COMPARATIVE ANALYSIS OF AUTOREGRESSIVE INTEGRATED MOVING AVERAGES MODEL ON RESAMPLED AND NON-RESAMPLED INFLATION TIME SERIES DATA FOR FORECASTING.

BY

DAVID, Peter Oaya

NSU/M.Sc/STA/0004/16/17

A DISSERTATION SUBMITTED TO THE SCHOOL OF POSTGRADUATE STUDIES, NASARAWA STATE UNIVERSITY KEFFI, IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER IN STATISTICS

DEPARTMENT OF STATISTICS

FACULTY OF NATURAL AND APPLIED SCIENCES

NASARAWA STATE UNIVERSITY, KEFFI NIGERIA

AUGUST, 2018

DECLARATION

I hereby declare that this dissertation "**Comparison of Autoregressive Integrated Moving Averages Model on Bootstrapped Samples and Non- Bootstrapped Samples of Inflation Data for Forecasting**" has been written by me and it is a report of my project work. It has not been presented in any previous application for Master Degree (MSc) in Statistics. All quotations are as indicated and sources of information are specifically acknowledged by means of references.

David Peter Oaya **Date**

NSU/M.Sc/STA/0004/16/17

CERTIFICATION

This dissertation "**Comparison Of Autoregressive Integrated Moving Averages Model On Bootstrapped Samples And Non- Bootstrapped Samples Of Inflation Data For Forecasting**" has been read and approved, having met the regulation governing the award of Masters of Science (M.Sc) degree in Statistics, School of Post Graduate Studies, Nasarawa State University, Keffi.

Dr. M.O Adenomon **Chairman, Supervisory committee**	**Date**
Dr. N. O. Nweze **Member, Supervisory committee**	**Date**
Dr. N.O Nweze **Head of Department**	**Date**
Dr(Mrs) M.U. Adehi **Internal Examiner**	**Date**
Prof. B.S. Jatau **Dean, Faculty of Natural & Applied Sciences**	**Date**
Dr. J. A. Ikughur **External Examiner**	**Date**
Prof. S. A. S. Aruwa **Dean, School of Postgraduate Studies**	**Date**

DEDICATION

This work is dedicated to God Almighty whose grace, favor and faithfulness I enjoyed throughout this program and to my lovely Mother Rebecca David Oaya for her love.

ACKNOWLEDGEMENTS

I wish to express my profound gratitude and appreciation to the following individuals whom their contributions have earnestly allowed this success to be part of history.

Firstly, to my supervisor Dr. N. O. Nweze for creating time, despite the accumulation of work as the head of department to go through my work. May Almighty God reward you for your input.

Also deserving my acknowledgement are my lecturers, Prof. C.H. Nwankwo, Dr. H. Yahaya, Dr. H.K. Oduwole and Mr. Chaku Emmanuel who piloted me through this breakthrough. I am also grateful to my course mates whose contribution to this work cannot be overlooked.

To my wonderful siblings and friends, Victor, Lydia, Usa, Peace, Pauline, Lawi, Umar, David, Michael, Joro, Emma who always understood with me most times I presented my needs, I am most grateful.

Much thanks to the staff and management of Yusal Teleview Nigeria limited and Mr. Daniel Richards for their financial Support.

ABSTRACT

The study compared the efficiency of establishing of variance obtained from Autoregressive Integrated Moving Averages (ARIMA) model for resampled inflation time series and actual inflation time series. Using the Akaike Information Criterion (AIC), most efficient forecast was obtained. Results shows that the models fit on resampled inflation series produced higher AICs than the actual data. Thus the models fit on the actual data are considered better for forecast.

TABLE OF CONTENTS

Title Page	i
Declaration	iii
Certification	iv
Dedication	v
Acknowledgments	vi
Abstract	vii
Table of Contents	viii

CHAPTER ONE
INTRODUCTION

1.1 Background of the Study	1
1.2. Statement of the Problem	2
1.3. Aim and Objectives of the Study	3
1.4. Statement of Hypothesis	3
1.5. Significance of the Study	3
1.6. Scope of the Study	3
1.7 Definition of Operational Terms	4

CHAPTER TWO

LITERATURE REVIEW

2.1 Concepual Framework	5
2.2 Autoregressive integrated moving average (ARIMA)	6
2.3. Information Criterion	8
2.4 Univariate Times series models	9
2.5 Application of Time Series	10
2.6 Structural Theory of Inflation	11

2.7 Empirical Literature of Inflation 12

CHAPTER THREE
RESEARCH METHODLOGY

3.1 Research Design 17

3.2. Methods of Data Collection 17

3.3 Bootstrap procedure 17

3.3.1Fixed Block Sampling 18

3.3.2 Stationary Block Sampling 18

3.3.3 Model-based Sampling 18

3.4 The ARIMA Process 19

CHAPTER FOUR

DATA PRESENTATION AND ANALYSIS

4.1 Data Presentation 21

4.2 Data Analysis and Results 21

4.2.1 Test for Stationarity using the ADF (Identification Statge of The ARIMA) 22

4.2.2 ARIMA Model Estimation 22

4.3. Forecasting With Arima Model 23

4.3.1 ARIMA Model of Resampled (Bootstrapped) Inflation Data 27

4.4 Findings 29

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary 30

5.2 Conclusions 30

5.3 Recommendations 30

5.4 Limitations of the Study 31

5.5 Contributions to Knowledge 31

5.5 Area of future Research 31

REFERENCES 32

APPENDIX 34

CHAPTER ONE

INTRODUCTION

1.2 Background of the Study

Model selection is an important aspect of statistics because it allows researchers to test the relative quality of one model over another. In this work, the Autoregressive Integrated Moving Averages (ARIMA) model will not be compare to a different model but to the same of its kind with the difference in the type of data used to fit in the model. The study will focus on resampled inflation time series data and non-resampled inflation time series data and use to develop a model, the basic method for the selection in this work is the Akaike information criterion.

The The Auto Regressive Integrated Moving Averages (ARIMA) model is a class of statistical models for analyzing and forecasting time series data. It explicitly caters to a suite of standard structures in time series data, and as such provides a simple yet powerful method for making skillful time series forecasts. ARIMA is an acronym that stands for Autoregressive Integrated Moving Average. It is a generalization of the simpler Autoregressive Moving Average and adds the notion of integration.

