	3.280081 (0.90367) [3.62973]	0.614159 (0.59765) [1.02763]	1.111784 (0.44159) [2.51768]	-0.544106 (2.97185) [-0.18309]	-0.037621 (0.17474) [-0.21530]	-0.129055 (0.17630) [-0.73200]	-0.699306 (0.49867) [-1.40234]
D(D_LNVS(-2))	2.786710 (0.71850) [3.87852]	0.039829 (0.47518) [0.08382]	0.766040 (0.35110) [2.18180]	2.123832 (2.36289) [0.89883]	0.126770 (0.13893) [0.91244]	0.061062 (0.14018) [0.43561]	0.002079 (0.39649) [0.00524]
D(D_LNVS(-3))	1.735222 (0.58996) [2.94127]	-0.917924 (0.39017) [-2.35262]	-0.519745 (0.28829) [-1.80285]	-1.158833 (1.94016) [-0.59729]	-0.038301 (0.11408) [-0.33574]	-0.066277 (0.11510) [-0.57582]	-0.298636 (0.32555) [-0.91732]
C	0.027089 (0.02475) [1.09435]	0.002223 (0.01637) [0.13577]	0.008449 (0.01210) [0.69852]	-0.068048 (0.08140) [-0.83592]	-0.003969 (0.00479) [-0.82928]	-0.004945 (0.00483) [-1.02405]	-0.008888 (0.01366) [-0.65065]
R-squared	0.804013	0.945793	0.948202	0.801236	0.839185	0.835813	0.898002
Adj. R-squared	0.677198	0.910718	0.914686	0.672624	0.735129	0.729574	0.832003
Sum sq. resids	1.038395	0.454186	0.247960	11.23044	0.038827	0.039525	0.316205
S.E. equation	0.174760	0.115579	0.085399	0.574723	0.033793	0.034095	0.096437
F-statistic	6.340053	26.96473	28.29085	6.229866	8.064703	7.867297	13.60632
Log likelihood	33.27369	56.84105	74.09031	-34.58343	126.9339	126.4261	67.16133
Akaike AIC	-0.360480	-1.187405	-1.792643	2.020471	-3.646804	-3.628986	-1.549520
Schwarz SC	0.463909	-0.363016	-0.968253	2.844860	-2.822415	-2.804597	-0.725131
Mean dependent	0.001796	0.016909	0.013876	0.010718	7.93E-05	7.50E-05	0.002561
S.D. dependent	0.307591	0.386807	0.292376	1.004466	0.065661	0.065565	0.235285
Determinant resid covariance (dof adj.)		4.15E-18					
Determinant resid covariance		1.11E-19					
Log likelihood		677.5996					
Akaike information criterion		-17.88069					
Schwarz criterion		-11.85906					
Number of coefficients		168					

Annexure C: Scatter plots





Contact Information:

Physical Address:
75 Fox Street, Imbumba House,
Johannesburg
2107

Postal Address:
Private Bag X12,
Marshalltown,
Johannesburg,
2107

Switchboard:
011 227 9000

Email address:
GPTCommunications@gauteng.gov.za

Website Address:
www.gauteng.gov.za

PG Image:
Author Unknown, [ONLINE]. <https://gautengfilm.org.za/gauteng-film-commission-directory/listing/fox-street-johannesburg/>



GAUTENG
PROVINCIAL GOVERNMENT
REPUBLIC OF SOUTH AFRICA

GGT2030
GROWING GAUTENG TOGETHER