ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

Demand for Electricity in Ethiopia as Related to Consumption, Population and Gross Domestic Product

SAMUEL ABERA

November, 2007

Addis Ababa

Demand for Electricity in Ethiopia as Related to Consumption, Population and Gross Domestic Product

BY

SAMUEL ABERA

"A THESIS SUBMITTED TO GRAGUATE STUDIES PROGRAM IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF THE MASTER OF SCIENCE IN STATISTICS IN ADDIS ABABA UNIVERSITY"

November, 2007

ADDIS ABABA

ADDIS ABABA UNIVERSITY

GRADUATE STUDIES PROGRAM

Demand for Electricity in Ethiopia as Related to Consumption, Population and Gross Domestic Product

BY

SAMUEL ABERA

Approved by Board of Examiners

_____________________	___________________
Department Head	Signature
_____________________	___________________
Advisor	Signature
_____________________	___________________
Internal Examiner	Signature
_____________________	___________________
External Examiner	Signature

Table of Contents

ACKNOWLEDGEMENTS

I have benefited a lot from comments by Dr. M.K. Sharma, my thesis advisor and instructor in the department of statistics, AAU. Many other individuals were also helpful for the successful completion of this paper. Especially members of Ethiopian electric Power Corporation, power system planning division helped me very much in giving most of the data and by being the main initiator for the study area. The efforts made by all my relatives were interesting too. I thank all of them.

Samuel Abera

July, 2007

Acronyms

ACF: Autocorrelation function

ADF: Augmented Dicky-Fuller

AIC: Akaike Information Criteria

CON: Consumption of electricity

CSA: Central Statistical Agency

DMD: Demand of electricity

EEPCO: Ethiopia Electric Power Corporation

FEVD: Forecast Error Variance Decomposition

GDP: Gross domestic Product

IRF: Impulse Response Function

JB: Jarque-Bera test

L: Log Likelihood statistic

LR: Likelihood Ratio

LM: Lagrange Multiplier

MAPE: Mean Absolute Percentage Error

MOFED: Ministry of Finance and Economic Development

MSE: Mean Square Error

MVTSA: Multivariate Time Series Analysis

NBE: National Bank of Ethiopia

PACF: Partial Autocorrelation Function

POP: Population size of Ethiopia

PP: Phillips-Peron unit root test

SC: Schwarz Criterion

SVAR: Structural Vector Autoregression

VAR: Vector Autoregression

VEC: Vector Error Correction Model

ABSTRACT

Multivariate time series (MVTS) models are helpful to examine the relationships among many macroeconomic variables of a country. They give us information about the effect of time on some variables and the effect of one variable on another at the same time. The main objective of this study is to develop a MVTS model which can be used to describe and forecast demand (DMD) and consumption (CON) of electricity, population size (POP), and gross domestic product (GDP) of Ethiopia simultaneously based on forty one years national time series data. Different vector autoregressive (VAR) and vector error correction (VEC) models are applied to the data together with unit root test for stationarity and Johansen test for cointegration. Unit root test reveals that all the series are not stationary and the possibility of expressing all the variables using two cointegration equations is proved by cointegration analysis. The final result shows that a VEC model of lag one with two cointegration equations best fits the data among various statistical model selection criteria perspective. The accuracy of forecasts using this model is found to be high with mean absolute percentage error (MAPE) less than 5%. The forecasting ability of one variable as compared to another was verified by Granger causality tests. Short-tem and long-term interpretations from the VEC model were also given for the different effects among variables under study.

CHAPTER ONE

INTRODUCTION

1.1 Background

Electricity consumption in Ethiopia is among the lowest in the world. Using 1998 E.C. as reference, consumption of electricity is less than 4% of the total energy consumption. Though the history of electricity use in Ethiopia dates back about a century, the abundant hydro power potential available has not led to diversified industrial development in the past. In the last two decades the rate of energy demand has increased only by between 7 and 10 per cent and it hardly suggests that the long awaited turnaround has begun. The electricity coverage of the country is only 17 per cent which is again the lowest in the world.

Though Ethiopia is a country endowed with abundant resources such as hydropower, geothermal, natural gas, coal, solar and wind energy that could generate surplus electricity, almost all these resources are not yet tapped. This is partly due to the fact that the power scene in Ethiopia is full of contradictions emanating in part from unclear objectives and policies.

Over the last decade, one has witnessed the growth of the power sector and its impact on the economic development of a country. Electricity is considered as the backbone of industry and economic development in general. In order to bring all rounded economic development and to enhance industrialization starting from its 12 per cent contribution to a level of middle income countries which is the five year development plan

(Ethiopian Electric Power Corporation, Annual Report, 1998), expanding the power sector is vital. For efficient planning and operation of power system, every activity should be based on careful studies in various areas of the sector. One of the study areas is about demand and consumption of electricity.

The research utilizes multivariate time series analysis techniques to study the relationship among demand (DMD) and consumption (CON) of electricity, population size (POP), and economic development (GDP) of Ethiopia. It is based on yearly time series data collected from 1958 to 1998 in Ethiopian calendar.

Observing past outcomes of a phenomenon of interest in order to anticipate the future values is referred to as forecasting or predicting (Pole et al, 1994). If a complete and relevant model describing the studied phenomenon is known and if all relevant initial conditions are available then forecasting becomes a trivial task. However, when a model is unknown, incomplete or too complex, a typical alternative way is to build a model that takes into account past values of this phenomenon. In other words, the model is based on *what* the system does but not based on *how* and *why* the system does it. The past values of the phenomenon over time form the so-called a time series. The prediction of future values of a phenomenon doesn't have to depend on the past data exclusively. There might also be additional and more relevant information available in the present state of the environment. For example, prediction of water consumption does not depend on past values of water consumption only, but also on other relevant information available such as outdoor temperature, measurement time, season, and others. Such a multidimensional time series dependency is usually referred to as *multivariate time series (MVTS).*

In general, there are two main goals of time series analysis and prediction (Kout et al, 1989):

- Identifying the nature of the phenomenon represented by the sequence of observations in the past and by the sequence of other system variables (supporting attributes) in case of the multivariate time series and
- Forecasting – predicting future values of the phenomenon.

Both of these goals require that the pattern of observed time series data be identified and more or less formally described. Once the pattern is established, it can be interpreted and integrated with other data. As in most other statistical analyses, in time series analysis it is assumed that the data consist of a systematic pattern (usually a set of identifiable components) and random noise (error) that usually makes the pattern difficult to identify (Chatfield C., 1984).

This study is concerned with modeling multivariate time series data, which consist of simultaneous observations on several related variable of interest. Our main variable of study is consumption of electricity in Ethiopia, and consumption is supposed to be related with many factors. Of these factors the most imperative ones are demand of electricity, population growth and economic development of the country. We analyze data and develop a MVTS model which can adequately describe the innate relationships among the variables.

1.2 Purpose and Research Questions

Modeling MVTS data effectively is important for many decision- making activities. Especially in the Ethiopian power sector where things are so much complicated for decision making. We will examine in greater detail the different statistical techniques for analyzing multivariate time series data which consist of yearly demand and consumption of electricity, population growth and economic development of Ethiopia. In the power sector it is clear that studying the relationship among the above variables is important to improve the quality and quantity of the service provided by the sector. The following research questions will be addressed:

(1) What kind of relationships exist among demand and consumption of electricity, population growth and economic development in the Ethiopian context?

(2) How do we describe the relationship among the study variables (CON, DMD, POP, and GDP)? Or which MVTS model best describes the relationship among the variables and can be used for forecasting purpose?

(3) Are electricity demand and consumption affected by economic development and population growth? If so, what kind of effect do they have?

(4) As we said above the consumption of electricity has increased by between 10 and 17 per cent in the last two decades. Does this really show us the development of the power sector or is that only due to the increase in population?

(5) To what extent does electricity have an impact on the development of the economy as a whole?

1.3 Objectives of the Study

The general objective of this paper is to develop a MVTS model which explains the relationship among electricity demand and consumption, population size and GDP in Ethiopia that can be used for forecasting purpose in planning. The data consist of electricity demand and consumption of electricity, population growth and economic development of Ethiopia for forty one years (from 1958 to 1998 E.C.). The study has the following specific objectives:

- ✓ To show the various statistical techniques of analyzing MVTS data.
- ✓ To develop a model which can explain the data adequately.
- ✓ To highlight some of the key factors affecting electricity consumption in Ethiopia.

