

WOLKITE UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

Modeling and Forecasting the Export Price of Coffee in Ethiopia

By:

Name **ID No**

1. Lamesgn Gashaw…….NCSR/223/09

2. Bruktayet Degif ……...NCSR/531/09

Advisor: Sebsib Muanenda (Assi. professor)

A Senior Research paper Submitted to the Department of Statistics in Partial Fulfillment of the Bachelor of Science Degree in Statistics

June, 2019

Wolkite, Ethiopia

WOLKITE UNIVERSITY

COLLEGE OF NATURAL AND COMPUTATIONAL SCIENCES

DEPARTMENT OF STATISTICS

This is to certify that the paper prepared by Lamesgn Gashaw and Bruktayet Degif, entitled: modeling and forecasting the export price of coffee in Ethiopia and submitted in partial fulfillment of the requirements for the Degree of Bachelor of Science complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Advisor _____Signature _____Date _____

Examiner _____Signature _____Date _____

_____

Chair of Department or Graduate Program Coordinator

ACKNOWLEDGEMENT

First and for most, we would like to thanks our God to keep us full of health to accomplish, we would like to thank almighty to God for their permission to do our daily activity as well as breathing fresh air. Next, we would like deepest gratitude to our advisor Sebsib Muanenda (Assi. professor.) for his constructive idea, giving necessary constructive comments and suggestions from the beginning until the end of this study. Come up with this research outcome and with who help us for this achievement. All our effort has been successful and it has happened with help of God. Finally, we would like to thank our families that helped us a lot of indifferent direction from the beginning until now.

Table of content

ACKNOWLEDGEMENT i

Table of content ii

List of Figures v

List of Tables vii

ACRONYMS vii

Abstract ix

CHAPTER ONE 1

1. INTRODUCTION 1

1.1. Background of the Study 1

1.2. Statement of the problem 4

1.3. Objective of the study 5

1.3.1. General objective 5

1.3.2. Specific objective 5

1.4. Significance of the study 5

CHAPTER TWO 6

2. LITERATURE REVIEW 6

2.1. Coffee market in Ethiopia 6

2.2. Theoretical and empirical studies on Export price of coffee 7

CHAPTER THREE 13

3. Data and methodology 13

3.1. Data and variable of the study 13

3.2. Methodology 13

3.2.1. Univariate Time Series Analysis 14

3.2.2. Stationary test 15

3.2.3. Unit Root Test 15

3.2.4. Augmented Dickey–Fuller (ADF) Unit Root Test 16

3.2.5. Phillips-Perron (PP) Unit Root Test 17
3.3. Time domain approach and ARIMAX mode 18
3.3.1. Autoregressive (AR) Process 18
3.3.2. Moving average (MA) processes 19
3.3.3. Autoregressive Moving Average (ARMA) Processes 20
3.3.4. Autoregressive Integrated Moving Average (ARIMA) Processes 21
3.3.5. Autoregressive Integrated Moving Average with exogenous variable (ARIMAX) Processes 22
3.3.6. Building ARIMA-X Model 22
3.4. Model Identification 22
3.5. Model selection criteria 23
3.5.1. Akaike Information Criterion (AIC) 24
3.5.2. Bayesian Information Criterion (BIC) 24
3.6. Parameter Estimation 25
3.7. Diagnostic Test 25
3.7.1. Ljung-Box test of serial correlation 26
3.7.2. Jarque-Bera test of normality 26
3.8. Forecasting 27
3.8.1. Minimum mean square error forecasts 27
3.8.1. Forecasting accuracy measures 28
CHAPTER FOUR 30
4. DATA ANALYSIS 30
4.1. Descriptive statistics and time plot the variables 30
4.2. Unit Root Test of Stationary 31
4.3. ARIMAX model 34
4.3.1. Model selection 34
4.3.2. Model estimation 36

4.3.3. Diagnostic Checking 38
4.3.4. Forecasting 39
CHAPTER FIVE 41
5. Conclusion and Recommendation 41
5.1. Conclusion 41
5.2. Recommendation 41
Reference 42
Appendix 46

List of Figures

Figure4.1 Time plot of original series export price coffee in Ethiopia. 31

Figure 4.2 Time plot of differenced series of export coffee price..... 46

Figure 4.3 Time plot of original series of exchange rate, food price and non-food price data in Ethiopia. 46

Figure 4.4 Time plot of differenced series of exchange rate, food price and non-food price data in Ethiopia. 47

Figure 4.5 Correlogram of monthly export price of coffee at first difference up to 20 lags..... 49

Figure 4.6 The residual, actual and fitted plot of ARIMAX (4, 1,4) model of coffee price. 50

Figure 4.7 Autocorrelation (ACF) and partial autocorrelation (PACF) of residuals for ARIMA (4, 1, 4) 50

Figure 4.8 Serial correlation Breush- godfrey LM test for residual of ARIMAX (4, 1,4) model. 51

figure 4.9Heteroskedasticity Test of Breusch-Pagan-Godfrey test for estimated ARIMAX (4, 1,4) model..... 53

Figure 4.10Jarkue- Bera normality test for the residuals of ARIMAX (4, 1, 4) 54

Figure 4.11 Out-off-sample forecast of monthly export price of coffee. ……..…….39

List of Tables

Table 2.1. Ethiopia's Coffee Production 7

Table3.1. Behavior of the ACF and PACF for ARMA Models..... 23

Table 4.1: descriptive statistics of export price of coffee, exchange rate, food price and non-food price index..... 30

Table 4.2:Unit root test results (At level)..... 32

Table 4.3: Unit root test results (At first difference)..... 33

Table 4.4: The ARIMAX model selection criteria for monthly differenced export price of coffee data using AIC, BIC and serial correlation. 35

Table 4.5: Forecast accuracy measure of $\boldsymbol{ARIMAX(0,1,1)}$ and $\boldsymbol{ARIMAX(4,1,4)}$ models of export price of coffee. 36

Table 4.6: Parameter estimation of ARIMAX (4, 1, 4)..... 37

Table 4.7 Forecast accuracy measure of ARIMAX (4, 1, 4) models of export price of coffee..40

Table4.8. Forecasted values of ARIMAX (4,1,4) models of export price of coffee..... 54

ACRONYMS

ACF	Autocorrelation function
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
ANN	Artificial Neural Network
AR	Autoregressive Processes
ARCH	Auto Regressive Conditional Heteroscedastic
ARIMA	Autoregressive Integrated Moving Average
ARIMA -X	Autoregressive Integrated Moving Average with exogenous variable
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CSA	Central Statistical Agency
CTDMA	Coffee and Tea Development and Marketing Authority
ECX	Ethiopian commodity exchange
GARCH	Generalized Auto Regressive Conditional Heteroscedastic
GDP	Gross Domestic Production
ICO	International Coffee Organization
JB	JarqueBera Statistic
LM	Lagrange Multiplier
MA	Moving Average process
MAE	Mean Absolute Error

MGARCH	multivariate Generalized Auto Regressive Conditional Heteroscedastic
MPAE	Mean Percentage Absolute Error
MPE	Mean Percentage Error
MSE	Mean Square Error
MY	Manual year
NBE	National bank of Ethiopia
OLS	Ordinal least square
PACF	partial Autocorrelation function
PE	Percentage Error
PP	Phillips and Perron test
RMSE	Root Mean Squared Error
SNNPR	Southern Nations, Nationalities, and People Regions
SPSS	statistical package for social since
USD	United States Dollar

Abstract

Coffee is great significant in world economy as the largest single commodity entering the international trade after petroleum and petroleum products. Coffee has been and remains the leading cash crop and export commodity of Ethiopia. The main objective of this study is to model and forecast export price of coffee in Ethiopia based on the past historical data from July, 2002 to January, 2019 using ARIMAX model.In this study used only secondary data and the data obtained from National bank of Ethiopia (NBE).The variables of interest in this study are export price of coffee, Real exchange rate, food price Index and non food price Index. Unit root test results indicate that the series become stationary at first difference. Among all candidate ARIMAX models, the best model for the series under consideration is ARIMAX (4, 1, 4) for export price of coffee in Ethiopia. Furthermore, Breusch-Godfrey Serial Correlation Lagrange multiplier (LM) test and Jarque-Bera statistic tests are showed that the disturbance terms are white noise and a series is normally distributed. Among the selected variable exchange rate and food price index has significant effect on export price of coffee in Ethiopia. Finally, out-of sample forecasts are made using ARIMAX (4, 1, 4) model the result indicates that the export price of coffee has decreasing trend in the future. The government supports to increase the exchange rate and reduced food price inflation of the country in order to increase the export price of coffee in Ethiopia and the export performance of the country.

Key word: - Coffee, ARIMAX, Forecasting, Ethiopia

CHAPTER ONE

1. INTRODUCTION

1.1. Background of the Study

Coffee is great significant in world economy as the largest single commodity entering the international trade after petroleum and petroleum products. The most important are coffee Arabica (Arabica coffee) and coffee canaphora (Robusta coffee). Foreign trade payments have a significance role in the development plans of many developing countries. For most of these countries, export represents an important share of the total Development. Trades in the foreign sales are important in forecasting of overall growth of a country. For some developing countries, export trade is such an important factor that an estimate of the foreign exchange earnings represents a first step in the formulation of development plans. As the export of a given country continue to expand, new profit opportunities develop not only to provide imports for export sector but also to take the advantage of external economic benefits (Hultman, 1967). Coffee is an engine of economic growth and development for a country. The benefit of country can get from international trade is largely dependent on exports. Coffee export is one of the most important activities in economic growth and development of a country. It plays a crucial role in creating employment opportunities, increasing production capacity and reducing production cost through its direct and indirect impact on domestic activities. Its contribution is also highly significant as a major source of foreign exchange to finance import of commodities that are not domestically produced but which are of vital importance for domestic production. In addition it strengthens the capacity of the government to provide public goods by increasing government revenue (Abbot and Make, 1979).

Coffee consumption in Africa is changing, especially in the urban areas. The rise in the middle class income has also helped to increase consumption. African coffees producers are expected to consider the price they can get on specialty coffee market and develop their own brand and coffee origin to grab the demand for specialty coffee beans globally. Moreover they can undertake innovative marketing strategy to increase the domestic demand for coffee consumers in Africa can be encouraged not only to consume instant coffee but also other coffee related products. Under developed countries are largely dependent on the export sector. Their exports

are mainly products that have high fluctuations in both prices and volumes in the international market. Among the primary commodity exports of least developed countries, coffee is the second most internationally traded product next to oil, accounts for a high percentage of their export. It is the main source of foreign exchange and export revenue of these countries but like any primary commodity exports price is characterized by high fluctuations (Roger, 1992)

Ethiopia is one of the least developed countries whose exports are mainly primary commodities. Among the few exports, coffee plays a major role in our country as a source of foreign exchange earnings, government revenue and employment opportunities. Thus, coffee constitutes the single most export item in Ethiopian exports. Coffee has been and remains the leading cash crop and export commodity of Ethiopia. It has accounted on average for about5% of gross domestic product (GDP), 10% of total agricultural production and 60% of total export earnings for the past three to four decades (Mekonen, 2009 and Zerihun, 2008). According to National Bank of Ethiopia, the annual average volume of coffee export was 160 thousand tons between 2000 and 2012, and the highest volume of exports was registered in 2012(192 thousand tons). Ethiopia is struggling to raise hard currency reserves due to high demand for energy imports, lack of port access and few exportable commodities besides coffee. Additionally, over the same time period declines in global coffee prices incentivized local producers to hold their supplies to drive up prices and wait until currency devaluation made it favorable to sell into international markets. The sum of these beliefs caused vast delays in the shipment of product (Zakaria, 2000).

In this studyfocused on modeling and forecasting the export price of coffee. Coffee is a major export commodity and an important source of employment in Ethiopia. Coffee is also the most traded agricultural commodity in the Ethiopian commodity exchange (ECX) both in terms of volume and value, significant reforms have occurred in the coffee sector in the past years. Certainly, higher price interdependency between domestic and international price. The objective of the ECX is being achieved in terms of stronger linkages with external markets and improved market information, but does not necessarily mean that it is a welfare improvement system, particularly among the most vulnerable groups of farmers.

Financial time series are often considered in practice, one of the most popular and frequently used stochastic time series models is the Autoregressive Integrated Moving Average (ARIMA) model (Box and Jenkins, 2003). The following modeling approach established by Box and

Jenkins. In this study we apply a number of Time Series models, including an ARIMA and ARIMAX model. The models would be estimate monthly export price data. The basic assumption made to implement this model is that the considered time series is linear and follows a particular known statistical distribution, such as the normal distribution. ARIMA model has subclasses of other models, such as the Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) models. The popularity of the ARIMA model is mainly due to its flexibility to represent several varieties of time series with simplicity as well as the associated Box-Jenkins methodology for optimal model building process. ARIMA model is theoretical superiority over other forecasting methods. In practice the models did not routinely outperform other time series methods (Makridakis and Collegues, 1984). ARIMA with exogenous variable (ARMAX) is a special time series model that includes both an autoregressive term ("AR"), a moving average term ("MA") and additional exogenous variables ("X").

1.2. Statement of the problem

Many countries focus on encouraging the export of goods and services, but for their comparative advantageous they export commodities. In our country there are so many problems that hinder the performance of the economy form operating at its potential. Coffee export is the backbone of our country in the case of trade market and it needs higher effort of good export policy for economic growth of the country. The combined effect of the problems has always been reflected on the unstable export earnings of Ethiopia. World production of coffee, coffee supply and domestic production of coffee in Ethiopia has an effect on its export price due to low productivity, poor market infrastructure and traditional marketing channels among others. These causes heavy and sudden fluctuation in export price of coffee and create serious problem in national income and investment and then also creates sever adverse impacts on the overall growth of developing countries like Ethiopia. Several studies have been conducted on coffee in Ethiopia; Amare (2009), Gebreyesus (2015) and Manual *et al. (*2015), Tesfahun*et al. (*2018). Most of these studies were focused on the production related problem of coffee ignoring the export price of coffee and analysis the impact of various variables on the export price of coffee. But in this study try to identify the major determinants for export price of coffee by using time series data analysis using ARIMAX model. Generally, this research is trying to answer the following question.

- What are the major determinants for coffee export of the country?
- What is appropriate ARIMAX model for data on export price of coffee?
- What is look like the future export price of coffee in Ethiopia?

1.3. Objective of the study

1.3.1. General objective

The General objective of this study is to model and forecast export price of coffee in Ethiopia based on the past historical datausing ARIMAX model.

1.3.2. Specific objective

Specific objectives of this study are:

- To identify the major determinants for coffee export price of the country.
- To analyze the key macroeconomic factors that has significant impact on the export price of coffee in Ethiopia.
- To fit appropriate ARIMAX model for data on export price of coffee.
- To predict and forecast the future pattern export price of coffee in Ethiopia.