Inflation is a critical component of any country's economy as it affects the lives of citizens with high impact. Therefore to forecast inflation one must be certain about the prediction so as to get close to accurate forecast. The value of the naira has been depreciating over years. Its purchasing power has been weakening as such it has been buying fewer and fewer goods and services each succeeding year. There has been sustainable rise in the prices of goods and services in Nigeria over the years. This situation is what economists describe as inflation. A highly volatile rate of inflation has the potential to do great economic harm. With nominal contracts, uncertainty about future prices is likely to entail higher risk. These costs mean that

for a given average inflation rate, higher inflation volatility can depress economic growth (Elder, 2004; Fatás and Mihov, 2005; Grier and Grier, 2006).

There is a large body of economic theory and econometric evidence dealing with impact of monetary policy on inflation volatility (for example, Gali and Monacelli, 2005; Leduc and Sill, 2007). Historically, weak institutions have been blamed for both high average inflation and high inflation volatility in developing countries, and close attention has been paid to the consequences for inflation volatility of alternative macroeconomic policy regimes (Bleaney and Fielding, 2002; Rother, 2004; Aisen and Veiga, 2006).

Inflation affects the economy negatively in many ways. It causes problems for economic activities, for example, it discourages investment limits export, and also creates problems for the economy in terms of income distribution and erodes purchasing power. Inflation derails the purchasing power of persons and the business entities, it brings to the minimum the quantity of goods and services that their income can buy, and this render the income of individuals meaningless. This situation obviously brings about a fall in the standard of living of the people and creates severe problems for the growth and development of the economy.

To adequately forecast this type of a time series data, it is important that the model to used should produce the most desired results.

1.2. Statement of the Problem

There is keen interest on what spending will be in the nearest future, as this solely depends on the income of people. The need to fit a model on resampled and non-resampled time series data is important to get accurate forecast and check the precision between the variance of these models.

1.3. Aim and Objectives of the Study

1.3.1 Aim

The aim of this study is to compare the efficiency of the resampled and actual inflation time series data.

1.3.2 Objective

i. To fit ARIMA model on inflation time series data

ii. To fit ARIMA model on resampled time series data

iii. To compare the efficiency of the models obtained from the resampled inflation time series data and non-resampled data.

1.4. Statement of Hypothesis

The research is aimed at bootstrapping previous data to verify the following statement;

> The ARIMA model for bootstrapped inflation samples is not more effective for forecasting than the ARIMA model for the non-bootstrapped inflation samples.

1.5. Significance of the Study

The study is relevant in comparing the ARIMA model fit on resampled time series data and non-resampled time series data for better forecast.

1.6. Scope of the Study

This study shall cover the period of 2014 – 2018.

1.7 Definition of Operational Terms

Inflation Rate: Inflation as measured by the consumer price index reflects the annual percentage change in the cost to the average consumer of acquiring a basket of goods and services that may be fixed or changed at specified intervals, such as yearly. The Laspeyres formula is generally used.

Exchange Rate: Official exchange rate refers to the exchange rate determined by national authorities or to the rate determined in the legally sanctioned exchange market. It is calculated as an annual average based on monthly averages (local currency units relative to the U.S. dollar).

Stationarity

This is a process that has a constant mean and variance that do not change over time and the process do not have a trend.

Bootstrapping: In statistics, bootstrapping is any test or metric that relies on random sampling with replacement. Bootstrapping allows assigning measures of accuracy (defined in terms of bias, variance, confidence intervals, prediction error or some other such measure) to sample estimates

Autoregressive *(*AR*)* Process: Statistical forecasting model in which future values are computed only on the basis of past values of a time series data.

CHAPTER TWO

LITERATURE REVIEW

2.1 Conceptual Framework

Singh and Xie (2008) stated that to understand bootstrap, suppose it were possible to draw repeated samples (of the same size) from the population of interest, a large number of times. Then, one would get a fairly good idea about the sampling distribution of a particular statistic from the collection of its values arising from these repeated samples. But, that does not make sense as it would be too expensive and defeat the purpose of a sample study. The purpose of a sample study is to gather information cheaply in a timely fashion. The idea behind bootstrap is to use the data of a sample study at hand as a "surrogate population", for the purpose of approximating the sampling distribution of a statistic; i.e. to resample (with replacement) from the sample data at hand and create a large number of "phantom samples" known as bootstrap samples.

The bootstrap technique introduced by Efron (1979) could possibly be a potential alternative in estimation and inference from time series models in finite samples. However, in time series regressions, the standard bootstrap resampling method designed for independent and identically distributed *(iid)* errors is not applicable because in most situations the assumption of *iid* errors is violated. The term "resampling" has been applied to a variety of techniques for statistical inference, among which stochastic permutation and the bootstrap are the most characteristic. There are at least four major types of resampling methods which include the randomization exact test, cross-validation, jackknife, and bootstrap.

Although today they are unified under a common theme, it is important to note that these four techniques were developed by different people at different periods of time for different purposes.

R.A. Fisher (1935) noted that with a large sample the exact Fisher test is not feasible because of the computational difficulty (before the age of powerful computers). Hence, in his later years Fisher lost interest in the permutation method because there were no computers in his days to automate such a laborious method. As a remedy, people suggested that a randomly-generated selected subset of the possible permutations could provide the benefits of the permutation test without excessive computational cost.

2.2 Autoregressive integrated moving average ARIMA

The Auto Regressive Integrated Moving Averages (ARIMA) An ARIMA model is a class of statistical models for analyzing and forecasting time series data.

Autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. Both of these models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). ARIMA models are applied in some cases where data are not stationary, to make these data stationary, the differencing is applied one or multiple times before fitting the ARIMA model. The differencing explains the “integrated” in the ARIMA.

It explicitly caters to a suite of standard structures in time series data, and as such provides a simple yet powerful. Contreras et al. (2003) used ARIMA models to predict next day electricity prices; they have found two ARIMA models to predict hourly prices in the electricity markets of Spain & California. The Spanish model needed 5 hours to predict future prices as opposed to the 2 hours needed by the Californian model.

Kumar et al (2004) used ARIMA model to forecast daily maximum surface ozone concentrations in Brunei Darussalam. They found that ARIMA (1, 0, 1) was suitable for the surface 0_3 data collected at the airport in Brunei Darussalam. Tsitsika et al. (2007) used ARIMA

model to forecast pelagic fish production. The final model selected were of the form ARIMA (1, 0, 1) and ARIMA (0, 1, 1). Azad et al. (2011) used ARIMA model in forecasting Exchange Rates of Bangladesh. By using Box Jenkins methodology they tried to find out best Model for forecasting.