1.4 Data Sources and Limitations

Our data for the study are multivariate in nature, which means the data consist of many variables. For each variable of study the data were collected based on a fixed interval of time period (year). They are found from three sources. We shall discuss the nature and source of data with the corresponding limitations for each of the variables of study below.

i. Electricity Yearly Demand (DMD): is measured as the sum of maximum (peak) load over the months of the year and the unit of measurement is Mega Watt (MW). The source of data is the Ethiopian Electric Power Corporation (EEPCO) power system planning division, statistics and documentation section. Since it is based on actual observations it is supposed to be highly reliable data.

ii. Electricity Yearly Consumption (CON): This is the total amount of power in Mega-watt-hour (MWh) consumed by all customers of EEPCO per year.

iii. Population (POP): is the total number of individuals living in Ethiopia as of July 7 (Hamle 1) in any of the years from 1958 to 1998 E.C. Since census is conducted every ten year, clearly the figures here are estimates projected each year. The data are taken from the Central Statistical Agency (CSA) of Ethiopia. And hence, the limitation is that the data are estimated observations but not actual.

iv. Gross Domestic Product (GDP) is the real gross domestic product measured in Millions of Birr of Ethiopia. These values are calculated at the current factor cost. They are found from Ministry of Finance and Economic Development (MoFED) Annual report and Annual report magazines of the National Bank of Ethiopia (NBE) (1965 – 2006).

All the data obtained are collected on yearly basis, from 1958 E.C. to 1998 E.C. Due to lack of data it was not possible to incorporate more variables like annual disposable income and urbanization.

CHAPTER TWO

LITERATURE REVIEW

Forecasting in time series is a common problem. Different approaches have been investigated in time series prediction over the years [Weigend and Gershenfeld, 1993]. The methods can be generally divided into global and local models. In global model approach only one model is used to characterize the phenomenon under examination. The local models are based on dividing the data set into smaller sets, each being modeled with a simple model. The global models give generally better results with stationary time series. Stationary series are series that do not change with time. The first approaches were devoted to linear models for which the theory is known and many algorithms for model building are available. The most widely used linear regression methods in univariate case have been the autoregressive (AR), moving average (MA) and autoregressive moving average (ARMA) models [Box, et al 1994]. An example of more complex regression method is the multivariate adaptive regression splines [Friedman, 1991]. Using a statistical approach, the integrated autoregressive moving average (ARIMA) methodology has been developed [Box, et al, 1994]. The methodology is useful for fitting a class of linear time series models.

Simple extensions of these models for the MVTS case were also developed. These include the vector autoregressive (VAR), vector error correction (VEC) and vector autoregressive moving average (VARMA) models. VAR models are popular in economics and marketing (Sims, 1980). The definitive technical reference for VAR models is Lütkepohl (1991), and updated surveys of VAR techniques are given in Watson (1994) and Waggoner and Zha (1999). Applications of VAR models to financial data are given in Hamilton (1994), Campbell, et al (1997), Cuthbertson (1996), Mills (1999) and Tsay (2001). In addition to data description and forecasting, the VAR model is also used for structural inference and policy analysis. In structural analysis, certain assumptions about the causal structure of the data under investigation are imposed, and the resulting causal impacts of unexpected shocks or innovations to specified variables on the variables in the model are summarized. These causal impacts are usually summarized with impulse response functions and forecast error variance decompositions.

Multivariate autoregressive moving average (VARMA) models have also been used for modeling multiple time series, in order to incorporate relationships between series as well as within series. There is a considerable literature on inference for these models using frequentist approach, such as least squares or maximum likelihood methods (Reinsel, 1993). A Bayesian modeling framework has wide acceptance which has the advantage of being able to incorporate priori information in a natural way. Bayesian inference has been facilitated by the use of Markov chain Monte Carlo algorithms (Gelfand and Smith, 1990).

In a number of ways statisticians have addressed the restriction of linearity in the Box-Jenkins approach. A large number of nonlinear time series models are available. The stochastic approach to nonlinear time series that can fit nonlinear models to time series data was developed [Tong, 1990]. This is the reason why many nonlinear methods have become widely applicable. More recently, machine learning techniques (mainly neural networks) have been studied as an alternative to these nonlinear model-driven approaches. The process of constructing the relationships between the inputs and output variables is addressed by certain general-purpose 'learning' algorithms. Neural networks represent an attractive approach for time series prediction problems (Lapedes 1987, Casdagli, 1992, Drossu, 1996, Thiesing, 1997 and Street, 1998). Neural networks can construct approximations for unknown function by learning from examples. Multilayered perception (MLP) using the back-propagation algorithm is the most popular approach. Neither neural networks nor non-linear time series models are our concern now. We shall only see the application of linear MVTS, namely VAR and VEC models in the Ethiopian context using forty one years time series data on demand and consumption of electricity, population size and gross domestic product.

Various studies are conducted on the application of MVTS models on demand and consumption of electricity. For example, Holtedahl and Joutz (2004) used VAR model in order to model residential electricity consumption in KWh, real price of electricity per KWh, real disposable income per capita and urbanization in Taiwan using 40 years time series data and found out a very good result. Though such studies are common in the developed and some developing countries, it is hardly possible to see them in most underdeveloped countries like Ethiopia.

Some studies were also attempted in the Ethiopian context. Samuel Faye (2002) studied Household's consumption pattern and demand for energy in urban Ethiopia using some econometric models and tried to compare different towns according to their energy consumption level. More recently, Emanuel Gebreyohannes (2006) developed a non-linear approach to modeling and forecasting the residential electricity consumption in Ethiopia.

A good understanding of electricity demand and consumption will help to calculate what the capital requirements may be, and to estimate potential environmental impact (Bo Q. Lin, 2003). The theoretical study of demand forecasting of electric power systems started in the mid-20th century. Before that, because the scale of power systems was limited, the study of demand forecasting had not taken shape. It was not until the 1980s that the theoretical study of medium to long-term electricity demand forecasting began, and a series of forecasting methods, such as AR algorithm, MA algorithm, general exponential smoothing algorithm, ARMA (ARIMA) algorithm, began to be successively developed and widely accepted in electricity demand forecasting of power systems, particularly in developed countries (Zhang Chun, 1987). The increasing desire to obtain reliable forecasts is not only for power utilities but also for governments, because of impacts on economic growth and environmental protection. The reliability and confidence levels associated with these forecasts are more crucial in fast-growing areas experiencing a phenomenal growth in electricity demand. The confidence levels associated with classical forecasting techniques, when applied to forecasting problems in mature and stable power systems, are unlikely to be similar to those of dynamic and fast-growing utilities. In general, it is more difficult and challenging to forecast electricity demand and consumption in a fast-growing economy, where structural changes could have a significant impact on electricity demand. This is attributed to differences in the nature of growth, socioeconomic conditions, occurrence of special events, and subsidized energy tariffs (Baraket and Eissa, 1989). However, it is possible to ascertain the accuracy, suitability, and credibility of established classical forecasting techniques while searching for more improvements by taking into account the nature of growth, socioeconomic conditions, occurrence of special events, and subsidized energy tariffs.

Accurate projection of electricity demand is a precondition for successfully implementing power system planning, which in turn will have significant impact on future GDP growth. According to the definition of a demand function, electricity demand, in general, is determined by some main factors including gross domestic product (GDP), prices, and population. But for non-developed countries like Ethiopia, price does not have as such significant impact on demand and consumption of electricity because it is highly subsidized by the government. This study utilizes some MVTS techniques like VAR and VEC models on demand and consumption of electricity in relation to GDP and population size in the Ethiopian power sector.

CHAPTER 3

TECHNIQUES FOR ANALAZING MULTIVARIATE TIME SERIES DATA

3.1. Vector Auto Regressive Models

The vector autoregression (VAR) model is one of the most successful, flexible, and easy to use model for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model to dynamic multivariate time series. The VAR model has proven to be especially useful for describing the dynamic behavior of economic and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model. In addition to data description and forecasting, the VAR model is also used for structural inference and policy analysis. In structural analysis, certain assumptions about the causal structure of the data under investigation are imposed, and the resulting causal impacts of unexpected shocks or innovations to specified variables on the variables in the model are summarized. These causal impacts are usually summarized with impulse response functions and forecast error variance decompositions. This chapter focuses on the analysis of covariance stationary multivariate time series using VAR models. The analysis of non-stationary multivariate time series using VAR models that incorporate co-integration relationships will also be discussed at the end of the chapter..