1.4. Significance of the study

Coffee export in Ethiopia is an engine for economic growth and development.This study wouldbecontribute the proper understanding of the past behaviours of coffee export in Ethiopia. And planning (estimating) the future pattern of coffee export price in Ethiopia.Generally this study wascontributing the following significance.

- It is important to an indicator about the past coffee export pattern of Ethiopia.
- Helping for policy makers to revise polices and strategies concerning foreign trade.
- It contribute to assess the export price of coffee in the past period.

CHAPTER TWO

2. LITERATURE REVIEW

2.1. Coffee market in Ethiopia

Coffee is the most important or legal trade commodity after oil in the world. It is the primary export of many developing countries accounting for as much as one third of export earning in several Latin America and African countries. Besides, coffee is a major source of foreign exchange in follows. It is also estimated that 95% of the total coffee output is produced by small scale producers and the remaining 5% by comparatives and private commercial farmers (Worako*et al,* 2008). This reflects the importance of the sector in improving the livelihood of the poor and contributing to the overall economic growth of the country. The performance of the coffee sector still remains lagging. Poorly instituted models of production (Zewdu*et al*, 2010), high marketing costs (Rashed et al, 2010), volatile prices (Gemech and Struthers, 2007), inadequate market infrastructure (Rashid et al, 2010, Gabre-Madhin, 2009), and unorganized commodity marketing approach (Gabre-Madhin, 2001; Gemech and Struthers, 2007) are indicators of market inefficiency that significantly limit coffee growers' share from the value of exports and total earnings from the sector.Coffee production in Ethiopia is almost exclusively situated in the two regions of Oromia and the Southern Nations, Nationalities, and People Regions (SNNPR) in the south and west of the country. Smallholder farmers produce 95% of Ethiopia's coffee (Tefera, 2013). In 2012, Ethiopia exported 3.2 million bags, making it the most important African coffee exporter and the tenth largest exporter in the world (ICO, 2013). Its share of the international coffee trade that year was about 3% Coffee is the most important export product of the country, accounting for about a quarter of the value of all exports in 2012. Coffee is cultivated by over 4 million primarily smallholder farming households (CSA, 2013). From a historical perspective, Ethiopia's coffee production over the last five years has been relatively stagnant for the reasons mentioned above. Production has hovered around 6.5 million bags (390,000 metric tons) and area planted has grown very little. See table 2.1. The newly re-established Coffee and Tea Development and Marketing Authority (CTDMA) hopes to change

this past narrative and jumpstart production by primarily using the existing land that is already under coffee cultivation.

Table 2.1. Ethiopia's Coffee Production

	2012/2013	2013/2014	2014/2015	2015/2016	2016/2017	2017/2018
Production (1000 MT 60-kg bags)	6,500	6,345	6,475	6,510	6,520	6,545
Area (1000 HA)	518	519	525	528	529	532

Coffee production for manual year (MY)17/18 (Oct-Sep) is forecast to remain almost unchanged from the previous year at slightly more than 6.545 million bags (393,000 metric tons). Area planted and harvested for this period are likewise nearly the same as the previous year. Post is making slight upward revisions to the official USD estimates for MY16/17 and MY15/16. Coffee stocks are forecast at 20,000 bags in MY17/18. Stocks for MY16/17 are lowered to 25,000 bags from the current USD official estimate (Michael and Counselor, 2017).

2.2. Theoretical and empirical studies on Export price of coffee

Several studies have investigated the relative accuracy of alternative export price forecasting Models. There are a number of different volatility estimating and forecasting models such as ARMA, ARIMA (Box and Jenkins, 2003), ARIMAX, GARCH family and stochastic volatility (SV) models. Among those models, the family of GARCH-type models introduced by(Engle, 1982) and then developed by many other researchers is the most common approach and proved to be sufficient to capture the properties of return series.

Gemech and Struthers (2007) presented the results of empirical investigation on the effects of market reform programs and their impacts on the volatility of coffee prices in Ethiopia. The study covers the period from 1982 to the end of 2001, though its main focus is on the period after the beginning of the reforms in 1992. Using the ARCH modeling techniques, they argued that there is evidence that Ethiopia experienced a significant increase in coffee price volatility after the adoption of the market-oriented reform programs. The econometric evidence presented in the

paper, suggested that the market reform programs were significant cause of the increased volatility after 1992. The authors also suggested a number of possible policy responses to the price volatility, including strengthening the role of smallholder producer associations; and the possibility of adopting appropriate risk management strategies. As a conclusion they stated that, the fact that Ethiopia is a price-taker in global market and is therefore prone to external shocks in coffee prices over which it has little control or influence means that the country will continue to be highly vulnerable to global shock.

Amare (2009) Univariate Time Series Analysis discusses the GARCH model fitting and volatility forecasting of export prices data. Ethiopia from January 1, 1997 to May 30, 2008. The data consists of 2567 daily observations, obtained from Ethiopian Custom Authority. The full samples are split into two parts: in sample data, in order to estimate the parameters of models and out sample data, in order to make forecasts. In sample data for coffee export is composed of 2223 observations from 01/01/1997 until 30/6/2007. Out sample data for coffee export is composed of 344 observations from 01/01/1997 until 30/6/2007. We have chosen to confine our analysis on total export prices of coffee export. The implied goal is to fit an appropriate GARCH model, to find out if the mentioned export prices are volatile or not and to forecast the volatility for some future times. ARMA(1,1) is found as the most appropriate model for the conditional mean equation of coffee export prices and it is also found that GARCH(2,1), for modeling volatility of coffee. Moreover, the results suggest that the export prices volatility is persistence indicating that past volatility is important in predicting (forecasting) future volatility.

Worako*et al*. (2011) used the (ARCH)/ (GARCH) approach to quantify the volatility of the price of coffee in Ethiopia. A distinction was made between producer, wholesale and export prices in order to compare the price risk faced by the respective participants in the coffee chain. The volatility in coffee prices within Ethiopia was also compared to the volatility levels in Brazilian coffee prices, since Brazil is a major coffee producing country in the world. Coffee prices within Ethiopia were found to be more volatile than in Brazil. Producer prices were found to be the most volatile, followed by whole sale prices and export prices respectively.

Nelson (2015) Kenya operates under a floating exchange rate system where the exchange rate of the country is determined through forces of demand and supply for the local currency. This means that the local currency keeps fluctuating against other world currencies and for this case,

the currencies of the primary market for the coffee sector in Kenya. This study adopted the use of secondary data to achieve the stated research objective. Coffee export earnings from Commercial Bank of Kenya were analyzed together with the Exchange rates (KshsVs USD) obtained from Central Bank of Kenya for the period January 2005 to December 2014. The model adopted for this study included inflation indices and foreign direct investment as a percentage of GDP statistics to derive a wholesome understanding of how these factors affect or relate to coffee export earnings in Kenya. Multiple regressions were employed to determine the relationship between coffee export earnings and the exchange rates, inflation indices and foreign direct investment as a percentage of GDP for the period of 2005-2014. The results proved that fluctuations in exchange rates and Foreign Direct Investments as a percentage of the GDP largely affected coffee export earnings in Kenya.

Manual *et al*. (2015) by using MGARCH models the data used for the analysis are monthly average producer, auction and international price for the period January 1992 through June 2013 in Ethiopia. The producer and auction price data include the five major Ethiopian coffee verities by origin of growing region namely: Sidama, Yirgachefe, Harar, Nekemt and Jimma. Sidama,andYirgachefe are part of SNNP producing region while Harar, Nekemte and Jimma are part of Oromia producing region. The five varieties constitute about 95% of the total coffee produced and exported in the country. Ethiopian coffee prices have become more integrated to world price after the implementation of the Ethiopian commodity exchange (ECX). We follow multivariate GARCH approach to evaluate the degree of correlation and volatility transmission across international auction and producer price returns. The results indicate that not all coffee regions have necessarily become more connected to world markets at least not in the first five years after the implementation of the mandatory regulation. We also do not find higher implementation across regional coffee price.

Naveena (2015) In the present study forecasting models based on Exponential Smoothing Model, Autoregressive Integrated Moving Average (ARIMA), Generalized Auto Regressive Conditional Heteroscedastic (GARCH) model and Artificial Neural Network (ANN) was developed for the export of Indian coffee. Results showed that ANN model forecasts were considerably more accurate than those of other models. Yearly data collected to forecast the export of 2011 to 2013. The Root Mean Square Error (RMSE) and Mean Absolute Percent

error (MAPE) were all lower on average for the artificial neural network model than for the other model. Based on results regarding the performance of different models for one and two step ahead forecasting of Indian coffee export, ANN performed better than other models. For three step ahead forecasting of Indian coffee export ARIMA performed better than other models. The ANN model in overall has provided a better forecast accuracy in terms of conventional RMSE and MAPE values as compared to the other models. The study has suggested that best model for Arabica and Robusta coffee projection plays a vital role in the adjustments of supply and demand in the future.

Gebreyesus (2015) investigating the major determinant of coffee export supply in Ethiopia for the period of 1981 to 2011. The data analysis use vector autoregressive and error correction approach to identify the major determinant. The study used export supply of coffee as dependent variable and Real export price of coffee, Real exchange rate, domestic production of coffee measuring in ton, world production of coffee, road Network measured in kilometers, and openness to trade were independent variables. The estimation of the specified model, test for stationarty were carried out using the Augmented Dickey-fuller tests. The results from the unit root testing revealed that all the variables used in the estimation are integrated of order one. The order of vector autoregressive (VAR) was identified using Schwaz information criterion, Hannan-Quinn Information criterion (HQIC), Akaike information criterion (AIC), and Final predicted error (FPE) criteria. And the result related to use one lag. It further used general causality test so as to find direction of causality between coffee export supply and some of the independent variables. The result of the study indicated that real export price of coffee, domestic production of coffee export, physical infrastructure, and world supply of coffee affect export supply significant.

Bonnell (2015) Time series models for per capita coffee consumption in the United States and average monthly prices for ground roast coffee in the Northeast. The United States were developed using a nine-year series from 1980 to 1988. The univariate Box-Jenkins Autoregressive Integrated Moving-Average (ARIMA) procedure in the Forecast Pro software was applied to the series. An ARIMA process is an algebraic statement exhibiting how a time series variable can be statistically related to its past values. Based on the Schwarz Criterion, an ARIMA (1, 0, 1) was selected to best describe coffee prices in the Northeast and the United

States. The monthly average coffee prices are therefore affected by coffee prices one month back and the corresponding random shocks in the model with the same lag period. The ARIMA procedure was unfit in describing the behavior of per capita coffee consumption in the United States. The consumption forecasts were unreliable because the model generated the same monthly consumption value for the forecast period 1988 to 1990. No alternative procedure was applied in estimating a per capita consumption model for coffee.

Tesfahun*et al. (*2018) by using kaleman filtering algorithm the daily closed price of Ethiopia coffee from 25 June 2008 to 5 January 2017 recorded by Ethiopia commodity exchange (ECX). Market is considered to study coffee price moments the ACF correlogram of coffee price from 2008 to 2017 indicates that the nature of the data is non-stationary. This allows us to use single time series state space model. The kaleman filtering is used to analysis this price through continuous prediction and updating, the algorithm test estimates the coffee price with a minimum error. Moreover, we forecast the coffee price for 5 January 2017 and the forecast value of coffee price on the specified data is 20.2015 US$/17kg and the actual coffee price on the same data is 20.5400 US$/17kg, which is a very good estimate. In general the kaleman filtering algorithm is applied in the field of estimating and forecasting coffee price.

Janvier and Mung (2017) applied ARIMA model. Unit price of coffee export which is the cost of coffee exported per kilogram is one of the indicators to see the economic stability of exported coffee in Rwanda. The value of unit price of coffee export in Rwanda on January 2010 - December 2017 is unstable. Unit price which unstable will impede (slow) the economic development in Rwanda, therefore need to be undertaken against the value of the modeling unit price in the future with a method of ARIMA. The Box-Jenkins methods used in this study is a literature method i.e. authors collect, select and analyze readings related to the issues examined and methods documentation. Based on the result the appropriate model to forecasting unit price of coffee export in Rwanda is ARIMA (1, 1, 1). The results of the forecasting unit price of coffee export on January 2018 to December 2019 are stable enough. The lowest unit price occurred on February 2018 while the highest occurred on December 2019.

Naveena*et al.* (2017) by using Hybrid ARIMA-ANN models the price series on Robusta coffee covered monthly data from January, 1995 to February, 2016(3 month dataset used for validation of result).These series vary from 29 to 155 Rs/Kg so it illustrate the

complexity and variation of Robusta coffee price in Indian. compared with ARIMA and ANN model to evaluate the past behavior of a time series data, in order to make inferences about its future behavior for Robusta species of Indian coffee. Indian Robusta Coffee has made a slot for itself in the world market, particularly for its decent blend up quality. the forecasting performance of these models are evaluated and compared by using common criteria's such as; Root Mean Square Error, Mean Absolute Percentage Error. Key findings reveal the superiority of Hybrid ARIMA-ANN model than in other Models, for forecasting of Indian Robusta coffee price. The model was then used to forecast 3 month out-of-sample set. After going through several stages ARIMA (0, 1, 1) model was found to be the best among the family of ARIMA models.

Novanda*et al*. (2018) by using ARIMA model to analyze and compare among forecasting techniques for selecting the best model. And predict the volatility of coffee commodity prices in Indonesia. The study uses secondary data at two different markets, world and domestic markets. Three forecasting techniques apply in this research, namely, MA, ARIMA and Decomposition. The most appropriate model to forecast world and domestic coffee prices are the ARIMA model. This conclusion is based on the lowest MAPE, MAD and MSD values. After differencing on world price data and domestic commodity coffee prices, ARIMA model for world price data and domestic coffee commodity prices are estimated. Comparing all output results, it is found that the best model of each data price of coffee is ARIMA model with order of (1, 1, 0). While for domestic coffee price, the best model is ARIMA (1, 1, 0).

CHAPTER THREE

3. Data and methodology

3.1. Data and variable of the study

The study used only secondary data and the data obtained from National bank of Ethiopia (NBE). The purpose of this study is to analysis and forecast the export price of coffee in Ethiopia. In order to propose suitable specifications, it is reasonable to investigate the nature of the data set. The characteristics of the data and their descriptive statistics partly indicate the appropriate models which should be employed. Monthly time series data on export price of coffee were used. The variables of interest in this study are export price of coffee, Real exchange rate ,food price Index and non food price Index. Export price of coffee is to be use as a dependent variable. Real exchange rate, food price Index and non food price Index are exogenous variables. June used as a base month in each year to calculate price Indexand non food price Index.

3.2. Methodology

The methodology adopt in this study was follows ARIMAX modeling approach. Before fitting time series models, apply stationary test so as statistical properties such as mean, variance, autocorrelation function are all constant over time. Most statistical forecasting methods are based on the assumption that the time series can be rendered approximately stationary through the use of differencing. Thus, the following step would be conducting unit root tests such as the Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1979) and Phillips-Perron (1988) tests. And then we estimate the autoregressive integrated moving average (ARIMA) and autoregressive integrated moving average with exogenous variable (ARIMAX). And apply diagnostic tests of the models. Finally, the appropriate time series models would be used to forecast the export price of coffee in Ethiopia.

