ISSN (e): 2250 – 3005 || Volume, 08 || Issue,11|| November – 2018 ||
International Journal of Computational Engineering Research (IJCER)

Forecasting Analysis of Coffee Export by multivariate Timeseries Models of Vector Autoregressive and Cointegration: A case study of Ethiopia

[1]Gebretsadik Gebru&[2]R.Ramakrishna
[1]Lecture,Department of Statistics,CNCS,Mekelle University,P.O.Box231,Mekelle, Tigria, Ethiopia.
Email:gebretsadik.gebru@mu.edu.et
[2] Associate Professor,Department of Statistics,CNCS,Mekelle University, P.O.Box231,Mekelle, Tigria, Ethiopia.Email:dr.rramakrishna@mu.edu.et
Corresponding Author:GebretsadikGebru

ABSTRACT: Ethiopia is known as the birth place of Coffee Arabica. Coffee has been and remains the leading cash crop and export commodity of Ethiopia. It has accounted on average for about 5% of gross domestic product (GDP), 10% of total agricultural production and 60% of total export earnings for the past three to four decades. Coffee is the most widely consumed stimulant beverage in Ethiopia and about 50% of the total produce is consumed locally. The annual per capital consumption of coffee in Ethiopia is about 2.4 kilograms. This is comparable to the consumption level of the leading coffee consuming countries. There are two major varieties of coffee, namely coffee Arabica and Robusta coffee. Ethiopia produces only Arabica coffee, which is believed to have originated in the rain forests of south western Ethiopia. The general objective of this study is to develop a multivariate time series model which explains the relationship among Monthly export volume of coffee and Free-On-Board (FOB) price of export of coffee, producer price and world price in Ethiopia that can be used for forecasting purpose in planning.

Keywords: Coffee, Multivariate Time series analysis,Volume, Free-On-Board Price, Producer price

Date of Submission: 23-05-2018 Date of acceptance: 31-12-2018

I. INTRODUCTION

Coffee grows in many tropical and sub-tropical part of the world. Latin America, Africa, Asia and Oceania accounts for 60, 30, and 10 percent of the total world coffee production, respectively. About 18 countries around the world derive 25 or more percent of their export earnings either from coffee, tea, or cocoa. Ethiopia is known as the birth place of Coffee Arabica. Coffee has been and remains the leading cash crop and export commodity of Ethiopia. It has accounted on average for about 5% of gross domestic product (GDP), 10% of total agricultural production and 60% of total export earnings for the past three to four decades.

Coffee is one of the highest valued commodities in international trade, with annual export revenues worth around $10 billion on average and annual retail sales of approximately $50 billion. It is a highly labor-intensive industry employing an estimated 100 million people in over 60 developing countries, where it is often a vital source of export revenues and income to producers, many of whom are smallholders. Over a million coffee farming households and about 25% of the total population of the country are dependent on production, processing, distribution, and export of coffee. Coffee is the major agricultural export crop, providing currently 35% of Ethiopia's foreign exchange earnings, down from 65% a decade ago because of the slump in coffee prices since the mid1990. Coffee is the most widely consumed stimulant beverage in Ethiopia and about 50% of the total produce is consumed locally. The annual per capital consumption of coffee in Ethiopia is about 2.4 kilograms. This is comparable to the consumption level of the leading coffee consuming countries. One can bravely say that coffee in Ethiopia is not produced only for export purposes, but also as highly prized and much favored traditional beverages. Given the importance of coffee for environment, economy and culture of the nation, there is a need for research and development to fully exploit the potential that exists in the country. For this, partnership between research organizations and end users like private companies is also of paramount importance.[1][4][12]

This study is concerned with modeling multivariate time series data using VAR and co-integration analysis, which consists of simultaneous observations on four related variables of interest. Our main variables of study are volume, FOB price, producer price, and world price of export of coffee Arabic in Ethiopia. We analyze data and develop a multivariate time series (MVTS) model which can adequately describe the innate relationship among the four study variables.Several studies about coffee export and related variables are done utilizing univariate time series analysis. Univariate time series analysis is important but it is inadequate for the analysis of interaction and co-movement of several time series simultaneously. In contrast, MVTS analysis involves a vector of time series that will be modeled simultaneously. MVTS deals with the interaction, co-movement and bi-directional causality of several time series. This study will examine the different statistical techniques for analyzing multivariate time series data which consists of monthly volume of export of coffee, FOB price of coffee export, producer price, and world price of coffee. In the Ethiopian coffee export sector it is clear that studying the relationship among the above variables is very important to improve the quality and quantity of coffee export. The general objective of this study is to develop a multivariate time series model which explains the relationship among volume and FOB price of export of coffee, producer price and world price in Ethiopia that can be used for forecasting purpose in planning.[5][8][9][10]

The study is organized into four chapters. Following the introductory chapter one, chapter twodiscusses the methodology and sources of data used in the study. Chapter three is followed bymodel estimation and interpretation of results and forecasting. Finally, chapter four presents discussion and recommendations of the study.