Let $Y_t = (y_{1t}, y_{2t}, ... y_{nt})'$ denote an $(n \times 1)$ vector of time series variables. The basic p-lag vector autoregressive (VAR (p)) model has the form

$$\mathbf{Y}_t = \mathbf{c} + \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \mathbf{\Pi}_2 \mathbf{Y}_{t-2} + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{t-p} + \varepsilon_t, \; t = 1, \ldots, T \quad \textbf{(3.1)}$$

where Π_i's are $(\boldsymbol{n} \times \boldsymbol{n})$ coefficient matrices, c is an $(\boldsymbol{n} \times 1)$ vector of constants and ε_t is an $(\boldsymbol{n} \times 1)$ unobservable zero mean white noise vector process (serially uncorrelated or independent) with time invariant covariance matrix Σ.

In lag operator notation, the VAR (p) is written as

$$\Pi(L)\mathbf{Y}_t = c + \varepsilon_t \quad \textbf{(3.2)}$$

where $\Pi(\boldsymbol{L}) = \boldsymbol{I}_n - \Pi_1 \boldsymbol{L} - \ldots - \Pi_p \boldsymbol{L}^p$. The process is stationary if the roots of

$$\det(\boldsymbol{I}_n - \boldsymbol{\Pi}_1 \boldsymbol{Z} - \ldots \boldsymbol{\Pi}_p \boldsymbol{Z}^p) = 0 \quad \textbf{(3.3)}$$

lie outside the complex unit circle (have modulus greater than one), or, equivalently, if the eigenvalues of the companion matrix

$$\mathbf{F} = \begin{pmatrix} \mathbf{\Pi}_1 & \mathbf{\Pi}_2 & \cdots & \mathbf{\Pi}_n \\ \mathbf{I}_n & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_n & \mathbf{0} \end{pmatrix} \quad \textbf{(3.4)}$$

have modulus less than one, otherwise the process is said to be non-stationary. Assuming that the process has been initialized in the infinite past, then a stable VAR (p) is stationary with time invariant means, variances and autocovariances.

If $\boldsymbol{Y}_t$ in (3.1) is covariance stationary, then the unconditional mean is given by

$$\mu = (\mathbf{I}_n - \mathbf{\Pi}_1 - \cdots - \mathbf{\Pi}_p)^{-1}\mathbf{c} \quad \textbf{(3.5)}$$

The mean adjusted form of the VAR (p) is then

$$\mathbf{Y}_t - \mu = \mathbf{\Pi}_1(\mathbf{Y}_{t-1} - \mu) + \mathbf{\Pi}_2(\mathbf{Y}_{t-2} - \mu) + \cdots + \mathbf{\Pi}_p(\mathbf{Y}_{t-p} - \mu) + \varepsilon_t \quad \textbf{(3.6)}$$

The basic VAR (p) model may be too restrictive to represent sufficiently the main characteristics of the data. In particular, other deterministic terms such as a linear time trend or seasonal dummy variables may be required to represent the data properly. Additionally, stochastic exogenous variables may be required as well. The general form of the VAR (p) model with deterministic terms and exogenous variables is given by

$$\mathbf{Y}_t = \mathbf{\Pi}_1 \mathbf{Y}_{t-1} + \mathbf{\Pi}_2 \mathbf{Y}_{t-2} + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{t-p} + \mathbf{\Phi} \mathbf{D}_t + \mathbf{G} \mathbf{X}_t + \varepsilon_t \quad \textbf{(3.7)}$$

where $\mathbf{D_t}$ represents $(\boldsymbol{l} \times 1)$ vector of deterministic components, $\boldsymbol{X_t}$ represents an $(\boldsymbol{m} \times 1)$ vector of exogenous variables, and Φ and G are $(l \times l)$ and $(m \times m)$ parameter (coefficient) matrices, respectively.

3.2 Forecasting in VAR

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 matrix Π of the VAR (p) process are assumed to be known and there are no deterministic terms or exogenous variables. The best linear predictor, in terms of minimum mean squared error (MSE), of Y_{t+1} or 1-step forecast based on information available at time T is

$$\mathbf{Y}_{T+1|T} = \mathbf{c} + \mathbf{\Pi}_1 \mathbf{Y}_T + \cdots + \mathbf{\Pi}_p \mathbf{Y}_{T-p+1} \quad \textbf{(3.8)}$$

Forecasts for longer horizons h (h-step forecasts) may be obtained using the chain-rule of forecasting as

$$\mathbf{Y}_{T+h|T} = \mathbf{c} + \Pi_1 \mathbf{Y}_{T+h-1|T} + \cdots + \Pi_p \mathbf{Y}_{T+h-p|T} \tag{3.9}$$

The h-step forecast errors may be expressed as

$$\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T} = \sum_{s=0}^{h-1} \Psi_s \varepsilon_{T+h-s} \tag{3.10}$$

where the matrices Ψ_s are determined by recursive substitution

$$\Psi_s = \sum_{j=1}^{p-1} \Psi_{s-j} \mathbf{\Pi}_j \tag{3.11}$$

with $\Psi_0 = \boldsymbol{I}_n$ and $\Pi_i = 0_{n\times n}$ for $\boldsymbol{j} > \boldsymbol{p}$. The forecasts are unbiased since all of the forecast errors have expectation zero and the MSE matrix for $\boldsymbol{Y}_{t+h|T}$ is

$$\begin{aligned} \Sigma(h) &= MSE\left(\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T}\right) \\ &= \sum_{s=0}^{h-1} \Psi_s \Sigma \Psi_s' \end{aligned} \tag{3.12}$$

Now consider forecasting $\boldsymbol{Y}_{T+h}$ when the parameters of the VAR (p) process are estimated using multivariate least squares. The best linear predictor of $\boldsymbol{Y}_{T+h}$ is now

$$\hat{\mathbf{Y}}_{T+h|T} = \hat{\mathbf{\Pi}}_1 \hat{\mathbf{Y}}_{T+h-1|T} + \cdots + \hat{\mathbf{\Pi}}_p \hat{\mathbf{Y}}_{T+h-p|T} \tag{3.13}$$

where $\hat{\Pi}_j$ are the estimated parameter matrices. The h-step forecast error is now

$$\mathbf{Y}_{T+h} - \hat{\mathbf{Y}}_{T+h|T} = \sum_{s=0}^{h-1} \Psi_s \varepsilon_{T+h-s} + \left(\mathbf{Y}_{T+h} - \hat{\mathbf{Y}}_{T+h|T}\right) \tag{3.14}$$

and the term $(\mathrm{Y}_{T+h} - \hat{\mathrm{Y}}_{T+h|T})$ captures the part of the forecast error due to estimating the parameters of the VAR. The MSE matrix of the h-step forecast is then

$$\hat{\Sigma}(h) = \Sigma(h) + MSE\left(\mathbf{Y}_{T+h} - \hat{\mathbf{Y}}_{T+h|T}\right) \tag{3.15}$$

3.3 Structural 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 dynamic properties of a VAR (p) are often summarized using various types of structural analysis. The three main types of structural analysis summaries are:

(1) Granger causality tests

(2) Impulse response functions

(3) Forecast error variance decompositions.

The following sections give brief descriptions of these summary measures.

3.3.1 Granger Causality Tests

One of the main uses of VAR models is forecasting. The structure of the VAR model provides information about a variable's or a group of variables' forecasting ability for other variables. The following intuitive notion of a variable's forecasting ability is due to Granger (1969). If a variable, or group of variables, y_1 is found to be helpful for predicting another variable, or group of variables, y_2 then y_1 is said to Granger-cause y_2; otherwise it is said to fail to Granger-cause y_2. Formally, y_1 fails to Granger-cause y_2 if for all $s > 0$ the MSE of a forecast of $y_{2,t+s}$ based on $(y_{2,t}, y_{2,t-1}, \ldots)$ is the same as the MSE of a forecast of $y_{2,t+s}$ based on $(y_{2,t}, y_{2,t-1}, \ldots)$ and $(y_{1,t}, y_{1,t-1}, \ldots)$. Clearly, the notion of Granger causality does not imply true causality. It only implies forecasting ability.

We can test for Granger non-causality in general for n variable VAR (p) models. For example, consider a VAR (p) model with $n = 3$ and $\boldsymbol{y}_t = (\boldsymbol{y}_{1t}, \boldsymbol{y}_{2t,} \boldsymbol{y}_{3t})$. In this model, y_2 does not Granger-cause y_1 if all of the coefficients on lagged values of y_2 are zero in the equation for y_1. Similarly, y_3 does not Granger-cause y_1 if all of the coefficients on lagged values of y_3 are zero in the equation for y_1. We may refer Lütkepohl (1991) or Hamilton (1994) for more details and examples.