Time series is broadly defined as any series of measurements taken at different times. It can be divided in to two major parts univariate and multivariate time series. Univariate time series analysis uses only the past history of the time series being forecast plus current and past

random error terms. Autoregressive integrated moving average (ARIMA) modeling is a specific subset of univariate modeling, in which a time series is expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a 'white noise' error term (the moving average component). On the other hand, multivariate time series analysis involves more than one time series data sets. Multivariate time series analysis is used when one wants to model and explain the interactions and co movements among a group of time series variables. This study is concern with modeling Univariate time series analysis.

3.2.1.Univariate Time Series Analysis

A time series Y_t is a process observed in sequence over time, t = 1,…,T . To indicate the dependence on time, we adopt new notation, and use the subscript t to denote the individual observation, and T to denote the number of observations. Because of the sequential nature of time series, we expect that Yt and $Y_{t-1}$are not independent, so classical assumptions are not valid. The primary model for univariate time series is Autoregressive (AR). At first sight it would not seem very sensible for economists to pay much attention to univariate Analysis. After all, economic theory is rich in suggestions for relationships between variables. Thus, attempting to explain and forecast a series using only information about the history of that Series would appear to be an inefficient procedure, because it ignores the optional information in related series. There are two possible rationales. The first is that a priori information about the possible relationships between series may not be well founded. In such case a purely statistical model relating current to previous values may be a useful summary device and may be used to generate reliable short-term forecasts. Alternatively, if theoretical speculations about the Economic structure are well founded, it can be shown that one manifestation of that structure Yields Autoregressive Moving Average (ARMA) equations for each of the dependent variable in the structure (Johnston and DiNardo, 1997).Researchers have observed that the large simultaneous-equations Macroeconomic models constructed in the 1960s frequently have poorer forecasting Performances than fairly simple, univariate time-series models based on just a few parameters and compact specifications. It is just this observation that raised to prominence of the univariate Time-series forecasting models pioneered by Box and Jenkins (1976).The theories of linear time series usually include stationarity, autocorrelation function, modeling, and estimation/forecasting (Tsay, 2005).

3.2.2. Stationary test

The foundation of time series analysis is stationary. A time series $\{Y_t\}$ is said to be strictly stationary if the joint distribution of $\{Y_1, Y_2, \dots Y_n\}$ is identical to that of $\{Y_{1+h}, Y_{2+h}, \dots Y_{t+h}\}$ where h is an arbitrary positive integers. In other words, strict stationary requires that the joint distribution of $\{Y, Y_{12} \dots Y_n\}$ is invariant under time shift. This is a very strong condition that is hard to verify empirically. A weaker version of stationary is often assumed. A time series $\{ Y_t\}$ is weakly stationary if both the mean of $Yt$and the covariance between Yt and $Y_{t+h}$are time invariant, where h is an arbitrary integer. Weak stationary implies that the time plot of the data was show that the T values fluctuate with constant variation around a fixed level. If a time series is not stationary, it is necessary to look for possible transformations that might induce stationary. In this study we will apply Unit root test to testing stationary.

3.2.3. Unit Root Test

The assumption of stationary is somewhat an unrealistic situation in most macroeconomic variables. A test for stationary that has become widely popular over the past several years is the Unit Root Test. A non-stationary process arises when one of the conditions for stationary does not hold. For non- stationary series, the effect of a shock never dies away. Moreover, a regression involving non-stationary variables leads to spurious regression; that is, one can regress completely unrelated series and find high R^2 (indicating how good one term is at predicting another) and the standard tests are not valid (Granger and Newbold,1986). Thus, the first step for an appropriate analysis is to determine whether the series is stationary or not. The widely used unit-root tests are the Augmented Dickey Fuller (ADF) and Phillips Perron (PP) tests. The following discussion outlines the basic features of unit root tests (Hamilton, 1994).

Consider an AR (1) process:

$Y_t = \alpha Y_{t-1} + \varepsilon_t$……………………………………………………………..………(1)

Where α is a parameter to be estimated and ε_t is assumed to be white noise. If $|\alpha| \geq 1, \{Y_t\}$ isanon-stationary series and the variance of Y increases with time and approaches infinity. On the other hand, if $| \alpha | < 1$, $\{Y_t\}$ is a stationary series. Thus, the hypothesis of Stationarity can be evaluated by testing whether the value of α is strictly less than one.

The hypotheses are:

H_O: α =1 (the series is not stationary)

H1: α <1 (the series is stationary)

3.2.4. Augmented Dickey–Fuller (ADF) Unit Root Test

The standard Dickey-Fuller test is conducted in the following manner: from equation (1) we have:

$$Y_t - Y_{t-1} = (\alpha - 1) Y_{t-1} + \varepsilon_t$$

This can be rewritten as:

$$\Delta Y_t = \pi Y_{t-1} + \varepsilon_t \quad \ldots\ldots(2)$$

Where π = α-1.The null and alternative hypothesis may be rewritten as:

H_O: α =1

H1: α <1 ……(3)

The test statistic is the conventional t-ratio for π:

$$T_\pi = \frac{\hat{\pi}}{\mathrm{se}(\hat{\pi})} \quad \ldots\ldots(4)$$

Where $\hat{\pi}$ the OLS is estimate of π and $\mathrm{se}(\hat{\pi})$ is the standard error of $\hat{\pi}$. Dickey and Fuller (1979) showed that, under the null hypothesis of a unit root, this statistic does not follow the conventional Student's t-distribution, and they derived asymptotic results and simulated critical values for various tests and sample sizes. MacKinnon (1996) implemented a much larger set of simulations than those tabulated by Dickey and Fuller. In addition, MacKinnon estimated response surfaces for the simulation results, permitting the calculation of Dickey-Fuller critical values and p-values for arbitrary sample sizes.

The simple Dickey-Fuller unit root test described above is valid only if the series is an AR (1) process. If the series is correlated at higher order lags, the assumption of white noise disturbances ε_t is violated. The ADF test constructs a parametric correction for higher-order

correlation by assuming that the series follows an AR (p) process, and adding lagged difference terms of the dependent variable Y_t to the right-hand side of the test regression:

$\Delta Y_t = \pi Y_{t-1} + \beta 1 \Delta Y_{t-1} + \beta 2 \Delta Y_{t-2} + \ldots + \beta p \Delta Y_{t-p} + ut$……………………………..(5)

This augmented specification is then used to test:

H_0: δ =0, H_1: δ < 0

An important result obtained by Fuller is that the asymptotic distribution of the t-statistic on δ is independent of the number of lagged first differences included in the ADF regression. Moreover, while the parametric assumption that y follows an autoregressive (AR) process may seem restrictive, Said and Dickey (1984) demonstrate that the ADF test remains valid even when the series has a moving average (MA) component, provided that enough lagged difference terms are augmented to the regression.

This augmented specification is then used to test for unit root using the t-ratio (equation 4). An important result obtained by Dickey-Fuller (1979) is that the asymptotic distribution of the t statistic is independent of the number of lagged first differences included in the ADF regression. Moreover, while the assumption that Y_t follows an AR process may seem restrictive, Said and Dickey (1984) demonstrate that the ADF test is asymptotically valid in the presence of a moving average component, provided that sufficient lagged difference terms are included in the test regression.

3.2.5. Phillips-Perron (PP) Unit Root Test

Phillips and Perron (1988) developed a number of unit-root tests that have become popular in the analysis of financial time series. This is alternative (nonparametric) method of controlling for serial correlation when testing for a unit root. The PP method estimates the non-augmented DF test and modifies the t -ratio of the α coefficient so that serial correlation does not affect the asymptotic distribution of the test statistic.The Phillips-Perron (PP) unit-root tests differ from the ADF tests mainly in how they deal with serial correlation and heteroskedasticity in the errors. In particular, where the ADF tests use a parametric autoregressive to approximate the ARMA structure of the errors in the test regression, the PP tests ignore any serial correlation in the test regression. The test regression for the PP tests is:

$\Delta Y_t = \sigma_0 + \sigma_1 t + \pi \Delta Y_{t-1} + \varepsilon_t$……………………………………………..……………(6)

Where ε_t is stationary and may be heteroskedastic. The PP tests correct for any serial correlation and heteroskedasticity in the errors ε_t of the test regression by directly modifying the test statistics t_π =0 and T_π.

Where $t_\pi = \frac{\hat{\boldsymbol{\rho}}}{\boldsymbol{se(\hat{\rho})}}$ **and** $\boldsymbol{T_\pi} = \frac{\hat{\pi}}{se(\hat{\pi})}$

Where $\hat{\boldsymbol{\rho}}$ and $\hat{\pi}$ are the OLS estimates of ρ and π, $se(\hat{\rho})$ and se($\hat{\pi}$) are the standard error of $\rho$and π, respectively.

3.3. Time domain approach and ARIMAX mode

The time domain approach is generally motivated by the presumption that correlation between adjacent points in time is best explained in terms of a dependence of the current value on past values. The time domain approach focuses on modeling some future value of a time series as a parametric function of the current and past values. Treating an asset return as a collection of random variables over time, we have a time series $\{Y_t\}$. Linear time series analysis provides a natural framework to study the dynamic structure of such a series.

One approach of time domain analysis is the Box-Jenkins (1970; see also Box et al., 1994) method which develops a systematic class of models called autoregressive integrated moving average (ARIMA) models. Conversely, the frequency domain approach assumes that the primary characteristics of interest in time series analyses are related to periodic or systematic sinusoidal variations found naturally in most data (Chatfield, 2000).

3.3.1. Autoregressive (AR) Process

A time series $\{Y_t\}$ is said to be an autoregressive process of order p. If it is a weighted linear sum of the past p values plus a random shock so that:

$Yt = \mu + \emptyset 1 Y_{t-1} + \emptyset 2 Y_{t-2} + \cdots + \emptyset p Y_{t-p} + \varepsilon t$……………………………………………(7)

Where ε_t is denotes a purely random process with zero mean and variance $\delta^2{}_\epsilon$. Using the backward shift operator B, such that $BY_t = Y_{t-1}$, the AR (p) model may be written more succinctly in the form:

$\emptyset(B)Yt = \varepsilon t$……(8)

Where $\emptyset$ (B) $=1-\emptyset_1 B-\emptyset_2 B^2-\cdots-\emptyset_p B^p$ is a polynomial in B of order p. The properties of AR processes defined by (7) can be examined by looking at the properties of the function$\emptyset$. As B is an operator, the algebraic properties of $\emptyset$have to be investigated by examining the properties of $\emptyset$ (Y), say, where Y denotes a complex variable, rather than by looking at $\emptyset$ (B).

The simplest example of an AR process is the first-order case given by

$Yt = \emptyset Y_{t-1} + \varepsilon t$……(9)

The time-series literature typically says that an AR (1) process is stationary provided that $|\emptyset| <1$. It is more accurate to say that there is a unique stationary solution of (9) which is causal, provided that $|\emptyset| <1$. The autocorrelation function (ACF) of a stationary AR (1) process is given by

$\rho^k = \emptyset^k$ for k=0, 1, 2...

For higher-order stationary AR processes, the ACF will typically be a mixture of terms which decrease exponentially or of damped. The ACF can be found by solving a set of difference equations called the Yule-Walker equations given by

$P^k = \emptyset 1\rho^{K-1}+\emptyset_2\rho^{k-2}+\cdots+\emptyset_p\rho^{k-p}$……(10)

For k =1,2,..., where $\rho_0= 0$. Notice that the AR model is typically written in mean-corrected form with no constant on the right-hand side of (7).This makes the mathematics much easier to handle.

A useful property of an AR (p) process is that it can be shown that the partial ACF is zero at all lags greater than p. This means that the sample partial ACF can be used to help determine the order of an AR process (assuming the order is unknown as is usually the case) by looking for the lag value at which the sample partial ACF 'Cuts off' (meaning that it should be approximately zero, or at least not significantly different from zero, for higher lags) (Hamilton, 1994).

3.3.2. Moving average (MA) processes

The present stationary observation (Yt) of the moving average process is a finite linear function of previous and current values of errors or white noises. A time series $\{Y_t\}$ is said to be a moving

average process of order q (abbreviated MA (q)) if it is a weighted linear sum of the last q random shocks so that

$Y_t = \mu + \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}$……(11)

.Where q is a non-negative integer, and μ, θ_1, $\theta_2 \ldots \theta_q$ are the constant and moving average coefficients, and $\{\varepsilon_t\}$ denotes a purely random process with zero mean and constant variance δ^2_ϵ. Equation (11) may alternatively be written in the form.

$Yt = \theta(B)\,\varepsilon_t$……(12)

Where $\theta(B) = (1 + \theta_1 B + B^2 + \cdots + \theta_q B^q)$ is a polynomial in B of order q. Note that some authors (including Box et al., 1994) parameterize an MA process by replacing the plus signs in (11) with minus signs, presumably so that θ(B) has a similar form to Ø (B) for AR processes, but this seems less natural in regard to MA processes. There is no difference in principle between the two notations but the signs of the θ values are reversed and this can cause confusion when comparing formulae. The autocorrelation function of a MA (q) process is zero beyond the order q of the process. In other words, the autocorrelation function of a moving average, MA (q), process has a cut off after lag q. The PACF will dampen exponentially. The ACF will be used to identify the order of the MA process (Chatfield, 2000)

3.3.3. Autoregressive Moving Average (ARMA) Processes

In some applications, the AR or MA models become cumber some because one may need a higher-order model with many parameters to adequately describe the dynamic structure of the series. To overcome this difficulty, the autoregressive moving-average (ARMA) models are introduced (Box, et al (1994)). Basically, an ARMA model combines the ideas of AR and MA models into compact form so that the number of parameters used is kept small. A mixed autoregressive moving average model with p autoregressive terms and q moving average terms $(ARMA(p,q))$can be written as:

$Y_t = \mu + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots + \emptyset_p Y_{t-p} - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \ldots - \theta_q\varepsilon_{t-q} + \varepsilon_t$……...(13)

Equation (13) may alternatively be written using the backward shift operator as:

$\emptyset\,(B)Y_t = \theta\,(B)\varepsilon_t$ ……(14)

Where$\emptyset(B)$ and $\theta(B)$are polynomials in B of finite order p and q, respectively. Equation (14) has a unique causal stationary solution provided that the roots of $\emptyset(B)$ = 0 lie outside the unit circle. The process is invertible provided that the roots of $\theta(B)$= 0 lie outside the unit circle. The importance of ARMA process is that many real world time series may be approximated in a mixed of AR and MA process.

