

https://dspace.org/

# Forecasting revenue using time series techniques in South Africa's commercial transactional banking industry

**Malepane, Benny Christian Thapedi**

**2022**
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/34744

Malepane, Benny Christian Thapedi. (2022). Forecasting revenue using time series techniques in South Africa's commercial transactional banking industry [Master's dissertation, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/




# Forecasting Revenue Using Time Series Techniques in South Africa's Commercial Transactional Banking Industry

**A research report submitted to the Faculty of Commerce, Law and Management, University of the Witwatersrand, in partial fulfilment of the requirements for the degree of Master of Management in Finance and Investments**

**by**

**Benny Christian Thapedi Malepane**

**(554202)**

**Supervisor: Dr Euphemia Godspower-Akpomiemie**

## DECLARATION

I, Benny Malepane declare that the work reported in this research is my own, except where otherwise indicated and acknowledged. It is submitted for the degree of Master of Management in Finance and Investments in the University of the Witwatersrand. This dissertation has not, either in whole or in part, been submitted for a degree or diploma to any other universities.

malepane 2022-06-17

Signature of candidate Date

February 2022

DEDICATION

This work is dedicated to my angel Bophelo who was born on 26 December 2021 and inspired me to look at life differently and give my best in everything.

## ACKNOWLEDGEMENT

I cannot express enough gratitude to my supervisor Dr Euphemia Godspower-Akpomiemie who guided me through this and offered constructive feedback which I learned a lot from. I offer my sincere gratitude.

Finally, my dear friend and brother Andrew Phogole for always encouraging me to never give up and keep believing that I will finally get the research completed. I am thankful.

## LIST OF ACRONYMS

| | |
|---|---|
| ADF: | Augmented Dicky Fuller |
| AIC: | Akaike Information Criterion |
| AR: | Autoregressive |
| ARIMA: | Autoregressive Integrated Moving Average |
| BIC: | Bayesian Information Criterion |
| BOF: | Benefits of Funds |
| CPI: | Consumer Price Index |
| COF: | Cost of Funds |
| CTI: | Cost to Income |
| GDP: | Gross Domestic Product |
| IFRS: | International Financial Reporting Standards |
| LMI: | Large Manufacturing Index |
| LRF: | Liquidity Risk Factor |
| LTD: | Loan to Deposit |
| MA: | Moving Average |
| MAPE: | Mean Absolute Percentage Error |
| NII: | Net Interest Income |
| NIM: | Net Interest Margin |
| NSFR: | Net Stable Funding Ratio |
| NIR: | Non-Interest Income |
| NPL: | Non-Performing Loan |
| ROA: | Return on Assets |

| | |
|---|---|
| ROE: | Return on Equity |
| RMSE: | Root Mean Square Error |
| VAR: | Vector Autoregressive |
| WACC: | Weighted Average Cost of Capital |

Contents

**DECLARATION** ..... II
**ACKNOWLEDGEMENT** ..... IV
**LIST OF ACRONYMS** ..... V
**LIST OF FIGURES AND TABLES** ..... X
**ABSTRACT** ..... XI
**CHAPTER 1: INTRODUCTION** ..... 1
**1.1 Introduction** ..... 1
**1.2 Context of the study** ..... 1
**1.3 Problem statement** ..... 3
**1.4 Purpose of the study** ..... 3
**1.5 Research question** ..... 3
**1.6 Significance of the study** ..... 4
**1.7 Delimitation of the study** ..... 4
**1.8 Assumptions** ..... 5
**1.9 Outline of the study** ..... 5
**CHAPTER 2: LITERATURE REVIEW** ..... 6
**2.1 Introduction** ..... 6
**2.2 Overview of the Banking Industry** ..... 6
**2.2.1 South African Banking Industry** ..... 6
**2.3 Theoretical Review** ..... 7
**2.3.1 Theory of financial intermediation** ..... 8
**2.4 Bank Revenue** ..... 8
**2.4.1 Net interest income from banking operations** ..... 8
**2.4.2 Non-interest income** ..... 9
**2.4.3 Trading income** ..... 11
**2.5 The cost base for banks** ..... 11
**2.5.1 Operational expenses** ..... 11
**2.5.2 Cost of credit** ..... 12
**2.5.3 Cost of capital** ..... 12
**2.5.4 Tax** ..... 12
**2.6 Bank performance measures** ..... 13

**2.6.1 Bank Performance measures** ..... 13
**2.6.2 Macro and micro economic factors and banking performance** ..... 16
**2.7 Bank Revenue Forecast** ..... 17
**2.8 Summary of review** ..... 19
**CHAPTER 3: RESEARCH METHODOLOGY** ..... 20
**3.1 Introduction** ..... 20
**3.2 Research design** ..... 20
**3.3 Description of data** ..... 20
**3.4 Data collection and sources** ..... 21
**3.5 Data analysis** ..... 21
**3.5.1 Unit root test** ..... 21
**3.5.2 Autoregressive model (A.R. model)** ..... 22
**3.5.3 Moving Average (MA) model** ..... 23
**3.5.4 Autoregressive integrated moving average (A.R.I.M.A) model** ..... 23
**3.5.5 Forecasting error** ..... 23
**3.5.6 Granger causality test** ..... 24
**3.5.7 Vector autoregressive (VAR) model** ..... 25
**3.6 Validity and reliability** ..... 25
**3.7 Ethical considerations** ..... 25
**3.8 Chapter summary** ..... 26
**CHAPTER 4: ESTIMATIONS AND RESULTS** ..... 27
**4.1 Introduction** ..... 27
**4.2 Augmented Dicky-Fuller (ADF) test** ..... 27
**4.2.1 Graphical representation of stationarity** ..... 28
**4.3. Autoregressive model (A.R. model) forecast** ..... 29
**4.3.1 Non-interest revenue (NIR) forecast** ..... 29
**4.4 Autoregressive integrated Moving Average (ARIMA model) forecast** ..... 29
**4.4.1 Non-interest revenue (NIR) forecast** ..... 29
**4.5 Vector autoregressive model (VAR model) forecast** ..... 30
**4.5.1 Non-interest revenue (NIR) forecast** ..... 30
**4.6 Granger Causality test** ..... 30
**4.7 Vector autoregressive (VAR) model** ..... 32
**4.7.1 Non-interest revenue (NIR) forecast** ..... 32
**4.8 Chapter Summary** ..... 32
**CHAPTER 5: SUMMARY AND CONCLUSION** ..... 33
**5.1 Introduction** ..... 33

**5.2 Summary of findings** ..... 33
**5.3 Policy implications** ..... 34
**5.4 Policy recommendations** ..... 35
**5.5 Conclusion** ..... 35
**5.6 Recommendation for further research** ..... 35
**REFERENCES** ..... 38
**APPENDIX** ..... 48
**Appendix I** ..... 48
**Appendix I A: ADF test results for total NIR (At level)** ..... 48
**Appendix I B: ADF test results for total NIR (At first difference)** ..... 49
**Appendix I C: ADF test results for inflation (At level)** ..... 50
**Appendix I D: ADF test results for inflation (At first difference)** ..... 51
**Appendix I E: ADF test results for prime rate (At level)** ..... 52
**Appendix I F: ADF test results for prime rate (At first difference)** ..... 53
**Appendix II** ..... 54
**Appendix II A: Forecast results for AR (1) model** ..... 54
**Appendix III** ..... 55
**Appendix III A: Forecast results for ARIMA (1,1,1) model** ..... 55
**Appendix IV** ..... 56
**Appendix IV A: Forecast results for MA (1) model** ..... 56
**Appendix V** ..... 57
**Appendix V A: Results for Granger causality test – Wild test** ..... 57
**Appendix IV** ..... 59
**Appendix IV A: Forecast results for VAR (1,1) model** ..... 59

## LIST OF FIGURES AND TABLES

Figure 1: Industry growth in the first quarter of 2021 compared to the fourth quarter of 2020

Figure 2: Graphical representation of stationary and no stationary graphs

Figure 3: The partial cross-correlation matrices

Table 1: ADF stationarity test at 5% level of significance

Table 2: RMSE, MAE and MAPE for AR (1)

Table 3: RMSE, MAE and MAPE for ARIMA (1,1,1)

Table 4: RMSE, MAE and MAPE for MA (1)

Table 5: RMSE, MAE and MAPE for VAR (1)

## ABSTRACT

Budgeting and forecasting is a crucial exercise carried by finance teams in organisations to predict their future performance. The use of time series models to forecast financial data has sparked interests from many researchers around the globe. Even though time series studied were carried on financial data, they haven't been widely applied to total Net Interest Revenue (NIR) from transactional banking. This highlighted a need to apply time series models on this data.

This research studies the efficiency of four time series models (AR, ARIMA, VAR and MA) on total NIR for transactional accounts in commercial banking. Macro-economic factors for instance, inflation and prime interest rates are also studied via the cointegration test to identify their influence on model performance using the VAR model. The best performing model is chosen based on statistical metrices, forecasting errors and significance of parameters.

The forecasting errors reveal that VAR is the best performing model amongst the four with MA coming second. This highlighted the need to consider inflation and interest rate movements when forecasting NIR on transactional accounts. However, the VAR models did not produce significant parameters, therefore, MA model had to be chosen as the overall best performing model as it produced better forecasting errors and significant parameters than the VAR model. As a result, this study recommends the use of MA models to forecast total NIR for transactional accounts in commercial banking.

This study recommends an amendment of certain policies within the banking sector in order to benefit from the results of this dissertation. The polices include budgeting, hiring and remuneration policy. Due to the government dependency on banking forecasts, it was also recommended that government adjusts its budgeting policy to stay consistent with the way banks forecast as they do their forecasts based on the results banks produce.

# CHAPTER 1: INTRODUCTION

## 1.1 Introduction

This section provides the research background, problem statement and the objectives of the study. Significance and delimitation of the study are also provided in the section. More so, the research outline, detailing what each chapter of the study entails is also provided in this section.

## 1.2 Context of the study

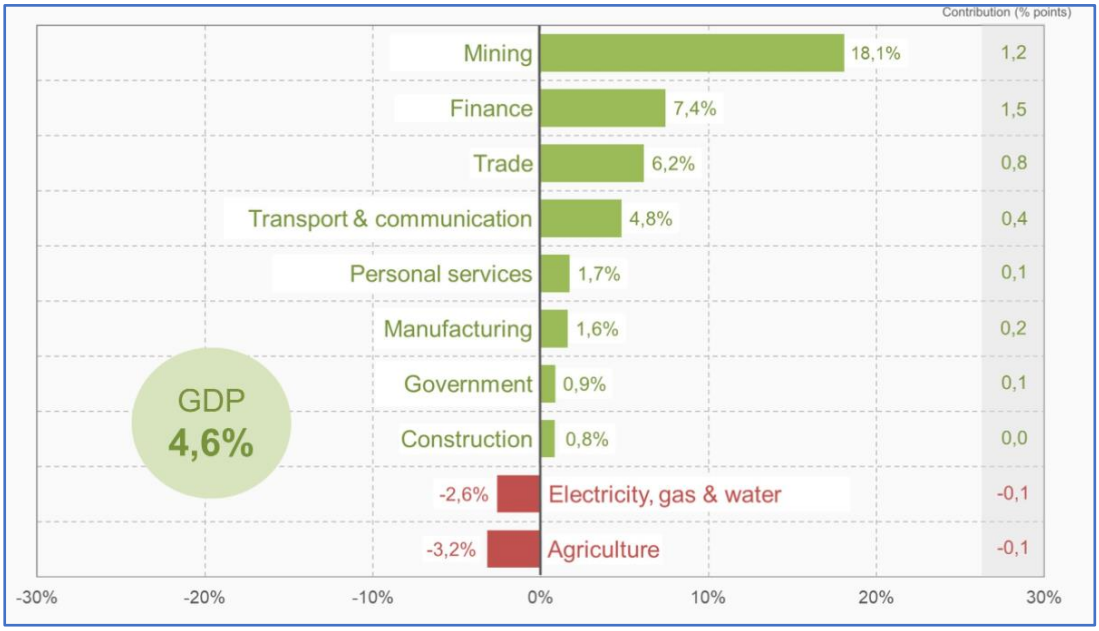
South Africa is considered an emerging market, but it has a well-developed banking system (Choudhry, 2015). Banks are commonly well managed, and they have sophisticated risk management and corporate management strategies relative to other African countries' financial institutions. They serve as agents of liquidity ensuring that there is enough liquidity in the financial system. They play a role of acting as financial intermediaries by enabling investments by firms and encouraging economic growth. They allow firms and individuals to trade locally and internationally. Finally, they create jobs (Choudhry, 2015).

There are different segments within a bank, for instance, retail banking, commercial banking and corporate banking. Retail banking is related to banking products and services offered to individual customers, whereas commercial and corporate banking relates to products and services offered to small and large businesses. This study only focused on commercial banking

Commercially banking is largely exposed to risks such as credit, operational and market risks. In additional, transactional volumes are huge compared to retail and corporate banking. As a result, the NIR base in large compared to the other segments, this means there is more transactional forecast data that can be utilised to develop time series models that can forecast credible results.

In the first quarter of 2021, the finance sector grew by 7.4% in South Africa as illustrated in figure 1. This shows how significant the finance industry is to the overall growth and Gross Domestic Product (GDP) of the country.

Figure 1: Industry growth from 2020Q4 to 2021Q1




Source: Stats' (2021)[1]

There are different revenue streams generated by the banks. The first one is the Net Interest Income (NII), which is the spread between benefits of funds (interest earned on customer loans) and cost of funds (interest paid to customer deposits). There is also the Non-Interest Revenue (NIR) which is the revenue that is not NII. This is transactional revenue that arises from activities such as cash deposits, cash withdrawals, payments, transfers and many more. The focus of this dissertation is on projection of NIR. NIR is related to transactional banking as it is fee income and this dissertation primarily focuses on it. The different revenue streams are further discussed in Chapter 2.

There are different areas in the banking institution that make use of projections to set up the banking strategy, such as the area of pricing, revenue, as well as costing. The projections can give a bank a sense of where its business will be in the future and which pricing strategies need to be adopted to get there. As a result, it is very important that the forecasts are as accurate as possible to avoid surprises and poor decision making. It is for this reason that a time series modelling is chosen for this research.

[1] www.statssa.co.za

## 1.3 Problem statement

Forecasting has been the biggest challenge in transactional banking for some time due to a lack of proper models that have been developed to project future revenue. Most of the budgets come from the finance department in every financial institution and this involves high level spreadsheets. These projections may not be accurate due to high level assumptions made, as well as lack of consideration on other factors that affect revenue. These factors include GDP, interest rates and inflation amongst others. This is a key challenge in banking industry, given that these numbers are used by senior management to make managerial decisions. The risk is that the bank can embark on a strategic journey that may lead to losses. For example, the forecasts may show that the bank will be profitable in the next financial year, therefore take a decision to embark on a strategy to increase market share, which might turn out to be unwise.

There are several factors which also expose major banks' transactional revenue such as cash deposit and cash withdrawal to risk. The most important of such factors is digitization. There are several Fintech banks that have penetrated the banking sector, including South Africa, which has led to an increase in competition. Examples of these FinTech banks are Tymebank, Discovery bank and Bank Zero. Recently, there has been a significant drop in some of the banking fees on transactions customers carry out on digital platforms. These fee cuts mean that the major banks' transactional revenue will be drastically reduced overtime. It is for this reason that projections need to adequately reflect these factors and hence appropriate models need to be chosen to perform the task.

## 1.4 Purpose of the study

The aim of this study is to forecast the revenue in commercial banking, using tentative time series techniques.

1. Autoregressive model (A.R.), Moving Average model (M.A), Autoregressive Integrated Moving Average (A.R.I.M.A) and Vector Autoregressive (V.A.R) model are the time series models tested in forecasting the revenue.
2. The best performing model is chosen as a forecast model for commercial banking revenue.

## 1.5 Research question

To achieve the above objective, the outlined research question is addressed.

What is the appropriate revenue forecasting model to be adopted by the commercial banks?

## 1.6 Significance of the study

The success of developing a projection model will play an important role in transactional banking to bridge the gap of lack of forecasting models. This will also help to ensure that those in the forefront of decision making are able to make better decisions as they will have accurate supporting information.

The banking environment is also dynamic. There are innovations that are developed over time that makes banking easier and cheaper. Therefore, models that can easily integrate factors such as innovation, which impacts transactional revenue are crucial.

The model ultimately chosen to forecast will also help future scholars to understand which models are efficient in terms of modelling time series data of similar nature like transactional revenue. Most importantly, the findings of this research will benefit those who are interested to explore further models that might be more suitable in transactional banking and encourage more scholars to explore this area.

## 1.7 Delimitation of the study

This study is only limited to one major bank in South Africa, therefore the data is not an industry wide data which might be more insightful. The scope is also only limited to non-interest income (NIR) which is not the only revenue stream that bank generates. There are other revenues, such as net interest income (NII) and trading income which are not part of the scope.