The goodness of fit test of the optimum ARIMA (0, 3, and 1) model showed non-significant autocorrelation in the residuals of the model. Datta (2011) used ARIMA model in forecasting inflation in the Bangladesh Economy. He showed that ARIMA (1, 0, 1) model fits the inflation data of Bangladesh satisfactorily. Al-Zeaud (2011) used ARIMA model in modeling and forecasting volatility. The result shows the best ARIMA models at 95% confidence interval for banks sector is ARIMA (2, 0, and 2) model. Uko et al. (2012) examined the relative predictive power of ARIMA, VAR and ECM models in forecasting inflation in Nigeria. The result shows that ARIMA model is a good predictor of inflation in Nigeria and serves as a benchmark model in inflation forecasting.

This acronym is descriptive, capturing the key aspects of the model itself. Briefly, they are:

i. **AR**: *Autoregression*. A model that uses the dependent relationship between an observation and some number of lagged observations.
ii. **I**: *Integrated*. The use of differencing of raw observations (e.g. subtracting an observation from an observation at the previous time step) in order to make the time series stationary.
iii. **MA**: *Moving Average*. A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

Each of these components are explicitly specified in the model as a parameter. A standard notation is used of ARIMA (p,d,q) where the parameters are substituted with integer values to

quickly indicate the specific ARIMA model being used. The parameters of the ARIMA model are defined as follows:

p: The number of lag observations included in the model, also called the lag order.

d: The number of times that the raw observations are differenced, also called the degree of differencing.

q: The size of the moving average window, also called the order of moving average.

A linear regression model is constructed including the specified number and type of terms, and the data is prepared by a degree of differencing in order to make it stationary, i.e. to remove trend and seasonal structures that negatively affect the regression model.

A value of 0 can be used for a parameter, which indicates to not use that element of the model. This way, the ARIMA model can be configured to perform the function of an ARMA model, and even a simple AR, I, or MA model.

Adopting an ARIMA model for a time series assumes that the underlying process that generated the observations is an ARIMA process. This may seem obvious, but helps to motivate the need to confirm the assumptions of the model in the raw observations and in the residual errors of forecasts from the model.

2.3 Information Criterion

The Akaike information criterion (AIC) is a measure of the relative goodness of fit of statistical mode, it suggested measuring the goodness of fit for some particular model by balancing the error of against the number of parameters in the model. It provided the measures of information lost when a given model is yet to describe reality. AIC values provides a means for model

selection and cannot say anything about how well a model fits the data fits the data in an absolute sense.

2.4 Univariate Times series models

Howrey (1980) stated that the time series models are based on a relatively weak nonparametric formulation of the model. There was more emphasis on the data analysis for simplification of the model. Different patterns or trends can be seen in the time series data. The time series is influenced by several factors like random components, seasonal components, cyclic components etc. In time series model, the past behavior of a time series is examined to extrapolate something about its future behavior instead of searching for consequences of one or more variables on the forecast variable. The random component in the time series may cover the influence of other components and make it difficult to be able to describe the observed trends or patterns in the data. Removal or reduction of random components from the time series will result in better forecasting or interpretation of series.

Importance of Time Series

(i) **To analyze the behavioral pattern of collected samples:** When analysis is done on past data, it helps describe the effects of external forces on the data. Investigations of such assist us to understand and set aside various set of factors in the problem under study. For example, a time series analyst can study the past behavior of rainfall in a particular geographical location over a given period of time and predict what rainfall will be like in the same area taking into consideration the factors affecting it.

(ii) **Evaluation of current trend:** the review and evaluation of progress made on the basis of plans are done on the basis of time series data. For instance, the

progress of fifteen years plans by a Government can be evaluated on the growth of the employment rate or inflation.

(iii) **Forecasting ability:** prediction is one of the paramount importances of time series that is projecting into the future of what the data under consideration will behave. For example, the analysis of time series relating to inflation would serve as a basis for forecasting future inflation rate.

(iv) **It helps in making comparatives studies:** when data are arranged in chronological order, comparison can be made between one-time period and another is facilitated. For example, a time series analyst can study series on inflation and unemployment rates of a country and then compare the forecast with the inflation and unemployment rates published by the Central Bank, World Bank, or National Bureau of Statistics. It also helps in making regional comparison amongst data collected on the basis of time.

2.5 Application of Time Series

Stofer et al (2010) explained that time series can be applied in diverse field of life which include the economics, where we are continually exposed to daily stock market quotations or monthly unemployment and employment figures, loss and profit making of our business on a daily basis. The social scientist follows population series, such as birth rates and death rates or school enrollment and the ratios of male to women in schools, community, Local Government Area, State, a geopolitical zone and possibly Country. In Medicine blood pressure measurements traced over time could be useful for evaluating drugs used in treating hypertension, thypoid, malaria fever, HIV/AIDS etc, and most intensive sophisticated applications of time series methods have been used to solve problems in the physical and environmental sciences, which also accounts for the basic

engineering flavor permeating the language of time series analysis. More modern investigations may center on whether a warming is present in global temperature measurements or whether levels of pollution may influence daily mortality in a particular place. Common features in a time series characteristic known as the power spectrum are used to help computers recognize and translate speech. Geophysical time series such as those produced by yearly depositions of various kinds can provide long-range proxies for temperature and rainfall. Seismic recordings can aid in mapping fault lines or in distinguishing between earthquakes and nuclear explosions.

2.6 Structural Theory of Inflation

Nnamocha (2002) defined inflation as a sustained rise in the general level of prices. Inflation affects an economy by increasing the opportunity cost of holding money, that is, the uncertainty over future rise in prices which may discourage savings and investment. High inflation is often associated with excessive growth of money supply. Though, money supply does not necessarily cause inflation when it grows in the same proportion as output. However, moderate inflation rates may be attributed to fluctuations in real demand for goods and services or changes in available supplies such as in periods of scarcity.