II. DATA AND METHODOLOGY

2.1 Data

Our data for the study consists of four variables collected based on a fixed interval of time period (monthly). A brief description of each is presented below.

i. Volume (net weight) of Export of Coffee (VOL): is the sum of net weight of all coffee Arabic that is exported monthly to the destination countries. The source of the data is the Ethiopian Revenue and Customs Authority (ERCA) planning, monitoring and evaluation section.

ii. FOB Price of Coffee Export (FO*B):* FOB price here refers to the price of total volume (net weight) of monthly export of coffee which includes price of coffee, cost of transportation to port, plus cost of loading onto ship. The unit of measurement is USD per kilogram and the source of data is the ERCA planning, monitoring and evaluation section.

iii. Producer Price (PP): this is the price at which the producers (owners) sold coffee to the exporters. The unit of measurement is USD per kilogram and the source of data is the Central Statistical Agency (CSA).

iv. World Coffee Price (WOP): This is the monthly price of coffee over the world and its unit of measurement is USD per kilogram. The source of data is the International Coffee Organization (ICO) website.

All the data obtained are collected on monthly basis from September 2006 to July 2011.

2.2 Methodology

Time series is broadly defined as series of measurements taken sequentially across time. It can be divided in to two major parts: univariate and multivariate time series. Univariate time series uses only the past history of the time series being forecast plus current and past random error terms. 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. The methodology adopted in this study follows the vector autoregressive (VAR) model and vector error correction model (VECM).

2.2.1 Vector Autoregressive (VAR) Models

The VAR model is one of the most successful, flexible and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model.

2.2.2 The Stationary Vector Autoregressive Model

Let $Y_t = (y_{1t}, y_{2t}, \ldots, y_{nt})^T$ denotes an$(n \times 1)$vector of time series variables. The basic p - lag vector autoregressive VAR (p) model has the form:

$Y_t = c + \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \ldots + \pi_p Y_{t-p} + \varepsilon_t$, $t = 1, 2, \ldots$ T ....(2.1)

where $c$denotes an $(n \times 1)$ vector of constants and π_j an $(n \times n)$matrix of autoregressive coefficients, $j = 1, 2, \ldots, p$ and ε_t is an $(n \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated) with time invariant covariance matrix Σ:

The general form of the VAR (p) model with deterministic terms and exogenous variables is given by

$Y_t = \pi_1 Y_{t-1} + \pi_2 Y_{t-2} + \ldots + \pi_p Y_{t-p} + \Phi D_t + G X_t + \varepsilon_t$ ....(2.2)

whereD_t represents an $(I \times 1)$matrix of deterministic components, X_t represents an $(n \times 1)$vector of exogenous variables and $\Phi$and$G$ are parameter matrices.

2.2.3 Testing Stationarity: Unit-Root test

The assumption of stationary is somewhat unrealistic situation in most macroeconomic variables.
The most popular ones are Augmented Dickey- Fuller (ADF) test due to Dickey and Fuller (1979, 1981), and the Phillip-Perron (PP) test due to Phillips (1987) and Phillips and Perron (1988). The following discussion outlines the basic features of unit root tests (Hamilton, 1994).
Consider an AR (1) process:

$Y_t = \rho Y_{t-1} + X_t^{'}\delta + \varepsilon_t$ ....(2.3)

where $X_t$are optional exogenous regressors which may consist of constant or a constant and trend, $\rho$and $\delta$are parameters to be estimated and ε_t is assumed to be white noise.

If $|\rho|\geq 1$, Y is a non-stationary series and the variance of Y increases with time and approaches infinity. On the other hand, if $|\rho|<1$, Y is a stationary series. Thus, the hypothesis of (trend) stationarity can be evaluated by testing whether the absolute value of ρ is strictly less than one. The hypotheses are:
H_0: The series are not stationary (ρ=1)
H_1: The series are stationary (ρ<1)

2.2.3.1 Augmented Dickey-Fuller (ADF) Unit-Root Test

The standard Dickey-Fuller test is conducted in the following manner: from equation (2.3) we have:
$Y_t - Y_{t-1} = (\rho - 1)Y_{t-1} + x_t^{'}\delta + \varepsilon_t$. This implies that

$\Delta Y_t = \alpha Y_{t-1} + X_t^{'}\delta + \varepsilon_t$ ....(2.4)

where $\alpha = \rho - 1$. The null and alternative hypothesis may be written as:
$H_0 : \alpha = 0$
$H_1 : \alpha < 0$..........(2.5)

The test statistic is the conventional $t - ratio$for α:

$t_\alpha = \hat{\alpha}/_{(se(\hat{\alpha}))}$ ....(2.6)

where $\hat{\alpha}$ is the estimate of α and $se(\hat{\alpha})$is the standard error of $\hat{\alpha}$.

2.2.3.2 Phillips-Perron (PP) Unit-Root Test

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 autoregression 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

$\nabla Y_t = a_0 + a_2 t + \pi Y_{t-1} + \varepsilon_t.$ ....(2.7)

where $\varepsilon_t$is $I(0)$ and may be heteroskedastic. The PP tests correct for any serial correlation and heteroskedasticity in the errors $\varepsilon_t$of the test regression by directly modifying the Dicky-Fuller test statistics $t_{\pi=0}$and $T_{\pi=\hat{\pi}}$ where

$$t_{\pi=\hat{\pi}} = \hat{\pi}/_{(se(\hat{\pi}))}$$

where $\hat{\pi}$ is the estimate of π and $se(\hat{\pi})$is the standard error of $\hat{\pi}$.[2][3]

2.2.4 Estimation of the Order of the VAR

The lag length for the $VAR(p)$ model may be determined using model selection criteria. The general approach is to fit $VAR(p)$ models with orders $p = 0, \ . \ . \ . \ , \ p_{max}$ and choose the value of p which minimizes some model selection criteria. Model selection criteria for VAR (p) models have the form:

$IC(p) = ln|\hat{\Sigma}_p| + C_T \, . \varphi(n.p)$ ....(2.8)

where
IC = Information Criteria, $\hat{\Sigma}_p = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t^{'}$ is the residual covariance matrix from a VAR (p) model, C_T is a sequence indexed by the sample size T, and $\varphi(n.p)$ is a penalty function which penalizes large $VAR\ (p)$ models.
The three most common information criteria to determine the order of VAR models are the Akaike (AIC), Schwarz – Bayesian (BIC) and Hannan – Quinn (HQ):

$AIC(p) = ln|\hat{\Sigma}_p| + \frac{2}{T}pn^2$ ....(2.9)

$BIC(p) = ln|\hat{\Sigma}_p| + \frac{\ln T}{T}pn^2$ ....(2.10)

$HQ(p) = ln|\hat{\Sigma}_p| + \frac{2\ln\ln T}{T}pn^2$ ....(2.11)

The AIC criterion asymptotically overestimates the order with positive probability (not zero), where as the BIC and HQ criteria estimate the order consistently under fairly general conditions if the true order p is less than or equal to p_{max} . For a model to be best it should have the smallest information criteria.

2.2.5 Co-integration Analysis
2.2.5.1 Co integration
If two or more series are individually integrated (i.e. in the time series sense) but some linear combination of them has a lower order of integration, then the series are said to be co-integrated. The three main methods for testing co-integration are:
1. The Engle-Granger two-step method
2. The Johansen procedure and
3. Phillips-Ouliaris Cointegration Test
In practice, co integration is used for such series of integrated I (1) in typical econometric tests, but it is more generally applicable and can be used for variables integrated of higher order to detect correlated accelerations or other second-difference effects. The procedure begins with unrestricted VAR involving potentially non stationary variables. The $VAR\ (p)$ model can be re-written into VECM form as:
$\Delta Y_t = \pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t$ t = 1, 2, …, T ....(2.12)
where π and the short-run parameter Γ_i, i= 1, 2, …, p-1 are $p \times p$matrices of coefficients.[6]

Testing for co integration using Johansen's methodology
The starting point in Johansen's procedure (1988, 1991) in determining the number of co integrating vectors is the VAR representation of Y_t. It assumes a vector autoregressive model of order p and is expressed as follows:
$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \ldots + A_p Y_{t-p} + BX_t + \varepsilon_t$ ....(2.13)
where $Y_t$is a p-vector of non-stationary I(1) variables, $X_t$is a d vector of deterministic variables and ε_t is a vector of innovations.

2.2.6 Vector Error Correction Modeling (VECM)
If such a stationary or I(0) linear combination exists, the non-stationary (with a unit root), time series are said to be co integrated. When the variables are co integrated, the corresponding error correction representations must be included in the system.
$\Delta Y_t = \pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_\text{i} \Delta Y_{t-i} + BX_t + \varepsilon_t$..(2.14)
where
$\pi = -I_n + \sum_{i=1}^{p} A_i,\ \Gamma_\text{i} = -\sum_{j=i+1}^{p} A_j$and I_n is an identity matrix.
The above specification of VECM contains information on both the short and the long-run adjustment to changes in $y_t$via estimating Γ and Π, respectively. Matrix Π can be decomposedas $\Pi = \alpha\beta'$, where α is $n \times r$ matrix of speed of adjustments, and β is an $n \times r$ matrix of parameters which determines the co-integrating relationships matrix of long- run coefficients such that $\beta' y_{t-k}$represent the multiple co-integration relationships.

2.2.7 Model Checking
A wide range of procedures is available for checking the adequacy of VAR and VECMs. They should be applied before a model is used for specific purpose to ensure that it represents the data adequately.

2.2.7.1 Test of Residual Autocorrelation
Two types of tests for residual AC are quite popular in applied work, Breusch-Godfrey LM tests and portmanteau tests. They are both based on statistics of the form
$Q = T\hat{c}'\hat{\Sigma}^{-1}\hat{c}$………………………………………………………………… (2.15)
where$\hat{\Sigma}$is a suitable scaling matrix. In other words, they are based on the residual auto covariance's. The choice of scaling matrix $\hat{\Sigma}$determines the type of test statistic and its asymptotic distribution under the null hypothesis of no residual AC. We will consider both types of tests in turn.

Portmanteau autocorrelation test
Suppose $y_t = (y_{1t}, \ldots, y_{kt})'$ is k-dimensional vector of observable time series variables with r<k co-integration relations. From equation (2.15) the residual auto covariance is
$\hat{C}_j = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_{t-j}$ ………………………………………………………………….(2.16)
where $\hat{\varepsilon} = \Delta y_t - \pi Y_{t-1} - \sum_{i=1}^{p-1} \Gamma_\text{i} \Delta Y_{t-i} - BX_t$ from VECM form.

Autocorrelation LM Test
This test was developed by Breusch and Godfrey in 1978. Assume a VAR model for the error
ε_t given by
$\varepsilon_t = D_1 u_{t-1} + \ldots + D_h u_{t-h} + \upsilon_t$ ....(2.17)
The quantity $\upsilon_t$denotes a white nose error term. Thus, to test autocorrelation in $u_t$we test

H_0: $D_1 = \cdots = D_h = 0$ against
H_1: $D_j \neq 0$ for at least one j < h
We use the LaGrange multiplier method to perform the test. The Lagrange Multiplier (LM) test for p^{th} order serial correlation is computed first by estimating an auxiliary regression where the OLS residuals are regressed on the variables in the original model plus *p* lagged residuals.