Macroeconomic factors that are considered in this study are limited to, consumer price index (CPI)[2] and prime rate. Factors such as unemployment rate which are generally known to impact the banking sector performance are omitted. This study did not factor in the impact of GDP, which impacts commercial banking performance as well. The GDP data was obtained at a quarterly basis whereas the total NIR, inflation and prime interest rate are monthly. Aligning the four data sets was challenging and it was decided that GDP is left out of this study.

[2] Consumer Price Index (CPI) is a measure of inflation. Therefore, CPI and inflation are used interchangeably in this dissertation.

## 1.8 Assumptions

It is assumed that transactional data utilized in this study was sufficient to derive a concrete conclusion. This is expected since transactional banking is vast in nature and has different types of fees that are charged that form part of total NIR. As a result, the data was expected to be sufficient and credible.

## 1.9 Outline of the study

This dissertation is outlined in three chapters, where the first chapter details the introduction and objectives of this research. Chapter 2 focuses on literature review where previous studies related to this topic are investigated and linked to this study. Chapter 3 focuses on the methodology which involves research design, which entails the data to be used, its sources and what kind of models were adopted and why. Chapter 4 presents the model results and a conclusion on the results is made in Chapter 5.

# CHAPTER 2: LITERATURE REVIEW

## 2.1 Introduction

This chapter presents past studies that are related to this research topic. This chapter mostly focuses on studies that are linked to time series forecasting and time series models recommended by various studies. Theory of financial intermediation, different revenue streams in banking, as well as performance indicators, are also discussed.

## 2.2 Overview of the Banking Industry

The Banking industry has improved substantially after the 2007-2008 financial crisis. According to Caplen (2020), the total global assets have increased to $124 trillion since the 2007-2008 financial crisis. Therefore, with so many assets to take care of, the big banks across the globe are developing new features to increase customer retention and attract new ones. In addition to this, start-ups and digital banks with disruptive technologies are penetrating the banking industry to provide a new banking experience in the form of innovative solutions. The influx of the new Fintech banks has resulted in an increase in competition, where traditional banks decided to either drag the market share with new start-ups or go into collaboration.

Recently, the common trend in banking industry is the transformation from traditional banking that involves branch networks to digital banking. Especially internet banking and mobile banking. This is especially true for millennials as they are more tech savvy than the old generation. This has led to an increase in competition and more funding for tech start-ups. The disruptiveness of the start-up has therefore put the fee income from major banks under a threat[3].

### 2.2.1 South African Banking Industry

The banking sector in South Africa is oligopolistic as it is dominated by four big banks (Ifeacho & Ngalawa, 2014). These include First National Bank (FNB), Standard Bank, ABSA, and Nedbank. Capitec has recently joined the market as the fifth biggest bank in the country. The South African banking sector is relatively developed and regulated

[3] This information was obtained from www.cbinsights.com

compared to other African banking sectors (Quiding, 2006). And this is one of the reasons why South African banks survived the 2007-2008 financial crisis (Choudhry, 2015).

Although the South African banking industry is oligopolistic, the industry has improved since 1994 (Ifeacho & Ngalawa, 2014). There has been an increase in the number of new entrants offering banking products such as loans and advances. These new entrants also focus on offering credit and banking solutions to the previously unbanked markets and lower income segments, to reduce financial exclusions. This is effectively increasing competition in the banking sector.

South Africa is also experiencing similar trends as the rest of the world in terms of digital transformation. The South African banking sector experienced several new Fintech banks entering the market such as Tymebank, Discovery and Bankzero. Bankzero was recently launched, and it is expected to disrupt the banking industry through tech innovations. These new players increased competition as drastic fee cuts and new product development by big banks have been observed.

## 2.3 Theoretical Review

The major role of banks in an economy is to carry out financial intermediations function. They act as agents of liquidity as they bring liquidity into the financial system. They accept deposits from their clients and lend them to individuals, firms and other financial institutions (Choudhry, 2015; Fama, 1980; Oldfield & Santamero, 1995). Intermediation of banks allow them to provide the ability to turn assets into cash at a relatively low cost. Banks can collect deposits from a large pool of depositors which allows them to reduce their overall costs of investments as they are able to lend these funds at higher rates.

Banks also play a role of financial intermediation by facilitating investments by firms and to enable job creation and economic growth. They bring together the consumers of capital and providers of funds. They can act as agent on behalf of their clients or as principals steering financial transactions for their own account. Banks also make it possible for international trade to occur. They allow funds to be transferred from one country to the next and allow the transfer of information amongst different markets and customers (Allen et al., 2008).

### 2.3.1 Theory of financial intermediation

Due to banks' intermediation function identified and detailed above, the theory that underpinned this research is the theory of financial intermediation. Financial intermediation theory posits that the financial system is efficient if it has an ability to minimise transactional costs, information asymmetry and the risks it faces (Allen & Santomero, 1997). Historically banks are the biggest role players of intermediation in the finance sector in terms of originating funds in the form of deposits, distributing these funds in terms of loans and servicing the funds (Fama, 1980; Oldfied & Santomero , 1995). There are also limitations of financial intermediations, according to Pyle (1971). Higher transaction costs and information asymmetry are disadvantages of intermediation.

The Arrow-Debreu model contradicts that banks play the role of financial intermediation in the financial market. Arrow-Debreu model demonstrated that in resource allocation, there is no information asymmetry and transaction costs are minimum, therefore the markets are perfect (Arrow & Debreu, 1954). Individuals and firms can be able to trade directly with each other without the involvement of an intermediary (Gianakopulos, 1989; Scholtens & Van Wensveen, 2000).

The crucial task for financial intermediaries is to distribute risk amongst market participants at low costs (Merton, 1989; Merton, 1995b). There are different preferences between risk managers, others prefer stable earnings for personal reasons (Pagano,2001) whereas others prefer optimal risk management strategies (Marcus, 1984).

## 2.4 Bank Revenue

Banks have different sources of revenue and are outlined below:

### 2.4.1 Net interest income from banking operations

Banks accept deposits from its customers and use them to lend money to its customers who need cash. When banks perform this activity, they attempt and hope to make a positive spread. This is the difference between the interest they charge on the loans and the interest they pay to depositors. The interest rate that is paid to depositors is called the cost of funds (COF) and the interest rate that is charged to loan holders is called the benefits of funds (BOF). Therefore, the spread is essentially the difference between the

BOF and COF which is called the net interest income (NII) after allowing for expected bad debt charge. NII turns to be the largest source of revenue for most banks.

This income is sensitive to market and credit risk. If customers fail to pay their loans, the bank's NII will reduce. If the interest rates decrease, the bank's NII will also decrease if the loans they have written are linked to a base rate that reprices frequently. Most lending is linked to prime rate in South Africa.

There are some costs which affect this revenue and are accounted for in the calculation. This is the loan provision that is held for bad debt and this is discussed in the next section.

According to Angbazo (1997), NII is positively linked to capital base, credit risk management quality and non-interest-bearing reserves but negatively related to liquidity of the bank. Theoretical models of NII as suggested by Ho & Saunders (1981), Mchane & Sharpe (1985) and Allen (1988) proved that NII is a function of interest rate risk and bank related factors which impact pricing of bank products, including implicit interest payments and opportunity cost of non-interest-bearing assets. Reduction in credit interest on funding due to decrease in rate negotiations also influences NII (Berger & Udell, 1992). Studies by Nuemark & Sharpe (1992) and Peterson & Rajan (1994) also reveal the same factors. This revenue is excluded in this dissertation as the focus is only on NIR which is also called fee income.

### 2.4.2 Non-interest income

Non-interest income is any income that doesn't involve interest rates. This is the income that the bank generates from charges on transactions and services. The main sources of this income are cash deposits and withdrawals, insurance income, commissions income and trading income which are discussed in section 2.5.3.

According to Stiroh (2004) NIR is an important aspect to diversify a bank's revenue. It is important for banks to amalgamate both types of income (NII and NIR) to reduce their overall earnings volatility. DeYoung & Roland (2001) also proved that fee-based income has an impact on the overall earnings volatility of a bank.

Lepetit et al. (2008) clarifies that growth in bank risk is not due to trading activities but due to fee and commission charging exercises. They confirm that an increase in fee income reduces risk by enhancing operational risk of banks. Sanya & Wolfe (2011)

analysed the effect of NIR on bank performance and insolvency risk and concluded that NIR reduces solvency risk. Brunnermeier at al. (2012) discovered that systematic risk of banks increases due to increase in different sources of NIR. Banks with advanced operational risk management rely heavily on NIR than banks with less developed operational risk management techniques (DeYoung & Rice, 2004).

#### 2.4.2.1 Fee income

Fee income is made up of the following:

**Cash deposits** - income that a bank generates when a cash deposit is made into a customer's transactional account

**Cash withdrawal** - income that a bank makes when a customer withdrawals money from their bank account

**Payments** - income that a bank makes when a customer makes a payment from their bank account to a third-party bank account. There are different forms of payments and different channels used to perform this particular transaction such as electronic or online payment and mobile payment. This is the biggest area where there is a lot of innovation and competition.

**Collections** - this is the type of income that the bank generates by collecting funds from third parties on behalf of their clients or themselves. These are also called debit orders.

**Insurance income** - this income involves insurance premiums that customers pay for cover such as life insurance or car insurance. Banks have also evolved into insurance in the past few years and offer similar types on insurance products and hence the income.

**Commission** - banks also act as distribution channels to offer certain products on behalf of other firms and charge commission for it. For example, this is where one can buy an insurance cover by simply using their banking app.

**Payroll solutions** - This is the type of income a bank generates by offering payroll solutions for their clients. The bank pays the client's employees on their behalf and charge a fee in the process.

There are many different types of NIR revenue, and the list is endless. The above are the most common ones. There are also costs that the banks incur in their pursuit of

generating revenue and these costs are discussed in the next section. This study covered the main revenue items that make up NIR.

### 2.4.3 Trading income

Finally, there is trading income which also falls under NIR. The trading income comes from the trading book which is different from the banking book. Trading book features the contracts that the bank gets into as part of its trading operations. The value of the bank's assets and liabilities are marked to market in the sense that if one of the market variable changes, the value of the trading book will also change. For example, if a bank purchases an asset worth R200 today and the value of the asset drops to R150 tomorrow, then the bank records a trading loss of R50.00. Trading book involves the trading of assets such as derivatives which include options and swaps. This revenue was excluded in this dissertation because it is not a massive NIR item and due to difficulty in obtaining trading income raw data.

## 2.5 The cost base for banks

Costs determine the efficiency of the bank (Adjei-Frimpong et al., 2014). There are interesting findings on costs efficiency and the size of the bank. Tecles & Tabak (2010) proved that larger banks are cost efficient whereas Isik & Hassan (2002) and Altunbas et al. (2007) found that the size of the bank has a negative relationship with cost efficiency.

Banks incur different type of costs and are outlined by Choudhry (2015) as follows:

### 2.5.1 Operational expenses

A bank's operational expenses incorporates administration costs as they run the biggest branch network and process transactions, for example, debit orders, staff salary and sales commission. Staff salaries turn to be the biggest costs to banks.

Typical operation expenses of a bank include the following:

- Marketing
- Staff salary
- Rentals
- Telecommunication
- IT and system infrastructure
- Credit scoring and assessment

- Depreciation of assets

### 2.5.2 Cost of credit

When the bank writes a new loan, it doesn't expect that it is going to receive all the money back. As a result, they hold impairments to cover expected losses. This is a regulatory requirement under international financial reporting standard (IFRS 9). This impairment charge impacts the NII of the bank as it is deducted as an expense on the bank's income statement. Past studies such as Staikouras et al. (2008) and Yildirim & Philippatos (2007) observed a negative relationship between cost efficiency and loan loss provisions. On the contrary, Altunbas et al. (2007) reported a positive association between loan provisions and cost efficiency.

### 2.5.3 Cost of capital

A bank's cost of capital is the minimum required return by shareholders. It is a weighted average of a bank's cost of debt and cost of equity. It reflects the proportion of capital contributed by debt and equity investors. It is also known as WACC. Banks would turn to have a lower cost of capital relative to other financial intermediaries as they have a higher degree of leverage and are highly geared. Having debt means that the bank can benefit from tax advantage as interest payment is tax deductible. Previous studies show mixed results between bank capitalization and cost efficiency. Casu & Girardone (2004) and Yildirim & Philipatos (2007) have reported a positive association between level of capitalization and efficiency. On the other hand, Altunbas et al. (2007) and Kablan (2010) identified a negative relationship. The negative relationship can be attributed to the fact that capital influences costs through its use as a source of funding for loans (Ariff & Luc, 2008; Staikouras et al., 2008).

### 2.5.4 Tax

The corporate tax in South Africa is 28% and any bank is liable to pay it on the pre-tax income. Taxation has a negative impact on the economic growth of South Africa (Dladla & Khobai, 2018). This was also established by Atems (2015), Stoilova (2017) and Badri & Allahyari (2013). Costs were excluded from the analysis because this study purely focuses on obtaining a better forecast model for NIR and not costs.

## 2.6 Bank performance measures

### 2.6.1 Bank Performance measures

Bank performance is a well-researched area and has received much attention in the past years (Kumbirai & Webb, 2010). So many scholars have focused their research on bank performance, including (Yeh, 1996; Webb, 2003; Lacewell, 2003; Halkos & Salamouris, 2004; Tarawneh, 2006). However, increase in the bank performance investigations was triggered by the recent global financial crisis of 2007-2008. The financial crisis was largely attributed to excessive risk taking by banks (Peni & Vähämaa, 2012).

The aim of bank performance measure is to distinguish between performers and non-performers (Berger & Humphrey, 1997). There are two types of approaches used to measure bank performance. They are the accounting approach and the econometric techniques. The accounting techniques make use of accounting ratios to measure performance which are primarily what is broadly used in the industry (Ncube, 2009). Econometric technique measures include those developed by Berger & Humphrey (1997), non-parametric data environmental analysis and parametric stochastic frontier approach. Other performance measures such as capital adequacy and liquidity can also help regulators to intervene at an early stage to limit potential crisis (Casu at al., 2006).

The measures of performance are discussed below.

#### 2.6.1.1 Profitability

This is the main measure of performance of commercial banks and is made up of the following components

**Return on assets (ROA):** ROA is represented as post tax profit as a percentage of total assets**,** it indicates the net income the bank generates per value of an asset. It measures the ability of the bank to acquire deposits and invest them in high return generating assets (Ahmed, 2009). The higher the ROA the more profitable the bank is.

**Return on equity (ROE):** ROE is represented as post tax profit as a percentage of total equity capital. It measures the net income that shareholders get per value of equity invested in the business. The higher the ROE the more profitable the bank to its shareholders.

**Cost to income ratio (CTI):** CTI is represented as total cost as a percentage of total income. It is the measure of the bank's cost relative to its income. It measures how much cost is generated per unit of income. The lower the CTI the more cost efficient is the bank.

#### 2.6.1.2 Liquidity

According to Basel Committee on Banking Supervision (1997) liquidity risk arises because of the bank's failure to fund an increase in assets or to cover short term liabilities. When a bank has a liquidity problem, it cannot be able to raise funding or convert assets into cash. Liquidity can be divided into funding liquidity and market liquidity.[4]Samad (2004) states that liquidity is the blood and life of a commercial bank.

Many previous studies focused on the effect of liquidity on the profitability or net interest margin (e.g., Shen et al., 2001; Barth et al., 2003; Demirgüç-Kunt et al., 2003; Kosmidou et al., 2005; Naceur & Kandil, 2009). There are mixed findings on the relationship between liquidity of a bank and its profitability. Molyneux & Thornton (1992) and Barth et al. (2003) reported a positive association between liquidity and profitability whereas Bourke (1989), Demirgüç-Kunt & Huizinga (1999), Kosmidou (2005,2008) reported the opposite.

Regulators have previously put more emphasis on credit and operational risk and less focus on liquidity risk. Liquidity stress can have negative repercussions on banks as confirmed by the 2007-2008 financial crisis (Matz, 2008).

While liquidity measurement of a bank concentrated on liquidity ratios as outlined below, Poorman & Blacke (2005) demonstrated that it wasn't sufficient to only focus on these ratios to measure the performance of the bank. Chen et al. (2018) analysed other alternatives of measuring liquidity of banks and found that liquidity risk incorporates components of external funding factors, regulatory and supervisory factors, and macro-economic factors.

The following liquidity risk metrics are used as a measurement of liquidity performance:

[4] Funding liquidity refers to the bank's inability to meet short term obligations when they become due because it cannot be able to raise funding. Market liquidity refers to a bank's inability to covert assets into cash without significantly lowering the price as there is not enough dept in the market for that particular asset.

**Loan to deposit ratio (LDT):** is represented as total loans as a percentage of total deposits. It is the measure of the relationship between deposits and customer lending. It measures the self-sustainability of the bank.