3.3.2 Impulse Response Functions

Any covariance stationary VAR (p) process has a Wold representation of the form (Stock and Watson, 2001), (Zha, 1998)

$$\mathbf{Y}_t = \mu + \varepsilon_t + \Psi_1 \varepsilon_{t-1} + \Psi_2 \varepsilon_{t-2} + \cdots \quad \textbf{(3.16)}$$

where the (n × n) moving average matrices Ψs are determined recursively using (3.11). This is important to interpret the (i, j)-th element, ψ_{ij}^{s} of the matrix Ψs as the dynamic multiplier or impulse response

$$\frac{\partial y_{i,t+s}}{\partial \varepsilon_{j,t}} = \frac{\partial y_{i,t}}{\partial \varepsilon_{j,t-s}} = \psi_{ij}^{s}, \quad i, j = 1, \ldots, n \quad \textbf{(3.17)}$$

However, this interpretation is only possible if $\text{var}(\varepsilon_t) = \Sigma$ is a diagonal matrix so that the elements of ε_t are uncorrelated. One way to make the errors uncorrelated is to follow Sims (1980) and estimate he triangular structural VAR (p) model.

$$\begin{aligned}
y_{1t} &= c_1 + \gamma'_{11}\mathbf{Y}_{t-1} + \cdots + \gamma'_{1p}\mathbf{Y}_{t-p} + \eta_{1t} \quad \textbf{(3.18)} \\
y_{2t} &= c_1 + \beta_{21}y_{1t} + \gamma'_{21}\mathbf{Y}_{t-1} + \cdots + \gamma'_{2p}\mathbf{Y}_{t-p} + \eta_{2t} \\
y_{3t} &= c_1 + \beta_{31}y_{1t} + \beta_{32}y_{2t} + \gamma'_{31}\mathbf{Y}_{t-1} + \cdots + \gamma'_{3p}\mathbf{Y}_{t-p} + \eta_{3t} \\
&\vdots \\
y_{nt} &= c_1 + \beta_{n1}y_{1t} + \cdots + \beta_{n,n-1}y_{n-1,t} + \gamma'_{n1}\mathbf{Y}_{t-1} + \cdots + \gamma'_{np}\mathbf{Y}_{t-p} + \eta_{nt}
\end{aligned}$$

In matrix form, the triangular structural VAR (p) model is

$$\mathbf{B}\mathbf{Y}_t = \mathbf{c} + \mathbf{\Gamma}_1\mathbf{Y}_{t-1} + \mathbf{\Gamma}_2\mathbf{Y}_{t-2} + \cdots + \mathbf{\Gamma}_p\mathbf{Y}_{t-p} + \boldsymbol{\eta}_t \quad \textbf{(3.19)}$$

where

$$\mathbf{B} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -\beta_{21} & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\beta_{n1} & -\beta_{n2} & \cdots & 1 \end{pmatrix} \tag{3.20}$$

is a lower triangular matrix with 1's along the diagonal. The algebra of least squares will ensure that the estimated covariance matrix of error vector η_t is diagonal. The uncorrelated/diagonal errors η_t are referred to as structural errors.

The triangular structural model (3.18) imposes the recursive causal ordering

$$y_1 \rightarrow y_2 \rightarrow \cdots \rightarrow y_n \tag{3.20}$$

The ordering (3.20) means that the contemporaneous values of the variables to the left of the arrow → affect the contemporaneous values of the variables to the right of the arrow but not vice-versa. These contemporaneous effects are captured by the coefficients β_{ij} in (3.18). For example, the ordering $y_1 \rightarrow y_2 \rightarrow y_3$ imposes the restrictions: y_{1t} affects y_{2t} and y_{3t} but y_{2t} and y_{3t} do not affect y_1; y_{2t} affects y_{3t} but y_{3t} does not affect y_{2t}. Similarly, the ordering $y_2 \rightarrow y_3 \rightarrow y_1$ imposes the restrictions: y_{2t} affects y_{3t} and y1t but y_{3t} and y_{1t} do not affect y_2; y_{3t} affects y_{1t} but y_{1t} does not affect y_{3t}. For a VAR (p) with n variables there are n! possible recursive causal orderings. Which ordering to use in practice depends on the context and whether prior theory can be used to justify a particular ordering. Results from alternative orderings can always be compared to determine the sensitivity of results to the imposed ordering (Lee, 1992).

Once a recursive ordering has been established, the Wold representation of Y_t based on the orthogonal errors η_t is given by

$$\mathbf{Y}_t = \mu + \Theta_0\eta_t + \Theta_1\eta_{t-1} + \Theta_2\eta_{t-2} + \cdots \qquad \textbf{(3.21)}$$

where $\Theta_0 = B^{-1}$ is a lower triangular matrix. The impulse responses to the orthogonal shocks η_{jt} are

$$\frac{\partial y_{i,t+s}}{\partial \eta_{j,t}} = \frac{\partial y_{i,t}}{\partial \eta_{j,t-s}} = \theta_{ij}^{s},\ \ i,j = 1,\ldots,n; s > 0 \qquad \textbf{(3.22)}$$

where θ_{ij}^{s} is the $(\boldsymbol{i},\boldsymbol{j})^{th}$ element of Θ_s. A plot of θ_{ij}^{s} against s is called the orthogonal impulse response function (IRF) of y_i with respect to η_j.With n variables there are n^2 possible impulse response functions.

In practice, the orthogonal IRF (3.22) based on the triangular VAR (p) (3.18) may be computed directly from the parameters of the non triangular VAR (p) (3.1).

3.3.3 Forecast Error Variance Decompositions

The forecast error variance decomposition (FEVD) answers the question: what portion of the variance of the forecast error in predicting $y_{i,T+h}$ is due to structural shock η_j? Using the orthogonal shocks η_t the h-step ahead forecast error vector, with known VAR coefficients, may be expressed as

$$\mathbf{Y}_{T+h} - \mathbf{Y}_{T+h|T} = \sum_{s=0}^{h-1} \Theta_s \eta_{T+h-s} \qquad \textbf{(3.23)}$$

For a particular variable $y_{i,T+h}$, this forecast error has the form

$$y_{i,T+h} - y_{i,T+h|T} = \sum_{s=0}^{h-1} \theta_{i1}^{s}\eta_{1,T+h-s} + \cdots + \sum_{s=0}^{h-1} \theta_{in}^{s}\eta_{n,T+h-s} \qquad \textbf{(3.24)}$$

Since the structural errors are orthogonal, the variance of the h-step forecast error is

$$var(y_{i,T+h} - y_{i,T+h|T}) = \sigma^2_{\eta_1} \sum_{s=0}^{h-1} (\theta^s_{i1})^2 + \cdots + \sigma^2_{\eta_n} \sum_{s=0}^{h-1} (\theta^s_{in})^2 \tag{3.25}$$

where $\sigma^2_{\eta j} = \text{var}(\eta_{jt})$. The portion of $(y_{i,T+h} - y_{i,T+h|T})$ due to shock η_j is then

$$FEVD_{i,j}(h) = \frac{\sigma^2_{\eta_j} \sum_{s=0}^{h-1} (\theta^s_{ij})^2}{\sigma^2_{\eta_1} \sum_{s=0}^{h-1} (\theta^s_{i1})^2 + \cdots + \sigma^2_{\eta_n} \sum_{s=0}^{h-1} (\theta^s_{in})^2}, \quad i,j = 1, \ldots, n \tag{3.26}$$

In a VAR with n variables there will be n^2 $FEVD_{i,j}(h)$ values. It must be kept in mind that the FEVD in (3.25) depends on the recursive casual ordering used to identify the structural shocks η_j and is not unique. Different casual ordering will produce different results.

3.4 Vector Error Correction and Cointegration Theory

3.4.1 VEC Models

The finding that many time series may contain a unit root has spurred the development of the theory of non-stationary time series analysis. Engle and Granger (1987) pointed out that a linear combination of two or more non-stationary series may be stationary. If such a stationary, or I(0), linear combination exists, the non-stationary (with a unit root), time series are said to be cointegrated. The linear combination which is stationary is called the cointegrating equation and may be interpreted as a long-run equilibrium relationship between the variables. For example, consumption and income are likely to be cointegrated. If they were not, then in the long-run consumption might drift above or below income, so that consumers were irrationally spending or piling up savings.