Inflation is the rate at which general prices of goods and services rise persistently over a period of time consequent upon continual rise in demand for goods and services relative to shortages in the supply of the goods and services. In order words, inflation is the persistent rise in the general price level over a period of time. Inflation has serious implication for the function of money as a medium of exchange and store of value. ₦120.00 could you a tin of Peak milk in 2013, but in 2014 you now need ₦130.00 to buy the same tin of Peak milk at the same or lesser quality. If this signify general price rise for

goods and services in the economy, this means that the resulting inflation rate for the CPI in one year is 8.33 percent, meaning that the general price level for typical Nigerian consumers rose by 8.33 percent. Hence persistent increase in general price level, means that each unit of the currency (Naira) will buy fewer goods and services. Inflation reflects a reduction in the purchasing power per unit of naira, which is a loss of real value in the medium of exchange and unit of account within the country

2.7 Empirical Literature of Inflation

Kasidi (2013), studying the impact of inflation on economic growth in Tanzania analyzed time-series data for the period 1990 to 2011 using the Ordinary Least Square (OLS). The study findings suggested that inflation has a negative impact on economic growth in Tanzania, and also that there was no long run relationship between inflation and economic growth during the period of study in Tanzania. Doguwa (2013), estimating the existence and level of inflation threshold in the relationship between inflation and economic growth in Nigeria, used three different approaches to estimate the threshold level of inflation above which money is not super-neutral. Vaona (2012) studying "inflation and growth in the long-run: a new Keynesian theory and further semiparametric evidence" concluded that inflation have negative linear impact on economic growth. Umaru and Zubairu (2012) examined the impact of inflation on economic growth and development in Nigeria between 1970-2010 in their study "Effect of Inflation on the Growth and Development of the Nigerian Economy: An Empirical Analysis" came to the conclusion that inflation possessed a positive impact on economic growth through encouraging productivity and output level and on evolution of total factor productivity. Marbuah, (2010) in the study, "The Inflation-Growth Nexus: Testing for Optimal Inflation for Ghana", investigated the relationship between inflation and

economic growth to ascertain whether a significant threshold effect existed in the case of Ghana over the period 1955-2009. The study found evidence of significant threshold effect of inflation on economic growth. The evidence showed both a minimum and maximum inflation threshold levels of 6% and 10% respectively. The study also found that adjusting for structural break in the model increases the effect of inflation on growth at a robust threshold level of 10% by a factor of 1.8 or approximately 81%. The study recommended the continue pursuit of inflation targeting framework by keeping inflation targets below 10% for beyond 10% threshold, inflation can be detrimental to Ghana's growth prospects. Hasanov, (2010), employed annual data set on growth rate of real GDP, Consumer Price Index, Inflation and growth rate of real Gross Fixed Capital Formation to investigate whether there was any threshold effect of inflation on economic growth over the period of 2001-2009 in the study, "Relationship between Inflation and Economic Growth in Azerbaijani Economy. Is there any Threshold Effect?" Estimated threshold model indicated that there was non-linear relationship between inflation and economic growth in the Azerbaijani economy and threshold level of inflation for GDP growth was 13 percent. Inflation rate lower than 13 percent reflected statistically significant positive effect on GDP growth but this positive relationship became negative when inflation exceeded 13 percent. He added that, economic growth was expected to decline by about 3 percent when inflation increased above the 13 percent threshold. Frimpong and Oteng-Abayie, (2010) in their study When is Inflation harmful? Estimating the Threshold Effect for Ghana found a threshold effect of inflation on economic growth of 11 percent for Ghana over the period 1960-2008 though failing the test of significance at that level. They also estimated a robust 11 percent threshold inflation level with close coefficients

after dropping growth rate of aggregate labour force and money supply growth which were found to be insignificant in the OLS models.

Espinoza (2010) examined threshold effect of inflation on GDP Growth by using a panel data of 165 countries including Oil Exporting Countries and Azerbaijan over the period of 1960-2007. Their study found that for all country groups' threshold level of inflation for GDP growth was about 10 percent (with the exclusion of industrialized countries where threshold level was much lower). Estimated results suggested that inflation from higher than 13 percent decreases real non-oil GDP by 207 percent per year. Lastly, review of literature on money supply and exchange rate influence on economic growth and inflation. Bick (2009) modelled a large panel-dataset of 124 industrialized and developing countries over the period from 1950 to 2004 in the study Inflation and Growth: New Evidence from a Dynamic Panel Threshold Analysis. Using a dynamic panel threshold model to shed light on the impact of inflation on economic growth, they found an inflation target of about 2 percent for industrialized countries and 17 percent for developing economies. Below the 17 percent threshold, the impact of inflation on economic growth remained insignificant, thus failing to support the growth-enhancing effects of inflation on economic growth in non-industrialized economies.

Sergii (2009) found that growth - inflation interaction was strictly concave with some threshold level of inflation. Inflation threshold level is estimated using a non-linear least squares technique, and inference made by applying a bootstrap approach. The main findings were that inflation rate above 8 percent tend to slow down economic growth while below 8 percent promotes economic growth. Schiavo (2007) used both nonparametric and semiparametric instrumental variable estimators, which have the

advantage of letting the data, speak as much as possible. They showed that, for developed countries, low inflation rates have hardly any real effect and high inflation rates have negative real effects. For developing countries, they found that too high variability does not allow reaching clear-cut results.

Using Libyan annual data for the period 1964 – 2010, Cevik and Teksoz (2013) adopted the cointegration and error correction models to investigate inflation dynamics. The study found inflation inertia to be a key determinant of consumer price inflation in Libya. The result also indicated that government spending, money supply growth, global inflation, exchange rate pass-through and imposition and subsequent removal of international sanctions played central roles in the Libyan inflation process. Kabundi (2012) employed single-equation error correction model based on the quantity theory of money to identify the main factors underlying inflation in Uganda. The study showed that both external and domestic factors affects inflation in Uganda, amongst which are money growth, world food prices, domestic supply and demand effects in the agricultural sector, energy prices and inflation inertia.

Asogu (1991) focused on the econometric investigation of the nature and causes of inflation in Nigeria. The study found that increase in real GDP or supply situation, especially food, and low cost of production of consumables tended to ameliorate inflation. He added that increase in government expenditures–deficits financing, tend to increase the money supply and worsen the depreciation of the exchange rate, which in turns intensify the inflationary pressure. The author noted that the monetary model does not adequately explain the inflation process in Nigeria. Moser (1995) developed an error

correction model of the inflation process in Nigeria based on money market equilibrium condition to analyze the dominant factors influencing inflation in Nigeria using annual data spanning from1960 to 1993. The author found that monetary expansion, driven mainly by expansionary fiscal policies, explains to a large degree the inflationary process in Nigeria. Other factors were devaluation of the Naira and agro-climatic conditions.

Faria and Carneiro (2001) examined the inflation-growth nexus from the perspective of an economy suffering from high and persistent inflation. He studied the case of Brazil and found empirical evidence for a negative effect of inflation on output in the short run. Khan and Senhadji (2001) estimated a panel regression with data from 140 countries and spanning about 40 years to investigate the nonlinear relationship between inflation and economic growth. Having established the presence of nonlinearity, they found a threshold range of 1-3 per cent for industrial economies and 11 – 12 per cent for developing economies. The estimated relationships were found to be robust to different estimation procedures, alternative specifications, changes in threshold levels and different data frequency. Mallik and Chowdhury (2001) empirically examined the relationship between inflation and GDP growth for four South Asian countries (Bangladesh, India, Pakistan and Sri Lanka) using co-integration and error correction models. They found evidence of a long-run positive relationship between GDP growth and inflation. They also discovered significant feedbacks between inflation and economic growth and concluded that the sensitivity of inflation to changes in growth rates is larger than that of growth to changes in inflation rates.