3.3.4. Autoregressive Integrated Moving Average (ARIMA) Processes

In practice, many time series are non-stationary and we cannot apply stationary AR, MA or ARMA processes directly. One possible way of handling non-stationary series is to apply differencing so as to make them stationary. Differencing of a series can transform a non stationary series to a stationary series. One advantage of differencing over de-trending to remove trend is that no parameters are estimated in the differencing operation. One disadvantage, however, is that differencing does not yield an estimate of the stationary process Y_t. If an estimate of Y_t is essential, then de-trending may be more appropriate. If the goal is to make the data stationary, then differencing may be more appropriate. The first difference operator ∇is defined by:

$$\nabla Yt = Yt - Y_{t-1}$$

$$\nabla Yt = (1 - B)Y_t$$

So that ∇can be expressed in terms of the backward shift operator B. In general, higher order differencing can be expressed as:

$$\nabla nY_t = (1 - B)nY_t$$

A series that is stationary without any differencing is said to be integrated of order zero (denoted by I (0)), and a series which is stationary after being differenced d times is said to be integrated of order d (denoted by I (d)).

If the original data series is differenced d times before fitting an ARMA(p, q) process, then the model for the original undifference series is said to be an ARIMA(p, d, q) process where the letter 'I' in the acronym stands for integrated and d denotes the number of differences taken. Mathematically, equation (14) is expressed in the following way:

$\emptyset\ (B)\nabla dYt = \theta\ (B)\varepsilon t$ …….(15)

Where ∇= 1 –B and d is the order of integration (Chatfield, 2000).

3.3.5. Autoregressive Integrated Moving Average with exogenous variable (ARIMAX) Processes

ARMAX is a special time series model that includes both an autoregressive term ("AR"), a moving average term ("MA") and additional exogenous variables ("X").The ARIMA-X model is an extension of Autoregressive Integrated Moving Average (ARIMA) model. An ARMAX model simply adds in the covariate on the right hand side:

$Yt = \alpha X_t + \mu + \emptyset_1 Y_{t-1} + \emptyset_2 Y_{t-2} + \ldots + \emptyset p Y_{t-p} - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} + \varepsilon_t$…….(16)

Where Yt, is the dependent variable in period t, and $\mu, \alpha, \emptyset_{p,}, \theta_q$ are the constant and coefficients, ε_t is the white noise process and Xt is the exogenous variable at time t. The presence of lagged values of the response variable on the right hand side of the equation mean thatα can only be interpreted conditional on the value of previous values of the response variable, which is hardly intuitive.The equation (16) is the general form of $ARIMAX\ (p, d, q)$ model.

3.3.6. Building ARIMA-X Model

There are a few basic steps to fitting ARIMAX models to time series data. These steps involve plotting the data; possibly transform the data, identifying the appropriate model, parameter estimation, diagnostics and forecasting. The original Box-Jenkins modeling procedure involves an iterative three-stage process of model selection, parameter estimation and diagnostic checking. But, further explanations of the process by Makridakis*et al*.(1998) often add a preliminary stage of data preparation and a final stage of model application (forecasting and evaluation of the forecast).

3.4. Model Identification

The first step in ARIMA modeling is identification of the appropriate model. ARIMA procedures are mainly based on the analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which are essential diagnostic instruments to identify the dependence structure of the series. The autocorrelation of a series $\{y_t\}$ at lag m is estimated by:

$$\hat{\rho}_m = \frac{\sum_{j=1}^{m-1}(Y_t-\bar{Y})(Y_{t-m}-\bar{Y})}{\sum_{j=1}^{m-1}(Y_t-\bar{Y})^2} \quad \text{.....(17)}$$

Where $\bar{Y}$ is the sample mean of Y_t. This is the correlation coefficient for values of the series m periods apart.

The partial autocorrelation function at lag m is the correlation between Y_t and Y_{t-m} after removing the effect of the intervening variables Y_{t-1}, Y_{t-2}, Y_{t-m+1} The partial autocorrelation function, donated by π_m is calculated by iteration method (Shumway and Stoffer, 2010) as:

$\Pi_0 = 1$

$\Pi_1 = \rho_1$

$\pi_m = \frac{\rho_m - \sum_{j=1}^{m-1}\pi_{m-1,j}\rho_{m-j}}{1-\sum_{j=1}^{m-1}\pi_{m-1,j}\rho_{m-j}}$ m>1

Where ρ_m the autocorrelation at lag order m and $\pi_{m,j}$ is given by the recursive equation:

$\pi_{m,j} = \pi_{m-1,j} - \pi_m\pi_{m-1,\,m-j}$, j=1, 2, 3,…,m-j

Table3.1. Behaviour of the ACF and PACF for ARMA Models

Model	ACF	PACF
AR(p)	Infinite: damps Out	Finite: cut off after lag p
MA(q)	Finite: cut off after lag q	Infinite: damps Out
ARMA(p, q)	Infinite: damps Out	Infinite: damps Out

3.5. Model selection criteria

Model selection is one of the fundamental tasks of scientific inquiry. The idea here is to fit all ARMA (p, q) models with orders $p<p_{max}$ and $q<q_{max}$ choose the values of p and q which minimize some model selection criteria. The use of a model-selection criterion, rather than a series of hypothesis tests have some reasons.

- The first reason is model-selection criterion gives a numerical-valued ranking of all models, so that the analyst can see if there is a clear winner or, alternatively, if there are several close competing models. This enables the analyst to assess the strength of evidence when comparing any pair of models, and this is more enlightening than the results of testing one particular model, which might simply be recorded as 'significant' or 'not significant'.
- A model-selection criterion can be used to compare non-nested models, as would arise, for example, when trying to decide whether to compute forecasts using an ARIMA, neural network or econometric model.
- Hypothesis tests require the specification of appropriate null hypotheses, which effectively means that some models are given more prior weight than others. Tests also assume that a true model exists and that it is contained in the set of candidate models.

3.5.1. Akaike Information Criterion (AIC)

Given a set of model for the data, the preferred model is the one with the minimum AIC value. Hence AIC not only rewards goodness of fit, but also includes a penalty that is an increasing function of the number of estimated parameters. This penalty discourages over fitting (increasing the number of free parameters in the model improves the goodness of the fit, regardless of the number of free parameters in the data-generating process). It is given by:

$$\mathbf{AIC = -2 \log (L_k) + 2k} \quad \text{.....(18)}$$

Where k is the number of independent parameters in the model (k= 1+p + q), and $L_k$is the maximized value of the likelihood function for the estimated model. Thus the AIC essentially chooses the model with the best fit, as measured by the likelihood function, subject to a penalty term that increases with the number of parameters fitted in the model.The best model is the one which minimizes the AIC value.

3.5.2. Bayesian Information Criterion (BIC)

An alternative widely used criterion is the Bayesian Information Criterion (BIC) which essentially replaces the term 2k in the AIC with the expression klogT. The BIC, like the AIC, penalizes the addition of extra parameters more severely than the AIC, and should be preferred to the ordinary AIC in time-series analysis, especially when the number of model parameters is

high compared with the number of observations. The BIC was introduced by Schwarz (1978), and is defined as:

$$\mathbf{BIC = -2\ log L_k + k log\ T} \quad \text{..... (19)}$$

Various simulation studies have tended to verify that BIC does well at getting the correct order in large samples, whereas AIC tends to be superior in smaller samples where the relative number of parameters is large (McQuarrie and Tsai (1998)) as pointed out by Shumway and Stoffer (2010). The best model is the one which minimizes the BIC value.

3.6. Parameter Estimation

Ordinal least square estimation was commonly applied to estimate parameter. Once the degree of differencing has been determined, we proceed to select the autoregressive and moving average orders by examining the sample autocorrelations and sample partial autocorrelations. After choosing the most appropriate model, the model parameters are estimated using several estimation procedures. In general, nonlinear estimation procedure is used to estimate ARIMA model parameters. That is, the least squareestimation method is used to test the parameters of the ARIMA model**.**

3.7. Diagnostic Test

Once a model has been identified and the parameters estimated, diagnostic checks will be apply to the fitted model. These include tests of serial correlation and normality of the residuals. After fitting a provisional time series model, we can assess its adequacy in various ways. We check residuals of estimate of ARIMA model(s) to see if they are white noise; Autocorrelation Function (ACF) pick best model with behaved residuals. Once we have identified a tentative model the next step is to determine the adequacy of the models.

Adequate model:

i) The errors are random.

ii) All parameter estimated are significantly different from zero.

iii) The model has the smallest root mean squared error

3.7.1. Ljung-Box test of serial correlation

Before an estimated model is used for statistical inference, the residuals must be examined for the presence of serial correlation. The Q-statistic is often used as a test of whether the series is white noise. The Q-statistic at lag m is a test statistic for the null hypothesis that there is no autocorrelation up to order m and is computed as:

$$Q = T(T+2)\sum_{j=1}^{M} \frac{\hat{\rho}_j^{\,2}}{T-j} \quad \text{……………………………………………………………..……(20)}$$

Where $\hat{p}j$the j-th lags autocorrelation and T is the number of observations. If the series is not based upon results of ARIMA estimation, then under the null hypothesis, Q is asymptotically distributed as chi-square with degrees of freedom equal to (m-k), where k denotes the number of parameters. If the series represents the residuals from ARIMA estimation, the appropriate degrees of freedom should be adjusted to represent the number of autocorrelations less the number of AR and MA terms.

3.7.2. Jarque-Bera test of normality

One of the most commonly applied tests for normality is the Jarque-Bera (JB) test. JB uses the property of a normally distributed random variable that the entire distribution is characterized by the first two moments the mean and the variance. The standardized third and fourth moments of a distribution are known as its skewness and kurtosis. Skewness measures the extent to which a distribution is not symmetric about its mean value, and kurtosis measures the peakedness or flatness of the distribution of the series.

Jarque and Bera (1981) formalize these ideas by testing whether the coefficients of skewness and excess kurtosis are jointly zero. If $\hat{\mu}_3$ and $\hat{\mu}_4$ are estimates of the third and the fourth central Moments, respectively, $\bar{y}$ is the sample mean and $\hat{\sigma}^2$ is an estimate of the second central moment, the sample skewness and kurtosis coefficients are computed as:

$$\hat{b}_1 = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{1}{T}\frac{\frac{1}{T}\sum_{i=1}^{T}(yi-\bar{y})^3}{\sqrt{\{\frac{1}{T}\sum_{t=1}^{T}(y-\bar{y})2\}^3}} \text{ and } \hat{b}_2 = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{T}\sum_{i=1}^{T}(yi-\bar{y})^4}{\sqrt{\{\frac{1}{T}\sum_{t=1}^{T}(y-\bar{y})2\}^2}} \quad \text{……………………………..(21)}$$

The JB test statistic is defined as:

$$JB = T\left(\frac{\hat{b}1}{6} + \frac{\hat{b}_2}{24}\right) \text{…………………….……………………………………………………..(22)}$$

Where T is the sample size, under the null hypothesis this test statistic asymptotically follows a chi-square distribution with two degrees of freedom. Thus, large values of this test statistic relative to the quintiles from the chi-square distribution with two degrees of freedom lead to rejection of the null hypothesis (Brooks, 2002).

3.8. Forecasting

The last step of the ARIMA modeling process is forecasting. In forecasting, the goal is to predict future values of a time series variable, Y_{t+k}, k =1, 2 . . . based on the data collected to the present, Y = $\{Yt, Yt-1, \ldots, Y1\}$.Throughout this section, we would assume $Y_t$is stationary and the model parameters are known. Forecasts of a given variable are based on a model fitted only to present and past observations of a given time series, so that Y_{T+k}depends only on the values of Y_n, Y_{n-1}, Y_{n-2},… possibly augmented by a simple function of time, such as a global linear trend. The last step in ARIMA modeling is forecasting. One of the reasons for the popularity of the ARIMA modeling is its success in forecasting.

3.8.1. Minimum mean square error forecasts

The best linear predictor in terms of minimum means squared error (MSE) of Y_t,…Y_{t-1}or 1-step forecast based on information available at time T. Stationary ARIMA process may generally be rewritten as in the form:

$\emptyset$ (B)$\nabla^d Yt = \theta$ (B)ε_t……(23)

Where $\emptyset$ (B) =1−$\emptyset_1$B−$\emptyset_2$B^2−⋯−$\emptyset_p$ B^p is a polynomial in B of order p is a stationary AR operator and θ(B) = $(1 + \theta 1B + B^2 + \cdot\cdot + \theta_q B^q)$ is a polynomial in B of order q is an invertible MA operator. Suppose at time t we have theY_t , $Y_{t-1,}$ $Y_{t-2,\ldots}$ observations and wish to forecast the h-step ahead future value Y_{t+h}as a linear combination of the observations. Let F_t be the information set at time t.

For t is replaced by t + h, h>1 we have

$Y_{t+h} = \emptyset_1 Y_{t+h-1} + \emptyset_2 Y_{t+h-2} + \ldots + \emptyset p + dY_{t+h-p-d} - \theta 1 \varepsilon_{t+h-1} - \theta_2 \varepsilon_{t+h-2} - \ldots - \theta q \varepsilon_{t+h-q} + \varepsilon_{t+h}$.... (24)

3.8.1. Forecasting accuracy measures

In most forecasting situations, accuracy is treated as the overriding criterion for selecting a forecasting method. In many instances, the word "accuracy" refers to the goodness of fit, which intern refers to how well the forecasting model is able to reproduce the data that are already known. To the consumer of forecasts, it is the accuracy of the future forecast that is most important. If Ytis the actual observation for the period t and Ft is the forecast for the sample period, then the error defined as:

$V_t = Yt - Ft$ ……..(25)

Usually, Ft is calculated using data $Y_1, \ldots, Y_{t-1}$.It is a one step forecast because it is forecasting one period ahead of the last observation used in the calculation .Therefore; we describe Y_t as a one step forecast error. It is the difference between the observation$Y_t$and forecast made using all observations up to but not including Y_t.

There are observations and forecasts for T time periods, then there would be T error terms, and the following standard statistical measures can be defined:

Mean error (ME) = $\frac{1}{T}\sum_{t=1}^{T} Vt$ ……..(26)

Mean absolute error (MAE) = = $\frac{1}{T}\sum_{t=1}^{T} |Vt|$ …….(27)

Mean squared error (MSE) = = $\frac{1}{T}\sum_{t=1}^{T} V_{t1}{}^{2}$ ……..(28)

Equation (25) can be used to compute the error for each period. These can then be averaged as in equation (26) to give the mean error. However, the ME is likely to be small since positive and negative errors tend to offset one another. It does not give much indication as to the size of the typical errors. Therefore, the MAE is defined by first making error positive by taking its absolute value, and then averaging the results. The idea behind the definition of MSE is similar. Here the errors are made positive by squaring each one, and then the squared errors are averaged. The MSE has advantage of being more interpretable and is easier to explain to non-specialist.

Each of these statistics deals with measures of accuracy whose size depends on the scale of the data. Therefore, they do not facilitate comparison across different time series and for different

time intervals. To make comparisons we need to work with relative or percentage error measures. First let us define a relative or percentage error as

$$PE_t = \left(\frac{Yt - Ft}{Yt}\right) X100 \quad \text{.....(29)}$$

Then the following two relative measures are frequently used:

Mean percentage error (MPE) = $\frac{1}{T}\sum_{t=1}^{T} PE_t$ .....(30)

Mean percentage Absolute error (MAPE): The first one is the Mean Absolute Percentage Error (MAPE). It is scale independent and its main issue is to be undefined when the denominator is null. This may happen often with intermittent data. Its equation is given below.