**1-week and 1-month liquidity rations** generally measures the gap risk. They demonstrate the net cash flow as a percentage of total liabilities for a certain maturity bucket. These considers the liquidity that will be realised should certain assets be sold during a stressed period. They are normally compared to regulatory limits and the bank needs to ensure that regulatory limits aren't bridged.

**Liquidity risk factor** is represented as average tenor of assets as a percentage of average tenor of liabilities. It is a static snapshot which demonstrate the aggregate size of the liquidity gap. The average tenor is calculated as a weighted average of both the behavioural or contractual tenor using the nominal or actual amount drawn as weights. The ration may be calculated in years or days. The higher the liquidity risk factor, the higher the liquidity gap and risk run by the bank.

**Concentration report and funding source report** shows the extent to which the bank is reliant on a single source of fund. An excessive reliant on single source of deposits is an early warning that should the funder withdraws their monies, the bank will be left with higher liquidity exposure.

**Inter-entity lending report** shows the extent to which a subsidiary to a banking group is reliant on group funds. The more reliance the more the liquidity risk.

There are regulatory metrices which are required to be kept within the required limit as per Basel III guidelines.

**Net Stable funding ratio (NSFR):** NSFR is represented as available amount of stable funding as a percentage of required amount of stable funding. It aims to encourage banks to focus on stable medium to long term deposits. This ratio should always be above 100%.

**Liquidity risk ratio (LCR):** is represented as stock of highly liquid assets as a percentage of total net outflow over the next 30 calendar days. It aims to encourage banks to always have enough liquidity reserves to meet outflows during a period of stress. This ratio should be above 100% all the time.

### 2.6.2 Macro and micro economic factors and banking performance

#### 2.6.2.1 Micro factors

Rina (2004) explored the micro factors that impact bank performance. It was discovered that operational efficiency, net interest margin (NIM), non-performing loans (NPLs), Loan to deposit (LDR) and capital adequacy are the main determinants of bank performance.

NPLs refers to the default book. Loans are classified into three different stages based on the performance of the borrower (Beerbaum & Ahmad, 2015). Stage one represents loans that are up to date with payments, stage two represents loans that missed one or two instalments and finally stage three represents loans that missed at least three instalments. According to Basel Committee on Banking Supervision (2004) stage three loans represents the NPL book. When customers miss their payments, the bank will eventually flag them as defaulted customers and start with the collections process. If there is a collateral attached to the loan, the bank will sell the collateral and recover some of the outstanding loan. If there is no collateral, the bank will lose the outstanding loan amount. Therefore, unsecured loans have higher loss ratios than secured loans. Higher NPL ratios results in lower NIM. There are several studies which prove that higher NPL book reduces profitability of a bank (e.g., Achou & Tenguh, 2008; Kolapo et al., 2012; Musyoki & Kadubo, 2012).

#### 2.6.2.2 Macro factors

GDP, interest rates, inflation and unemployment rates have been found to be amongst macroeconomic factors impacting banking performance (Kosmidou at al., 2005; Choudhry, 2015) . Firm's profitability will rise during economic expansion as more people have jobs and are also willing to buy banking products such as credit and savings products. Kosmidou at al. (2005) reported a positive relationship between bank profitability and GDP. Banks offer loans that are linked to a base rate. If the base rate decrease, it means that the bank's interest income will decrease as well. In the case of South Africa, most loans are linked to the prime lending rate. A classic example are the mortgages. If there is any movement of the base rate, the bank's performance will be impacted either upwards or downwards depending on the direction of the movement. The study of Aburime (2009), Pasioueas & Kosmidou (2007) and sufian et al. (2008) found a positive and statistically significant relationship between the bank's profitability

and interest rates. However, other studies reported a negative correlation such as Bourke (1989).

Inflation is the basis on bank's pricing model. The starting point of a bank's pricing is to look at current inflation level. Higher inflation suggests that banking prices should increase on products and services and vice versa. The research of Molyneux & Thornton (1992) demonstrated a positive relationship between inflation and profitability whereas Bourke (1989) derived a negative relationship. The negative relationship between inflation and profitability is largely attributed to the bank's ability to predict inflation (Bourke, 1989). The higher unemployment rate means that customers don't have enough money to service their existing loans and they will therefore default on their obligations. Their disposable income will also reduce and therefore savings product demand will decline. This will affect bank performance (Choudhry, 2015). According to Dardac & Moinescu (2007), unemployment rate appears to be the most significant macroeconomic variable for explaining systematic short-term movements in performance for institutions offering credit.

### 2.7 Bank Revenue Forecast

The literature regarding forecasting different revenue streams in Banking has drawn attention from different experts such as Actuaries, Economists, and researchers worldwide. This is expected because of the way banking has taken a drastic turn especially given the recent pandemic from 2019, which influenced the transactional behaviour of clients. Most entities are moving towards digital transactions rather than the traditional services, such as cash deposits and withdrawals. Given this, it is expected that the future revenue of the bank will change because digital banking turns to be cheaper.

Wallander (1999) explains the downside of budgeting or forecasting revenue in banking. The author explains that budgeting is an outdated way of driving a bank in the right directions. The author further affirms that budgeting might not be effective in forecasting revenue as it might lead to wrong decision making by senior stakeholders in the bank.

Wallander (1999) explains the different ways that can be adopted to measure the performance of the Bank. One of the most popular tools is Balanced Scorecard, where different matrices are utilised to measure the performance of the Bank. However, the

analysts have applied caution while applying the tools that are relevant to that bank. Wallander (1999) concludes that Budgeting can be applicable in other types of business but not mostly banks. For example, organisations that get revenues from appropriations and rather than markets. This is mostly the case in public sectors, as well as human resource (HR) departments in private companies such as banks.

Maghfuriyah et al. (2019) studied the importance of market structure and behaviour on Islamic banking. They used structure conduct performance to forecast the growth of Islamic banking in both the short term and the long term by utilising the error correction model. They found that market structure drives performance because it describes the segmentation, level of diversification and the cost structure in the banking industry. It also describes the variables that drives competition within the banking industry.

Structure conduct performance is a theory which explains that market structures occurred due to policies being designed and adopted in a particular sector (Beck et al., 2015). Simatele (2015) studied the relationship between the structure of the bank, performance, and competition. Based on the structure conduct, the hypothesis of the study was that competition in the banking sector was very low as the industry was dominated by four major banks. At the time of the study, 80% of market share was held by the big four banks.

There are some policy designs that might reduce the level of competition in an industry. Examples include mergers and acquisition or policies that leads to barricades in new market entrance. This can result in market concentration i.e., the banking industry being dominated by a few banks. This subsequently results in monopolistic behaviour in the sector resulting in the big players making very excessive profits.

In relation to South African banking sector, which is currently dominated by four major banks, there has been a level of increase in competition, especially since Capitec Bank joined the “Big four”, and of recent influx of FinTech start-ups. The market structure and the level of competition in the South African banking sector will impact their future income.

Streimikiene et al. (2018) forecasted the tax revenue using three different time series techniques, including, Autoregressive model (A.R. with seasonal dummies), Autoregressive Integrated Moving Average (A.R.I.M.A), and Vector Autoregressive

(V.A.R) model. The study analysed the efficiency of these three models, and which one was better predicting the tax revenue.

Large scale manufacturing index (L.S.M), which is a close proxy to GDP, Consumer Price Index (CPI) and international oil prices were also modelled using multivariate analysis to identify interactions with the tax revenue. Streimikiene et al. (2018) found that the ARIMA model was a better predictor for Pakistan's tax revenue.

Just like previous scholars, for instance Streimikiene et al. (2018), Maghfuriyah et al. (2019), this study is adopting similar techniques to forecast the transactional revenue in a commercial banking space, as well as the impact of macro-economic factors in the revenue forecast, for example, revenue turns to go up during the economic boom, hence higher GDP.

Streimikiene et al. (2018) found that both models achieve similar results, but because of the simplicity possessed by the AR, it was chosen as the best model for the predictions. This study also tested different models to forecast revenue and a better model was recommended.

### 2.8 Summary of review

This section covered the review of literature relating to this research topic. Evidence from literature has shown that time series modelling has been used in various sectors for forecasting, most importantly, it has been applied in commercial banks to predict revenue (Streimikienie et a., 2018; Maghfuriyah et al., 2019; Zhang, 2003 & Navarro-Esbin et al., 2002). However, transactional data has not been applied on various time series model to test their accuracy. This study explored different time series models on a South African commercial bank and recommend a suitable model for revenue forecast.

# CHAPTER 3: RESEARCH METHODOLOGY

## 3.1 Introduction

This section presents the methods deployed in this study, including research design, population, and sample, as well as the method of data collection and analysis. Limitations of this study and ethical consideration are also presented.

Many studies have been conducted in different fields to project different items such as tax revenue, energy consumptions and cash flows in pursuit of determining the best time series model to adopt. Stiemekienie (2018) adopted three models, ARIMA, VAR and AR and made effort to choose which model was ideal for tax revenue forecast. Other studies such as Jain et al. (2007) recommends the use of hybrid models to help capture all the characteristics of the data.

## 3.2 Research design

This study is quantitative in nature, therefore adopts quantitative research design in investigating the appropriate revenue forecasting model to be adopted by the commercial banks. Leedy & Ormrod (2010) states that statistics established from a quantitative analysis can be used to establish relationships amongst variables. The aim of this study is to establish which time series model is suitable to project revenue in transactional banking. Therefore, a quantitative research design is adopted.

## 3.3 Description of data

For the purpose of this study, monthly time series revenue data from a South African bank, for a period of five years (2016 to 2020) was used. This period has been chosen to ensure that there is relevant data to develop associations between the variables of interest.

This study only focused on data relating to one of the major banks in South Africa. This study is limited to one bank because there has not been enough exploration of time series in revenue projections/forecasting, therefore this was a starting point and future studies can incorporate other banks. Data for interest rates and inflation also covered the same frequency and period for consistency with the revenue data.

## 3.4 Data collection and sources

The data required for this study was collected from a major bank and Statistic South Africa's website. The revenue data was collected from a major bank, while the macroeconomic (inflation and interest rate) data was be collected from the reserve bank website. The period of the data is 2016 to 2020.

## 3.5 Data analysis

This research seeks to identify the relationship between two macro-economic factors and bank's transactional revenue. The analysis uses previous time series data of revenue and macro-economic factors and try to establish to what extend does the independent variables (macro-economic variables, explain the dependent variable, banks' revenue)).

This study forecasts the transactional banking revenue for the period 202101-202110 using four different time series models (AR, ARIMA, MA and VAR), also analyse the impact of inflation and prime interest rate on transactional revenue. The model that gives the most accurate results was recommended as a forecasting model for revenue.

Simple data checks were employed and the validity of data was checked before actual modelling was carried out. This includes checks such as reasonability of the numbers and consistency over time, checking for any missing values and missing data. An eyeball test was conducted across all data sets to ensure that there were no missing values and formats of the data were consistent across all data points.

The analysis of the data begins with testing for stationary of each time series to ensure that the data is stationary, using unit root test. Augmented Dicky Fuller test is used for this purpose. The other tests include testing for relationships between the variables with the revenue. Cointegration analysis and Granger Causality tests are used to test for relationships between variables.

Finally, the revenue was estimated using the identified four models. The robustness of the models are compared through root mean squared error test to pick the better model for forecasting revenue.

### 3.5.1 Unit root test

The generalised equation for unit root test is given as follows:

$$\Delta y_t = \alpha_0 + \alpha_1 y_{t-1} + \sum_{i=1}^{n} \alpha \, \Delta y_t + e_t \quad (1)$$

Where: t is the time period, y denotes the time series, n is the optimum number of lags, $\alpha_0$ is the constant term and e represents the error term.

### 3.5.2 Autoregressive model (A.R. model)

The random walk is one of the most crucial and simplest models in econometrics (Jackson, 2013; Nassar et al.,2004). It is also known as the autoregressive model or simply the AR (1) process. The simplest form of the model is given as follows:

$$y_t = \Psi_1 y_{t-1} + u_t \quad (2)$$

Equation 2 is called a random walk with no drift. If a constant is added to equation 2, then the model becomes a random walk process with drift. AR (p) is said to be stationary if the $\Psi L^{-1}$ converges to zero. A stationary AR(p) model can be expressed as follows:

$$\Psi(L) y_t = u_t$$

$$y_t = \Psi L^{-1} u_t \quad (3)$$

If the lag length increases, then the autocorrelation decays to zero. For a stationary AR (p) process, the following are the roots of the equation:

$$1 - \Psi_1 z - \Psi_2 z^2 - \ldots - \Psi_p z^p = 0 \quad (4)$$

All roots placed on the outer circle are > 1. If the roots of the random walk lie on the circle, we have a non-stationary random walk process and can be computed as follows:

$$y_t = y_{t-1} + u_t$$

$$y_t = L y_t + u_t$$

$$y_t (1-L) = u_t$$

$$1 - z = 0$$

$$z = 1 \quad (5)$$

from equation 5, 1-z is called the characteristic equation and the root z lies on the circle. Finally, the mean and variance of an AR (1) process are given as follows:

$$E(y_t) = \frac{\mu}{1-\Psi_1} \text{ and } Var(y_t) = \frac{\sigma^2}{(1-\Psi_1{}^2)} \quad (6)$$

### 3.5.3 Moving Average (MA) model

The moving average is expressed as follows:

Let $y_t$ = μ + $u_t$ +$\phi_1 u_{t-1}$ + …….. + $\phi_q u_{t-q}$ (7)

Is a qth order moving average model, denoted as MA(q).

An MA is simply linear combination of the white noise process and has been recommended by several studies such as (Lee et al, 2012) and (Hallim, 1981-1986). MA was also chosen for its simplicity in forecasting (Wabomba, 2016). $y_t$ depends on the current and previous values of a white noise disturbance term. The mean and variance of an MA(q) process is given as follows:

E ($y_t$) = μ and Var ($y_t$) = (1 + $\phi_1^2$ + ……+ $\phi_q^2$)$\sigma^2$

### 3.5.4 Autoregressive integrated moving average (A.R.I.M.A) model

ARIMA is a generalised form of the autoregressive moving average (ARMA) in econometrics time series modelling. These from of models are recognized to be the best in estimating financial future values (Box & Jenkins, 1970; McKenzie, 1984; Kalpakis & Gada, 2001). The equation of stationary ARIMA (p,d,q) time series is a sequence of random variables ( $Y_t$ ).

The equation of the ARIMA time series is as follows:

$Y_t$ - $\phi_1 Y_{t-1}$ – L - $\phi_p Y_{t-p}$ = $Z_t$ + $\Theta_1 Z_{t-1}$ + L + $\Theta_q Z_{t-q}$ (8)

$Z_t$ denotes a sequence of uncorrelated random terms which have zero mean and a constant variance. They can be demonstrated as:

$\{Z_t\}$ ~ WN (0, $\sigma^2$) (9)

### 3.5.5 Forecasting error

There is always an error in forecasting financial data. The forecasting errors were used to recommend a suitable model as guided by several studies (Mckenzie, 2011, Leys et al., 2013; Moon &Yao, 2011; Myttenaere, 2016). There are three techniques used to measure the error. Namely, the root mean square error (RMSE), mean absolute percentage error (MAPE), and mean absolute error (MAE). All these techniques are used to test the accuracy of forecasts. The equations of these techniques are given as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \grave{y}_i)^2}{n}} \quad (10)$$

$$\text{MAE} = \frac{\sum_{i=1}^{n}|y_i - \grave{y}_i|}{n} \quad (11)$$

$$\text{MAPE} = \frac{\sum_{i=1}^{n}\left|\frac{y_i - \grave{y}_i}{y_i}\right|}{n} \text{ x } 100\% \quad (12)$$

Where n is the number of predicted values, $y_i$ is the actual observed value and $\grave{y}_i$ is the predicted value at the i-th iteration.

### 3.5.6 Granger causality test

The granger causality test is used to establish relationships between two variables according to Granger (1969). Many studies have already been carried out that uses the granger causality test such as Goebel et al. (2003), Chen et al. (2009) and Mcfarlin (2009). It tries to establish whether one variable can forecast another factor or variable in time series analysis.

The pairwise granger causality is expressed as follows:

$$Y_t = \sum_{n=1}^{p} A_n X_{(t-p)} + \sum_{n=1}^{p} B_n Y_{(t-p)} + CZ_t + E_t \quad (13)$$

$$X_t = \sum_{n=1}^{p} A'_n Y'_{(t-p)} + \sum_{n=1}^{p} B'_n X'_{(t-p)} + C'Z_t + E'_t \quad (14)$$

X and Y are two time series at time period t. $X_{(t-p)}$ and $Y_{(t-p)}$ are two time series at period t-p. Here p represents number of lagged orders of the time period. $A_n$ and $A'_n$ are ratified trace coefficients. $B_n$ and $B'_n$ are known as autoregression model. $Z_t$ is considered as co-variables at period t. Finally, the value of the F statistics is calculated using the equation below:

$$F = \frac{(RSSR - RSSU)/m}{RRSU/(n-k)} \quad (15)$$

Equation 15 follows the $F_{m,n-k}$ distribution where k = m+n++1. If the estimated value of F is greater than the critical value of F, then we reject the null hypothesis and concludes that $X_t$ causes $Y_t$.