2.2.7.2 Normality of the Residuals
Normality tests whether the residuals of the regression are normally distributed or not. The null hypothesis is that the residuals are normally distributed. Several tests for normality are available but the most commonly used test for normality of regression disturbances is due to Jarque and Bera (1980). The JB test statistic is:
$JB = T\left[\frac{\hat{b}_1}{6} + \frac{\hat{k}^2}{24}\right]$ …………………………………………………………………… (2.18)
where $\hat{b}_1$ and $\hat{k}$ are the sample skewness and kurtosis coefficients, respectively.
This test statistic is asymptotically distributed as $\chi^2(2)$under the null hypothesis; thus large values of this test statistic relative to the quantiles from the $\chi^2(2)$ distribution lead to rejection of the null hypothesis.[2][3][6]

2.2.8 Forecasting
Forecasting is one of the main objectives of multivariate time series analysis. Forecasting from a VAR model is similar to forecasting from a univariate AR model and the following gives a brief description. Consider first the problem of forecasting future values of $Y_t$when the parameters Π of the VAR (p) process are assumed to be known and there are no deterministic terms or exogenousvariables. The best linear predictor in terms of minimum means squared error (MSE) of Y_{t+1}or 1-step forecast based on information available at time T is:
$Y_{T+1/T} = C + \Pi_1 Y_T + \Pi_2 Y_{T-1} + \cdots + \Pi_p Y_{T-p+1}$ ....(2.19)
for $T \geq p.$

2.2.9. Measures of Forecasting Accuracy
In most forecasting situations, accuracy is treated as the overriding criterion for selecting a forecasting method. To the consumer of forecasts, it is the accuracy of the future forecast that is most important.
If Y_{jt}, j=1, 2, . . ., k is the actual observation for the period t and F_{jt} is the forecast of Y_{jt}, then the residual is defined as:
$\hat{\varepsilon}_{jt} = Y_{jt} - F_{jt}$ ....(2.20)
Usually F_{jt} is calculated using dataY_{j1} ,Y_{j2} , ... Y_{jt-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 $\hat{\varepsilon}_t$as a one step forecast error. It is the difference between the observation Y_{jt} and forecast made using all observations up to but not including Y_{jt}.
If there are observations and forecasts for T time periods, then there will be T error terms, and the following standard statistical measures can be defined:
$Mean\ Error\ (ME) = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{jt}$ .... (2.21)
$Mean\ Absolute\ Error\ (MAE) = \frac{1}{T}\sum_{t=1}^{T}|\hat{\varepsilon}_{jt}|$..(2.22)
$Mean\ Squared\ Error\ (MSE) = \frac{1}{T}\sum_{t=1}^{T} \hat{\varepsilon}_{jt}^2$ ....(2.23)
First let us define a relative or percentage error as
$PE_t = \left(\frac{Y_{jt} - F_{jt}}{Y_{jt}}\right) \times 100$ ....(2.24)
Then the following two relative measures are frequently used:
$Mean\ Percentage\ Error\ (MPE) = \frac{1}{T}\sum_{t=1}^{T} PE_t$...(2.25)
$Mean\ Percentage\ Absolute\ Error\ (MPAE) = \frac{1}{T}\sum_{t=1}^{T}|PE_t|$...(2.26)
Alternatively, Theil's U statistic can be used as a measure of forecasting accuracy. Like
MAPE statistic, high values suggest poor performance in the forecast. Theil's U can be estimated as:
$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(Y_{jt} - F_{jt})^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n} F_{jt}^{\ 2}} + \sqrt{\frac{1}{n}\sum_{t=1}^{n} Y_{jt}^{\ 2}}} \quad (2.27)$$
The scaling of U is such that it will always lie between 0 and 1. If U = 0, $Y_{jt} = F_{jt}$ for all forecasts and there is a perfect fit; if U = 1 the predictive performance is not good.[11]

2.2.10Structural Vector Autoregressive (SVAR) Analysis
The general VAR (p) model has many parameters and they may be difficult to interpret due to complex interactions and feedback between the variables in the model. As a result, the

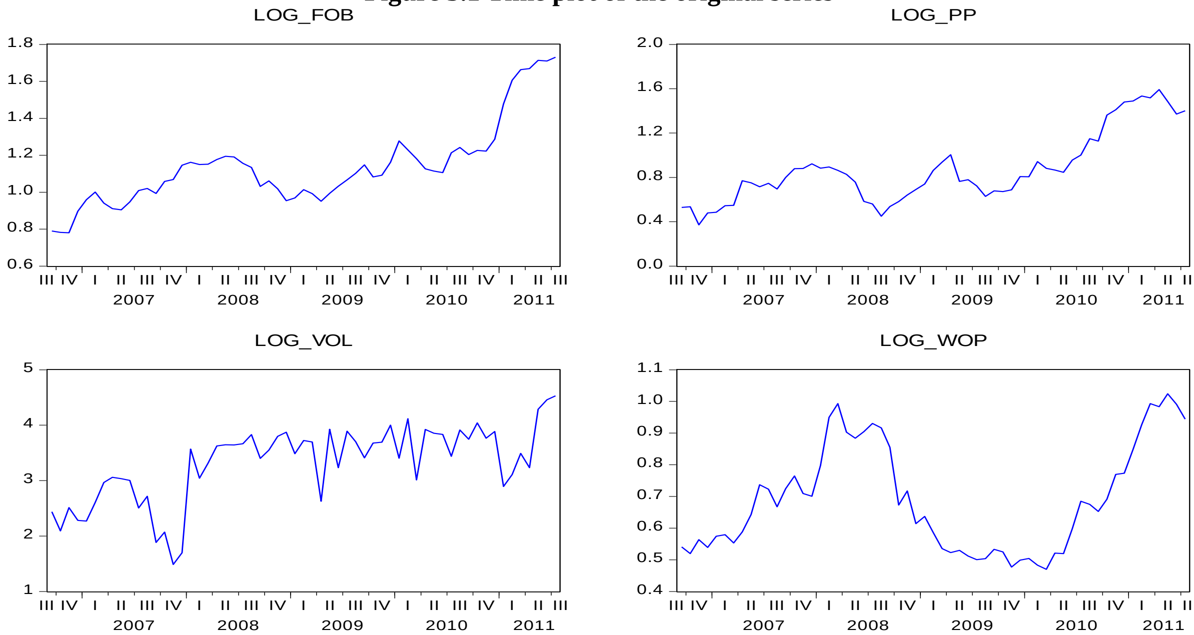
bydynamic properties of a VAR (p) are often summarized using various types of structural analysis. , which are granger causality test and impulse response function.