MAPE = $\frac{1}{T}\sum_{t=1}^{T} |PE_t|$ .....(31)

Equation (29) can be used to compute the percentage error for any time period .These can be averaged as in equation (30) to give the mean percentage error. However, as with the ME, the MPE is likely to be small since positive and negative PEs tends to offset one another. Hence the MAPE is defined using absolute values of PE in equation (32). Alternatively, Theil's U statistic can be used as a measure of forecasting accuracy. Like MAPE statistics, high values suggest poor performance in the forecast. However, and unlike MAPE, the U-Theil corrects the performance scale that MPAE had. Theil's U can be estimated as:

$$U = \frac{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Yt - Ft)2}}{\sqrt{\frac{1}{T}\sum_{t=1}^{T}(Yt)^2} + \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Ft)^2}}$$

The scaling of U is such that it will always lie between 0 and 1. If U = 0 Yt = Ft for all forecasts

And there is a perfect fit; if U = 1 the predictive performance is not good.

CHAPTER FOUR

4. DATA ANALYSIS

4.1. Descriptive statistics and time plot the variables

The data analysis is carried out using E-views 7 software the widows based forecasting and econometrics analysis package. The data set consists of monthly export prices of coffee in Ethiopia (in million of US dollar), exchange rate (in birr per million of US dollar), food price index and non-food price index in Ethiopia for the period spanning from July, 2002 to January, 2019.The total number of observations for each variable is 199 monthly data. In the estimation purpose used the data from the period July, 2002 to December, 2016 the total number of these observations are 174. And from January, 2017 to January, 2019 used as out of sample forecast the total numbers of observations are 25. Summary statistics of the export price of coffee and selected variables are presented in the table below.

Table 4.1: descriptive statistics of export price of coffee, exchange rate, food price and non-food price index.

Series	Mean	median	Maximum	Minimum	Std. Dev.
Coffee price	3.091	3.125	5.660	1.190	1.125
Exchange rate	13.708	12.537	22.345	8.568	4.989
Food price	50.817	42.050	105.400	20.300	28.075
Non-food price	49.427	40.650	100.000	19.900	27.034

From the above result, the mean values of export price of coffee (in million of US dollar), Exchange rate (in birr per million of US dollar), food price and non-food price index over the study period are3.091, 13.708, 50.817and 49.427, respectively. The minimum and maximum values of the export price of coffee was 1.190and 5.660, respectively. The standard deviation of the export price of export price of coffee is 1.125.

4.1.1. Time plot of original series

The time plot of the series under consideration are shown in Figure 4.1. From the figures we can observe that the export price of coffee exhibits an upward trend up to October 2011 and from December, 2013 to February, 2015. And downward trend from the period October, 2011 to December 2013 and February, 2015 to December, 2015. Finallyobservesome upward and downward from the period December 2015 to December, 2016. Moreover, in appendix figure 4.3 shows Exchange rate, food price index and non-food price index show an increasing trend. The Y-axis is the value of coffee price and X-axis indicates month and year.

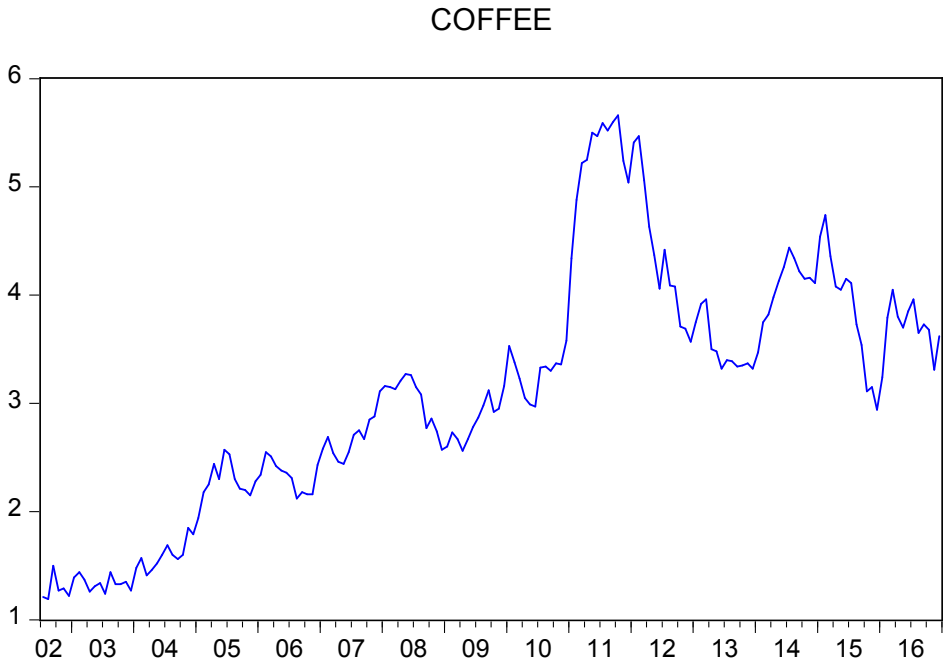


Figure4.1 time plot of original series export price coffee in Ethiopia.

4.2. Unit Root Test of Stationary

It is the formal method to test the Stationarity of the series. Unit root tests are used to test the null hypothesis that an observed trend in a time series is erratic and not predictable versus the alternative that the trend is smooth and deterministic. They can be viewed as a statistical test to see a series should be differenced or linearly de-trended to achieve Stationarity. Before construct the main models, the first logical step is to check the Stationarity of the average monthly export

price of coffee. That is, variables have to be tested for the presence of unit root(s) and the order of integration of each series has to be determined. The Stationarity of each series can be tested using the Augmented Dickey-Fuller (ADF) test and the Phillips and Perron (PP) test. The null hypothesis of the test the series is not stationary and the alternative hypothesis the series is stationary. The results of ADF and PP tests with intercept but no trend and with intercept and trend both at level and first difference for each series are presented in Tables 4.2 and 4.3.

Table 4.2:Unit root test results (At level)

Include test equation with intercept					Include test equation with trend and intercept			
series	Test statistic		P-value		Test statistics		p-value	
	ADF	PP	ADF	PP	ADF	PP	ADF	PP
Coffee price	-1.890	-1.801	0.337	0.379	-2.183	-1.978	0.495	0.609
Exchange rate	1.380	1.380	0.998	0.998	-2.117	-2.112	0.533	0.535
Food price	2.703	2.212	1.000	0.999	-2.799	-2.7197	0.199	0.230
Non food price	4.133	3.282	1.000	1.000	-2.667	-2.604	0.252	0.279
Critical values	1%	-3.468			-4.012			
	5%	-2.878			-3.436			
	10%	-2.576			-3.142			

The critical values used for the tests are the MacKinnon (1996) one-sided critical values and one-sided p-values. Test results presented in Table 4.2 indicate that p-value of all four series greater than the selected significant levels. And also the absolute values of the test statistic are less than absolute values of the critical values.The null hypothesis that the series in levels contain unit root

could not be rejected this means the series is not stationary at level. This non-stationary series indicates mean, variance and covariance are not constant over time. Then the series is non-stationary, it is necessary to look for possible transformations that might induce Stationarity. In practice, econometricians usually transform financial prices in to return forms. This is because often return series are found to be stationary such that analysis is possible. In this study applied to their first differences in order to determine the order of integration of the non stationary time series. First differencing removes a linear trend from the data.

Table 4.3: Unit root test results (At first difference)

Include test equation with intercept					Include test equation with trend and intercept			
Series	Test statistic		probability		Test statistics		p-values	
	ADF	PP	ADF	PP	ADF	PP	ADF	PP
Coffee price	-10.489	-10.489	0.00	0.00	-10.499	-10.507	0.00	0.00
Exchange rate	-11.848	-11.848	0.00	0.00	-12.086	-12.07	0.00	0.00
Food price	-10.713	-11.035	0.00	0.00	-11.330	-11.363	0.00	0.00
Non food price	-4.902	-12.564	0.00	0.00	-13.249	-13.358	0.00	0.00
Critical values	1%	-3.468			-4.012			
	5%	-2.878			-3.436			
	10%	-2.576			-3.142			

The results in Table 4.3 indicate that all p-value of the four series less than the selected significant levels. And also the absolute values of the test statistic are greater than absolute values the critical values. This indicates the null hypothesis of unit root is rejected for the first differences all the variables under consideration. In other word reject the null hypothesis means the series is stationary. This implies that the four time series are integrated of order one (1).

Therefore, the ADF and PP tests show that all series are non-stationary in levels but stationary in first differences. This first differenced of the series has whose statistical properties such as: mean, variance and covariance are constant over time.

4.3. ARIMAX model

An ARIMAX model was endeavored using the E-views statistical package. This model used three tools for modeling process. The first is autoregressive (AR) terms. The second tool is the integrated term. The forecasting model contains an integrated term it can handle a series that tends to drift over time. A pure first-order integrated process is called a random walk. The third tool is the moving average term. A moving average forecasting model uses lagged values of the forecast error to improve the current forecast. A first-order moving average term uses the most recent forecast error. There are a few basic steps to fitting ARIMAX models to time series data. Now, the model that we are looking at is ARIMAX (p, 1, q). The original Box-Jenkins modeling procedure involves an iterative three-stage process we have to model selection, parameter estimation and diagnostic checking. And a final stage of model application (forecasting and evaluation of the forecast).

4.3.1. Model selection

In this step, the tentative form of the ARIMAX model is determined. The order of differencing is determined and the next step in building an ARIMAX model for a series. In most a lower order ARIMA models are often considered. It is better to compare some AR, MA, ARMA and ARIMA models and take the best one. Computing different model selection criterion in order to suggest the best of all the ARIMAX model alternatives at the in- sample stage for coffee price is useful. Here, two criterions, Akaike Information Criterion (AIC) and Bayesian Information or Schwarz Criterion (BIC) are used. The presence of autocorrelation in the residual was tested using the Lagrange multiplier (LM) test for each of the mean equations considered. Only models with no remaining serial correlations are considered as candidate models.

Table 4.4:ARIMAX model selection criteria for monthly differenced export price of coffee data using AIC, BIC and serial correlation.

Model	AIC	BIC	Serial correlation
ARIMAX(0, 1, 1)	**-0.967510**	**-0.876375**	**NO**
ARIMAX(0, 1, 2)	-0.958152	-0.848789	NO
ARIMAX(0, 1, 3)	-0.951326	-0.823736	NO
ARIMAX(0, 1, 4)	-0.939767	-0.793950	NO
ARIMAX(1, 1, 0)	-0.962172	-0.870675	NO
ARIMAX(1, 1, 1)	-0.980451	-0.870655	NO
ARIMAX(1, 1, 2)	-0.974220	-0.846124	NO
ARIMAX(1, 1, 3)	-0.984973	-0.838578	NO
ARIMAX(1, 1, 4)	-0.976595	-0.811900	NO
ARIMAX(2, 1, 0)	-0.968057	-0.857823	NO
ARIMAX(2, 1, 1)	-0.972956	-0.844350	NO
ARIMAX(2, 1, 2)	-1.003184	-0.856206	NO
ARIMAX(2, 1, 3)	-0.996431	-0.831080	NO
ARIMAX(2, 1, 4)	-0.989653	-0.805930	NO
ARIMAX(3, 1, 0)	-0.975650	-0.846528	NO
ARIMAX(3, 1, 1)	-0.987691	-0.840124	NO
ARIMAX(3, 1, 2)	-1.015590	-0.849577	NO
ARIMAX(3, 1, 3)	-1.010628	-0.826169	NO
ARIMAX(3, 1, 4)	-1.013285	-0.810380	NO
ARIMAX(4, 1, 0)	-0.957442	-0.809281	NO
ARIMAX(4, 1, 1)	-0.972658	-0.805977	NO
ARIMAX(4, 1, 2)	-1.020883	-0.835682	NO
ARIMAX(4, 1, 3)	-0.995027	-0.791306	NO
ARIMAX(4, 1, 4)	**-1.032586**	**-0.810345**	**NO**

The appropriate model should have lowest values of Akaike Information Criterion (AIC) and Bayesian Information or Schwarz Criterion (BIC). Table 4.4 gives the alternative ARIMAX models with their AIC and BIC values and also serial correlation of the differenced series. Hence, considering thus Table the minimum AIC value and exhibit no serial correlation obtained at ARIMAX (4, 1, 4) and the minimum BIC value and exhibit no serial correlation obtained at ARIMAX (0, 1, 1) for differenced export price of coffee. To select the appropriate model from the two ARIMAX models used the forecasting accuracy methods based on the data from the period July, 2002 to December, 2016. The model has the smallest values of RMSE, MAE and Theil inequality coefficient. The results of thus values are displayed in table 4.5.

Table 4.5: forecast accuracy measure of $\boldsymbol{ARIMAX(0,1,1)}$ and $\boldsymbol{ARIMAX(4,1,4)}$ models of export price of coffee.

Criteria	ARIMAX(0,1,1)	ARIMAX(4,1,4)
Root Mean Squared Error	0.14398	0.136528
Mean Absolute Error	0.0.11396	0.108002
Theil Inequality Coefficient	0.741936	0.647279

From table 4.5 observed that$ARIMAX\ (4,1,4)$ has the smallest values in most of the forecast accuracy measures than $ARIMAX(0,1,1)$. Therefore, $ARIMAX(4,1,4)$ is the best model for the data under consideration. $ARIMAX(4,1,4)$, Indicates the model has: four lag of dependent variable (4), the variable being used is first difference stationary (1) and four lags of the error term (4). Hence the model given by the following expression is used to the main equation of the differenced export price of coffee.

$(1-\emptyset_1 B-\emptyset_2 B^2-\emptyset_3 B^3-\emptyset_4 B^4)\ \nabla Y_t = \alpha Xt + (1+\theta_1 B+\theta_2 B^2+\theta_3 B^3+\theta_4 B^4)\ \varepsilon_t$

Where Y_t, is represent the series of coffee price in period t, and $\mu, \alpha, \emptyset_p,, \theta_q$ are the constant and coefficients, ∇ is the first order of integration, ε_t is the white noise process and Xt is the exogenous variable at time t.

4.3.2. Model estimation

E-views estimate ARIMAX models using Least squares estimation. In our estimation specification include as many MA and AR terms in the equation. In box- Jenkins methodology parameter estimates are usually obtained by Least squares method which is asymptotically correct for any time series data. Hence, we used Least squares estimation method for monthly export price of coffee to estimate the parameters. As this stage it was done a test of the significant of the parameters. Table 4.6 below, the p-values obtained AR (1) in the table the final estimate of the parameters, namely in the amount of 1.212. Because of the p-value= 0.0008< 0.05 then the parameter AR (1) is significant. The obtained p-value is also an MA (1) in the table the final estimate of the parameter is -1.304. Because of the p-value= 0.0004< 0.05 then the parameter MA (1) is significant. Since the parameter of AR (1) and MA (1) significant then models ARIMAX (4, 1, 4) can be inserted in to likely model. The results are summarized in the following table below.