An alternative test used to test for relationship is the wild test which is adopted in this study.

### 3.5.7 Vector autoregressive (VAR) model

The Vector Autoregressive (VAR) is a time series model that captures the association between multiple quantities as they differ over time. VAR is chosen to model multivariate time series because it is very flexible and user friendly (Hyyarinen et al., 2003; Ehrmann et al., 2003; Bayraci et al., 2009).

Let $Y_t = (y_{1t}, y_{2t}, \ldots\ldots\ldots\ldots, y_{nt})$ represent (nx1) vectors of time series variables and p lags for the fundamental VAR model. Then VAR (p) model can be expressed as follows:

$$Y_t = c + \Pi_1 Y_{t-1} + \Pi_2 Y_{t-2} + \ldots. + \Pi_p Y_{t-p} + e_t \qquad (16)$$

Where t = 1,..,T. $\Pi_i$ are the (nxn) matrices of coefficients, and $e_t$ is the (nx1) non observable mean of white noise vector procedure, which is consecutively uncorrelated or independent with time invariant covariance matrix $\Sigma$.

The VAR (2) model can be expressed as follows:

$$\begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \pi^1{}_{11} & \pi^1{}_{12} \\ \pi^1{}_{21} & \pi^1{}_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} \pi^2{}_{11} & \pi^2{}_{12} \\ \pi^2{}_{21} & \pi^2{}_{22} \end{pmatrix} \begin{pmatrix} y_{1t-2} \\ y_{2t-2} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} \qquad (17)$$

In equation 17, cov $(e_{1t}, e_{2t}) = \sigma_{12}$ for t = s or 0 otherwise. It is also very crucial to appreciate that each equation has the same lagged estimates of $y_{1t}$ and $y_{2t}$.

## 3.6 Validity and reliability

The sources of the data are reliable. Detailed checks were also carried out on the data to ensure validity. Some basic checks such as searching for missing values, sense checking minimum and maximus, ranges and other data handling were conducted in detail to ensure that the findings of this study draws a valid conclusion on revenue forecasting.

## 3.7 Ethical considerations

This study is based on quantitative design, it did not involve contact with humans and animals during data collection and analysis process. Therefore, ethical requirements on human and animal contacts were not required. However, this research is dealing with banks internal data, but the name of the bank is strictly withheld during this research process. More so, the ethical requirement of Wits University regarding quantitative data analysis was strictly complied with. Ethical clearance was obtained from ethics department of the Wits Business School.

## 3.8 Chapter summary

This chapter touched on the research design and described the data used in this dissertation and its sources. The sources of the data were also explained. Different models chosen to undertake this study were elaborated in detail and justifications of their usage were provided.

# CHAPTER 4: ESTIMATIONS AND RESULTS

## 4.1 Introduction

This section starts by presenting the ADF test results as it is a necessary test to be done before any time series model can be fitted. Thereafter, results for AR (1), ARIMA (1,1,1) and MA (1) are presented together with the different model errors. This section further presents the results for the granger causality test for total NIR with inflation and prime interest rate. The results for VAR (1) together with model errors are presented after the granger causality test as this test is needed before fitting a VAR model. A recommendation is finally made on the suitable model for total NIR based on the results.

AR (1), ARIMA (1,1,1), MA (1) and VAR(1) models were estimated for total NIR. The data from 201701 to 202012 was used as in sample estimation data points. The variable was forecasted for a total of 10 months, from 202101 until 202110. The results are discussed in the next sections. The total NIR from out of sample data between 202101 and 202110 is compared to forecasts to derive the errors.

## 4.2 Augmented Dicky-Fuller (ADF) test

It is essential to test for stationarity of the data before we can apply any time series model to it. The ADF is the most used technique to test for stationarity. Table 1 shows that total NIR, inflation and prime rate have a unit root and are stationery across all different methods after transforming the data and applying first difference. However, the prime rate is not stationary under the trend method after taking first difference. Since the variable is stationary under zero mean and single mean after taking first difference, it is satisfactory that the time series is stationary and we can proceed applying models on the data. Therefore, the time series are of integrated of order I(1) because they are stationary after taking the first difference. Detailed results of the ADF can be found in Appendix I.

Table 1: ADF stationarity test at 5% level of significance.

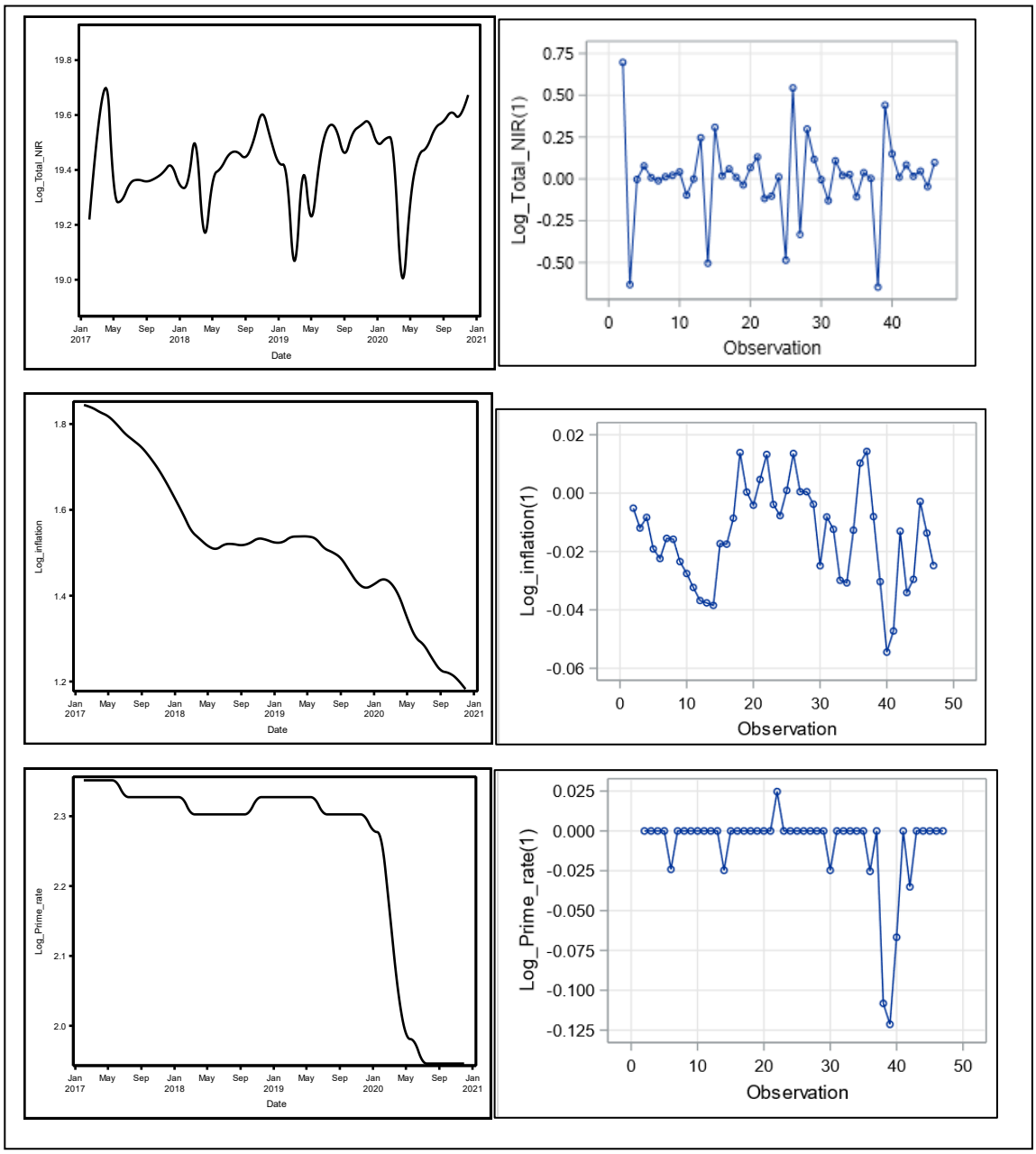
| | Zero Mean | | | | Single Mean | | | | Trend | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | At Level | | At Ist Difference | | At Level | | At Ist Difference | | At Level | | At Ist Difference | |
| Variable | C-Value | P-Value | C-Value | P-Value | C-Value | P-Value | C-Value | P-Value | C-Value | P-Value | C-Value | P-Value |
| Log_Total_NIR | 0.23 | 0.7474 | -13.43 | <.0001 | -6.41 | 0.0001 | -13.26 | 0.0001 | -6.58 | <.0001 | -13.24 | <.0001 |
| Log_Inflation | -5.81 | <.0001 | -2.06 | 0.0392 | 0.09 | 0.962 | -3.01 | 0.0411 | -0.79 | 0.9594 | -2.98 | 0.1489 |
| Log_Prime_rate | -2.18 | 0.0292 | -3.3 | 0.0015 | 1.04 | 0.9964 | -3.49 | 0.0124 | -0.63 | 0.972 | -3.61 | 0.0404 |

Source: Author's calculations.

### 4.2.1 Graphical representation of stationarity

The graphs in Figure 2 show that at first difference the time series become stationary. This is also proof that the time series are of integrated order I(1).

Figure 2: Graphical representation of stationary and no stationary graphs.




Source: Author's calculations.

### 4.3. Autoregressive model (A.R. model) forecast

Autoregressive model of order one is presented in this section together with the errors which compares actual and forecasted results.

#### 4.3.1 Non-interest revenue (NIR) forecast

AR (1) model was estimated for total NIR. The data from 201701 to 202012 was used as a base data for forecasting. The variable was forecasted for a total of 10 months, from 202101 until 202110. The AR (1) parameter is also statistically significant. Total forecasted revenue for the 10 months of forecast is R3.35 billion versus an actual amount of R3.13 billion. The total NIR forecasting errors were also calculated and the results are summarised in Table 2. For detailed results, see Appendix II.

Table 2: RMSE, MAE and MAPE for AR (1).

| | | |
|---|---|---|
| RMSE | R | 25 642 050 |
| MAE | R | 21 617 745 |
| MAPE | | 7.12% |
| Total Actual - Total Forecasted | -R | 216 177 454 |

Source: Author's calculations.

Notes: Total Actual minus Total Forecasted represents the difference between the forecasted revenue and Actual revenue.

### 4.4 Autoregressive integrated Moving Average (ARIMA model) forecast

Autoregressive integrated moving average model of order one is presented in this section together with the errors which compares actual and forecasted results.

#### 4.4.1 Non-interest revenue (NIR) forecast

ARIMA (1,1,1) model was estimated for total NIR. The data from 201701 to 202012 was used as base data for forecasting. The variable was forecasted for a total of 10 months, from 202101 until 202110. The MA (1,1) parameter is statistically significant whereas AR (1,1) parameter is insignificant in this model. Total forecasted revenue for the 10 months of forecast is R3.33 billion versus an actual amount of R3.13 billion. The total NIR forecasting errors were also calculated and the results are summarised in Table 3. For detailed results, see Appendix III.

Table 3: RMSE, MAE and MAPE for ARIMA (1,1,1).

| | | |
|---|---|---|
| RMSE | R | 23 973 618 |
| MAE | R | 20 196 603 |
| MAPE | | 6.65% |
| Total Actual - Total Forecasted | -R | 196 995 513 |

Source: Author's calculations.

Notes: Total Actual minus Total Forecasted represents the difference between the forecasted revenue and Actual revenue

## 4.5 Vector autoregressive model (VAR model) forecast

Vector autoregressive model of order one is presented in this section together with the errors which compares actual and forecasted results.

### 4.5.1 Non-interest revenue (NIR) forecast

MA (1) model was estimated for total NIR. The data from 201701 to 202012 was used as base data for forecasting. The variable was forecasted for a total of 10 months, from 202101 until 202110. The MA (1,1) parameter is statistically significant. Total forecasted revenue for the 10 months of forecast is R3.32 billion versus an actual amount of R3.13 billion. The total NIR forecasting errors were also calculated and the results are summarised in Table 4. For detailed results, see Appendix IV.

Table 4: RMSE, MAE and MAPE for MA (1).

| | | |
|---|---|---|
| RMSE | R | 22 675 657 |
| MAE | R | 18 910 900 |
| MAPE | | 6.24% |
| Total Actual - Total Forecasted | -R | 181 471 120 |

Source: Author's calculations.

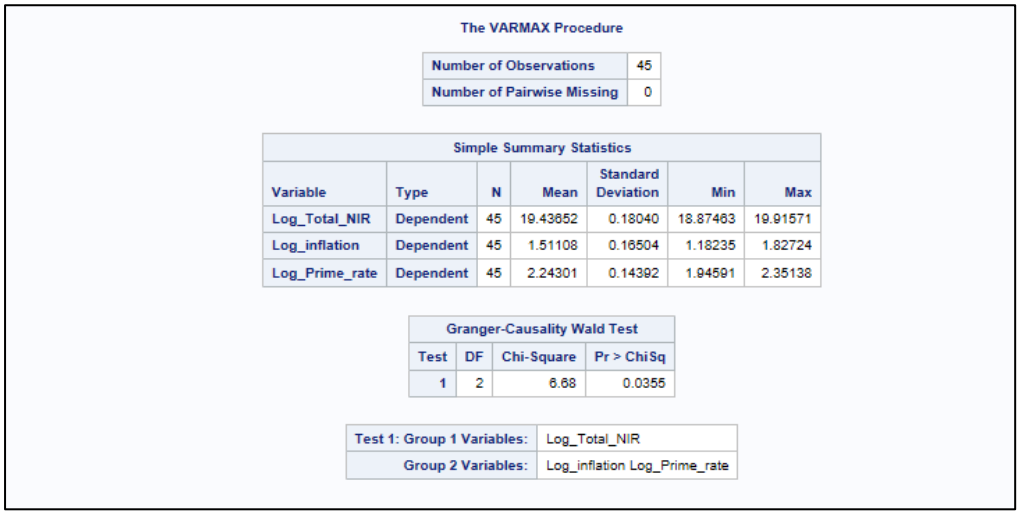
Notes: Total Actual minus Total Forecasted represents the difference between the forecasted revenue and Actual revenue.

## 4.6 Granger Causality test

It important to carry out a test to check if there is any relationship between the variables considered in the model before fitting a VAR model. A granger causality test using the wild test was applied on total NIR with inflation and prime rate in differencing form using lag 1 and 2. The null hypothesis of the test is that the total NIR is only influenced by itself and not by inflation and prime rate.

At 5% level of significance, the null hypothesis is rejected and it is concluded that total NIR is influenced by inflation and prime rate. Therefore, a VAR (1) was fitted, which may provide better results than the other fitted models. The results of the wild test are shown in Figure 3. Detailed results are presented in Appendix V.

Figure 3: Granger causality test using wild test for Total NIR with inflation and prime rate.



The VARMAX Procedure

| | |
|---|---|
| Number of Observations | 45 |
| Number of Pairwise Missing | 0 |

Simple Summary Statistics

| Variable | Type | N | Mean | Standard Deviation | Min | Max |
|---|---|---|---|---|---|---|
| Log_Total_NIR | Dependent | 45 | 19.43652 | 0.18040 | 18.87463 | 19.91571 |
| Log_inflation | Dependent | 45 | 1.51108 | 0.16504 | 1.18235 | 1.82724 |
| Log_Prime_rate | Dependent | 45 | 2.24301 | 0.14392 | 1.94591 | 2.35138 |

Granger-Causality Wald Test

| Test | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|
| 1 | 2 | 6.68 | 0.0355 |

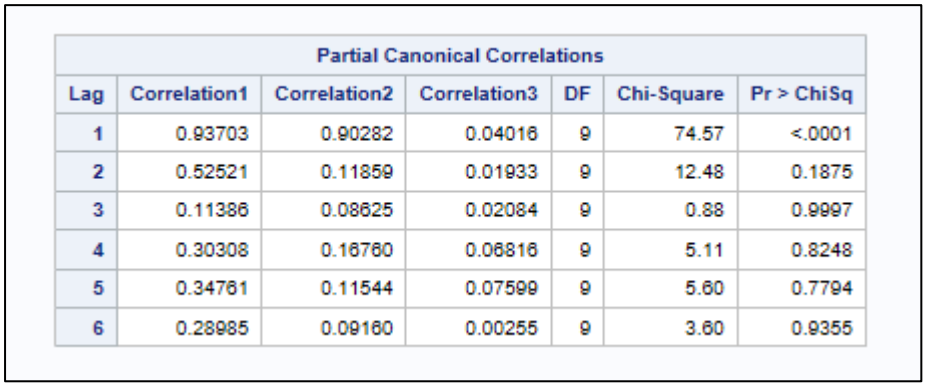
| | |
|---|---|
| Test 1: Group 1 Variables: | Log_Total_NIR |
| Group 2 Variables: | Log_inflation Log_Prime_rate |

Source: Author's calculations.