III. RESULTS

3.1 Descriptive Analysis and Time plot

EViews 7, the windows-based forecasting and econometric analysis package, was used to estimate the relationship among the volume of coffee export (VOL), free-on-board price (FOB), producer price (PP) and world price (WOP) in the case of Ethiopia. Our study data consists of monthly volume (net weight) of coffee export (in million of kilograms), monthly free-on-board price (in USD per kilogram), monthly producer price (in USD per kilogram) and monthly world price of coffee (in USD per kilogram). The time period covered is from September 2006 to July 2011. The time plot of each of the series is shown in Figure 3.1 below. From the time plot we can observe that all the series except world price show an increasing trend over the study period. World price of coffee has declined in 2008-2009 and rises up then after.

Figure 3.1 Time plot of the original series



The ADF and PP test shows that all series are non-stationary in levels and stationary in the first differences.[6]

3.2. VAR Model Specification

3.2.1. Estimating for Order of the VAR

Specifying the lag length has strong implications for subsequent modeling choices. For determining the appropriate lag length for the VAR model the Akaike information criterion (AIC), Schwarz information criterion (SC) and Hannan-Quin (HQ) information criteria were used. The results are shown in Table 3.1.
The AIC, SC and HQ tests suggest that the appropriate lag length for the VAR model is one (1).We specify the VAR as a four variable system for a sample period from September 2006 to July 2011.

Table 3.1: VAR lag order selection results

Lag	AIC	SC	HQ
1	-6.611245	-5.874584	-6.327144
2	-6.432789	-5.106799	-5.921407
3	-6.383314	-4.467996	-5.644651

From the above table we can observe that VAR (1) is the best since it has the minimum AIC, SC and HQ. [6]

3.2.2 Lag exclusion test

To check whether the chosen lag is optimal, Wald lag exclusion test is used. Given that VAR modeling requires uniform lag length for each variable, the result in Table 3.2 shows that the first lag is significant for all variables at the one percent level of significance. Therefore; VAR (1) is found suitable for the data set and hence could be adopted.

Table 3.2: VAR Lag Exclusion Wald Tests

	LOG_FOB	LOG_PP	LOG_VOL	LOG_WOP	Joint
Lag 1	1105.695	656.2622	66.78052	541.9360	2409.610
	[0.000000]	[0.000000]	[1.08e-13]	[0.000000]	[0.000000]
Df	4	4	4	4	16

Chi-squared test statistics for lag exclusion: numbers in [ ] are p-values

The results of the estimated VAR model are presented are statistically significant at the 5% level of significance. By using vector Auto regression estimates, we identified that free-on-board price is significantly explained by its own past and by producer price lagged by one period. This implies that a one dollar increase in a onetime lagged producer price leads to an increase of free-on-board price by an amount of $ 0.13. Volume of coffee export is significantly explained by its own past only. This indicates that Ethiopian coffee export has no significance relationship with producer price, free-on-board price and world price. World price and producer price are also significantly explained by their own past. [6]

3.2.3 Co integration analysis

Since the variables are integrated of order one, we proceed to test for co-integration. Johansen (1995) cointegration test is applied at the predetermined lag 1. In these tests, Trace statistic and Maximum Eigenvalue statistic are compared to special critical values. The maximum eigenvalue and trace tests proceed sequentially from the first hypothesis –no cointegration– to an increasing number of cointegrating vectors.

The results of cointegration tests forLog_FOB, Log_PP, Log_VOL and Log_WOP are reported in Table 3.3. The trace statistic indicates that there is one cointegrating vector in the system at the 95 percent confidence level (estimated LR statistic,50.69> 47.86 at 95 percent critical value).

Table 3.3: Johansen Cointegration test results (assumption: linear deterministic trend)

Number of Cointegrating Vector	Eigenvalue	Trace Test			Maximum Eigenvalue Test		
		Statistic	0.05 Critical Value	Prob.**	Statistic	0.05 Critical Value	Prob.**
None *	0.374670	50.69621	47.85613	0.0264	26.29067	27.58434	0.0725
At most 1	0.284347	24.40554	29.79707	0.1838	18.73533	21.13162	0.1048
At most 2	0.094000	5.670211	15.49471	0.7341	5.528069	14.26460	0.6742
At most 3	0.002535	0.142141	3.841466	0.7062	0.142141	3.841466	0.7062

Normalized cointegrating coefficients (standard error in () and t-statistic in [])

FOB	PP	VOL	WOP
1.000000	-0.917434*	-0.013367*	-0.189454*
	(0.14929)(0.00566)(0.07642)		
	[-6.14537]	[-2.36183]	[-2.47911]