Table 4.6: Parameter estimation of ARIMAX (4, 1, 4)

Variable	**Coefficient**	**Std. error**	**t-statistics**	**p-value**
Constant(C)	-0.072	0.037	-1.946	0.0535
Exchange rate(X1)	0.023	0.008	3.098	0.0023*
Food(X2)	-0.023	0.013	-1.801	0.0737*
Non-food(X3)	0.0193	0.013	1.487	0.1389
AR(1)	1.212	0.356	3.405	0.0008*
AR(2)	-0.596	0.674	-0.884	0.3782
AR(3)	-0.407	0.672	-0.606	0.5452
AR(4)	0.603	0.317	1.904	0.0588
MA(1)	-1.304	0.362	-3.602	0.0004*
MA(2)	0.611	0.721	0.848	0.3980
MA(3)	0.378	0.715	0.528	0.5981
MA(4)	-0.659	0.351	-1.878	0.0623

An AR (4) process is stationary provided that $\left| \emptyset_1+\emptyset_2 + \emptyset_3+\emptyset_4 \right| < 1$. Since, the sum of the coefficient is 0.812. It is more accurate to say that there is a unique stationary condition

of AR(4) which is causal, provided that the sum is less than one. The condition $|\theta_1+\theta_2+\theta_3+\theta_4|<1$ is called the invertible condition for a MA (4) process. The sum of the coefficient is |-0.974|. Imposing an invertible condition ensures that there is a unique MA process for a given ACF. Therefore, the selected model is satisfying the stationary and invertible conditions.

From the table 4.6 observe the p-value of exchange rate is 0.0023 it is less than 0.05 (i.e. 0.0023<0.05) so, we conclude that increase exchange rate in million of US dollar in each month the export price of coffee is increased on averagely by 2.3% in each month. Therefore, exchange rate has significant effect on export price of coffee in Ethiopia. And also food price index is significant at 10% level of significant. The P-value of food price is 0.0737 it is less than 0.10 (i.e. 0.0737<0.10). So, we conclude that increase food price index in each month the export price of coffee is decreased on averagely by 2.3% in each month. Therefore, food price has significant effect on export price of coffee in Ethiopia.

4.3.3. Diagnostic Checking

The goal of ARIMAX model is a parsimonious representation of the process governing the residual. Various diagnostics tests were performed to check the appropriateness of the fitted models. In appendix figure 4.7 The Ljung-box Q-statistics test indicates that autocorrelation in the standardized residuals are not significantly zero for the first 20 lags for coffee price. Indicating that the residuals are uncorrelated (white noise) the flat correlogram of the residuals is most ideal. The correlogram of the residuals is flat which indicates that all information has been captured. In appendix figure 4.8 also Breusch-Godfrey Serial Correlation Lagrange multiplier (LM) test is used to test $ARIMAX\ (4,1,4)$ has no serial correlation.

In appendix figure 4.10 the histogram displays the frequency distribution of the residuals. It divides the series range (distance between its maximum and minimum values) into a number of equal length intervals or bins and displays a count of the number of observations that fall into each bin. The Jarque-Bera statistic tests used to test whether the residual of the fitted model is normally distributed or not. Under the null hypothesis of normality, the p-value is 0.672 this is greater than 0.05 (i.e. 0.672>0.05), indicate that the residuals of the fitted model is normally distributed. So, the forecast will be based on this model.

4.3.4. Forecasting

After estimated an equation with an ARIMAX error process, we can make forecasts from the model. E-Views automatically make use of the ARIMAX error model to improve the forecast. Essence of fitting an ARIMAX model is to forecast future values of series. The final selected ARIMAX (4, 1, 4) model is used to forecasting the export price of coffee. In this study Dynamic forecasting method were used. Dynamic method calculates forecasts for periods after the first period in the sample by using the previously forecasted values of the lagged left-hand variable. These are also called n-step ahead forecasts. In a properly specified model, the realized values of the endogenous variable will differ from the forecasts by less than plus or minus two standard errors 95 percent of the time. A plot of this 95% confidence interval is produced when you make forecasts in statistical package for social since (SPSS). The out-off-sample forecasting period is from January, 2017 to January, 2019.

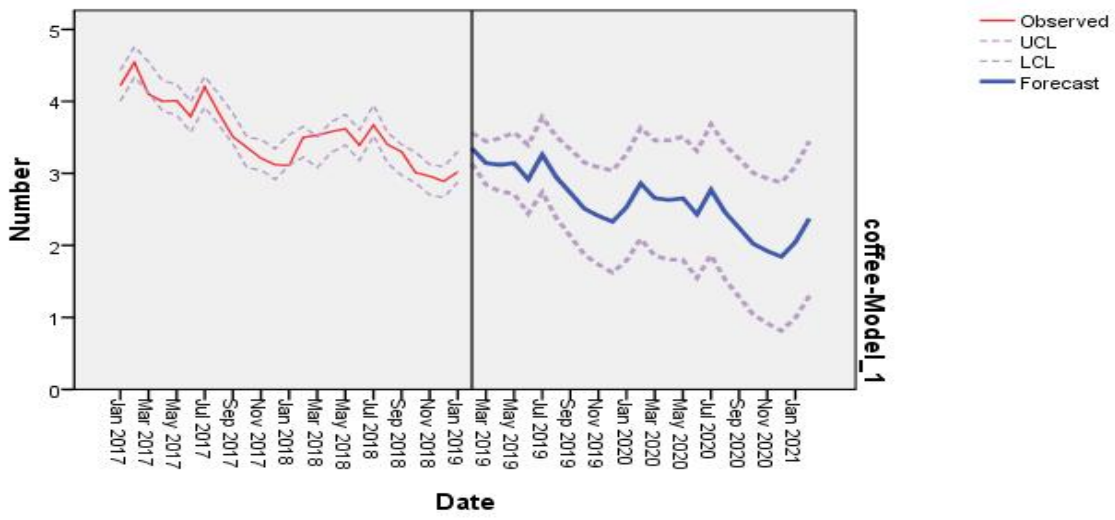


Figure 4.11 out-off-sample forecast of monthly export price of coffee.

From the above figure observed that export price of coffee in Ethiopia decrease from period to period in the future two consecutive years. From period February, 2019 to April 2019 observe coffee price is decrease and May, 2019 rise by small amount. Frome Jun, 2019 to December, 2019 decrease gradually and some up and dawn in the two consecutive months. Finally decrease Frome March, 2020 to December, 2020 and in the first two month of 2021 rise by small amount.

Table 4.7 Forecast accuracy measure of ARIMAX (4, 1, 4) models of export price of coffee.

Criteria	ARIMA(4,1,4)
Root Mean Squared Error	0.1306
Mean Absolute Error	0.0978
Mean Absolute Percentage Error	91.0094
Theil Inequality Coefficient	0.3209
Bias Proportion	0.0015
Variance Proportion	0.0815
Covariance Proportion	0.9170

Table 4.7 indicates the model has lowest root mean square (RMSE) and Mean Absolute Error (MAE) beast the forecasting. The values of RMSE and MAE are 0.1306 and 0.0978 respectively thus numbers are very small indicates the forecast is good. Theil inequality coefficient (U) is used to in the accuracy measures. The value of U is 0.3209 this is close to zero. Mining that there is a perfect fit the actual and the forecast values are the same. Bias proportion is equals to 0.0015or 0.15% this is also called systematic error. The gap between the actual mean and the forecast value is small. Variance proportion Indication of the ability of the forecasts to replication degree of variability in the variable to be forecast. The variance proportion is 0.0815 this value is small so the forecast is good. Covariance proportion measures unsystematic error. Ideally, this should have the highest proportion of inequality this value is 0.9170.

CHAPTERFIVE

5. Conclusion and Recommendation

5.1. Conclusion

The objective of this study was to apply univariate time series analysis to export price of coffee in Ethiopia using monthly data ranging from July 2002 to January 2019. The Stationarity of each series can be tested using the Augmented Dickey-Fuller (ADF) test and the Phillips and Perron (PP) unit root tests. Unit root test results indicate the series is not stationary at level become stationary at first order. Among all candidate ARIMAX models, ARIMAX (4, 1, 4) has the lowest Akaike Information Criterion (AIC). The best model for the series under consideration is ARIMAX (4, 1, 4) for export price of coffee in Ethiopia. Moreover, Least squares estimation indicates that the coefficients of the AR(1) and MA(1) terms are significant at 5% level of significance implying that the selected ARIMAX model is beast to forecast the future values of export price of coffee. Furthermore, Breusch-Godfrey Serial Correlation Lagrange multiplier (LM) test and Jarque-Bera statistic tests are show that the disturbance terms are white noise and a series is normally distributed. The ARIMAX (4, 1, 4) model analysis result shows from the selected variable exchange rate and food price index has significant effect on export price of coffee in Ethiopia. Whereas non food price index has not significant effect on export price of coffee. Finally, forecasts are made using ARIMAX (4, 1, 4) model. Moreover, the out-of-sample forecasts indicate that the export price of coffee has decreasing trend in the future.

5.2. Recommendation

Based on the finding we would like to recommend the government to increase coffee price. The government support to increase the exchange rate of the country. It is important to increase the export price of coffee in Ethiopia and the export price of the country. In addition to this the government controls and reduced food price inflation of the country in order to increase the export price of coffee in Ethiopia. Moreover, the government use the beast ways of reduced the factors to increasing export price of coffee and export performance of the country. In regard to internal impacts, policy makers and government administrative bodies can play active role in implementing appropriate policies, reducing institutional and structural obstacles related export price of coffee in Ethiopia.

Reference

1. AmareTerefe (2009). Application of GARCH Models in Forecasting the Volatility of Export Price in Ethiopia.
2. Box, G.E.P. and G.M. Jenkins (1970). Time series analysis: Forecasting and control, San Francisco: Holden-Day.
3. Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. San Francisco: Holden Day.
4. Box, G.E.P., G.M. Jenkins and G.C. Reinsel (1994). Time Series Analysis, Forecasting, and Control, 3rd ed. Englewood Cliffs, NJ: Prentice Hall.
5. Brooks, C. (2002). Introductory Econometrics for Finance. Cambridge University Press.
6. Chatfield, C. (2000). Time-Series Forecasting in London: Chapman & Hall/CRC.
7. CSA (Central Statistical Agency). 2013. Agricultural Sample Survey 2012/2013 (2005 E.C.), Report on area and production of major crops. Central Statistical Agency, Addis Ababa.
8. Dickey, D.A. and Fuller, W.A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root, Journal of the American Statistical Association, 74, 427–431.
9. Dodaro, S. (1993). Exports and growth: A reconsideration of causality. The journal of developing areas, 27 (2), 227-244.
10. Engle, R.F (1982). Autoregressive conditional Heteroskedsticity with Estimates of the variance of United Kingdom inflation. Journal of economic 50:987-1007.
11. Gebre-medhin (2001). Market institution, transaction costs, and social capital in the Ethiopian Grain market. Research report 124. International food policy research institute, Washington, DC.
12. Gebre-medhin, E.Z (2009). A market for all farmers: market Institution and small holder participation. Agriculture for development paper No AFD-0903. Is used July.
13. Gebreyesus T. (2015). Determinants of Coffee Export Performance in Ethiopia. College of Business and Economics, Department of Economics, Aksum University, Ethiopia,

14. Gemech, F. and Struthers, J. (2007). Coffee Prices Volatility in Ethiopia: Effects of Market Reform Programs, Journal of International Development J. Int. Dev. 19, 1131–1142, www.interscience.wiley.com, Published online in Wiley Inter Science.
15. Gemech, F., Struters, J. (2007). Coffee price volatility in Ethiopia: effects of market reform programs. Journal of international development 19, 1131-1142.
16. Granger and Newbold (1986). Forecasting Economic Time Series. Academic Press.
17. Hamilton, J.D. (1994): Time Series Analysis. Princeton University Press, Princeton.
18. Hultman, C.W. (1967). Exports and Economic Growth: A Survey, Land Economics, Vol.43, pp. 148 157.
19. International Coffee Organization (ICO) (2013). Historical Data. Summer 2013. Available at: http://www.ico.org/new_historical.asp [accessed 3 Sept 2013].
20. Janvier and Mung (2017). Forecasting the Unit Price of Coffee Export in Rwanda Using ARIMA Model.Jomo Kenyatta University of Agriculture and Technology, Kenya International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064.
21. Jarque, C. M. and Bera, A. K. (1981). An Efficient Large-Sample Test for Normality of
22. Johnston, J. and DiNardo, J. (1997): Econometric Methods, International Editions, McGraw Hill Book.
23. MacKinnon, J. (1996). Numerical Distribution Functions for Unit Root and Co integration Tests, Journal of Applied Econometrics, 11, 601-618.
24. Makridakis S. and collegeus (1984). Forecasting acorecy of time series methods, chichester: wiley.
25. Manual A. Hernandez, Solommon lemma and Shahidur Rashid (2015). The Ethiopian commodity exchange and the coffee market: are local prices more integrated to global market.
26. McQuarrie, A.D.R. and C-L. Tsai (1998). Regression and Time Series Model Selection, Singapore: World Scientific.
27. MekonenHailemichaelSalla (2009). Influence of Genotype, Location and Processing Methods on the Quality of Coffee (Coffee Arabica L). Hawassa University, Ethiopia.
28. Michael and Counselor, (2017). Government Institutes Reforms to Boost Coffee Exports. Addis Ababa Ethiopia.

29. Naveena (2015). Statistical Modeling for Forecasting of Price and Export of Indian Coffee. Department of Farming Engineering Institute of Agricultural sciences Banaras Hindu University Varanasi - 221 005.
30. Naveena K., Singh Subedar, RathodSantosha and Singh Abhishek (2017). Hybrid Time Series Modeling for Forecasting the price of Washed coffee (Arabica Plantation coffee) in India. Institute of Agricultural Sciences, Banaras Hindu University, Varanasi, 221005.
31. Nelson Mandela Abuka (2015). The Effects of Exchange Rate Fluctuations on Export Earnings of Coffee Industry in Kenya. Registration number; D63/72702/2014.
32. Novanda R, EkoSumartono, PutriSuciAsriani, EllysYuliarti, KetutSukiyono, BasukiSigitPriyono, Irnad, Reswita, MellySuryanty and Vera Octalia (2018). Comparison of Various Forecasting Techniques for Coffee Prices. Department of Agricultural Socio-Economic, Faculty of Agriculture, University of Bengkulu, Indonesia.
33. Observations and Regression Residuals, Economics letters, 7(4). 313-318.
34. Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. Biometrika, 75, 335-346.
35. Rasid S., Winter-Nelson A., Gorcia P. (2010). Purpose and potential for commodity Exchanges in African Economies. IFPRI Discussion paper No, 01035. International food policy Research institute, Washington DC.
36. Roger (1992).The economy of Ethiopia.StMartin press London.
37. Said, S.E. and D. Dickey (1984). Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika, 71, 599-607.
38. Shumway, R.H. and D.S. Stoffer (2010). Time Series Analysis and Its Applications with R Examples, Second Edition. New York, Springer.
39. Tefera, A. (2012). Ethiopia Coffee Annual Report. Rep. no. ET 1202. Global Agricultural Information Network (GAIN), US Department of Agriculture, 15 May 2012.
40. TesfahunBerhaneNurilingShibabaw, AemiroShibabaw, MolalignAdem, Abera, muhamed (2018). Forecasting the Ethiopian coffee price using kalman filtering Algorithm. Journal of resource and Ecology, 2018, 9(3):302-305.