Note: The test is conducted at 5% level of significance.

Furthermore, the partial cross-correlation matrices in Figure 4 are insignificant after lag 1 with respect to two standard errors. This indicates that an AR order of p=1 is an appropriate choice.

Figure 4: The partial cross-correlation matrices.



Partial Canonical Correlations

| Lag | Correlation1 | Correlation2 | Correlation3 | DF | Chi-Square | Pr > ChiSq |
|---|---|---|---|---|---|---|
| 1 | 0.93703 | 0.90282 | 0.04016 | 9 | 74.57 | <.0001 |
| 2 | 0.52521 | 0.11859 | 0.01933 | 9 | 12.48 | 0.1875 |
| 3 | 0.11386 | 0.08625 | 0.02084 | 9 | 0.88 | 0.9997 |
| 4 | 0.30308 | 0.16760 | 0.06816 | 9 | 5.11 | 0.8248 |
| 5 | 0.34761 | 0.11544 | 0.07599 | 9 | 5.60 | 0.7794 |
| 6 | 0.28985 | 0.09160 | 0.00255 | 9 | 3.60 | 0.9355 |

Source: Author's calculations.

## 4.7 Vector autoregressive (VAR) model

Vector autoregressive model of order one is presented in this section together with the errors which compares actual and forecasted results.

### 4.7.1 Non-interest revenue (NIR) forecast

VAR (1) model was estimated for total NIR. The data from 201701 to 202012 was used as base data for forecasting. The variable was forecasted for a total of 10 months, from 202101 until 202110. All the VAR (1) parameters are not statistically significant. Total forecasted revenue for the 10 months of forecast is R3.08 billion versus an actual amount of R3.13 billion. The total INR forecasting errors were also calculated and the results are summarised in Table 5. For detailed results, see Appendix VI.

Table 5: RMSE, MAE and MAPE for VAR (1).

| | | |
|---|---|---|
| RMSE | R | 14 367 301 |
| MAE | R | 12 640 019 |
| MAPE | | 4.01% |
| Total Actual - Total Forecasted | R | 52 092 415 |

Source: Author's calculations.

Notes: Total Actual minus Total Forecasted represents the difference between the forecasted revenue and Actual revenue.

## 4.8 Chapter Summary

The main aim of this research is to determine which time series is best suitable to model non-interest revenue (NIR) in transactional banking. The results of this study show that the forecasting errors for total NIR are lower under the VAR (1) model. Based on these statistics, it is fair to conclude that the VAR model is a better predictor for total NIR compared to the rest of the other models.

However, it was observed that the parameters produced by the VAR models were not statistically significant. This is a concern because it indicates that the results may be random, therefore, the second-best model based on error was chosen. The MA (1) model is the second-best model based on total non-interest revenue forecasting errors, also by being able to produce statistically significant parameters. Based on this conclusion, it can be argued that the total non-interest revenue forecast from 202101 to 202110 is R3.32 billion.

# CHAPTER 5: SUMMARY AND CONCLUSION

## 5.1 Introduction

This chapter is represented in five sections. The first section looks at the overall summary of findings, followed by policy implications and policy recommendations. Thereafter, an overall conclusion is drawn, finally recommendation for additional research is made.

## 5.2 Summary of findings

Transactional banking NIR is a key financial metrics in the bank as it contributes significantly to its overall profitability. This study presents some insights to comprehend different models available to forecast NIR in transactional banking. The objective of this study was to find a better model that can be utilised to predict NIR. Macro-economic variables that impact NIR were also analysed to determine the extent to which their presence improve predictability of NIR. Such macro-economic variables include inflation and the prime lending interest rate.

The findings showed that VAR (1) and MA (1) are better predictors of NIR than ARIMA (1,1,1) and AR (1) based on prediction errors produced by the model. This further proved that the inclusion of macro-economic factors, inflation, and interest rates is crucial when forecasting revenue. When further comparing VAR (1) and MA (1) using significance of parameters, MA (1) proved to be better and was ultimately chosen as the best model for NIR as it produced significant parameters, whereas this was not the case for ARIMA (1,1,1). MA (1) also proved to be advantageous to use than VAR (1) as it takes less time to run than VAR(1) as it involved different variables (interest rate and inflation) to be analysed whereas with MA(1) it was just the NIR that was forecasted.

The results of this study are consistent with studies that were conducted using these methods such as Valipour et al. (2012), Babu & Reddy (2014) and Lee & Tong (2011). Some other studies such as Streimikien et al. (2018), Maghfuriyah et al. (2019), Zhang (2003) and Navarro-Esbin et al. (2002) found the other models to be efficient. Streimikien et al. (2018) incorporated three macro-economic variables (GDP, inflation, and interest rate) to test efficiency of VAR on tax revenue and still found ARIMA to be better than VAR. This is consistent with this study as adding macro-economic factors doesn't necessarily make the VAR model better than the other ones.

### 5.3 Policy implications

Several policy implications can be drawn from the findings of this study. From budgeting perspective, the use of other methods for forecasting and budgeting has been widely accepted by many banks. The reason is that those methods are high level and quicker, they do not require intense modelling and help the banks reach quick budgeting decisions. But the findings of this research study reveal that MA (1) model is a more efficient model to forecast NIR based on errors and significance of parameters. This means that banks should consider this time series model and encourage its use, as it can lead to better estimates and better budget decisions. Should banks consider this, there will be an impact on their budgeting policies.

Other compelling reasons for banks to consider moving to time series modelling is that the forecast results impact the pricing, price negotiations and finance policies. The bank ultimately uses its forecasted revenue numbers to determine what prices to charge on different transactions and services offered to clients. The more accurate the forecasts are, the better the pricing. There are certain customers that have ability to negotiate their pricing with the banks and accuracy of revenue forecasts for these type of clients helps the bank offer better discounted pricing. Lastly, the finance team uses budget to derive future financial statements. The consideration of banks to move to time series modelling, more especially MA (1) will impact these policies.

The government is a big user of financial data from banks in forecasting its tax revenue. This impacts the fiscal policy as government revenue collection and expenditure influences the economy. Therefore, this means that the new models adopted by banks in forecasting their revenues will impact on government's tax revenue forecasts. The government can also be influenced to adopt similar forecast models to stay consistent with the banking industry.

Finally, different models require a different level of skillset. The current models adopted by many banks are less technical and can be conducted with ease by many employees. Time series modelling are mostly carried out by Quantitative Analysts, Actuaries, Statisticians, and Data Scientists. Therefore, moving to MA (1) model means that banks will need to alter their employment policies to target these skillsets. This will also impact

their remuneration policy as individuals with this kind of skills are currently expensive in the market as their supply is limited.

### 5.4 Policy recommendations

Banks have been undertaking forecasting exercises for many years. The finance department in many banking departments undertake this exercise as they are responsible to ensure cashflows are correctly accounted for in financial statements. Despite having undertook this exercise for long, the techniques used are very easy and high level. In the case of NIR forecasts, the transactional behaviour and seasonality impacts are not accounted for in detail. It is recommended that the budget policy makers support the creation of time series forecast models as MA (1) has proven to be efficient in forecasting NIR. This will provide more meaningful and data insights driven forecasts which in turn helps policy makers derive informed strategic decisions.

Banks should further support the development of time series models by investing in the skillsets needed to undertake the model development and management. This will require banks to invest in new systems that can be able to house these models and allow proper monitoring to be maintained. Banks will further be required to hire experts who can take a lead in implementation of these models.

### 5.5 Conclusion

Moving average model (MA model) has been recommended as the best time series model for forecasting non interest revenue (NIR). This recommendation impacts certain polices within the bank such as budgeting, finance, remuneration and hiring polices as discussed. The recommendation made on alterations in policy have a potential positive impact in better forecasts results and influence on strategic decision making. The recommendations should however be treated with care since they are based on data from only one bank from the financial sector.

### 5.6 Recommendation for further research

The data utilised in this study covered a period of five years. Time series requires a reasonable amount of data to be able to produce credible results. In the future, one can extend the period of the data and remodel, to see if the result could remain the same. This

study also looked at selected models, AR, ARIMA, MA and VAR. Other models can also be deployed to replicate these forecasts, to check if they will be more efficient than the ones used in this study.

GDP was left out in this study because the data points were produced quarterly whereas other data points are monthly, this has caused inconsistency and difficulty in incorporating this variable in this study. Future studies can include this variable in order to attempt to find a better model. Other macro-economic variables such as unemployment rate and economic growth were also omitted. Unemployment rate affects the ability for individuals to spent and buy financial products whereas economic growth influences companies' profitability. Future studies can look at these variables as well.

This study was only focusing on data from one major bank in South Africa. Other banks can be incorporated in this study to check the consistency of the models across banks. This study can also be extended to emerging market perhaps the findings might turn out different. This is possible since banks from emerging markets serve different clientele and operate in different economic conditions.

This study did not incorporate exogenous factors such political unrest, government policies, local and international turmoil, as well as unexpected pandemics such as Covid-19. Almost the rest of the world has experienced Covid-19 which led to a halt in economic activity as many countries were put under lock down restrictions. South Africa has also been affected by political unrests, which affected banking activities in some parts of the country. Future research studies may incorporate these exogenous factors.

This study concentrated on total NIR and did not split it in its individual parts such as cash withdrawals, cash deposits or payments to name a few. Applying this study to different revenue streams may lead to a different recommendation for each. In addition to this, other revenue streams such as NII were not part of this research and their inclusion may result in a completely different outcome.

Another important factor in banking is competition, through digitization. The FinTech banks increased competition in the South African banking sector, as a result, a decline in banking fees occurred. Such impact was not factored into in this study. Future scholars can incorporate the effect of FinTech and the result may show a different outcome on the choice of a model for total NIR.

Finally, the model selection criterion was based on statistical errors and significance of parameters derived from the model. There are many other model selection criteria that can be used in addition to modelling errors and significance of parameters. This includes, but not limited to, AIC, BIC, and residuals. Future studies can consider the use of these statistics as a decision factor.

## REFERENCES


Adjei-Frimpong, K., Gan, C., & Hu, B. (2014). Cost efficiency of Ghana's banking industry: A panel data analysis. The International Journal of Business and Finance Research, 8(2), 69-86.

Aburime, U.T. (2009). Impact of political affiliation on bank profitability in Nigeria. African Journal of Accounting, Economics, Finance and Banking Research, 4(4), 61-75.

Achou, T. F., & Tenguh, N. C. (2008). Bank performance and credit risk management. Master's degree Project, Finance University of Skodve School of Technology and Society.

Allen, F., & Santomero, A. M. (1997). The theory of financial intermediation. *Journal of banking & finance*, *21*(11-12), 1461-1485.

Allen, F., Carletti, E., & Gu, X. (2008). The roles of banks in financial systems. Oxford handbook of banking, 32-57.

Allen, L. (1988). The determinants of bank interest margins: A note. *Journal of Financial and Quantitative analysis*, *23*(2), 231-235.Angbazo, L. (1997). Commercial bank net interest margins, default risk, interest-rate risk, and off-balance sheet banking. Journal of Banking & Finance, 21(1), 55-87.

Ariff, M., & Luc, C. (2008). Cost and profit efficiency of Chinese banks: A non-parametric analysis. China economic review, 19(2), 260-273.

Arrow, K. J., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. Econometrica: Journal of the Econometric Society, 265-290.

Atems, B. (2015). Another look at tax policy and state economic growth: The long-run and short-run of it. Economics Letters, 127, 64-67.

Babu, C. N., & Reddy, B. E. (2014). A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data. Applied Soft Computing, 23, 27-38.

Barth, J. R., Nolle, D. E., Phumiwasana, T., & Yago, G. (2003). A cross-country analysis of the bank supervisory framework and bank performance. Financial Markets, Institutions & Instruments, 12(2), 67-120.

Basel Committee on Banking Supervision, (2004),.Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Bank for International Settlements.

Basel Committee on Banking Supervision, (1997).Core Principles for Effective Banking Supervision. Bank for International Settlements.

Bastan, M., Ramazani, K. R., Delshad, S. S., & Akbarpour, S. (2017). Revenue structure of mobile banking: a system dynamics model. In *International Symposium on Industrial Engineering and Operations Management (IEOM)*.

Bayraci, S., Ari, Y., & Yildirim, Y. (2011). A Vector Auto-Regressıve (VAR) Model for the Turkish Financial Markets.

Beck, T., Senbet, L., & Simbanegavi, W. (2015). Financial inclusion and innovation in Africa: An overview. *Journal of African Economies*, *24*(suppl_1), i3-i11.

Beerbaum, D., & Ahmad, S. (2015). Credit risk according to IFRS 9: Significant increase in credit risk and implications for financial institutions. Sammar, Credit Risk According to IFRS, 9.

Berger, A. N., & Humphrey, D. B. (1997). Efficiency of financial institutions: International survey and directions for future research. *European journal of operational research*, *98*(2), 175-212.

Berger, A. N., & Udell, G. F. (1992). Some evidence on the empirical significance of credit rationing. Journal of Political Economy, 100(5), 1047-1077.

Bourke, P. (1989). Concentration and other determinants of bank profitability in Europe, North America and Australia. Journal of Banking & Finance, 13(1), 65-79.

Brunnermeier, M. K., Dong G.N., & Palia, D. (2012). Banks'Non-Interest Income and Systemic Risk (AFA 2012 Chicago Meetings Paper). Princeton University.

Caplen, B. (2020). Top 1000 World Banks 2020 Africa Press Release: African banks are small but profitable.www.thebanker.com.

Casu, B., & Girardone, C. (2004). Financial conglomeration: efficiency, productivity and strategic drive. Applied Financial Economics, 14(10), 687-696.

Casu, B., Girardone, C., & Molyneux, P. (2006). *Introduction to banking* (Vol. 10). Pearson education.

Chen, G., Hamilton, J. P., Thomason, M. E., Gotlib, I. H., Saad, Z. S., & Cox, R. W. (2009, April). Granger causality via vector auto-regression tuned for fMRI data analysis. In Proc Intl Soc Mag Reson Med (Vol. 17, p. 1718).

Chen, Y. K., Shen, C. H., Kao, L., & Yeh, C. Y. (2018). Bank liquidity risk and performance. Review of pacific basin financial markets and policies, 21(01), 1850007.

Choudhry, M. (2015). *The Moorad Choudhry Anthology: Past, Present and Future Principles of Banking and Finance*. John Wiley & Sons.

DARDAC, N., & MOINESCU, B. (2007). The performance of credit institutions in the context of increased macroeconomic unbalances. In international research conference change leadership in Romania's new economy (p. 96) California State University of San Bernardino, USA.

Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, *74*, 902-924.

Demirgüç-Kunt, A., Laeven, L., & Levine, R. (2003). The impact of bank regulations, concentration, and institutions on bank margins. Concentration, and Institutions on Bank Margins (April 15, 2003).

De Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean absolute percentage error for regression models. Neurocomputing, 192, 38-48.

DeYoung, R., & Rice, T. (2004). Noninterest income and financial performance at US commercial banks. Financial review, 39(1), 101-127.

DeYoung, R., & Roland, K. P. (2001). Product mix and earnings volatility at commercial banks: Evidence from a degree of total leverage model. Journal of Financial Intermediation, 10(1), 54-84.

Dladla, K., & Khobai, H. (2018). The impact of taxation on economic growth in South Africa.

Ehrmann, M., Ellison, M., & Valla, N. (2003). Regime-dependent impulse response functions in a Markov-switching vector autoregression model. Economics Letters, 78(3), 295-299.

Fama, E. F. (1980). Banking in the Theory of Finance. *Journal of monetary economics*, *6*(1), 39-57.

Geanakoplos, J. (1989). Arrow-Debreu model of general equilibrium. In *General Equilibrium* (pp. 43-61). Palgrave Macmillan, London.

Granger, C. W. J. (1969). Investigating Casual Relations by Econometric Models and Cross-Spektral Methods Econometrica 37.

Goebel, R., Roebroeck, A., Kim, D. S., & Formisano, E. (2003). Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. Magnetic resonance imaging, 21(10), 1251-1261.

Hallin, M. (1981). Nonstationary first-order moving average processes: the model-building problem (No. 2013/2199). ULB--Universite Libre de Bruxelles.

Halkos, G. E., & Salamouris, D. S. (2004). Efficiency measurement of the Greek commercial banks with the use of financial ratios: a data envelopment analysis approach. Management accounting research, 15(2), 201-224.

Hallin, M. (1986). Non-stationary q-dependent processes and time-varying moving-average models: invertibility properties and the forecasting problem. Advances in Applied Probability, 18(1), 170-210.

Ho, T. S., & Saunders, A. (1981). The determinants of bank interest margins: theory and empirical evidence. Journal of Financial and Quantitative analysis, 16(4), 581-600.