A vector error correction (VEC) model is a restricted VAR that has cointegration restrictions built into the specification, so that it is designed for use with non-stationary series that are known to be cointegrated. The VEC specification restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationships while allowing a wide range of short-run dynamics. The cointegration term is known as the error correction term since the deviation from long-run equilibrium is corrected gradually through a series of partial short-run adjustments.

3.4.2 Testing for Cointegration

Given a group of non-stationary series, we may be interested in determining whether the series are cointegrated, and if they are, identify the cointegrating (long-run equilibrium) relationships. We can interpret the long-run paths of cointegrated variables as interdependent. Application of cointegration test in estimation are analyzed by Johansen and Juselius (1990). VAR-based cointegration tests using the methodology developed by Johansen (1988) is the most common method. Johansen's method is to test the restrictions imposed by cointegration on the unrestricted VAR involving the series. It applies the maximum likelihood method to determine the presence of cointegrating vectors in non-stationary time series. The trace test and eigenvalue test are used to determine the number of cointegrating vectors. This implies a stationary long-run equilibrium relationship between the variables. The maximum lag length of the VAR model which is used in Johansen Procedure is determined by the Likelihood Ratio (LR) statistics.

Consider a VAR of order p

$$Y_t = A_1 Y_{t-1} + \cdots + A_p Y_{t-p} + \varepsilon_t \qquad \textbf{(3.27)}$$

where Y_t is a k-vector of non-stationary I(1) variables, and ε_t is a vector of innovations. We can rewrite the VAR as:

$$\Delta Y_t = \Pi Y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta Y_{t-i} + \varepsilon_t \qquad \textbf{(3.28)}$$

where $\Pi = \sum_{i=1}^{p} A_i - I$, $\Gamma_i = -\sum_{j=i+1}^{p} A_j$ and Δ is the difference operator.

Granger's representation theorem asserts that if the coefficient (parameter) matrix Π has reduced rank r<k, then there exist $k \times r$ matrices θ and ϕ each with rank r such that $\Pi = \theta \phi'$ and $\phi' Y_t$ is stationary. r is the number of cointegrating relations (the cointegrating rank) and each column of ϕ is the cointegrating vector. The elements of θ are known as the adjustment parameters in the vector error correction model. Johansen's method is to estimate the Π matrix in an unrestricted form, then test whether we can reject the restrictions implied by the reduced rank of Π.

If you have k endogenous variables, each of which has one unit root, there can be from zero to k-1 linearly independent, cointegrating relations. If there are no cointegrating relations, standard time series analyses such as the (unrestricted) VAR may be applied to the first-differences of the data. Since there are k separate integrated elements driving the series, levels of the series do not appear in the VAR in this case. Conversely, if there is one cointegrating equation in the system, then a single linear combination of the levels of the endogenous series $\phi' Y_{t-1}$, should be added to each equation in the VAR.

Each column of the ϕ matrix gives an estimate of a cointegrating vector. The cointegrating vector is not identified unless we impose some arbitrary normalization. We can adopt the normalization so that the r cointegrating relations are solved for the first r variables in the Y_t vector as a function of the remaining k-r variables.

When multiplied by a coefficient for an equation, the resulting term $\theta\phi' Y_{t-1}$, is referred to as an error correction term. If there are additional cointegrating equations, each will contribute an additional error correction term involving a different linear combination of the levels of the series.

3.5 Lag Order Selection in VAR/VEC Models

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

$$IC(p) = \ln|\bar{\Sigma}(p)| + c_T \cdot \varphi(n, p) \tag{3.29}$$

where $IC = \text{Information Criteria}$, $\bar{\sum}(p) = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$ is the residual covariance matrix from a VAR(p) model, c_T is a sequence indexed by the sample size T, and φ(n, p) is a penalty function which penalizes large VAR(p) models.

The three most common information criteria to determine the order of lag are the Akaike (AIC), Schwarz-Bayesian (BIC) and Hannan-Quinn (HQ):

$$\begin{aligned} AIC(p) &= \ln|\bar{\Sigma}(p)| + \frac{2}{T}pn^2 \\ BIC(p) &= \ln|\bar{\Sigma}(p)| + \frac{\ln T}{T}pn^2 \\ HQ(p) &= \ln|\bar{\Sigma}(p)| + \frac{2\ln\ln T}{T}pn^2 \end{aligned}$$

(3.30)

The AIC criterion asymptotically overestimates the order with positive probability (not zero), whereas 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.

CHAPTER FOUR

ANALYSIS, RESULTS AND DISCUSSIONS

4.1 Descriptive Analysis and Time Plot

It is good to first analyze the descriptive aspect of a series before making any kind of inferential statistics. Our study data consist of yearly demand of electricity in MW, yearly consumption of electricity in GWH, mid year population in millions and nominal GDP in millions of Birr of Ethiopia. The time period covered is from 1958 E.C. to 1998 E.C. The raw data are given in Table 1 of Appendix A. Table 4.1 shows the summary results of each of the series. As a first step of time series analysis, the time plots of each of the series are also given on Fig. 1 of Appendix 1. From the time plot we can observe that all the series show an increasing trend over the study period. There exists a positive correlation among the variables in the study. This may be true because, the increase in electricity consumption can lead to an increase in gross domestic product (GDP) or the increase in population size can result in an increase in demand of electricity. The correlation matrix (in the order DMD, CON, POP, and GDP) down here shows these large positive correlations among the series. Over the time period considered, consumption of electricity grew faster than the other variables (93 %) while population size showed a minimum increase (69 %).

$$\begin{pmatrix} 1 & 0.9968 & 0.9648 & 0.9663 \\ 0.9968 & 1 & 0.9762 & 0.9773 \\ 0.9648 & 0.9762 & 1 & 0.9757 \\ 0.9663 & 0.9773 & 0.9757 & 1 \end{pmatrix}$$

Table 4.1 Summary Result of the Series

	CON	GDP	DMD	POP
Mean	845.1234	10161.52	209.2841	44.73610
Median	731.7300	10326.50	167.7900	44.65000
Maximum	2361.290	22179.08	614.2700	75.07000
Minimum	166.3300	2979.200	43.30000	23.14000
Percent Inc.	1319.6417	644.4623	1318.6374	224.4166
St.Dev.	564.7181	5282.377	144.5237	16.34712
Skewness	0.794582	0.372248	0.953063	0.237657
Kurtosis	2.857264	2.298558	3.199567	1.721511

4.2 Nature of the Data

4.2.1 Test of Randomness

A series is said to be random if the observations are independent and could not have occurred in any order in time. Test of randomness is used to explain whether the series is time dependent, that is non-random or time independent, that is random. The series has to be non-random in order to apply time series analysis. There are several tests of randomness but the Rank test is the one used when a trend is suspected to exist in the series. In our case we shall use this test for each of the series.

Rank Test

This method of testing randomness involves counting the number of cases in which each observation is greater than the observation(s) which have occurred before it. That is, for a given series $y_1, y_2 \cdots y_n$, count the number of cases where $y_j > y_i,\ for\ j > i,\ i = 1, 2, \cdots n$; n is the total number of observations. In other words the method depends on the count of all positive differences and involves checking $\frac{1}{2}n(n-1)$ pairwise differences. Let p denote the counted number. An appropriate test statistic, known as Kendall's τ, is

$$\tau = \frac{4p}{n(n-1)} - 1, \ \ -1 \le \tau \le 1 \qquad \textbf{(4.1)}$$

If a given series is random then, $E[\tau] = 0$ and $V[\tau] = \frac{2(2n+5)}{9n(n-1)}$. For large n, the distribution of τ is assumed to be approximately normal (Smith, 1999). So, testing the hypothesis of random versus non-random series is equivalent to testing the hypothesis $H_0: \tau = 0$ versus $H_1: \tau \neq 0$. We reject H_0 at 95% level of confidence if $|\tau|$ is not in the confidence interval $\left(0, \ \frac{2(2n+5)}{9n(n-1)}\right)$.

For each of our series the value of τ and the corresponding confidence interval are given below. In all cases we reject the null hypothesis and conclude that the series are non-random. In other words, all the series are really time series and hence, we can apply any time series techniques to analyze the data.

Table 4.2 Result for Rank Test.

Series	n	p	τ	C.I.
DMD	41	680	0.6585	(0, 0.0118)
CON	41	728	0.7756	(0, 0.0118)
POP	41	642	0.5658	(0, 0.0118)
GDP	41	620	0.5121	(0, 0.0118)

4.2.2 Unit Root Test

Before we apply the different techniques of multivariate time series analysis, we need to check for stationarity of each variable under study since many of the methods assume that each data set is stationary with respect to the mean and the variance. As we can easily see from the time plots (Fig. 1 of Appendix B) all the series are not stationary. To test the presence of stationarity formally we can apply the unit root test. There are various kinds of unit root tests (Green 2000). The most popular of these tests are the Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test. The ADF and PP tests differ mainly in how they treat serial correlation in the test regressions. The following table shows the computed values of the two test statistics and p-value from these two tests for each of the variables.