41. Tsay, R.S. (2010). Analysis of Financial Time Series, Third Edition. New Jersey, John Wiley & Sons, Inc.
42. Worako, H.D van Schalkwyk, Z.G Alemu& G Ayele (2008). Producer price and price transmission in a deregulated Ethiopian coffee market.
43. Worako, T. K., Jordaan, H., & Van Schalkwyk, H. D. (2011). Instigating volatility in coffee prices along the Ethiopian coffee value chain. Agrekon, 50(3), 90-108. https://doi.org/10.1080/03031853.2011.617865
44. ZakariaSamiaGutu (2000). Policy options for Ethiopia's coffee export. Addis Ababa, Ethiopia.
45. ZelalemBefekadu(1997). Evaluation of the system of coffee marketing in Ethiopia.
46. ZerihunG.Alemu and TadeseK.Worrako (2008). Price Transmission and Adjustment in Ethiopian Coffee Market. University of the Free State.
47. Zewdu A., Bamlaku A., Alemub M. (2010). Agricultural productivity growth and poverty reduction in rural Ethiopia. Institute for land, water and society, Charles Sturt University, WaggaWagga campus, NSW, Australia 2010.

Appendix

Figure 4.2 time plot of differenced series of export coffee price.

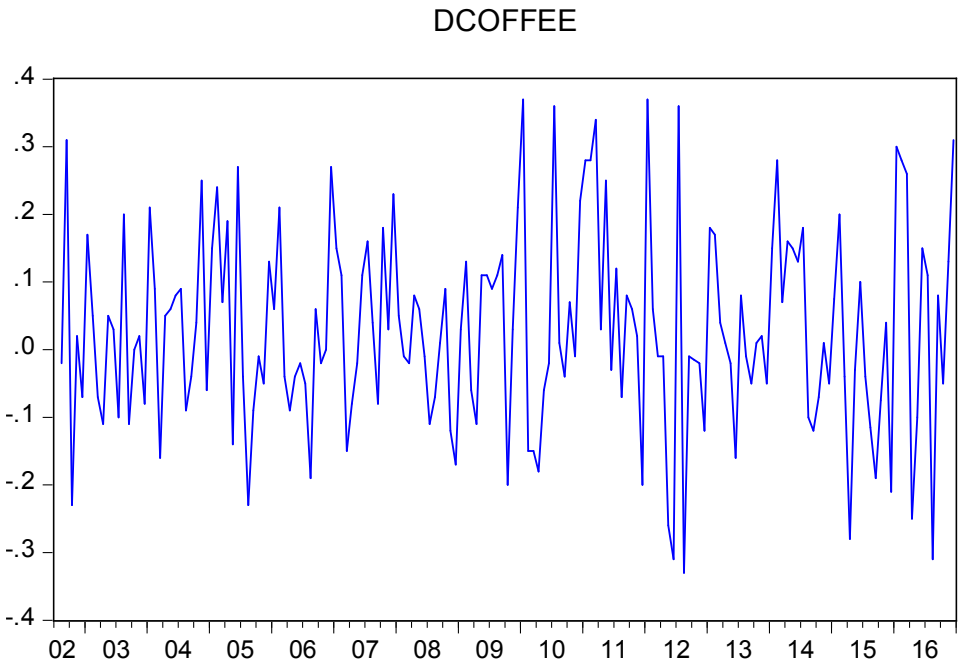
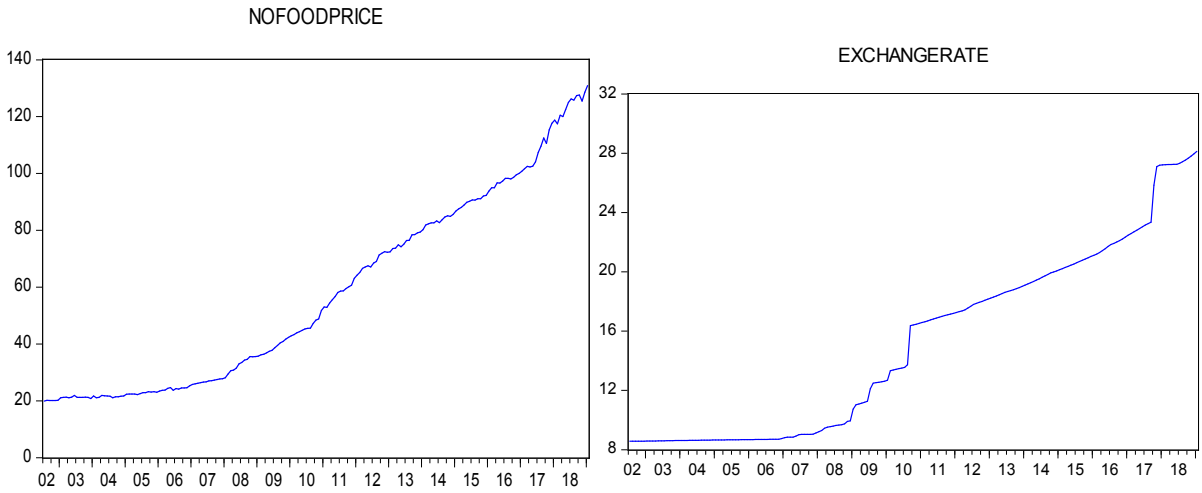


Figure 4.3 time plot of original series of exchange rate, food price and non-food price data in Ethiopia.



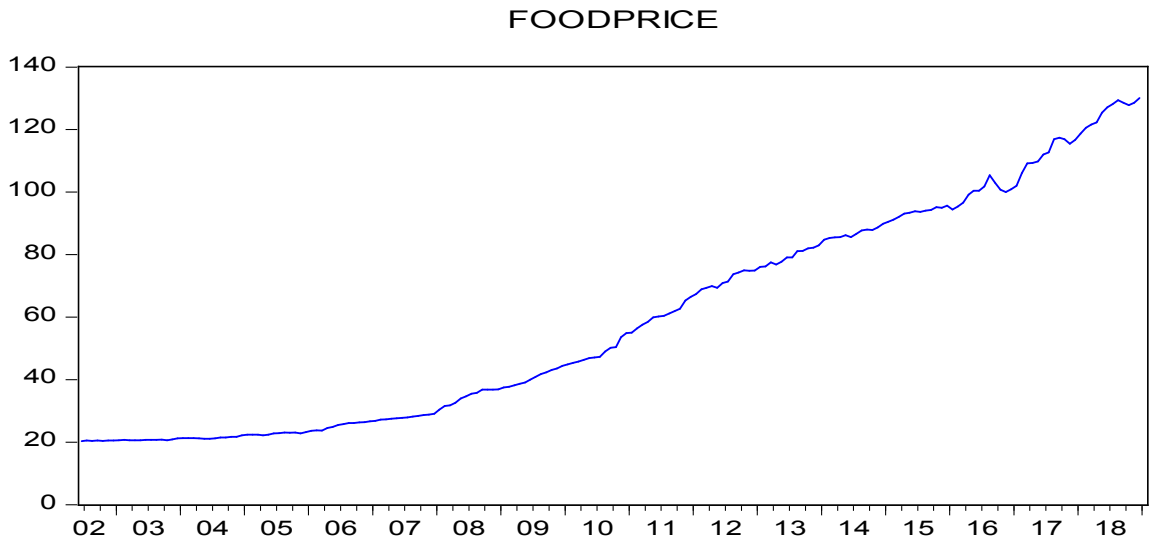
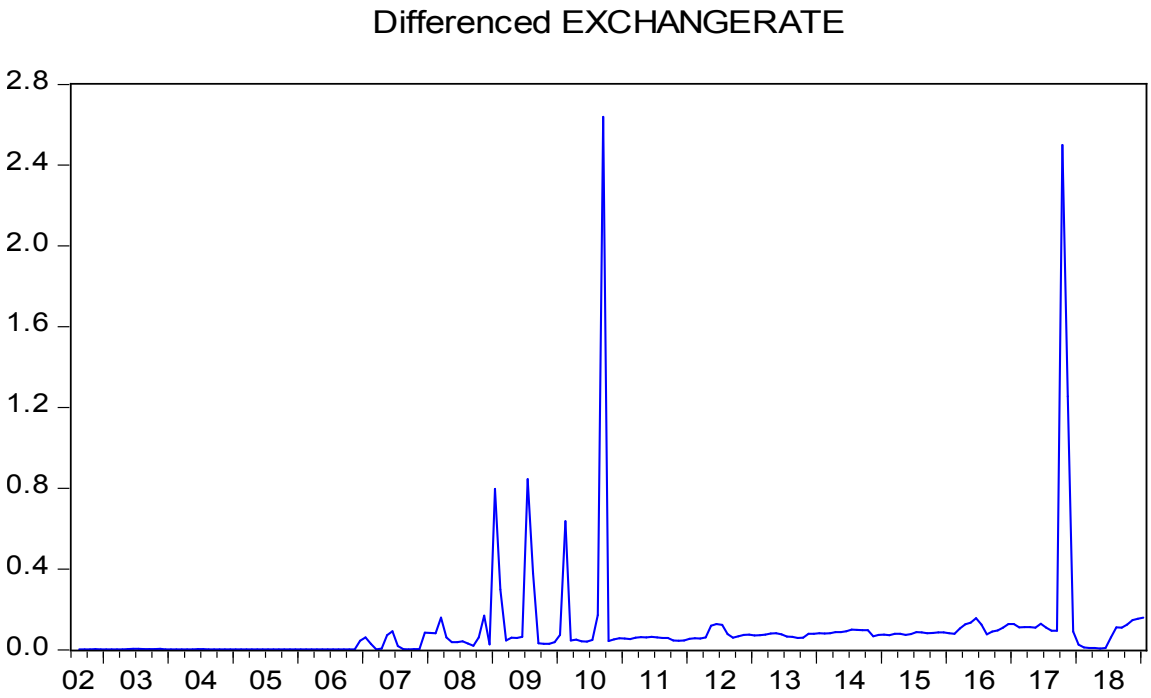
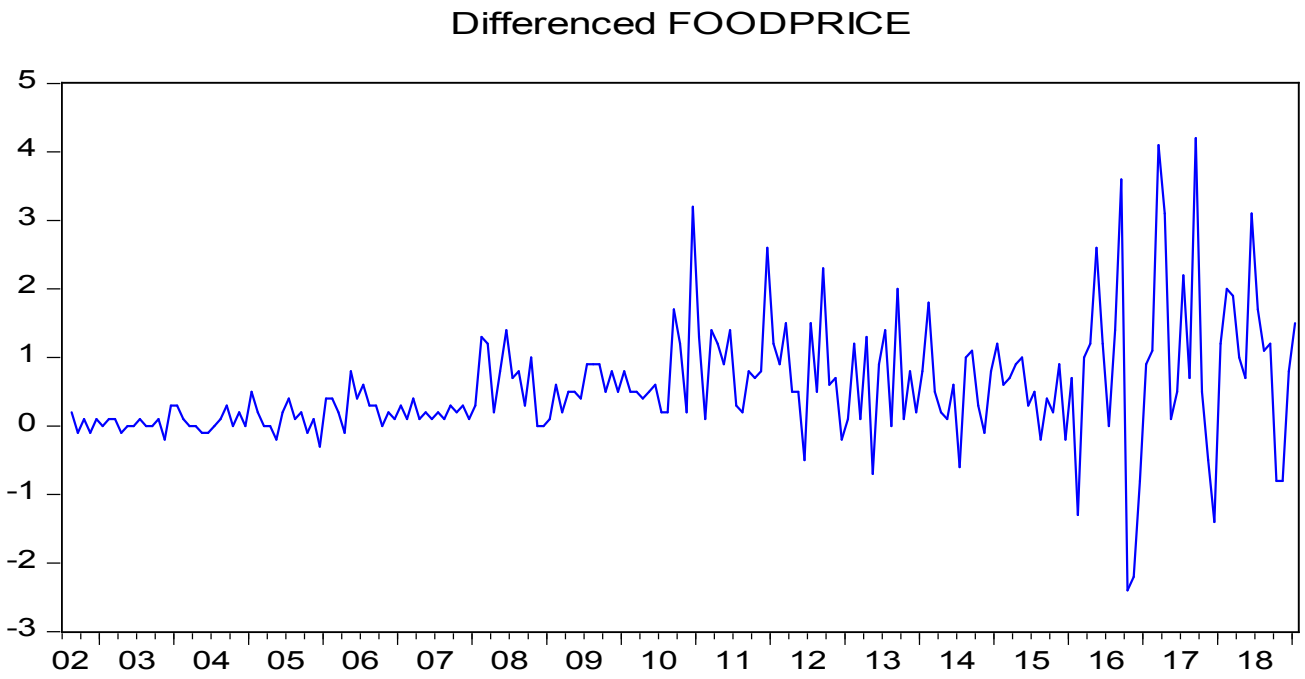


Figure 4.4 time plot of differenced series of exchange rate, food price and non-food price data in Ethiopia.