* denotes rejection of the hypothesis at the 0.05 level

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

The main purpose of cointegration analysis is to get a stationary series from two or more non-stationary series. The resulting stationary series is written as a linear combination of the non-stationary series under study. In our case, we found out that there is one stationary cointegrated series from the four non-stationary series. If we denote this stationary series by Z then using the results obtained from Table 3.3 we have the following.

$$Z_t = logFOB_t - 0.91743logPP_t - 0.01337logVOL_t - 0.18945\log WOP_t - 0.654292$$

(0.000) (0.023) (0.017)

The result tells us that Z is stationary despite the fact that all the four series are non-stationary. Since all of the variables are significant at the conventional significance levels, we can infer from this result that there exist long-run causal relationships among FOB, PP, VOL and WOP.This long-run model is:

$$logFOB_t = 0.654292 + 0.91743logPP_t + 0.01337logVOL_t + 0.18945\log WOP_t$$

(6.14537) (2.36183) (2.47911)

From the long run equation above the value 0.92indicates that a one dollar increase in producer price induces, on average, an increase of about $ 0.92 in free-on-board price in the long-run. Similarly, a one dollar increase in world price leads to increase by about $ 0.19 in the free-on-board price. On the other hand, a one kilogram increase in volume of coffee export induces on average an increase of about $ 0.01 in free-on-board price in the long-run.[6]

3.3. Model Estimation

The responses of Log_FOB, Log_PP, Log_VOL and Log_WOP to short-term output movements are captured by the $\Gamma_i$coefficient matrices. The αcoefficient vector reveals the speed of adjustment to theequilibrium which measures the deviation from the long-run relationship among the price-volume relationship of coffee export.

Table 3.4: Vector Error Correction Estimates

Standard errors in ( ) & t-statistics in [ ]

Error Correction	D(LOG_FOB)	D(LOG_PP)	D(LOG_VOL)	D(LOG_WOP)
CointEQ1	-0.147401* (0.04205) [-3.50566]	-0.12819* (0.06309) [2.03186]	0.201158 (0.43728) [0.46003]	-0.126959* (0.04733) [-2.68258]
D(LOG_FOB(-1))	0.337402* (0.11963) [2.82038]	0.042776 (0.23461) [0.18094]	0.742911 (1.24410) [0.59715]	0.159835 (0.13465) [1.18704]
D(LOG_PP(-1))	-0.127559 (0.08182) [-1.66107]	0.005829 (0.16169) [0.03605]	-0.315149 (0.85088) [-0.37038]	-0.056057 (0.09209) [-0.60870]
D(LOG_VOL(-1))	-0.010053 (0.01207) [-0.83282]	-0.001074 (0.02386) [-0.04502]	-0.501495* (0.12554) [-3.99467]	0.011959 (0.01359) [0.88017]
D(LOG_WOP(-1))	0.08227 (0.11843) [0.69417]	-0.025611 (0.23404) [-0.10943]	0.027347 (1.23161) [0.02220]	0.064196 (0.13330) [0.48160]
C	0.014560 (0.00669) [2.17643]	0.014632 (0.01322) [1.10679	0.052970 (0.06957) [0.76137]	0.004765 (0.00753) [0.63280]

The coefficients from the Table 3.4, measures the short-run adjustments of the deviations of the endogenous variables from their long- run values.Therefore, we can see that free-on-board price is significantly affected by its lagged value in the short-run. On the other hand, the insignificant coefficient of world price implies there is no price transmission from world price to free-on-board price. Furthermore, 14.74% of the short run disequilibria in FOB is adjusted within one month.Similarly,12.82% of the short run disequilibria in producer price is adjusted within one month. On the other hand, volume of coffee export is significantly affected by its lagged value in theshort run.World price is affected by neither of FOB, PP and VOL in the short runand 12.69% of its short run disequilibria is adjusted within one month.[6]

3.4 Structural Analysis

3.4.1 Granger Causality Test and Impulse-Response Functions

Granger causality test is considered a useful technique for determining whether one time series is good for forecasting the other. The pair wise Granger-causality tests shows that producer price granger causes free-on-board price. This indicates that, the change in producer price leads to change in the free-on-board price. That is, producer price provides important information to forecast future value of the free-on-board price. All the other pairs do not granger cause each other. For example, world price does not granger cause free-on-board price and producer price. This is an indication that there is no transmission of price signals from the world market to the local market.

The impulse response function indicates that free-on-board price innovations have a positive impact on producer price. This implies producer price positively affects free-on-board price. It exhibits a rising trend initially and reaches 0.07 and it stabilizes at around 10 month time horizon. Furthermore, VOL and WOP are almost not affected by one SD change.

Similarly, a one standard deviation shock applied for producer price has a positive impact on free-on-board price. Impulse responses for volume of coffee export have initially negative effect on free-on-board price and then have positive effect after around 2 month time horizon. It has also initially negative effect on producer price and then has positive effect after 6 month time horizon. Furthermore, it has negative effect on world price. Finally, world price innovation has a negative effect on free-on-board price and volume of coffee export, and has a positive effect on producer price.