Table 4.3 ADF and PP Unit Root Test Result for Original Series

Variable	Test Statistic		P-value		Remark
	ADF	PP	ADF	PP	
DMD	5.230	3.705	1.000	1.000	Not stationary
CON	5.556	3.234	1.000	1.000	Not stationary
POP	1.994	0.846	0.9986	0.9985	Not stationary
GDP	1.809	1.651	0.9984	0.9988	Not stationary

As we can clearly see from the above table, all the variables do not satisfy the unit root process and hence all are not stationary. But after differencing the series with appropriate order of differencing (order 2 for DMD and CON and order 1 for POP and GDP), the series in question become stationary. The next table shows the result of the ADF and PP test result after differencing the series. The time plots of the series after differencing are given in Appendix B Fig. 2. One can easily appreciate the difference between these and plots of the non-stationary ones.

Table 4.4 ADF and PP Unit Root Test Result for Differenced Series

Variable	Order of differencing	Test Statistic		P-value		Remark
		ADF	PP	ADF	PP	
DMD	1	3.125	-6.213	0.183	0.212	Not -Stationary
	2	-9.782	-44.417	0.000	0.000	Stationary
CON	1	3.763	-8.234	0.195	0.246	Not -Stationary
	2	-8.917	-52.914	0.000	0.000	Stationary
POP	1	-5.144	-32.544	0.000	0.000	Stationary
GDP	1	-5.189	-30.738	0.000	0.000	Stationary

4.2.3 ACF and PACF

In any time series analysis, the basic tools in identifying an appropriate model are the autocorrelation and partial autocorrelation functions. Depending on the behaviors of these functions, that is, whether they decay exponentially or cut off after a certain lag, we can specify a suitable model for our data. For our study variables the ACF and PACF are given in Fig.3 of Appendix B. Studying the properties of these functions will help us determine the appropriate lags to be used for any VAR and VEC model. The ACF of the original series die out fairly quickly but the PACF neither cut off nor die out fairly quickly. This is another indication that the series are not stationary.

4.3. Selecting and Estimating VAR Models

Model Selection

We specify the VAR as a four variable system for a sample period from 1958 to 1998. The general form of the VAR model is

$$\boldsymbol{Y}_t = \boldsymbol{c} + \pi_1 \boldsymbol{Y}_{t-1} + \pi_2 \boldsymbol{Y}_{t-2} + \cdots \pi_p \boldsymbol{Y}_{t-p} + \varepsilon_t \qquad \textbf{(4.2)}$$

where

$$\boldsymbol{Y}_t = \begin{bmatrix} Y_{1t} \\ Y_{2t} \\ Y_{3t} \\ Y_{4t} \end{bmatrix} = \begin{bmatrix} DMD \\ CON \\ POP \\ GDP \end{bmatrix} = \begin{bmatrix} \text{Demand of Electricity} \\ \text{Consumption of Electricity} \\ \text{Population} \\ \text{Gross Domestic Product} \end{bmatrix}, \quad t = 1, 2, \cdots, 41$$

Based on the properties of ACF and PACF plus the result from the unit root test we can propose the following VAR models as candidates for describing the data from which we shall select a suitable model based on various methods of selection. Let VAR (2) and VAR (1) be the candidate

VAR models for our data. Using STATA package the summarized result in estimating the parameters of these models are given below.

Table 4.5 Lag order selection result table

Model	Lag Order(P)	Likelihood statistic (L)	Schwarz-Criterion (SC)	Akiake Information Criterion (AIC)
VAR(2)	2	-680.9584	38.30269	36.7671
VAR(1)	1	-710.0703	37.34796	36.50352

From the above result we can observe that, using the overriding criteria VAR (1) is the best since it has the minimum L, SC and AIC. Therefore, the estimated VAR model becomes

$$\hat{\boldsymbol{Y}}_t = \hat{\boldsymbol{c}} + \hat{\Pi}_1 \boldsymbol{Y}_{t-1} + \hat{\varepsilon}_t \tag{4.3}$$

The coefficient matrix Π_1 and vector of constants c were successfully estimated by the method of least square using STATA package. But some coefficients are found to be insignificant. These are the coefficient of GDP, POP and DMD on lagged values of DMD, CON and GDP. All the remaining coefficients are statistically significant at 95% level of confidence. Substituting the estimated values of the parameters and assuming that $\boldsymbol{E}[\hat{\varepsilon}] = 0$, the final model is:

$$\begin{bmatrix} \hat{DMD} \\ \hat{CON} \\ \hat{POP} \\ \hat{GDP} \end{bmatrix} = \begin{bmatrix} \hat{Y}_{1t} \\ \hat{Y}_{2t} \\ \hat{Y}_{3t} \\ \hat{Y}_{4t} \end{bmatrix} = \begin{bmatrix} 29.0890 \\ 42.4233 \\ 2.6522 \\ 1498.3701 \end{bmatrix} + \begin{bmatrix} 0.7873 & 0.1334 & -1.75125 & 0 \\ 1.9953 & 0.5721 & 0 & 0.0129 \\ -0.0409 & 0.0121 & 0.8527 & 0.0004 \\ 0 & 4.2961 & -56.8595 & 0.8691 \end{bmatrix} \begin{bmatrix} Y_{1t-1} \\ Y_{2t-1} \\ Y_{3t-1} \\ Y_{4t-1} \end{bmatrix}$$

(4.4)

This can be explicitly written as

$$\hat{DMD}_t = 29.0890 + 0.7873\boldsymbol{DMD}_{t-1} + 0.1334\boldsymbol{CON}_{t-1} - 1.75125\boldsymbol{POP}_{t-1}$$

$$\hat{CON}_t = 42.4233 + 1.9953\boldsymbol{DMD}_{t-1} + 0.5721\boldsymbol{CON}_{t-1} + 0.0129\boldsymbol{GDP}_{t-1}$$

$$\hat{POP}_t = 2.6522 - 0.0409\boldsymbol{DMD}_{t-1} + 0.0121\boldsymbol{CON}_{t-1} + 0.8527\boldsymbol{POP}_{t-1} - 0.0004\boldsymbol{GDP}_{t-1}$$

$$\hat{GDP}_t = 1498.3701 + 4.2961\boldsymbol{CON}_{t-1} - 56.8595\boldsymbol{POP}_{t-1} + 0.8691\boldsymbol{GDP}_{t-1}$$

for $t = 1, 2...41$ **(4.5)**

From the model we can see that population size in Ethiopia is negatively related with demand of electricity and GDP. The insignificant coefficients indicate that the lagged variables GDP, POP and DMD do not cause or forecast the endogenous variable DMD, CON and GDP, respectively. As expected, demand and consumption of electricity have two way causations. In other words, high demand of electricity leads to high consumption and vice versa. Consumption of electricity and economic development (GDP) have also the same kind of relationships.

The accuracy of each of the above equations can also be verified based on the following result in Table 4.6 obtained from STATA in estimating the parameters of the model. One might observe that it is a good model in explaining the nature of the variation in the data.

Table 4.6 VAR (1) model estimation result

Equation	RMSE	R^2	χ^2	P-value
DMD	14.5639	0.6480	4340.731	0.0000
CON	41.8247	0.7688	7609.212	0.0000
POP	0.93761	0.8208	14124.87	0.0000
GDP	668.207	0.6129	2890.279	0.0000

Properties of the Disturbance Terms

It is also mandatory to study the nature of the residuals before asserting that the model is adequate. There are two basic issues here. The first is checking for the error terms whether some patterns remained unexplained, that is, testing for autocorrelation among the error terms, and testing for the normality of the residuals since most of our tests are based on the assumption that the error terms are normally distributed.

To test for error autocorrelation we used the Lagrange-Multiplier (LM) test which tests the null hypothesis H_0: No autocorrelation among residuals versus the alternative H_1: There is autocorrelation among residuals. The computed values of LM statistic at lag one is 20.0492 with 16 degrees of freedom informing us not to reject H_0. Therefore, there is no autocorrelation among the residuals in VAR (1) model. The graph of ACF of residuals (see Fig.4 of Appendix B) also reveals the same thing.