Hyvärinen, A., Zhang, K., Shimizu, S., & Hoyer, P. O. (2010). Estimation of a structural vector autoregression model using non-gaussianity. Journal of Machine Learning Research, 11(5).

Ifeacho, C., & Ngalawa, H. (2014). Performance of the South African banking sector since 1994. *Journal of applied business research (jabr)*, *30*(4), 1183-1196.

Isik, I., & Hassan, M. K. (2002). Cost and profit efficiency of the Turkish banking industry: An empirical investigation. Financial Review, 37(2), 257-279.

Jackson, L. B. (2013). Digital Filters and Signal Processing: With MATLAB® Exercises. Springer Science & Business Media.

Jain, A., & Kumar, A. M. (2007). Hybrid neural network models for hydrologic time series forecasting. *Applied Soft Computing*, *7*(2), 585-592.

Ketabforoush Badri, A., Dizaji, M., & Allahyari, S. (2014). An analysis of the effects of taxes and GDP on employment in Iran's economy. European Online Journal of Natural and Social Sciences, 2(3 (s)), pp-1546.

Khair, U., Fahmi, H., Al Hakim, S., & Rahim, R. (2017, December). Forecasting error calculation with mean absolute deviation and mean absolute percentage error. In Journal of Physics: Conference Series (Vol. 930, No. 1, p. 012002). IOP Publishing.

Kolapo, T. F., Ayeni, R. K., & Oke, M. O. (2012). CREDIT RISK AND COMMERCIAL BANKS'PERFORMANCE IN NIGERIA: A PANEL MODEL APPROACH. Australian journal of business and management research, 2(2), 31.

Kosmidou, K., Tanna, S., & Pasiouras, F. (2005). Determinants of profitability of domestic UK commercial banks: panel evidence from the period 1995-2002. In Money Macro and Finance (MMF) Research Group Conference (Vol. 45, pp. 1-27).

Kosmidou, K. (2008). The determinants of banks' profits in Greece during the period of EU financial integration. Managerial finance.

Kosmidou, K., Tanna, S., & Pasiouras, F. (2005, June). Determinants of profitability of domestic UK commercial banks: panel evidence from the period 1995-2002. In Money Macro and Finance (MMF) Research Group Conference (Vol. 45, pp. 1-27).

Kumbirai, M., & Webb, R. (2010). A financial ratio analysis of commercial bank performance in South Africa. *African Review of Economics and Finance*, *2*(1), 30-53.

Lacewell, S. K. (2003). Do efficient institutions score well using ratio analysis? An examination of commercial banks in the 1990s. Academy of Banking Studies Journal, 2, 17.

Lee, W. I., Chen, C. W., Chen, K. H., Chen, T. H., & Liu, C. C. (2012). A comparative study on the forecast of fresh food sales using logistic regression, moving average and BPNN methods. Journal of Marine Science and Technology, 20(2), 4.

Leedy, P. D., & Ormrod, J. E. (2005). *Practical research* (Vol. 108). Saddle River, NJ: Pearson Custom.

Lee, Y. S., & Tong, L. I. (2011). Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. Knowledge-Based Systems, 24(1), 66-72.

Lepetit, L., Nys, E., Rous, P., & Tarazi, A. (2008). The expansion of services in European banking: Implications for loan pricing and interest margins. Journal of Banking & Finance, 32(11), 2325-2335.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. Journal of experimental social psychology, 49(4), 764-766.

Maghfuriyah, A., Azam, S., & Shukri, S. (2019). Market structure and Islamic banking performance in Indonesia: An error correction model. *Management Science Letters*, *9*(9), 1407-1418.

McFarlin, D. R., Kerr, D. L., Green, D. E., & Nitschke, J. B. (2009). The effect of controllability on context dependent granger causality in snake phobia fMRI data at 400-ms Resolution. Soc. Neurosc, 289.

Marcus, A. J. (1984). Deregulation and bank financial policy. Journal of banking & finance, 8(4), 557-565.

Matz, L. (2008). Liquidity risk: New lessons and old lessons. Financial Managers Society Inc. White Paper.

Merton, R. C. (1989). On the application of the continuous-time theory of finance to financial intermediation and insurance. The Geneva Papers on Risk and Insurance-Issues and Practice, 14(3), 225-261.

Merton, R. C. (1995). A functional perspective of financial intermediation. Financial management, 23-41.

McKenzie, J. (2011). Mean absolute percentage error and bias in economic forecasting. Economics Letters, 113(3), 259-262.

McShane, R. W., & Sharpe, I. G. (1985). A time series/cross section analysis of the determinants of Australian trading bank loan/deposit interest margins: 1962–1981. Journal of Banking & Finance, 9(1), 115-136.

Molyneux, P., & Thornton, J. (1992). Determinants of European bank profitability: A note. Journal of banking & Finance, 16(6), 1173-1178.

Musyoki, D., & Kadubo, A. S. (2012). The impact of credit risk management on the financial performance of banks in Kenya for the period. International Journal of Business and Public Management, 2(2), 72-80.

Naceur, S. B., & Kandil, M. (2009). The impact of capital requirements on banks' cost of intermediation and performance: The case of Egypt. Journal of Economics and Business, 61(1), 70-89.

Nassar, S., Schwarz, K. P., & Noureldin, A. (2004). Modeling inertial sensor errors using autoregressive (AR) models. NAVIGATION, Journal of the Institute of Navigation, 51(4), 259-268.

Navarro-Esbrı, J., Diamadopoulos, E., & Ginestar, D. (2002). Time series analysis and forecasting techniques for municipal solid waste management. *Resources, conservation and Recycling*, *35*(3), 201-214.

Ncube, M. (2009). Efficiency of the banking sector in South Africa. *University of the Witwatersrand*, 1-44.

Neumark, D., & Sharpe, S. A. (1992). Market structure and the nature of price rigidity: evidence from the market for consumer deposits. The Quarterly Journal of Economics, 107(2), 657-680.

Oldfield, G. S., & Santomero, A. M. (1995). The place of risk management in financial institutions (Vol. 39). Wharton School, University of Pennsylvania.

Pagano, M. S. (2001). How theories of financial intermediation and corporate risk-management influence bank risk-taking behavior. Financial Markets, Institutions & Instruments, 10(5), 277-323.

Pasiouras, F. and Kosmidou, K. (2007). Factors influencing the profitability of domestic and foreign commercial banks in the European Union. Research in International Business and Finance, 21, 222-237.

Pena-Sanchez, Y., & Ringwood, J. (2017). A critical comparison of AR and ARMA models for short-term wave forecasting. In *Proceedings of the 12th European Wave and Tidal Energy Conference 27th Aug-1st Sept 2017* (No. 961, pp. 1-6). European Wave and Tidal Energy Conference 2017.

Peni, E., & Vähämaa, S. (2012). Did good corporate governance improve bank performance during the financial crisis?. *Journal of Financial Services Research*, *41*(1), 19-35.Petersen, M. A., & Rajan, R. G. (1994). The benefits of lending relationships: Evidence from small business data. *The journal of finance*, *49*(1), 3-37.Poorman, F. Jr., and Blake, J. (2005).Measuring and Modeling Liquidity Risk: New Ideas and Metrics. Financial Managers Society Inc. White Paper.

Quiding, C. (2010). *The cost-benefit of regulation in South African banking* (Doctoral dissertation, University of Pretoria).

Samad, A. (2004). Bahrain commercial bank's performance during 1994-2001. *Credit and Financial Management Review*, *10*(1), 33-40.

Sanya, S., & Wolfe, S. (2011). Can banks in emerging economies benefit from revenue diversification?. Journal of Financial Services Research, 40(1), 79-101.

Scholtens, B., & Van Wensveen, D. (2000). A critique on the theory of financial intermediation. *Journal of Banking & Finance*, *24*(8), 1243-1251.

Semih Yildirim, H., & Philippatos, G. C. (2007). Efficiency of banks: Recent evidence from the transition economies of Europe, 1993–2000. European Journal of Finance, 13(2), 123-143.

Shen, C. H., Kuo, C. J., & Chen, H. J. (2001). Determinants of net interest margins in Taiwan banking industry. Journal of Financial Studies, 9(1), 47.

Simatele, M. (2015). Market Structure and competition in the South African banking sector. *Procedia Economics and Finance*, *30*, 825-835.

Staikouras, C., Mamatzakis, E., & Koutsomanoli-Filippaki, A. (2008). Cost efficiency of the banking industry in the South Eastern European region. Journal of International Financial Markets, Institutions and Money, 18(5), 483-497.

Stiroh, K. J. (2004). Diversification in banking: Is noninterest income the answer? Journal of money, Credit and Banking, 853-882.

Stoilova, D. (2017). Tax structure and economic growth: Evidence from the European Union. Contaduría y administración, 62(3), 1041-1057.

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

Sufian, F. and Chong, R.R. (2008). Determinants of bank profitability in a developing economy: empirical evidence from Philippines. Asian Academy of Management Journal of Accounting and Finance, 4(2), 91-112.

Tarawneh, M. (2006). A comparison of financial performance in the banking sector: Some evidence from Omani commercial banks. International Research Journal of Finance and Economics, 3(3), 101-112.

Tecles, P. L., & Tabak, B. M. (2010). Determinants of bank efficiency: The case of Brazil. European Journal of Operational Research, 207(3), 1587-1598.

Thygerson, K., J. (1995). *Management of financial Institutions,* HarperCollins College Publishers.

Valipour, M., Banihabib, M. E., & Behbahani, S. M. R. (2012). Parameters estimate of autoregressive moving average and autoregressive integrated moving average models and compare their ability for inflow forecasting. J Math Stat, 8(3), 330-338.

Wabomba, M. S., Mutwiri, M. P., & Mungai, F. (2016). Modeling and forecasting Kenyan GDP using autoregressive integrated moving average (ARIMA) models. Science Journal of Applied Mathematics and Statistics, 4(2), 64-73.

Wallander, J. (1999). Budgeting—an unnecessary evil. *Scandinavian journal of Management*, *15*(4), 405-421.

Webb, R. (2003). Levels of efficiency in UK retail banks: a DEA window analysis. Int. J. of the Economics of Business, 10(3), 305-322.

Yeh, Q. J. (1996). The application of data envelopment analysis in conjunction with financial ratios for bank performance evaluation. Journal of the operational research society, 47(8), 980-988.

Moon, Y., & Yao, T. (2011). A robust mean absolute deviation model for portfolio optimization. *Computers & Operations Research*, *38*(9), 1251-1258.Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, *50*, 159-175.

Pyle, D. H. (1971). On the theory of financial intermediation. *The Journal of Finance*, *26*(3), 737-747.

# APPENDIX

## Appendix I

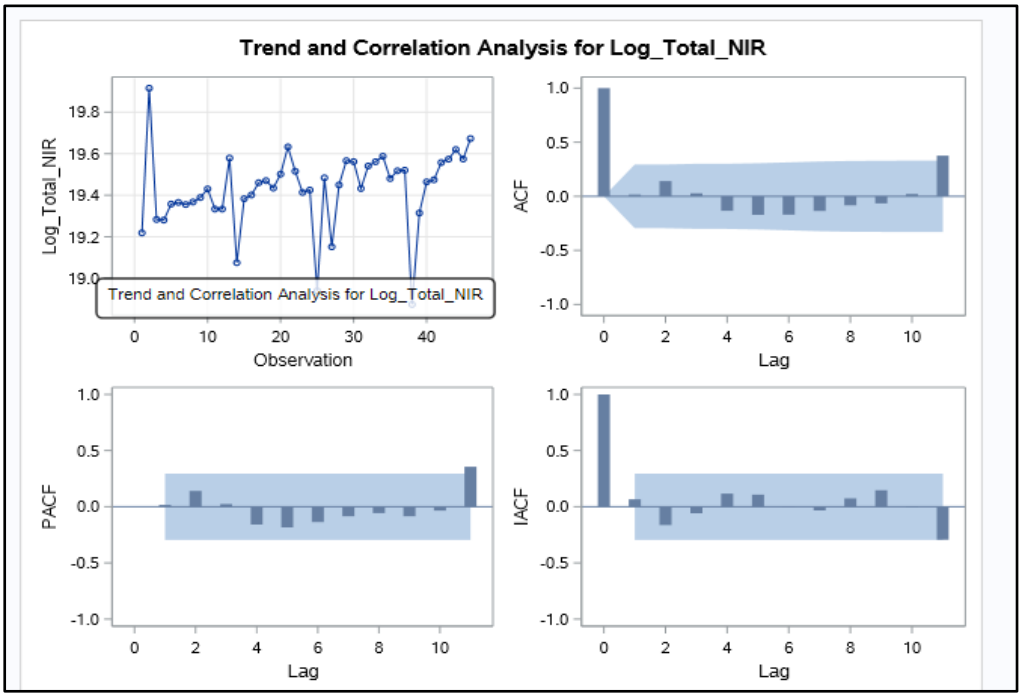
### Appendix I A: ADF test results for total NIR (At level)

The ARIMA Procedure

| Name of Variable = Log_Total_NIR | |
|---|---|
| Mean of Working Series | 19.4318 |
| Standard Deviation | 0.179254 |
| Number of Observations | 46 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 5.18 | 6 | 0.5210 | 0.017 | 0.141 | 0.028 | -0.134 | -0.172 | -0.170 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | 0.0197 | 0.6825 | 0.23 | 0.7474 | | |
| | 1 | -0.0028 | 0.6773 | -0.07 | 0.6561 | | |
| Single Mean | 0 | -44.1861 | 0.0004 | -6.41 | 0.0001 | 20.61 | 0.0010 |
| | 1 | -29.0323 | 0.0003 | -3.74 | 0.0063 | 7.01 | 0.0019 |
| Trend | 0 | -45.8332 | <.0001 | -6.58 | <.0001 | 21.64 | 0.0010 |
| | 1 | -33.8172 | 0.0005 | -4.24 | 0.0086 | 9.62 | 0.0010 |



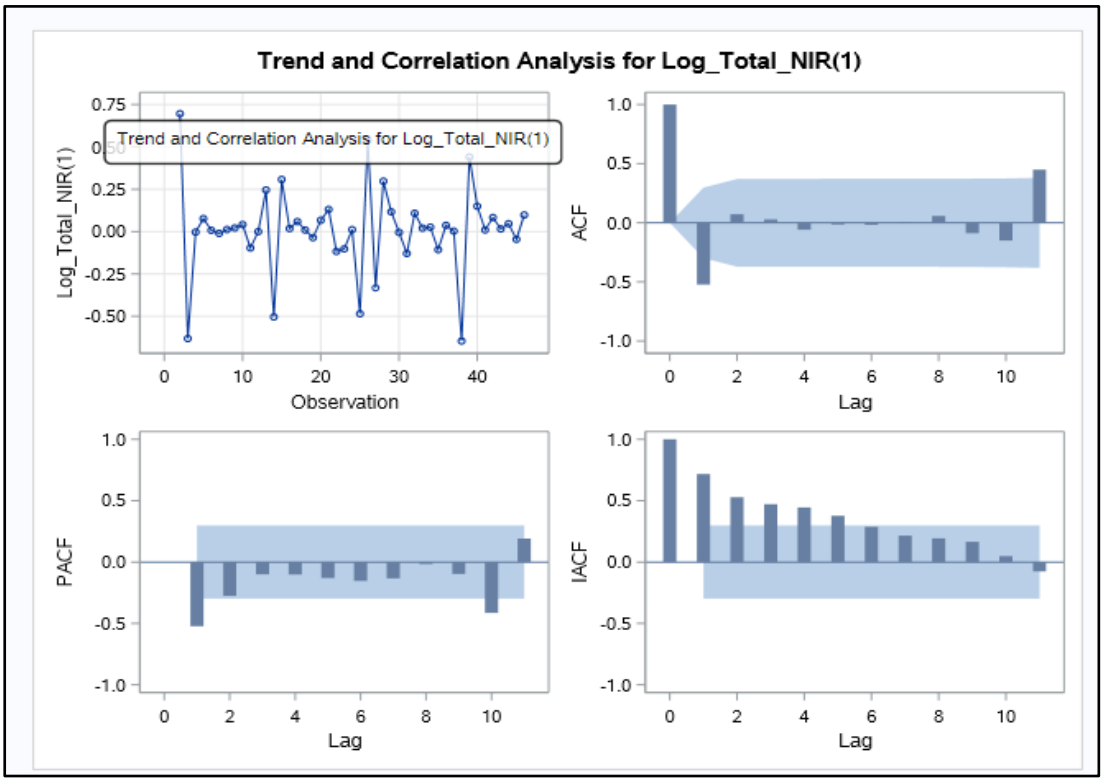
## Appendix I B: ADF test results for total NIR (At first difference)