Differenced FOODPRICE
5
4
3
2
1
0
-1
-2
-3
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18

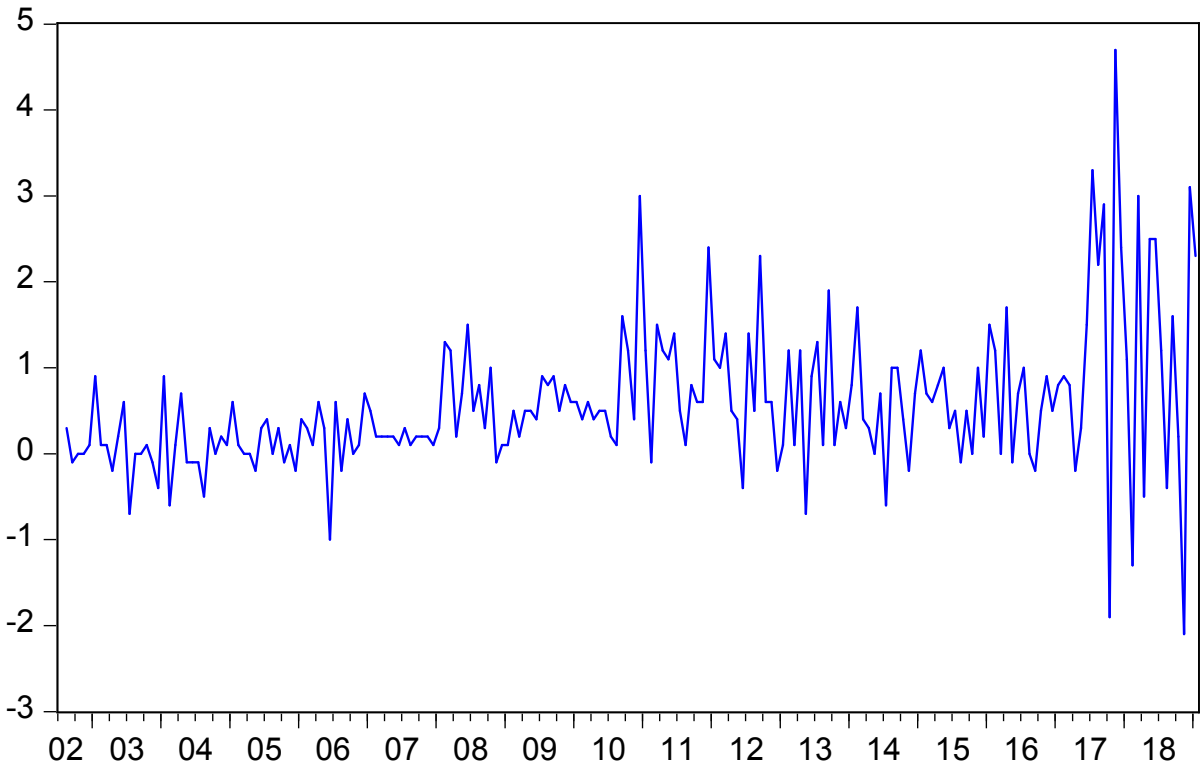

Differenced NONFOODPRICE
5
4
3
2
1
0
-1
-2
-3
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18

Figure 4.5correlogram of monthly export price of coffee at first difference up to 20 lags.

Date: 03/19/19 Time: 10:28

Sample: 2002M07 2016M12

Included observations: 173

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.|* |	.|* |	1	0.205	0.205	7.3717	0.007
.|. |	.|. |	2	0.027	-0.015	7.5019	0.023
*|. |	*|. |	3	-0.071	-0.076	8.3938	0.039
.|. |	.|* |	4	0.055	0.089	8.9291	0.063
.|. |	.|. |	5	0.046	0.020	9.3064	0.097
.|* |	.|* |	6	0.145	0.129	13.122	0.041
.|. |	.|. |	7	0.044	-0.002	13.475	0.061
.|. |	.|. |	8	-0.038	-0.055	13.743	0.089
**|. |	*|. |	9	-0.221	-0.199	22.791	0.007
*|. |	.|. |	10	-0.119	-0.056	25.406	0.005
.|. |	.|. |	11	-0.031	-0.007	25.588	0.007
.|* |	.|* |	12	0.193	0.183	32.598	0.001
.|. |	.|. |	13	0.067	0.018	33.436	0.001
*|. |	*|. |	14	-0.127	-0.142	36.514	0.001
**|. |	*|. |	15	-0.229	-0.131	46.581	0.000
*|. |	*|. |	16	-0.131	-0.070	49.888	0.000
.|. |	.|. |	17	-0.019	0.005	49.958	0.000
.|. |	.|. |	18	0.021	-0.048	50.042	0.000
.|. |	*|. |	19	-0.050	-0.083	50.533	0.000
.|. |	.|. |	20	-0.014	0.048	50.571	0.000

Figure 4.6the residual, actual and fitted plot of ARIMAX (4, 1,4) model of coffee price.

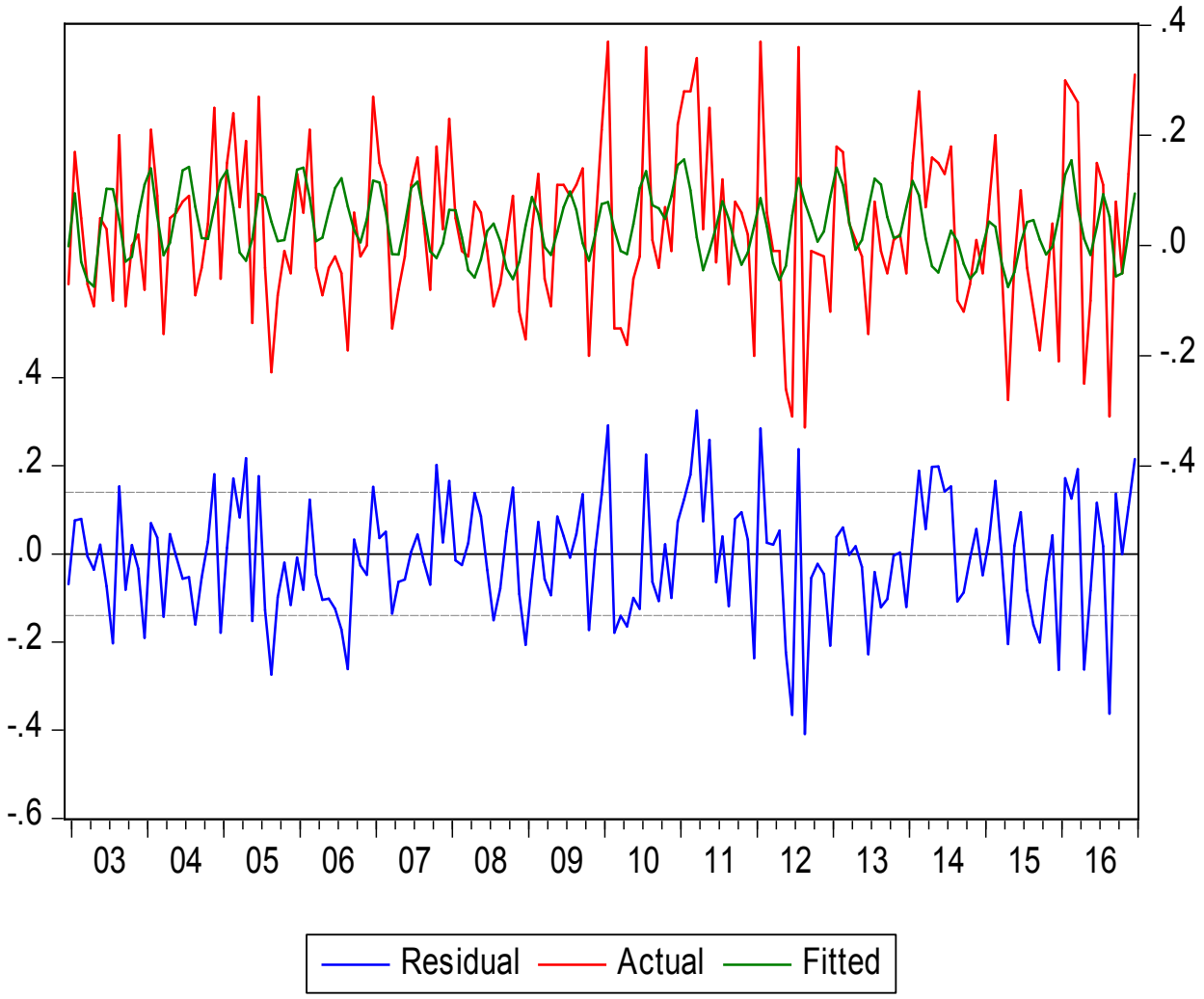


Figure 4.7 autocorrelation (ACF) and partial autocorrelation (PACF) of residuals for ARIMA (4, 1, 4)

Date: 03/21/19 Time: 02:01

Sample: 2002M12 2016M12

Included observations: 169

Q-statistic

probabilities

adjusted for 8
ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
.|* |	.|* |	1	0.123	0.123	2.5915	
.|* |	.|* |	2	0.119	0.105	5.0220	
.|. |	.|. |	3	-0.029	-0.057	5.1705	
.|* |	.|* |	4	0.083	0.083	6.3789	
.|. |	.|. |	5	-0.006	-0.016	6.3854	
.|. |	.|. |	6	0.014	-0.003	6.4216	
.|. |	.|. |	7	0.009	0.019	6.4358	
.|* |	.|* |	8	0.098	0.089	8.1444	
*|. |	*|. |	9	-0.114	-0.142	10.472	0.001
.|. |	.|. |	10	0.007	0.020	10.480	0.005
.|. |	.|. |	11	-0.044	-0.013	10.831	0.013
.|* |	.|. |	12	0.087	0.068	12.216	0.016
.|. |	.|. |	13	-0.026	-0.020	12.342	0.030
.|. |	.|. |	14	0.045	0.032	12.723	0.048
.|. |	.|. |	15	0.062	0.069	13.446	0.062
.|. |	.|. |	16	0.071	0.029	14.407	0.072
.|* |	.|* |	17	0.074	0.084	15.446	0.079
.|. |	.|. |	18	0.012	-0.036	15.475	0.116
*|. |	*|. |	19	-0.079	-0.098	16.671	0.118
.|. |	.|. |	20	-0.001	0.001	16.672	0.162

Figure 4.8 serial correlation Breush- godfrey LM test for residual of ARIMAX (4, 1,4) model.

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.546704	Prob. F(2,155)	0.5800
Obs*R-squared	0.390471	Prob. Chi-Square(2)	0.8226

Test Equation:

Dependent Variable: RESID

Method: Least Squares

Date: 03/21/19 Time: 01:55

Sample: 2002M12 2016M12

Included observations: 169

Presample missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001561	0.036918	0.042283	0.9663
EXCHANGE	-0.000430	0.007564	-0.056807	0.9548
FOOD	0.004609	0.013171	0.349951	0.7269
NONFOOD	-0.004675	0.013387	-0.349243	0.7274
AR(1)	-0.052396	0.358149	-0.146297	0.8839
AR(2)	0.066295	0.676134	0.098050	0.9220
AR(3)	-0.063937	0.674158	-0.094840	0.9246
AR(4)	0.013951	0.318505	0.043800	0.9651
MA(1)	-0.039046	0.384615	-0.101520	0.9193
MA(2)	0.078031	0.766382	0.101818	0.9190
MA(3)	-0.078178	0.759116	-0.102986	0.9181
MA(4)	0.038057	0.372574	0.102146	0.9188
RESID(-1)	0.140792	0.157313	0.894978	0.3722
RESID(-2)	-0.004796	0.136076	-0.035242	0.9719

R-squared	0.002310	Mean dependent var	-0.009226
Adjusted R-squared	-0.081367	S.D. dependent var	0.134576
S.E. of regression	0.139944	Akaike info criterion	-1.015947
Sum squared resid	3.035554	Schwarz criterion	-0.756666
Log likelihood	99.84753	Hannan-Quinn criter.	-0.910726
F-statistic	0.027612	Durbin-Watson stat	1.985456

Prob(F-statistic)	1.000000		

figure 4.9Heteroskedasticity Test of Breusch-Pagan-Godfrey test for estimated ARIMAX (4, 1,4) model.

Heteroskedasticity Test: Breusch-Pagan-Godfrey

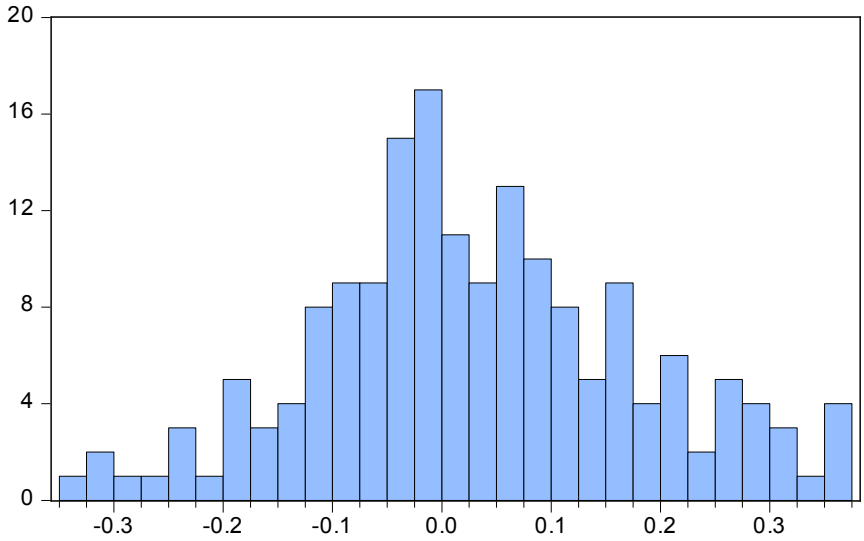
F-statistic	2.400958	Prob. F(3,165)	0.0697
Obs*R-squared	7.068904	Prob. Chi-Square(3)	0.0697
Scaled explained SS	6.373183	Prob. Chi-Square(3)	0.0948

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 03/21/19 Time: 01:53
Sample: 2002M12 2016M12
Included observations: 169

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.007399	0.012648	-0.585035	0.5593
EXCHANGE	0.003524	0.002501	1.409140	0.1607
FOOD	-0.001034	0.002536	-0.407763	0.6840
NONFOOD	0.000599	0.002623	0.228438	0.8196

R-squared	0.041828	Mean dependent var	0.018089
Adjusted R-squared	0.024407	S.D. dependent var	0.026224
S.E. of regression	0.025902	Akaike info criterion	-4.445617
Sum squared resid	0.110700	Schwarz criterion	-4.371537
Log likelihood	379.6546	Hannan-Quinn criter.	-4.415554
F-statistic	2.400958	Durbin-Watson stat	1.817052
Prob(F-statistic)	0.069659		

Figure 4.10Jarkue- Bera normality test for the residuals of ARIMAX (4, 1, 4)



Month	Out of sample Forecast value	Month	Forecast value
FEB 2019	3.35	FEB 2020	2.86
MAR 2019	3.14	MAR 2020	2.66
APR 2019	3.12	APR 2020	2.63
MAY 2019	3.14	MAY 2020	2.65
JUN 2019	2.92	JUN 2020	2.43
JUL 2019	2.26	JUL 2020	2.77
AUG 2019	2.95	AUG 2020	2.46
SEP 2019	2.73	SEP 2020	2.24
OCT 2019	2.51	OCT 2020	2.02
NOV 2019	2.41	NOV 2020	1.92
DEC 2019	2.33	DEC 2020	1.84
JANU 2020	2.53	JANU 2021	2.05
		FEB 2021	2.37

Table4.8. Forecasted values of ARIMAX (4,1,4) models of export price of coffee.

Declaration

We, the undersigned, declare that this thesis is our original work and has not been presented for a degree in any other university, and that all sources of materials used for the paper have been duly acknowledged.

Declared by:

Name: _____

Signature: _____

Date: _____

Name: _____

Signature: _____

Date: _____

Confirmed by Advisor:

Name: _____

Signature: _____

Date: _____