This study puts the countries on a knife edge as they struggle to achieve non-inflationary growth. The challenge for them, therefore, is to find a growth rate that is consistent with a

stable inflation rate, rather than beat inflation first to take them to a path of faster economic growth. Nell, (2000) examined the issue if inflation was detrimental to economic growth or not by using Vector Auto Regressive (VAR) technique

CHAPTER THREE

RESEARCH METHODLOGY

3.1 Research Design

The bootstrap is a non-parametric computer-intensive statistical technique that uses a unique finite sample to describe the variability of a statistic without making any distributional assumptions about the data. The bootstrap is especially useful in situations where finite sample theory is difficult or even impossible to derive or when only asymptotic theory is available. The latter being the case with time series data. Formally, the bootstrap consists of a methodology for estimating standard errors by repeatedly resampling with replacement from the original finite sample, which is believed to be a sample of independent and identically distributed (iid) observations from an unknown probability distribution. A discussion of the R tool used for the bootsrap resampling technique of the preferred method is given as well as the R codes for the ARIMA.

3.2. Methods of Data Collection

The data for this study shall be obtained from the Central Bank of Nigeria Statistical Bulletin 2018 publication.

3.3 Bootstrap procedure

For time series data we will use the `r-studio`, which has several methods for resampling the time series data. We will show three of them here. We first set that we want to obtain a statistic for 100 resampling of our data

3.3.1 Fixed Block Sampling

For fixed block sampling, you first generate a date to begin a block, and then select from that date to a point that gives a time series of length `block-length`. You then draw a new time point, select a new block, and add it to your time series. You repeat the process until you have a time series of the length of your original time series. You then calculate a statistic for each sample, and aggregate the results.

3.3.2 Stationary Block Sampling

For stationary block sampling, you first generate a date to begin a block, and then select a random length generated by a geometric distribution with mean, `block length`. You save that block of data, and then draw a new time point and block length, select it, add it to your time series. You repeat the process until you have a time series of the length of your original time series. Just like in the fixed block sampling, you repeat that process until you have all of your resampled time series, and aggregate the statistics for each one.

In this case we will set sim="geom" for stationary block sampling.

3.3.3 Model-based Sampling

For model-based resampling, the trick is to use your model to generate new time series, and adding noise with the resampled residuals from the fit. This method is particularly risky if you are not entirely sure about the underlying model for your data.

We first have to fit the model, and extract the residuals from the model. Then we create a function that can simulate ARIMA models according to our fit model.

3.4 The Arima Process

The Three Stages of ARIMA Modeling

The analysis performed by PROC ARIMA is divided into three stages, corresponding to the stages described by Box and Jenkins (1976). The IDENTIFY, ESTIMATE, and FORECAST statements perform these three stages, which are summarized below.

i. In the identification stage, you use the IDENTIFY statement to specify the response series and identify candidate ARIMA models for it. The IDENTIFY statement reads time series that are to be used in later statements, possibly differencing them, and computes autocorrelations, inverse autocorrelations, partial autocorrelations, and cross correlations. Stationarity tests can be performed to determine if differencing is necessary. The analysis of the IDENTIFY statement output usually suggests one or more ARIMA models that could be fit Options allow you to test for stationarity and tentative ARMA order identification.

ii. In the estimation and diagnostic checking stage, you use the ESTIMATE statement to specify the ARIMA model to fit to the variable specified in the previous IDENTIFY statement, and to estimate the parameters of that model. The ESTIMATE statement also produces diagnostic statistics to help you judge the adequacy of the model. Significance tests for parameter estimates indicate whether some terms in the model may be unnecessary. Goodness-of-fit statistics aid in comparing this model to others. Tests for white noise residuals indicate whether the residual series contains additional information that might be utilized by a more complex model. If the diagnostic tests indicate problems with the model, you try another model, then repeat the estimation and diagnostic checking stage.

iii. In the forecasting stage you use the FORECAST statement to forecast future values of the time series and to generate confidence intervals for these forecasts from the ARIMA model produced by the preceding ESTIMATE statement. An ARIMA (p, d, q) model can be checked if it is a good statistical fit for data or not, using Akaike Information Criterion (AIC) and Schwarz Criterion (SC) method. Autocorrelation (ACF) and partial autocorrelation (PACF) statistics help to determine the right parameters for ARIMA model (Real Options Valuation, 2007). Autoregressive integrated moving average which are used for short term memory system are characterized by using the last *i* values for making prediction in univariate time series analysis. In most statistical analysis the last *i* are given in order to predict the actual value at time $i + t$. The general expression for the ARIMA process $\{y\}$ is given by the equation below

$$\left(1 - \phi_\rho B\right)\nabla^d y_t = \left(I - \theta_q B\right)\varepsilon_t$$

where

$(1 - \phi_P B)\nabla^d y_t = \varepsilon_t$ is the autoregressive component

$y_t = \left(I - \theta_q B\right)\nabla^d \varepsilon_t$ is the moving average component

CHAPTER FOUR

DATA PRESENTATION AND ANALYSIS

4.1 Data Presentation

The data used for this analysis is a secondary data and can be found at the appendices page.

4.2 Data Analysis and Results

The monthly inflation data from the Central bank of Nigeria Statistical bulletin was transformed to a quarterly data as presented below;

Table 1. Table of Quarterly Non-Bootstrapped Inflation Data

S/N	DATE	INFRATE
1	2014, q1	7.83
2	2014, q2	8.03
3	2014, q3	8.37
4	2014, q4	8.00
5	2015, q1	8.37
6	2015, q2	8.97
7	2015, q3	9.30
8	2015, q4	9.41
9	2016, q1	11.26
10	2016, q2	15.26
11	2016, q3	17.53
12	2016, q4	18.45
13	2017, q1	17.92
14	2017, q2	16.53
15	2017, q3	16.01
16	2017, q4	15.73
17	2018, q1	13.33
18	2018, q2	12.05

The above table represents 18 entries from 2014-2018 but 2018 have only two quarters. The quarters were obtained by takin the mean of every 3 months to make one quarter so as to get precise amount of data for the bootstrap.