3.5 Results from Diagnostic Tests

Table 3.5: Results from the Diagnostic Tests

Test	F-statistic	Probability
1. Normality Jarque-Bera Statistic	2.373	0.305
2. Serial correlation Breusch-Godfrey serial correlation LM test	2.377	0.129
3. Autoregressive conditional heteroscedasticity	0.05	0.975

ARCH LM test		
4. Heteroscedasticity White heteroscedasticity test	1.998	0.0575
5. Stability Chow forecast test	1.938	0.126
6. Specification error Ramsey RESET test	0.647	0.425

From the Table 3.5, we can see that the model was tested for serial correlation, autoregressive conditional heteroscedasticity, specification error and stability. The results indicate that the model is well specified.[6]

3.6. Forecasting

One of the fundamental applications of time series analysis or developing a time series model is forecasting. The previous discussion confirms that VAR (1) model is a good model to describe the series. In this section we examine the forecasting accuracy of the fitted model and then make a forecast for August 2011 to July 2012.

3.6.1. Evaluation of forecast accuracy

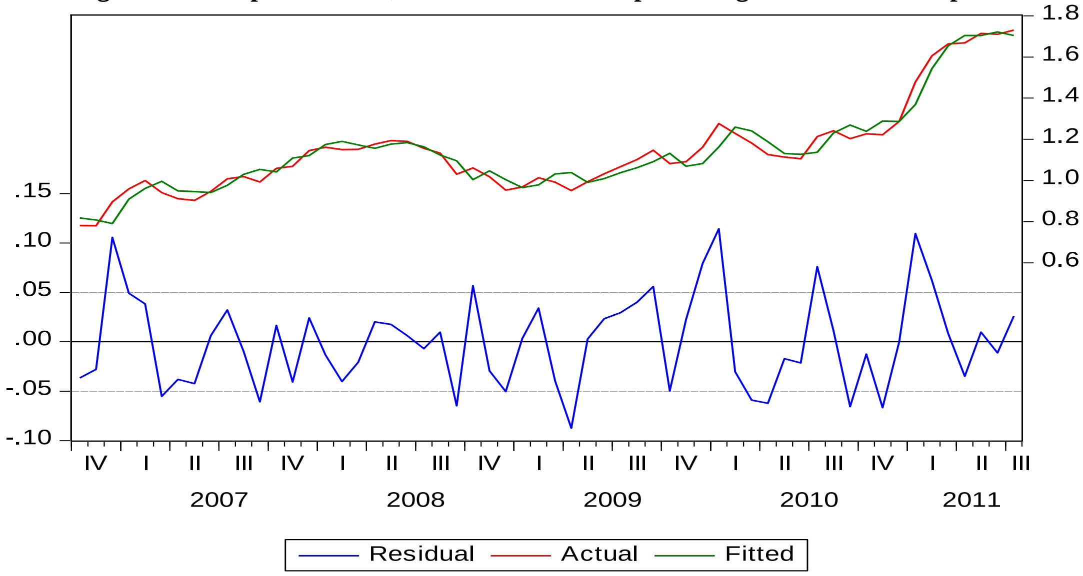
The mean square error (MSE), root mean square error (RMSE), mean absolute error (MAE) and Theil U statistics were used to assess the forecasting performance. The RMSE and MAE statistics are scale-dependent measures but allow a comparison between the actual and forecast values. The Theil-U statistics is independent of the scale of the variables and is constructed to lie between zero and one, zero indicating a perfect fit. Table 3.6 reports the forecasting accuracy statistics of the estimated model.

Table 3.6: Forecasting Accuracy statistic

Accuracy measure	Variables			
	Log_FOB	Log_PP	Log_VOL	Log_WOP
Root Mean Squared Error	0.070	0.174	1.479	0.099
Mean Absolute Error	0.058	0.149	0.362	0.078
Mean Absolute percent error	3.031	3.256	3.651	2.216
Theil Inequality Coefficient	0.011	0.019	0.023	0.013

For the VAR (1) model, the MAPE in forecasting Log_FOB, Log_PP, Log_VOL and Log_WOP are 3.03, 3.26, 3.65 and 2.22, respectively. These computed values show that the average percentage error for each of the equations used to forecast the study variables is less than 4%. TheTheil-U statistics is relatively close to zero, which indicates that the difference between the actual value and the predicted value is very small. The graph of the predicted values together with the actual observationsfor Log_FOB is given in figure 3.2 below.[6]

Figure 3.2: Graph of Actual, Fitted and Residual plot of log of free-on-board price



3.6.2. Out of sample forecasting analysis

Out of sample forecasted values for the series under study, using the vector autoregressive model, are presented in table 3.7 below.

Table 3.7: Forecasts from the VAR (1) models

Months	Log_FOB	Log_PP	Log_VOL	Log_WOP
Aug-11	5.482	4.629	66.874	2.577
Sep-11	5.66	4.754	78.275	2.644
Oct-11	5.771	4.8876	83.304	2.683
Nov-11	5.895	4.999	92.0698	2.697
Dec-11	6.028	5.119	93.012	2.752
Jan-12	6.108	5.237	94.817	2.809
Feb-12	6.272	5.352	95.982	2.866
Mar-12	6.420	5.465	98.523	2.923
Apr-12	6.569	5.577	99.396	3.07
May-12	6.718	5.676	101.707	3.115
Jun-12	6.893	5.776	102.688	3.154
Jul-12	7.032	5.878	103.536	3.195

The result indicates that the free-on-board price and producer price have high increasing trend. However, the volume of coffee export and world price exhibit slow increment rates.[6]

IV. DISCUSSION

The objective of this paper was to forecasting and multivariate time series analysis to price-volume relationship of coffee export in Ethiopia using monthly data ranging from September 2006 to July 2011. Over the time period considered, all the four series have an increasing pattern, that is, there is a sign of non-stationarity in each of the series. Formally, the data were tested for stationarity and all the four series were found to be non-stationary using Augmented Dickey-Fuller and Phillips-Perron unit root tests. Appropriate differencing made the series stationary.