The ARIMA Procedure

| Name of Variable = Log_Total_NIR | |
|---|---|
| Period(s) of Differencing | 1 |
| Mean of Working Series | 0.010085 |
| Standard Deviation | 0.249406 |
| Number of Observations | 45 |
| Observation(s) eliminated by differencing | 1 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 13.63 | 6 | 0.0340 | -0.523 | 0.073 | 0.028 | -0.059 | -0.014 | -0.018 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -67.0876 | <.0001 | -13.43 | <.0001 | | |
| | 1 | -103.269 | 0.0001 | -7.38 | <.0001 | | |
| Single Mean | 0 | -67.0805 | 0.0003 | -13.26 | 0.0001 | 88.04 | 0.0010 |
| | 1 | -103.363 | 0.0001 | -7.29 | 0.0001 | 26.60 | 0.0010 |
| Trend | 0 | -67.0353 | <.0001 | -13.24 | <.0001 | 88.19 | 0.0010 |
| | 1 | -105.629 | 0.0001 | -7.28 | <.0001 | 26.49 | 0.0010 |



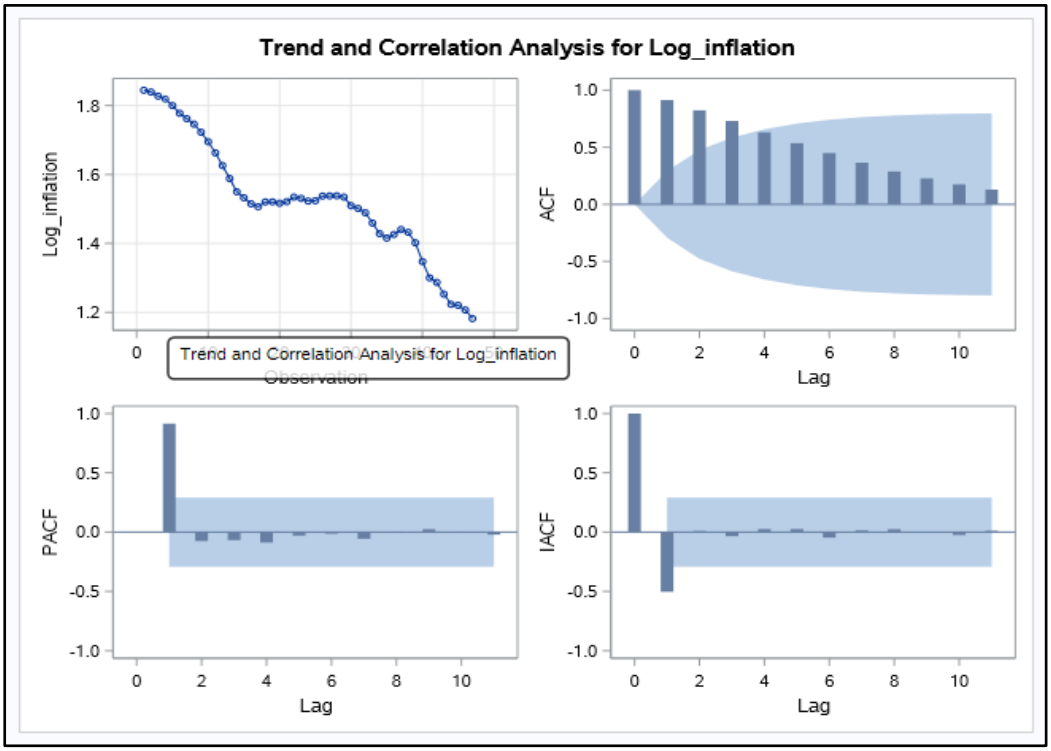
## Appendix I C: ADF test results for inflation (At level)

The ARIMA Procedure

| Name of Variable = Log_inflation | |
|---|---|
| Mean of Working Series | 1.525156 |
| Standard Deviation | 0.173071 |
| Number of Observations | 47 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 152.85 | 6 | <.0001 | 0.914 | 0.824 | 0.730 | 0.631 | 0.536 | 0.449 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -0.4261 | 0.5821 | -5.81 | <.0001 | | |
| | 1 | -0.4592 | 0.5746 | -2.20 | 0.0281 | | |
| Single Mean | 0 | 0.0643 | 0.9571 | 0.09 | 0.9620 | 16.94 | 0.0010 |
| | 1 | -0.8923 | 0.8913 | -0.57 | 0.8668 | 2.41 | 0.4743 |
| Trend | 0 | -1.6755 | 0.9745 | -0.79 | 0.9594 | 0.38 | 0.9900 |
| | 1 | -11.1756 | 0.3117 | -2.18 | 0.4891 | 2.38 | 0.7090 |



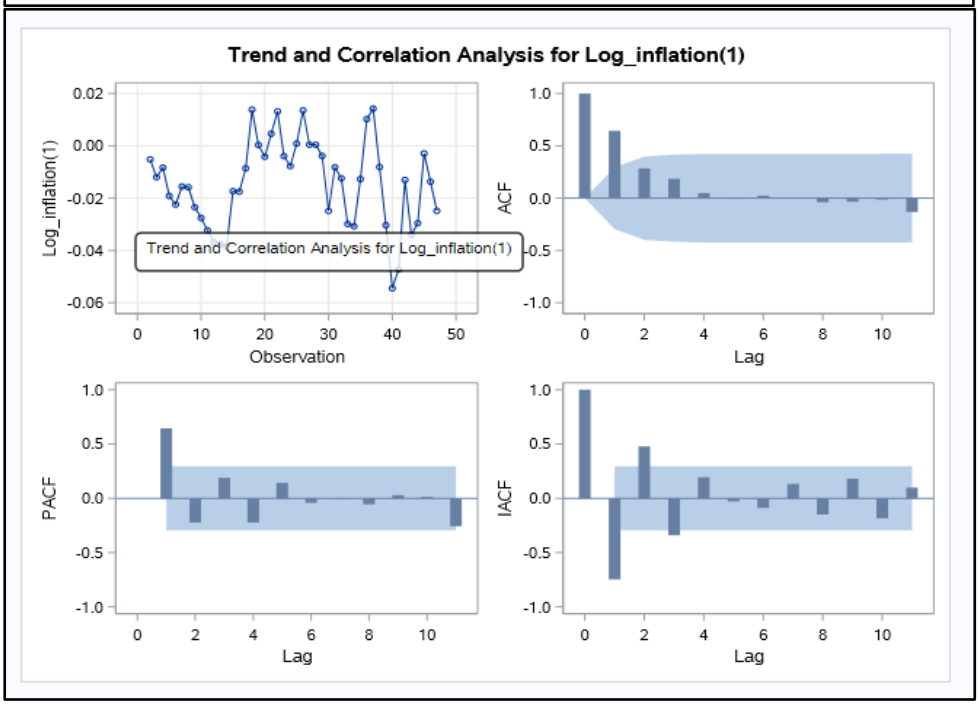
## Appendix I D: ADF test results for inflation (At first difference)

**The ARIMA Procedure**

| Name of Variable = Log_inflation | |
|---|---|
| Period(s) of Differencing | 1 |
| Mean of Working Series | -0.01439 |
| Standard Deviation | 0.016403 |
| Number of Observations | 46 |
| Observation(s) eliminated by differencing | 1 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 26.39 | 6 | 0.0002 | 0.645 | 0.285 | 0.185 | 0.050 | -0.005 | 0.023 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -8.4715 | 0.0387 | -2.06 | 0.0392 | | |
| | 1 | -11.1393 | 0.0164 | -2.23 | 0.0260 | | |
| Single Mean | 0 | -15.7257 | 0.0217 | -3.01 | 0.0411 | 4.56 | 0.0662 |
| | 1 | -24.8422 | 0.0011 | -3.38 | 0.0168 | 5.73 | 0.0270 |
| Trend | 0 | -15.7467 | 0.1151 | -2.98 | 0.1489 | 4.44 | 0.3208 |
| | 1 | -24.9345 | 0.0095 | -3.35 | 0.0720 | 5.61 | 0.0998 |



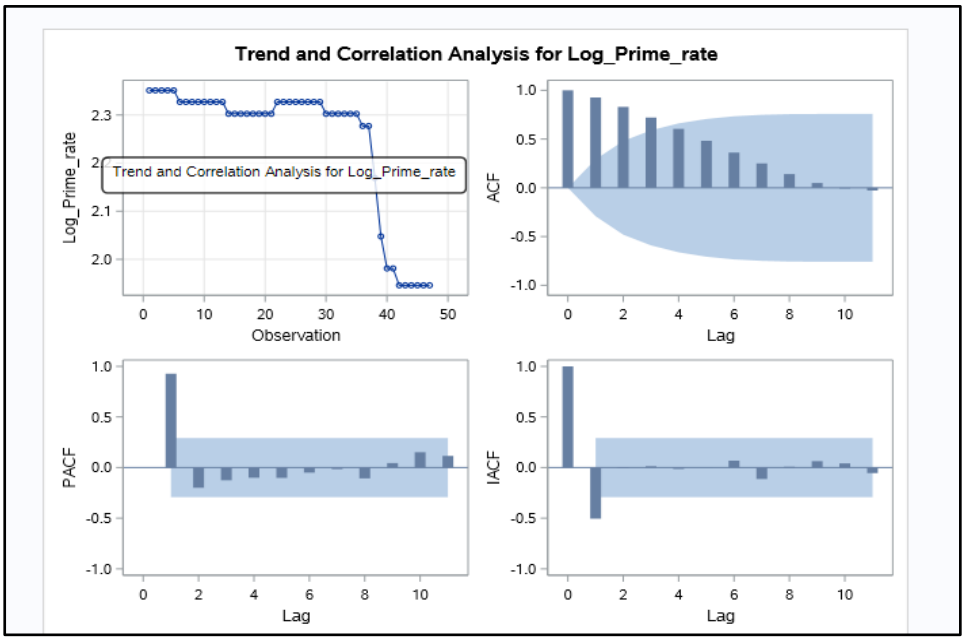
## Appendix I E: ADF test results for prime rate (At level)

The ARIMA Procedure

| Name of Variable = Log_Prime_rate | |
|---|---|
| Mean of Working Series | 2.247622 |
| Standard Deviation | 0.140961 |
| Number of Observations | 47 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 145.13 | 6 | <.0001 | 0.927 | 0.831 | 0.721 | 0.604 | 0.481 | 0.361 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -0.1743 | 0.6391 | -2.18 | 0.0292 | | |
| | 1 | -0.1834 | 0.6369 | -1.15 | 0.2231 | | |
| Single Mean | 0 | 1.3790 | 0.9922 | 1.04 | 0.9964 | 3.09 | 0.3076 |
| | 1 | -1.3211 | 0.8493 | -0.46 | 0.8894 | 0.73 | 0.8833 |
| Trend | 0 | -1.2278 | 0.9828 | -0.63 | 0.9720 | 2.20 | 0.7427 |
| | 1 | -6.3162 | 0.6957 | -1.65 | 0.7587 | 1.81 | 0.8171 |



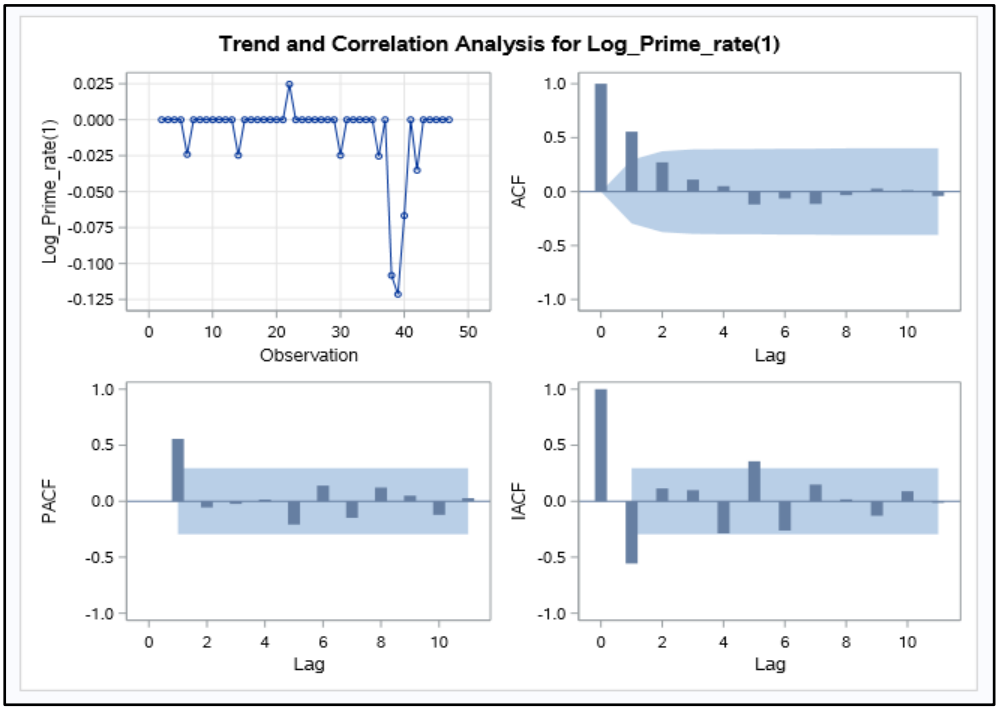
## Appendix I F: ADF test results for prime rate (At first difference)

The ARIMA Procedure

| Name of Variable = Log_Prime_rate | |
|---|---|
| Period(s) of Differencing | 1 |
| Mean of Working Series | -0.00881 |
| Standard Deviation | 0.026208 |
| Number of Observations | 46 |
| Observation(s) eliminated by differencing | 1 |

| Autocorrelation Check for White Noise | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| To Lag | Chi-Square | DF | Pr > ChiSq | Autocorrelations | | | | | |
| 6 | 20.63 | 6 | 0.0021 | 0.556 | 0.271 | 0.112 | 0.051 | -0.119 | -0.063 |

| Augmented Dickey-Fuller Unit Root Tests | | | | | | | |
|---|---|---|---|---|---|---|---|
| Type | Lags | Rho | Pr < Rho | Tau | Pr < Tau | F | Pr > F |
| Zero Mean | 0 | -17.8395 | 0.0018 | -3.30 | 0.0015 | | |
| | 1 | -18.2652 | 0.0015 | -2.95 | 0.0040 | | |
| Single Mean | 0 | -19.9075 | 0.0058 | -3.49 | 0.0124 | 6.11 | 0.0192 |
| | 1 | -21.7446 | 0.0031 | -3.18 | 0.0275 | 5.07 | 0.0445 |
| Trend | 0 | -21.6727 | 0.0250 | -3.61 | 0.0404 | 6.53 | 0.0591 |
| | 1 | -25.3351 | 0.0084 | -3.33 | 0.0753 | 5.56 | 0.1093 |



# Appendix II

## Appendix II A: Forecast results for AR (1) model

| Date | Actual | Forecasted | a = Actual - Forecasted | a squared | \|Actual - Forecasted\| | \|a/Actual\| |
|---|---|---|---|---|---|---|
| 01-Nov-21 | R 286 472 704 | R 329 353 303 | -R 42 880 599 | R 1 838 745 810 050 120 | R 42 880 599 | 0.149685 |
| 01-Dec-21 | R 306 701 544 | R 332 119 314 | -R 25 417 771 | R 646 063 062 783 229 | R 25 417 771 | 0.082875 |
| 01-Jan-22 | R 331 510 212 | R 332 283 094 | -R 772 882 | R 597 346 694 153 | R 772 882 | 0.002331 |
| 01-Feb-22 | R 320 633 079 | R 333 775 620 | -R 13 142 540 | R 172 726 366 063 205 | R 13 142 540 | 0.040989 |
| 01-Mar-22 | R 302 617 920 | R 334 589 664 | -R 31 971 744 | R 1 022 192 430 388 380 | R 31 971 744 | 0.105651 |
| 01-Apr-22 | R 322 478 410 | R 335 750 153 | -R 13 271 743 | R 176 139 149 252 149 | R 13 271 743 | 0.041155 |
| 01-May-22 | R 292 191 500 | R 336 733 741 | -R 44 542 241 | R 1 984 011 216 377 140 | R 44 542 241 | 0.152442 |
| 01-Jun-22 | R 316 635 212 | R 337 807 658 | -R 21 172 446 | R 448 272 471 317 321 | R 21 172 446 | 0.066867 |
| 01-Jul-22 | R 324 033 227 | R 338 835 451 | -R 14 802 224 | R 219 105 832 090 110 | R 14 802 224 | 0.045681 |
| 01-Aug-22 | R 331 683 532 | R 339 886 796 | -R 8 203 264 | R 67 293 546 980 606 | R 8 203 264 | 0.024732 |
| | | | | | | |
| **Total** | **R 3 134 957 339** | **R 3 351 134 794** | **-R 216 177 454** | **R 6 575 147 231 996 420** | **R 216 177 454** | **0.712408** |

| | | |
|---|---|---|
| n | | 10 |
| RMSE | R | 25 642 050 |
| MAE | R | 21 617 745 |
| MAPE | | 7.12% |
| Total Actual - Total Forecasted | -R | 216 177 454 |

**Conditional Least Squares Estimation**

| Parameter | Estimate | Standard Error | t Value | Approx Pr > \|t\| | Lag |
|---|---|---|---|---|---|
| MU | 1043384.0 | 4819699.2 | 0.22 | 0.8294 | 0 |
| AR1,1 | -0.51062 | 0.11813 | -4.32 | <.0001 | 1 |

| | |
|---|---|
| Constant Estimate | 1576155 |
| Variance Estimate | 2.885E15 |
| Std Error Estimate | 53713926 |
| AIC | 2115.956 |
| SBC | 2119.971 |
| Number of Residuals | 55 |

* AIC and SBC do not include log determinant.