4.2.1 Test for Stationarity using the ADF (Identification Statge of The ARIMA)

The Augmented Dickey–Fuller Test below shows the p-values at 5th and 6th differencing; however by practice if the p-value of an ADF test is less than 0.05 then the data is stationary at that point. So we shall take the `INFRATEdiff6 being the sixth` difference of our time series data for our analysis. The differencing is up to six times because of the transformation of the data from monthly to quarterly data.

```
> INFRATEdiff5=diff(INFRATE, differences=5)
> adf.test(INFRATEdiff5)

        Augmented Dickey-Fuller Test

data:  INFRATEdiff5
Dickey-Fuller = -1.1422, Lag order = 2, p-value = 0.8992
alternative hypothesis: stationary
```

```
> INFRATEdiff6=diff(INFRATE, differences=6)
> adf.test(INFRATEdiff6)

        Augmented Dickey-Fuller Test

data:  INFRATEdiff6
Dickey-Fuller = -3.6276, Lag order = 1, p-value = 0.04803
alternative hypothesis: stationary
```

4.2.2 ARIMA Model Estimation

Still using r studio, to fit ARIMA model for the data, we shall fit the model on the already differenced data which is our `INFRATEdiff6.`

we shall run the ARIMA code at with different values for the (p,q,d). The model that returns the best AIC (Akaike Information Criterion), the model with the list value of the Aic is considered the best model for forecasting.

R Studio ARIMA RESULTS

```
arima(INFRATEdiff6, order = c(0,0,0))

Call:
arima(x = INFRATEdiff6, order = c(0, 0, 0))

Coefficients:
      intercept
         0.6677
s.e.     4.0457

sigma^2 estimated as 130.9:  log likelihood = -30.85,  aic = 63.7
> arima(INFRATEdiff6, order = c(1,0,0))

Call:
arima(x = INFRATEdiff6, order = c(1, 0, 0))

Coefficients:
          ar1  intercept
      -0.5866     0.9789
s.e.   0.2510     2.0615

sigma^2 estimated as 77.52:  log likelihood = -28.96,  aic = 61.93



> arima(INFRATEdiff6, order = c(1,0,1))

Call:
arima(x = INFRATEdiff6, order = c(1, 0, 1))

Coefficients:
          ar1      ma1  intercept
      -0.4102  -0.9999     0.2755
s.e.   0.3052   0.3612     0.5911

sigma^2 estimated as 33.87:  log likelihood = -26.94,  aic = 59.89



arima(INFRATEdiff6, order = c(1,1,1))

Call:
arima(x = INFRATEdiff6, order = c(1, 1, 1))

Coefficients:
          ar1      ma1
      -0.5353  -0.9999
s.e.   0.2793   0.4707

sigma^2 estimated as 89.99:  log likelihood = -27.27,  aic = 58.55
```

Table 2. Table of AICs of Non-Bootstrapped ARIMA models

	ARIMA (0,0,0)	ARIMA (1,0,0)	ARIMA (1,0,1)	ARIMA (1,1,1)

log likelihood	-30.85	-28.96	-26.94	-27.27
AICs	63.7	61.93	59.89	58.55
BICs	65.85	64.16	62.20	60.38

From the Table above, it is noticeable that the computed ARIMA model with order (1,1,1) will give the best fit for our purpose of forecasting.

4.3 Forecasting With Arima Model

Here, we shall be using the FORECAST and PREDICT statements in r studio to obtain future points of our data set.

```
>fit_diff_arf <- forecast(arima.final, h =12)
> print(fit_diff_arf)
  Point      Forecast     Lo 80      Hi 80     Lo 95     Hi 95
2018 Q3     -3.8638781 -6.626671 -1.1010855 -8.089205 0.3614485
2018 Q4     -3.3286012 -6.975783  0.3185808 -8.906485 2.2492824
2019 Q1     -2.8673047 -7.052056  1.3174463 -9.267329 3.5327197
2019 Q2     -2.4697638 -7.012779  2.0732514 -9.417706 4.4781783
2019 Q3     -2.1271669 -6.918956  2.6646220 -9.455576 5.2012419
2019 Q4     -1.8319201 -6.800415  3.1365745 -9.430577 5.7667367
2020 Q1     -1.5774793 -6.673246  3.5182878 -9.370783 6.2158239
2020 Q2     -1.3582047 -6.546475  3.8300655 -9.292979 6.5765698
2020 Q3     -1.1692359 -6.425153  4.0866812 -9.207467 6.8689956
2020 Q4     -1.0063844 -6.311984  4.2992148 -9.120598 7.1078294
2021 Q1     -0.8660406 -6.208239  4.4761577 -9.036228 7.3041466
2021 Q2     -0.7450936 -6.114312  4.6241247 -8.956604 7.4664172
```

Now we shall do the prediction,

```
predicted <- predict(arima.final, n.ahead = 10)

predicted
```

Table 3.1

The table below is the predicted points of inflation 2018-2020 from the ARIMA model of Non resampled inflation data.

	Qtr1	Qtr2	Qtr3	Qtr4
2018			10.11479	10.88283
2019	11.40704	11.38808	11.47120	11.44769
2020	11.46626	11.45803	11.46293	11.46045

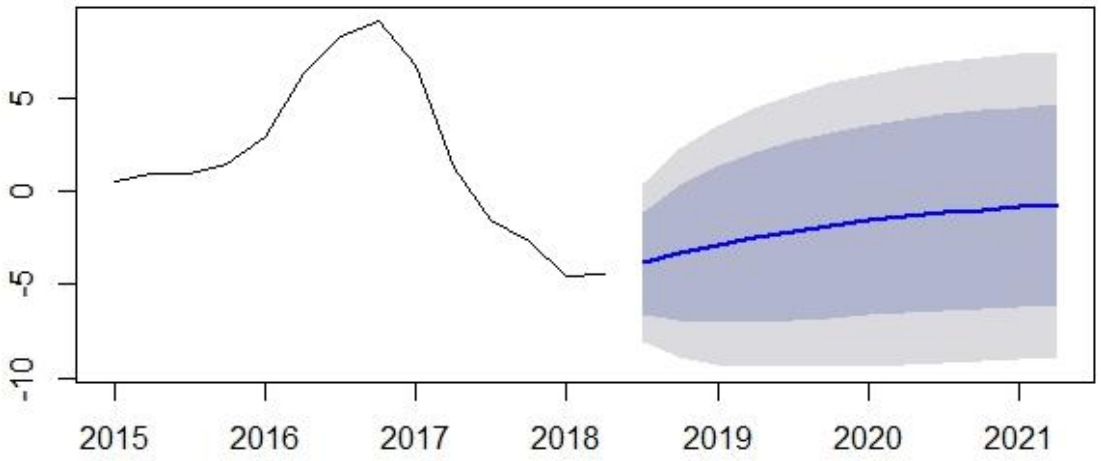


Fig 1. A graph showing the behavior of the inflation rates, the shaded areas are the confidence intervals for the forecasted values.