Different vector autoregressive models were tested using AIC, SC and HQ information criteria to fit the series. Among all candidate VAR models, VAR (1) was found to be the best to describe the data. Error diagnosis of this model showed that the disturbance terms are white noise and normally distributed. This model expressed each variable under study as a function of its lag and the lag of other variables. The granger causality test tells us that producer price granger cause's free-on-board price. That is, producer price provides important information to forecast future value of the free-on-board price. Furthermore, producer price does not granger causes world price and vice versa.

The VAR (1) model analysis result shows that free-on-board price is significantly explained by its own past and by lagged value of producer price. This implies that a one dollar increase in producer price leads to an increase of free-on-board price by an amount of $ 0.92. Similarly, volume of coffee export, producer price and world price are significantly explained by their own past values.

From the cointegration analysis result, the trace statistic indicates that there is one cointegrating vector in the system at the 95 percent confidence level. We can infer from this result that there exist long-run causal relationships among free-on-board price, producer price, volume of coffee export and world price.From the VEC model we can observe that free-on-board price is significantly affected by its lagged value in the short-run.Furthermore, 14.74% and 12.825% of the short run disequilibria in free-on-board and producer price is adjusted within one month, respectively.Similarly, volume of coffee export is significantly affected by its lagged value in the short run. World price is affected by neither of free-on-board price, producer price and volume of coffee export in the short run and 12.69% of its short run disequilibria is adjusted within one month.

IRF analysis based on VAR (1) model was also performed. The IRF analysis result shows that the response of a variable for a one standard deviation (SD) of its innovations change increases from time to time except for volume of coffee export. Free-on-board price has a positive response for a one SD change in producer price.

There are few studies that have been designed to identify empirically the relative impact of external and domestic factors contributing to volume of coffee export in Ethiopia. The results of this paper could help understand factors that govern the volume of coffee export in Ethiopia. Moreover, this study could be a good stepping-ground for other studies on agricultural marketing and marketing cooperatives. In brief, this research would be useful to cooperative societies, researchers, and governmental and nongovernmental organizations for policy formulation, planning and development of agricultural marketing for both coffee exporters' and producers' in Ethiopia, which helps to achieve the development goal of the country.[6]

ACKNOWLEDGMENTS

We are heartily thankful to Department of Statistics, CNCS, Mekelle University to provide good environment, invaluable guidance, support, and patience from the initial to the final level enabled us to develop a better understanding of this research paper.

REFERENCES

[1]. Arnold Jeremiah Godfrey, X. 2011. Market Reform and its Impact on the Price Transformation in the coffee supply chain: A case study of Colombia, Ghana and Ivory Coast. Cornell University- US America, a thesis presented to the Faculty of the Graduate School of Cornell University in Partial Fulfillment of the requirements for the Degree of Master of Science.

[2]. Dickey, D.A., Fuller, W.A.1979. Distribution of the estimators for autoregressive time series with a unit-root. Journal of the American Statistical Association, 74, pp.427-431.

[3]. Dickey, D.A., Fuller, W.A.1981. Likelihood ratio statistics for autoregressive time series with a unit-root. Econometrica, 49, pp.1057-1072.

[4]. Dodaro, S.1993. Exports and growth: A reconsideration of causality. The journal of developing areas, 27 (2), pp.227-244.

[5]. Gebremariam, B.1989. Economics of the Ethiopian coffee industry (Unpublished). Ministry of coffee and tea development, Addis Ababa-Ethiopia.

[6]. Gebretsadik,G. 2012.vector autoregressive and cointegration of coffee export: the case of Ethiopia, a Thesis Submitted to the Department of Statistics in partial fulfillment of the Degree of Master of Science ,Addis ababa University .

[7]. Granger, C.W.J.2009. Some thoughts on the development of co-integration. Journal of Econometrics, 158, pp.156-159.

[8]. Kuhlin, Josef, Modig, Aron, 2009. A Study of the Competition at the Kenyan and Tanzanian Coffee Auctions. School of Bussiness, Economics ans Law. University of Gothenburg.

[9]. Kang, H. & Kennedy, P.L.2009. Emperical evidence from a coffee paradox: an export supply/price asymmetry approach. Journal of Rural Development 32(3) pp.107-137.

[10]. Mengistu, E. S. 2010. Price transmission system in Ethiopian coffee Market. Swedish University of Agricultural Sciences. ISSN 1401-4084, degree thesis No 625.

[11]. Ramakrishna, R.2016. Neural Networks Forecasting Yield per Hectare of Jowar :A Case Study,Ceser Publications, IJEES,Volume37,Issue1,pp.55-62.

[12]. Tadese, K. Worako,H.D.Van Schalkwyk, Z.G.&Alemu,A. 2008. Producer Price and Price Transmission in a Deregulated Ethiopian Coffee Market. Agrekon, Vol 47, No 4,pp.492-508

[13]. Zerihun,G.,Alemu&Tadese,K.W.2008.Price Transmission and Adjustment in the Ethiopian Coffee Market Contributed Paper prepared for presentation at the International Association of Agricultural Economists Conference, Beijing, China, August pp16-22.

GebretsadikGebru"Forecasting Analysis of Coffee Export by multivariate Time series Models of Vector Autoregressive and Cointegration: A case study of Ethiopia "International Journal of Computational Engineering Research (IJCER), vol. 08, no. 11, 2018, pp01-11