The Jarque-Bera test is used to test the null hypothesis that the disturbances are normally distributed. The JB test statistic is given by the formula

$$JB = \frac{n}{6}\left[S^2 + \frac{(K-3)^2}{4}\right]$$

Under the null hypothesis JB has a χ^2-distribution with 2 degrees of freedom.

In the above formula, S is a measure of skewness and K is a measure of kurtosis and n is the sample size. Here, the calculated chi-square values of each equation are 4.018, 0.325, 2.712 and 0.054 each with 2 (1 d.f. for skewness and 1 d.f. for kurtosis) degrees of freedom which shows the non-rejection of the null hypothesis. This together with the normal probability plot of the residuals shown at Fig. 6 of Appendix B reveal that the disturbance terms are normally distributed.

4.4 Structural Analysis

As discussed in Chapter Three the main types of structural analysis are Granger Causality tests, Impulse-Response functions and Forecast Error Variance decompositions. The next subtopic is about applying these summary measures to our data.

Granger Causality Tests

The structure of the VAR model provides information about the forecasting ability of a variable or a group of variables. The Granger-causality tests help us to measure whether one variable can be used to forecast the other. For instance, if demand of electricity is found to be helpful for predicting another variable, say GDP, then demand of electricity is said to Granger-cause GDP; otherwise it is said to fail to Granger-cause GDP. The following output is from EViews package for pairwise Granger- causality tests among the variables. The result shows that DMD does not Granger cause GDP and vice versa and POP does not Granger cause CON. All the other pairs Granger cause each other. This implies that DMD cannot be used to forecast GDP and vice versa. Similarly, POP cannot forecast CON. Remember that Granger causality test measures forecasting ability of one variable to another; it doesn’t show true causation.

Table 4.7 Pairwise Granger Causality Tests at lag 1

Null Hypothesis:	F-Statistic	P-Value
CONS does not Granger Cause DMD	5.25147	0.00901
DMD does not Granger Cause CONS	8.45284	0.00613
POP does not Granger Cause DMD	4.03035	0.01667
DMD does not Granger Cause POP	3.08497	0.03230
GDP does not Granger Cause DMD	0.01098	0.91712
DMD does not Granger Cause GDP	0.83320	0.43425
POP does not Granger Cause CONS	3.79757	0.05894
CONS does not Granger Cause POP	2.61007	0.04397
GDP does not Granger Cause CONS	9.00126	0.00489
CONS does not Granger Cause GDP	5.18622	0.02864
GDP does not Granger Cause POP	10.2886	0.00276
POP does not Granger Cause GDP	4.26189	0.01186

Impulse Response Functions

An impulse response function traces the response of a variable of interest to an exogenous shock. Often the response is portrayed graphically, with horizon on the horizontal axis and response on the vertical axis. It traces the effect of a one standard deviation shock to one of the innovations on current and future values of the endogenous variables. A shock to the i^{th} variable directly affects the i^{th} variable, and is also transmitted to all of the endogenous variables through the dynamic structure of the VAR.

Let $\varepsilon_{1t}, \varepsilon_{2t}, \varepsilon_{3t}$, and ε_{4t} be the innovations of DMD, CON, POP and GDP, ($\boldsymbol{Y_{1t}, Y_{2t}, Y_{3t}}$, ***and*** $\boldsymbol{Y_{4t}}$) in the above VAR (1) model. A change in ε_{1t} will immediately change the value of current DMD. It also changes all future values of CON, POP, and GDP since lagged DMD appears in all equations.

We can use EViews to compute one impulse response function for each innovation and endogenous variable pair. In our case there are 16 potential impulse response functions. The combined graphs of these IRF functions for ten years period are given in Fig. 8 of Appendix B. From the graph we can observe the following:

- Over the ten periods considered the response of a variable for one standard deviation (SD) shock of its innovations increases from time to time except for POP.
- DMD has a positive and a negative response for a one SD change in GDP and POP, respectively, while CON being not affected by a one SD change.
- CON has a positive response for a one SD innovation change in DMD and GDP and a negative response for POP.
- One SD innovation of DMD and POP has almost equal and opposite impact on GDP.

Forecast Error Variance Decompositions

Variance decomposition provides a different method of depicting the system dynamics. Impulse response functions trace the effects of a shock to an endogenous variable on the variables in the VAR. By contrast, variance decomposition decomposes variation in an endogenous variable into the component shocks to the endogenous variables in the VAR. The variance decomposition gives information about the relative importance of each random innovation to the variables in the VAR.

Usually, we plot the decomposition of each forecast variance as line graphs. The variance decomposition is displayed as separate line graphs with the y-axis height measuring the relative importance of each innovation. We use the order DMD CON POP GDP. For the variable that comes first in the ordering, the only source of the one period ahead variation is its own innovation and the first number is always 100 percent.

The results of the forecast error variance decompositions of the four study variables over a time period of 10 years are shown in Table 4 of Appendix A and Fig. 9 of Appendix B.

- Demand shocks provide the biggest percentage of the short-run movements accounting for more than 64% of the forecast error variance of DMD and CON in the first 10 years though its relative importance declines over time. GDP and POP shocks increase their percentage as the contribution of DMD shock decreases. The percentage contribution of CON shock for its forecast variance error decreases from 42 % to less than 1% over the 10 year time.
- All the shocks have almost similar contribution (14% – 35%) for the forecast error variance of POP around the 10^{th} period while POP percentage shock declined instantly from the first to the 10^{th} period. The importance of DMD and GDP shocks becomes significant after the 5^{th} year.
- CON and POP shocks have much lower percentage contribution for the forecast error variance of GDP which account for less than 6%. This variance is highly contributed by GDP shocks accounting more than 67.8 % and 25.64%, respectively.

4.4 VEC Models and Cointegration Analysis

A vector error correction (VEC) model is a restricted VAR that has cointegration restrictions built into the specification, so that it is designed for use with non-stationary series that are known to be cointegrated. The VEC specification restricts the long-run behavior of the endogenous variables to converge to their cointegrating relationships while allowing a wide range of short-run dynamics. Since our data are non-stationary, it seems reasonable to apply cointegration analysis and VEC modeling.

Johansen Test for Cointegration

As discussed in the previous chapter Johansen test is the most widely used tool for testing for cointegration. EView statistical package can perform Johansen cointegration test. By assuming no deterministic trend with an intercept in the cointegration relation on the differenced series and a lag interval of two, we got the following output from the package.

Table 4.8 Johansen cointegration test

Eigenvalue	Likelihood Ratio	5 Percent Critical Value	1 Percent Critical Value	Hypothesized No. of CE(s)
0.493439	61.39088	47.21	54.46	None **
0.425412	34.86656	29.68	35.65	At most 1 *
0.268122	13.25661	15.41	20.04	At most 2
0.027389	1.083073	3.76	6.65	At most 3
*(**) denotes rejection of the hypothesis at 5%(1%) significance level				
L.R. test indicates 2 cointegrating equation(s) at 5% significance level				

To determine the number of cointegrating relations r, we can proceed sequentially from r = 0 to r = 3 (i.e. 4-1) until we fail to reject. The first row in Table 4.8 tests the hypothesis of no cointegration, the second row tests the hypothesis of one cointegrating relation, the third row tests the hypothesis of two cointegrating relations, and so on, all against the alternative hypothesis of full rank, that is, all series in the VAR are stationary. The test result indicates that two cointegrating equations are sufficient. The next output in Table 4.9 gives us the normalized cointegrating coefficients for these two cointegrating equations. The numbers in parentheses under the estimated coefficients are the asymptotic standard errors. Some of the normalized coefficients will be shown without standard errors. This will be the case for coefficients that are normalized to 1.0 and for coefficients that are not identified.

Table 4.9 Normalized Cointegrating Coefficients

CONS	DMD	GDP	POP	C
1.000000	0.000000	-0.189005	21.07665	135.3337
		(0.09031)	(24.4017)	
0.000000	1.000000	-0.061984	9.169314	10.86514
		(0.03130)	(8.45682)	

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 are two stationary cointegrated series from the four non-stationary series (DMD, CON, POP and GDP). If we let these two stationary series, which are simply linear combinations of some of the study variables, as CE1 and CE2, then using the result in Table 4.9 we have the following.

$$\boldsymbol{CE}1 = \boldsymbol{CON} - 0.189005\boldsymbol{GDP} + 21.07665\boldsymbol{POP} + 135.3337,$$

and

$$\boldsymbol{CE}2 = \boldsymbol{DMD} - 0.061984\boldsymbol{GDP} + 9.169314\boldsymbol{POP} + 10.86514. \quad \textbf{(4.6)}$$

Then the result shows that in the long run CER1 and CER2 are stationary despite the fact that all the four series are non-stationary. One can infer from this result that there exist long-run causal relationships among CON, GDP and POP on one side and among DMD, GDP and POP on the other.