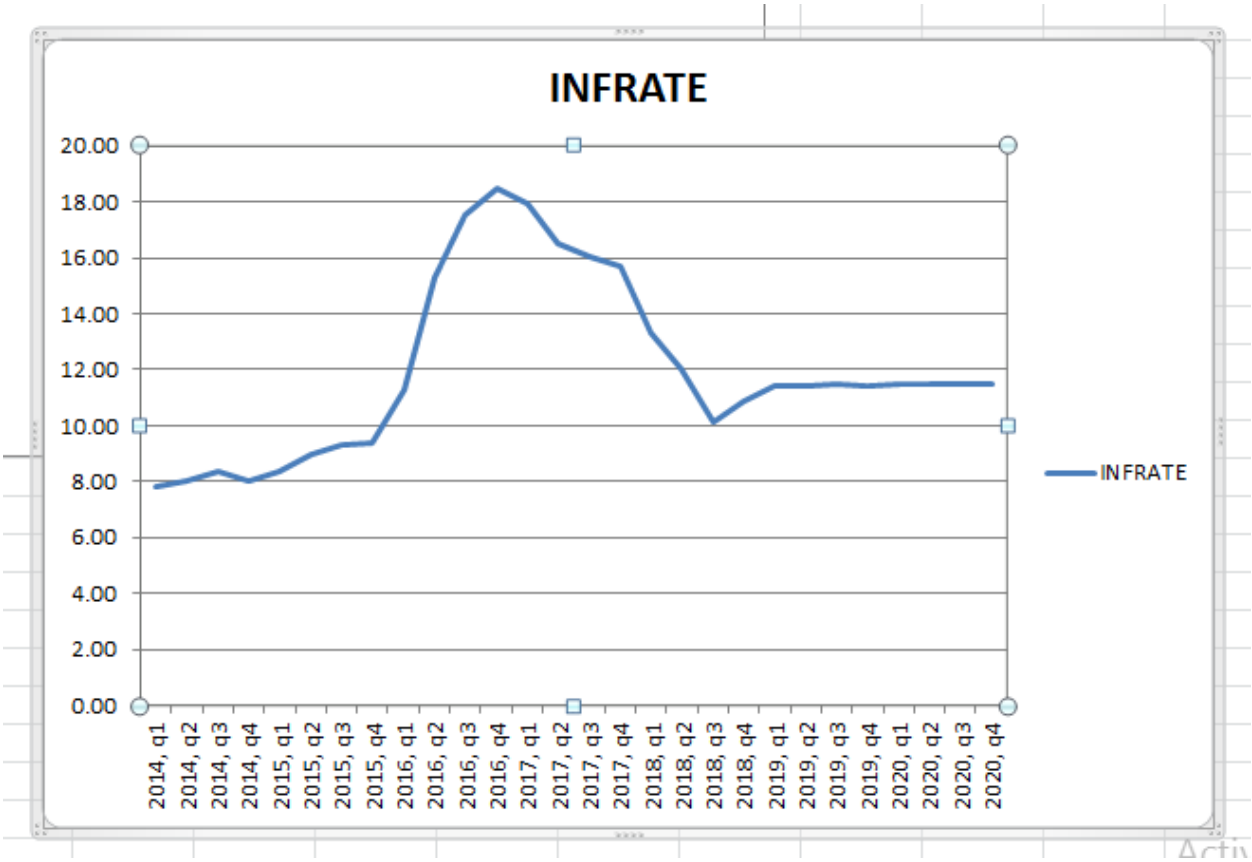


Fig 2. A graph of the forecasted inflation rates from the ARIMA model of Non resampled inflation data.

4.3.1 ARIMA Model of Resampled (Bootstrapped) Inflation Data.

We shall do a bootstrap resampling of the values, that is taking repeated samples with replacement and fit an ARIMA model to the bootstrapped values, and we can check the goodness of fit of the model. The values of the bootstrapped data where obtained using the boot function in R studio. The BTINFDATA represents the bootstrapped values.

Table 3 Table of Bootstrapped Quarterly Inflation Data

S/N	DATE	BTINFDATA
1	2014, q1	9.406667
2	2014, q2	8.033333
3	2014, q3	8.033333
4	2014, q4	15.726667
5	2015, q1	17.92
6	2015, q2	8.366667
7	2015, q3	13.33
8	2015, q4	9.406667
9	2016, q1	8
10	2016, q2	9.3
11	2016, q3	15.26
12	2016, q4	12.045
13	2017, q1	8
14	2017, q2	17.53
15	2017, q3	13.33
16	2017, q4	15.26
17	2018, q1	17.53
18	2018, q2	15.726667

After obtaining the bootstrapped samples, we still use the ADF test to check for stationarity, the bootstrapped data is stationary at the second difference from the result below;

```
> ndiffs(BTINFDATA, test="adf")
[1] 2
> BTINFDATAdiff=diff(BTINFDATA, differences=1)
> adf.test(BTINFDATAdiff)
```

```
        Augmented Dickey-Fuller Test

data:  BTINFDATAdiff
Dickey-Fuller = -2.8383, Lag order = 2, p-value = 0.253
alternative hypothesis: stationary

> BTINFDATAdiff2=diff(BTINFDATA, differences=2)
> adf.test(BTINFDATAdiff2)

        Augmented Dickey-Fuller Test

data:  BTINFDATAdiff2
Dickey-Fuller = -4.6661, Lag order = 2, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(BTINFDATAdiff2) : p-value smaller than printed p-value
```

The ARIMA model is fitted to the data at second difference and we obtained the below result;

```
arima(x = BTINFDATAdiff2, order = c(1, 1, 1))

Coefficients:
          ar1      ma1
      -0.4978  -1.0000
s.e.   0.2099   0.1772

sigma^2 estimated as 49.51:  log likelihood = -52.46,  aic = 108.92
```

Now comparing the AICs of the bootstrapped and non-bootstrapped ARIMA models

Table 4. Table of AICs and BICs from ARIMA models for Bootstrapped and Non Bootstrapped samples

	ARIMA (0, 0, 0)		**ARIMA (1,0,0)**		**ARIMA(1,0,1)**		**ARIMA(1,1,1)**	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
Bootstrapped	114.6	118.1414	111.07	115.3902	102.76	107.8495	108.92	113.0485

Non Bootstrapped	63.7	65.8599	61.93	64.16726	59.89	62.20322	58.55	60.38392

4.2.4 Findings

From the results in Table 4, it is clear that the AICs obtained from the ARIMA model fit on the resampled data is higher than that of the non-resampled data, the BICs obtained behave in the same way. This shows that the ARIMA model of non-resampled inflation time series data will produce the best forecast.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

This study examined the monthly inflation data. We examined the time series data by conducting stationarity test using Augmented Dickey-Fuller (ADF) unit roots test. The inflation data for the non-bootstrap samples became stationary at the sixth difference while that of bootstrap samples became stationary at second difference. The Autoregressive integrated moving averages (ARIMA) was applied to both the stationary data for bootstrap and non-bootstrap samples, our finding shows that the ARIMA model of the non-bootstrapped inflation data is more perfect for forecasting. The stationarity test is important in ARIMA so as to avoid wrong forecast.