**Correlations of Parameter Estimates**

| Parameter | MU | AR1,1 |
|---|---|---|
| MU | 1.000 | -0.014 |
| AR1,1 | -0.014 | 1.000 |

# Appendix III

## Appendix III A: Forecast results for ARIMA (1,1,1) model

| Date | Actual | Forecasted | a = Actual - Forecasted | a squared | \|Actual - Forecasted\| | \|a/Actual\| |
|---|---|---|---|---|---|---|
| 01-Nov-21 | R 286 472 704 | R 326 377 862 | -R 39 905 158 | R 1 592 421 618 245 880 | R 39 905 158 | 0.139298 |
| 01-Dec-21 | R 306 701 544 | R 327 432 637 | -R 20 731 093 | R 429 778 209 926 606 | R 20 731 093 | 0.067594 |
| 01-Jan-22 | R 331 510 212 | R 329 024 955 | R 2 485 257 | R 6 176 503 847 122 | R 2 485 257 | 0.007497 |
| 01-Feb-22 | R 320 633 079 | R 330 662 702 | -R 10 029 623 | R 100 593 333 711 162 | R 10 029 623 | 0.031281 |
| 01-Mar-22 | R 302 617 920 | R 332 304 289 | -R 29 686 370 | R 881 280 545 372 253 | R 29 686 370 | 0.098099 |
| 01-Apr-22 | R 322 478 410 | R 333 946 201 | -R 11 467 790 | R 131 510 217 576 108 | R 11 467 790 | 0.035561 |
| 01-May-22 | R 292 191 500 | R 335 588 140 | -R 43 396 640 | R 1 883 268 334 648 910 | R 43 396 640 | 0.148521 |
| 01-Jun-22 | R 316 635 212 | R 337 230 081 | -R 20 594 869 | R 424 148 642 308 471 | R 20 594 869 | 0.065043 |
| 01-Jul-22 | R 324 033 227 | R 338 872 023 | -R 14 838 795 | R 220 189 845 658 950 | R 14 838 795 | 0.045794 |
| 01-Aug-22 | R 331 683 532 | R 340 513 964 | -R 8 830 432 | R 77 976 537 607 481 | R 8 830 432 | 0.026623 |
| | | | | | | |
| **Total** | **R 3 134 957 339** | **R 3 331 952 852** | **-R 196 995 513** | **R 5 747 343 788 902 940** | **R 201 966 028** | **0.665311** |

| | | |
|---|---|---|
| n | | 10 |
| RMSE | R | 23 973 618 |
| MAE | R | 20 196 603 |
| MAPE | | 6.65% |
| Total Actual - Total Forecasted | -R | 196 995 513 |

**Conditional Least Squares Estimation**

| Parameter | Estimate | Standard Error | t Value | Approx Pr > \|t\| | Lag |
|---|---|---|---|---|---|
| MU | 1641941.5 | 527400.8 | 3.11 | 0.0030 | 0 |
| MA1,1 | 0.96500 | 0.05006 | 19.28 | <.0001 | 1 |
| AR1,1 | 0.08451 | 0.14827 | 0.57 | 0.5711 | 1 |

| | |
|---|---|
| Constant Estimate | 1503175 |
| Variance Estimate | 2.234E15 |
| Std Error Estimate | 47260724 |
| AIC | 2102.829 |
| SBC | 2108.851 |
| Number of Residuals | 55 |

* AIC and SBC do not include log determinant.

**Correlations of Parameter Estimates**

| Parameter | MU | MA1,1 | AR1,1 |
|---|---|---|---|
| MU | 1.000 | 0.664 | 0.238 |
| MA1,1 | 0.664 | 1.000 | 0.362 |
| AR1,1 | 0.238 | 0.362 | 1.000 |

# Appendix IV

## Appendix IV A: Forecast results for MA (1) model

| Date | Actual | Forecasted | a = Actual - Forecasted | a squared | \|Actual - Forecasted\| | \|a/Actual\| |
|---|---|---|---|---|---|---|
| 01-Nov-21 | R 286 472 704 | R 324 530 009 | -R 38 057 306 | R 1 448 358 517 144 290 | R 38 057 306 | 0.132848 |
| 01-Dec-21 | R 306 701 544 | R 326 110 640 | -R 19 409 096 | R 376 713 011 807 713 | R 19 409 096 | 0.063283 |
| 01-Jan-22 | R 331 510 212 | R 327 691 270 | R 3 818 942 | R 14 584 315 860 632 | R 3 818 942 | 0.011520 |
| 01-Feb-22 | R 320 633 079 | R 329 271 900 | -R 8 638 821 | R 74 629 229 652 501 | R 8 638 821 | 0.026943 |
| 01-Mar-22 | R 302 617 920 | R 330 852 531 | -R 28 234 611 | R 797 193 271 310 101 | R 28 234 611 | 0.093301 |
| 01-Apr-22 | R 322 478 410 | R 332 433 161 | -R 9 954 751 | R 99 097 062 893 121 | R 9 954 751 | 0.030870 |
| 01-May-22 | R 292 191 500 | R 334 013 791 | -R 41 822 291 | R 1 749 104 055 438 220 | R 41 822 291 | 0.143133 |
| 01-Jun-22 | R 316 635 212 | R 335 594 422 | -R 18 959 210 | R 359 451 646 099 750 | R 18 959 210 | 0.059877 |
| 01-Jul-22 | R 324 033 227 | R 337 175 052 | -R 13 141 825 | R 172 707 561 439 799 | R 13 141 825 | 0.040557 |
| 01-Aug-22 | R 331 683 532 | R 338 755 682 | -R 7 072 151 | R 50 015 318 352 572 | R 7 072 151 | 0.021322 |
| | | | | | | |
| **Total** | **R 3 134 957 339** | **R 3 316 428 460** | **-R 181 471 120** | **R 5 141 853 989 998 700** | **R 189 109 004** | **0.623654** |

| | | |
|---|---|---|
| n | | 10 |
| RMSE | R | 22 675 657 |
| MAE | R | 18 910 900 |
| MAPE | | 6.24% |
| Total Actual - Total Forecasted | -R | 181 471 120 |

**Conditional Least Squares Estimation**

| Parameter | Estimate | Standard Error | t Value | Approx Pr > \|t\| | Lag |
|---|---|---|---|---|---|
| MU | 1580630.3 | 496803.1 | 3.18 | 0.0024 | 0 |
| MA1,1 | 0.95298 | 0.04833 | 19.72 | <.0001 | 1 |

| | |
|---|---|
| Constant Estimate | 1580630 |
| Variance Estimate | 2.205E15 |
| Std Error Estimate | 46955167 |
| AIC | 2101.163 |
| SBC | 2105.178 |
| Number of Residuals | 55 |

* AIC and SBC do not include log determinant.

**Correlations of Parameter Estimates**

| Parameter | MU | MA1,1 |
|---|---|---|
| MU | 1.000 | 0.522 |
| MA1,1 | 0.522 | 1.000 |

# Appendix V

## Appendix V A: Results for Granger causality test – Wild test

| Partial Canonical Correlations | | | | | | |
|---|---|---|---|---|---|---|
| Lag | Correlation1 | Correlation2 | Correlation3 | DF | Chi-Square | Pr > ChiSq |
| 1 | 0.93703 | 0.90282 | 0.04016 | 9 | 74.57 | <.0001 |
| 2 | 0.52521 | 0.11859 | 0.01933 | 9 | 12.48 | 0.1875 |
| 3 | 0.11386 | 0.08625 | 0.02084 | 9 | 0.88 | 0.9997 |
| 4 | 0.30308 | 0.16760 | 0.06816 | 9 | 5.11 | 0.8248 |
| 5 | 0.34761 | 0.11544 | 0.07599 | 9 | 5.60 | 0.7794 |
| 6 | 0.28985 | 0.09160 | 0.00255 | 9 | 3.60 | 0.9355 |

The VARMAX Procedure

| | |
|---|---|
| Number of Observations | 45 |
| Number of Pairwise Missing | 0 |

| Simple Summary Statistics | | | | | | |
|---|---|---|---|---|---|---|
| Variable | Type | N | Mean | Standard Deviation | Min | Max |
| Log_Total_NIR | Dependent | 45 | 19.43652 | 0.18040 | 18.87463 | 19.91571 |
| Log_inflation | Dependent | 45 | 1.51108 | 0.16504 | 1.18235 | 1.82724 |
| Log_Prime_rate | Dependent | 45 | 2.24301 | 0.14392 | 1.94591 | 2.35138 |

| Granger-Causality Wald Test | | | |
|---|---|---|---|
| Test | DF | Chi-Square | Pr > ChiSq |
| 1 | 2 | 6.68 | 0.0355 |

| | |
|---|---|
| Test 1: Group 1 Variables: | Log_Total_NIR |
| Group 2 Variables: | Log_inflation Log_Prime_rate |

The VARMAX Procedure

| Type of Model | VAR(1) |
|---|---|
| Estimation Method | Least Squares Estimation |

**Model Parameter Estimates**

| Equation | Parameter | Estimate | Standard Error | t Value | Pr > \|t\| | Variable |
|---|---|---|---|---|---|---|
| Log_Total_NIR | CONST1 | 19.32406 | 2.73032 | 7.08 | 0.0001 | 1 |
| | AR1_1_1 | 0.03119 | 0.13819 | 0.23 | 0.8226 | Log_Total_NIR(t-1) |
| | AR1_1_2 | -0.45600 | 0.25528 | -1.79 | 0.0816 | Log_inflation(t-1) |
| | AR1_1_3 | 0.08352 | 0.29090 | 0.29 | 0.7755 | Log_Prime_rate(t-1) |
| Log_inflation | CONST2 | -0.56840 | 0.24206 | -2.35 | 0.0239 | 1 |
| | AR1_2_1 | 0.02173 | 0.01225 | 1.77 | 0.0838 | Log_Total_NIR(t-1) |
| | AR1_2_2 | 0.92841 | 0.02263 | 41.02 | 0.0001 | Log_inflation(t-1) |
| | AR1_2_3 | 0.10679 | 0.02579 | 4.14 | 0.0002 | Log_Prime_rate(t-1) |
| Log_Prime_rate | CONST3 | -0.40984 | 0.47107 | -0.87 | 0.3895 | 1 |
| | AR1_3_1 | 0.01770 | 0.02384 | 0.74 | 0.4623 | Log_Total_NIR(t-1) |
| | AR1_3_2 | 0.02186 | 0.04404 | 0.50 | 0.6225 | Log_inflation(t-1) |
| | AR1_3_3 | 1.01049 | 0.05019 | 20.13 | 0.0001 | Log_Prime_rate(t-1) |

**Covariances of Innovations**

| Variable | Log_Total_NIR | Log_inflation | Log_Prime_rate |
|---|---|---|---|
| Log_Total_NIR | 0.02539 | -0.00003 | 0.00176 |
| Log_inflation | -0.00003 | 0.00020 | 0.00010 |
| Log_Prime_rate | 0.00176 | 0.00010 | 0.00076 |

| Log-likelihood | 372.4068 |
|---|---|

| Information Criteria | |
|---|---|
| AICC | -681.454 |
| HQC | -696.904 |
| AIC | -708.814 |
| SBC | -676.698 |
| FPEC | 3.829E-9 |

# Appendix IV

## Appendix IV A: Forecast results for VAR (1) model

| Date | Actual | Forecasted | a = Actual - Forecasted | a squared | \|Actual - Forecasted\| | \|a/Actual\| |
|---|---|---|---|---|---|---|
| 01-Nov-21 | R 286 472 704 | R 300 301 040 | -R 13 828 336 | R 191 222 874 316 739 | R 13 828 336 | 0.048271 |
| 01-Dec-21 | R 306 701 544 | R 303 702 996 | R 2 998 548 | R 8 991 289 508 518 | R 2 998 548 | 0.009777 |
| 01-Jan-22 | R 331 510 212 | R 304 924 471 | R 26 585 740 | R 706 801 590 488 465 | R 26 585 740 | 0.080196 |
| 01-Feb-22 | R 320 633 079 | R 306 269 627 | R 14 363 453 | R 206 308 770 879 302 | R 14 363 453 | 0.044797 |
| 01-Mar-22 | R 302 617 920 | R 307 643 252 | -R 5 025 332 | R 25 253 961 308 350 | R 5 025 332 | 0.016606 |
| 01-Apr-22 | R 322 478 410 | R 309 050 342 | R 13 428 069 | R 180 313 032 502 806 | R 13 428 069 | 0.041640 |
| 01-May-22 | R 292 191 500 | R 310 491 719 | -R 18 300 219 | R 334 898 009 958 354 | R 18 300 219 | 0.062631 |
| 01-Jun-22 | R 316 635 212 | R 311 968 393 | R 4 666 819 | R 21 779 198 738 599 | R 4 666 819 | 0.014739 |
| 01-Jul-22 | R 324 033 227 | R 313 481 379 | R 10 551 849 | R 111 341 510 987 391 | R 10 551 849 | 0.032564 |
| 01-Aug-22 | R 331 683 532 | R 315 031 707 | R 16 651 825 | R 277 283 268 170 312 | R 16 651 825 | 0.050204 |
| | | | | | | |
| **Total** | **R 3 134 957 339** | **R 3 082 864 924** | **R 52 092 415** | **R 2 064 193 506 858 840** | **R 126 400 189** | **0.401425** |

| | | |
|---|---|---|
| n | | 10 |
| RMSE | R | 14 367 301 |
| MAE | R | 12 640 019 |
| MAPE | | 4.01% |
| Total Actual - Total Forecasted | R | 52 092 415 |

**The VARMAX Procedure**

| | |
|---|---|
| Type of Model | ARX(1,1) |
| Estimation Method | Least Squares Estimation |

**Model Parameter Estimates**

| Equation | Parameter | Estimate | Standard Error | t Value | Pr > \|t\| | Variable |
|---|---|---|---|---|---|---|
| Total_NIR | CONST1 | 395022078.62 | 83497066.373 | 4.73 | 0.0001 | 1 |
| | XL0_1_1 | 4399661.3001 | 122090387.42 | 0.04 | 0.9714 | Inflation(t) |
| | XL0_1_2 | 58695430.133 | 32662001.205 | 1.80 | 0.0801 | Prime_rate(t) |
| | XL1_1_1 | -9810979.226 | 113961755.99 | -0.09 | 0.9318 | Inflation(t-1) |
| | XL1_1_2 | -66303605.57 | 33323233.427 | -1.99 | 0.0537 | Prime_rate(t-1) |
| | AR1_1_1 | -0.04176 | 0.15981 | -0.26 | 0.7952 | Total_NIR(t-1) |

**Variance Estimate for the Innovation**

| Total_NIR |
|---|
| 2.2817408E15 |

| | |
|---|---|
| Log-likelihood | -814.964 |

**Information Criteria**

| | |
|---|---|
| AICC | 1646.955 |
| HQC | 1648.642 |
| AIC | 1643.928 |
| SBC | 1656.574 |
| FPEC | 2.586E15 |

**Forecasts**

| Variable | Obs | Time | Forecast | Standard Error | 95% Confidence Limits | |
|---|---|---|---|---|---|---|
| Total_NIR | 47 | 47 | 300301039.69 | 49722851.947 | 202846040.66 | 397756038.71 |
| | 48 | 48 | 303702995.79 | 49813165.300 | 206070985.85 | 401335005.73 |
| | 49 | 49 | 304924471.49 | 49839837.669 | 207240184.66 | 402608758.31 |
| | 50 | 50 | 306269626.77 | 49869174.343 | 208527841.12 | 404011412.42 |
| | 51 | 51 | 307643251.52 | 49900357.940 | 209840347.14 | 405446155.90 |
| | 52 | 52 | 309050341.52 | 49933501.820 | 211182476.33 | 406918206.71 |
| | 53 | 53 | 310491718.93 | 49968680.202 | 212554905.38 | 408428532.48 |
| | 54 | 54 | 311968392.81 | 50005973.694 | 213958485.36 | 409978300.26 |
| | 55 | 55 | 313481378.52 | 50045467.602 | 215394064.43 | 411568692.61 |
| | 56 | 56 | 315031706.77 | 50087252.249 | 216862496.28 | 413200917.27 |