5.2 Conclusions

Going by the values of the AICs obtained from ARIMA of different specification as seen in table 3.1. It is observed that the AICs obtained for using non-bootstrapped data are less than those from bootstrapped inflation data. Therefore we conclude that the forecast from the non-bootstrapped data will be adopted for the purpose of this study.

The result of the forecast shows that inflation drops at the 3rd and 4th quarters of 2018 and rise by one unit to maintain a slightly steady value throughout 2019 and begins to drop at the 4th quarter of 2020.

5.3 Recommendations

From this study and the data obtained, it is recommended to use non bootstrapped data for forecasting time series.

5.4 Limitations of the Study

The limitation of this study includes access to codes for running analysis on r-studio, also data was not easily accessible at the beginning.

5.5 Contributions to Knowledge

The study will help to educate researchers on the variations of ARIMA model obtained from resampled and non-resampled time series data.

REFERENCES

Asoguo, J.O. (1991): An Econometric Analysis of the Nature and Causes of Inflation in Nigeria. CBN Journal of Applied Statistics 3(2): 12 – 187

Cevik S. & Teksoz K. (2013): Lost in Transmission? The Effectiveness of Monetary Policy transmission in the Gulf Cooperation Council (GCC): 23-39

Datta K. (2012): ARIMA Forecasting of Inflation in the Bangladesh Economy. The IUP Journal of Bank Management. 10(4): 7-15

Doguwa S. (2013): Inflation and Economic Growth in Nigeria: Detecting the Threshold Level. CBN Journal of Applied Statistics, 3(2): 10-26

Efron B. (1979): Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics 11(1) 89_167

Elder J.(2004): Another Perspective on the Effects of Inflation Uncertainty; Journal of Money, Credit and Banking. 36(5): 8-19

Faria, J. & Carneiro, F. (2001): Does High Inflation Affect Growth in the Long and Short Run? Journal of Applied Economics. 4 (1): 89-105.

Grier S. (2006): Targeting interventions for ethnic minority and low-income populations. Future Child. 2006 Spring: 16(1):187-207.

Hasanov F. (2010): Relationship between inflation and economic growth in Azerbaijani economy: Is there any threshold effect? Asian Journal of Business and Management Sciences. 1(1): 1-11.

Howrey E. (1980): The Role of Time Series Analysis in Econometric Model Evaluation. National Bureau of Economic Research, Inc. 275-307

Jordi G. & Tommaso M. (2005): Monetary Policy and Exchange Rate Volatility in a Small Open Economy. The Review of Economic Studies. 72(3): 707–734.

Kabundi A. (2012): Dynamics of inflation in Uganda. ADB Working Paper Series -African Development Bank, Tunis, Tunisia. 152(1): 12-20

Kasidi F. & Kenani M. (2013): Impact of Inflation on Economic Growth: A Case Study of Tanzania, Asian Journal of Empirical Research, 3(4): 363-380.

Kumar K. (2004): Forecasting daily maximum surface ozone concentrations in Brunei Darussalam: ARIMA Modeling Approach Journal of the Air & Waste Management Association. 54(1): 20:29

Liv Q. Liu X, & Jiang B (2011): Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. National Institute for Communicable Disease Control and Prevention, China. 1(1): 25:30

Marbuah G. (2010): The Inflation-Growth Nexus: Testing for Optimal Inflation for Ghana. Journal of Monetary and Economic Integration. 11(2): 20-31

Nnamocha P. N. (2002): Public Finance for a Developing Economy, Bon Publishers Transmission Channels in the GCC Countries. International Monetary Fund. 12(191): 345-389

Tsitsikas V. E. (2007): Modeling and forecasting pelagic fish production using univariate and multivariate ARIMA models Fisheries Science 73(5): 979-988.

Uko A.K & Emeka N. (2012); Inflation forecasts with ARIMA, vector autoregressive and error correction models in Nigeria, European Journal of Economics, Finance and Administrative Sciences 3(3): 11-38

Umaru H. & Zubairu A. (2012): Effect of Inflation on the Growth and Development of the Nigerian Economy: An Empirical Analysis. International Journal of Business and Social Science, 3(10):165 – 186

Vaona A. (2012): Inflation and growth in the long-run: A new Keynesian theory and further semiparametric evidence. Cambridge University Press. 16(01): 94-132.

Xie M. & Singh K. (2008): Bootsrap: A statistical Method: Annals of Statistics, 9(6): 1-10

APPENDIX

YEAR	MONTH	INFLATION
2014	1	8
2014	2	7.7
2014	3	7.8
2014	4	7.9
2014	5	8
2014	6	8.2
2014	7	8.3
2014	8	8.5
2014	9	8.3
2014	10	8.1
2014	11	7.9
2014	12	8
2015	1	8.2
2015	2	8.4
2015	3	8.5
2015	4	8.7
2015	5	9
2015	6	9.2
2015	7	9.2
2015	8	9.3
2015	9	9.4
2015	10	9.3
2015	11	9.37
2015	12	9.55
2016	1	9.62
2016	2	11.38
2016	3	12.77
2016	4	13.72
2016	5	15.58
2016	6	16.48
2016	7	17.13
2016	8	17.61
2016	9	17.85
2016	10	18.33
2016	11	18.48
2016	12	18.55
2017	1	18.72
2017	2	17.78

2017	3	17.26
2017	4	17.24
2017	5	16.25
2017	6	16.1
2017	7	16.05
2017	8	16.01
2017	9	15.98
2017	10	15.91
2017	11	15.9
2017	12	15.37
2018	1	15.13
2018	2	14.33
2018	3	13.34
2018	4	12.48
2018	5	11.61

Source, CBN Statistical bulletin.