Estimating VEC Models

VEC specification only applies to cointegrated series. As tested above the series we have is cointegrated with order two. We also use EView to estimate a VEC model for the series. The specifications of a VEC model is the same as that of VAR except that we need to also specify the number of cointegration equations, *r*. Estimation of a VEC model proceeds by first determining one or more cointegrating equations using the Johansen procedure. The first difference of each endogenous variable is then regressed on a one period lag of the cointegrating equation(s) and lagged first differences of all of the endogenous variables in the system. With lag level of one and two order of cointegration, EView gives the summary output given in Table 4.10. We assume that there is trend in the data

Table 4.10 Estimated Parameter Result for VEC Model

Equation	No. Param.	RMSE	R^2	χ^2	P-value
ΔDMD	6	14.5119	0.9109	66.28312	0.0000
ΔCON	6	42.8408	0.9248	119.7035	0.0000
ΔPOP	6	.907665	0.9272	92.93633	0.0000
ΔGDP	6	643.965	0.8963	25.31422	0.0000
AIC = 36.42187, L = -710.4374, SC = 36.29742					

One can see from the above table that the VEC model applied with lag one and two cointegration equations is fair enough to describe the data under study. Note that the first column in the table indicates that the model is applied on the first order differenced series. The parameters of each of the equations above were also estimated and given as follows. The t-statistics are shown in parenthesis.

$$
\begin{aligned}
\Delta \boldsymbol{DMD}_t = & -\underset{(1..2941)}{0.1586}(\boldsymbol{DMD}_{t-1} - 11.2763\,\boldsymbol{POP}_{t-1} + 0.0250\,\boldsymbol{GDP}_{t-1} + 50.6674) \\
& + \underset{(4..3267)}{0.0953^{*}}(\boldsymbol{CON}_{t-1} - 40.7417\,\boldsymbol{POP}_{t-1} + 0.0749\,\boldsymbol{GDP}_{t-1} + 245.4457) \\
& + \underset{(6.0342)}{0.7473^{*}}\Delta \boldsymbol{DMD}_{t-1} - \underset{(2.6730)}{0.0800^{*}}\Delta \boldsymbol{CON}_{t-1} + \underset{(0.8246)}{1.5435}\Delta \boldsymbol{POP}_{t-1} + \underset{(0.3512)}{0.0070}\Delta \boldsymbol{GDP}_{t-1} + \underset{(1.1253)}{14.274} \\
\Delta \boldsymbol{CON}_t = & \ \underset{(3.2531)}{2.0490^{*}}(\boldsymbol{DMD}_{t-1} - 11.2763\,\boldsymbol{POP}_{t-1} + 0.0250\,\boldsymbol{GDP}_{t-1} + 50.6674) \\
& - \underset{(-1..9846)}{0.4634}(\boldsymbol{CON}_{t-1} - 40.74170\,\boldsymbol{POP}_{t-1} + 0.0750\,\boldsymbol{GDP}_{t-1} + 245.4457) \\
& + \underset{(-4.0012)}{2.2016}\Delta \boldsymbol{DMD}_{t-1} + \underset{(7.2005)}{0.1032}\Delta \boldsymbol{CON}_{t-1} - \underset{(1.0563)}{6.6439}\Delta \boldsymbol{POP}_{t-1} + \underset{(0.9824)}{0.0153}\Delta \boldsymbol{GDP}_{t-1} + \underset{(2.6542)}{54.874} \\
\Delta \boldsymbol{POP}_t = & -\underset{(1.7832)}{0.0457}(\boldsymbol{DMD}_{t-1} - 11.2763\,\boldsymbol{POP}_{t-1} + 0.0250\,\boldsymbol{GDP}_{t-1} + 50.6674) \\
& + \underset{(0.3651)}{0.0158}(\boldsymbol{CON}_{t-1} - 40.7417\,\boldsymbol{POP}_{t-1} + 0.0749\,\boldsymbol{GDP}_{t-1} + 245.4457) \\
& + \underset{(1.0238)}{0.02573}\Delta \boldsymbol{DMD}_{t-1} - \underset{(-0.7462)}{0.0039}\Delta \boldsymbol{CON}_{t-1} - \underset{(3.6599)}{0.1683}\Delta \boldsymbol{POP}_{t-1} - \underset{(2.9684)}{0.0004}\Delta \boldsymbol{GDP}_{t-1} + \underset{(0.6823)}{1.2983} \\
\Delta \boldsymbol{GDP}_t = & -\underset{(1.6320)}{0.1300}(\boldsymbol{DMD}_{t-1} - 11.2763\,\boldsymbol{POP}_{t-1} + 0.0250\,\boldsymbol{GDP}_{t-1} + 50.6674) \\
& + \underset{(3.1045)}{0.9556}(\boldsymbol{CON}_{t-1} - 40.7417\,\boldsymbol{POP}_{t-1} + 0.0749\,\boldsymbol{GDP}_{t-1} + 245.4457) \\
& + \underset{(1.0745)}{9.6016}\Delta \boldsymbol{DMD}_{t-1} - \underset{(0.6013)}{0.4483}\Delta \boldsymbol{CON}_{t-1} - \underset{(9.3818)}{100.7214}\Delta \boldsymbol{POP}_{t-1} + \underset{(3.7625)}{0.1534}\Delta \boldsymbol{GDP}_{t-1} + \underset{(-2.9748)}{479.997}
\end{aligned}
\tag{4.7}
$$

N.B: * indicates that the coefficient is statistically significant at 5% level.

In the equations, current changes in endogenous variable are functions of current changes in the first difference of all the variables and the two cointegration relations. Specifically, the coefficient of the first order differenced variable captures any immediate effect it has on the dependent variable of the equation, described as a contemporaneous or short-term effect. But each coefficient inside the brace (cointegration vector), reflects the equilibrium effect. It is the causal effect that occurs over future time periods, often referred to as the long-term effect that the variable has on the dependent variable of the equation. Finally, the long-term effect occurs at a rate dictated by the value of the coefficient of the cointegrated vector. For example, considering the first equation we can have the following interpretations. One can make similar inference for the others.

- ⇨ The coefficient of the lagged first order difference of DMD (equals 0.7473) is statistically significant at 5%. This indicates that DMD has a positive short-run (immediate) impact on the change in demand of electricity.
- ⇨ Similarly, the coefficient of the change in first order lagged value of CON (-0.08) is statistically significant at 5% which indicates that CON has a negative short-term impact on change in DMD.
- ⇨ POP and GDP do not have contemporaneous or short-term effect on DMD since the coefficients of their first order lagged difference are statistically non-significant.
- ⇨ Since the coefficient of the second cointegrating vector is significant we can conclude that CON, POP and GDP have a long-run or equilibrium effect on demand of electricity. Here, the long-run effect of CON and GDP is positive while that of POP is negative.

Error diagnosis for this model was performed. The results show that the VEC model disturbance terms are also white noise and are normally distributed.

Now, we have two models, VAR (1) and VECM with lag one and two order of cointegration, which can describe our series. Therefore, our last task will be to select the better of these two models.

Since VAR and VEC models belong to the same class of models we can use the AIC, L and SC criteria. The values of these statistics for VAR (1) were 36.50352, -710.0703 and 37.34796, respectively, while for the VEC model were 36.42187, -710.4374, and 36.29742. The AIC, L and SC criteria show us that the VEC model is slightly better than the VAR model. Accordingly, we can adopt the VEC model for prediction and forecasting purposes. The predicted series using this model are given in Table 2 (Appendix A) and the corresponding plots with the actual observations are shown in Fig. 6 of Appendix B.

CHAPTER FIVE

FORECASTING

Forecasting Using VEC Models

One of the fundamental applications of time series analysis or developing a TS model is forecasting. In our case, we have modeled the data and concluded that VEC model of order one with two rank of cointegration is the best to describe the series. It has a measure of goodness-of-fit, R^2 = 0.91, 0.92, 0.92 and 0.89 for DMD, CON, POP and GDP equations respectively. The plot of predicted versus actual observations, given in Fig. 6 of Appendix B shows that the model is good enough to describe the series accurately. But how good is this model in forecasting future values?

A one period ahead forecast of the next ten years based on 1998 for each series is given in Table 3 of Appendix A. As a measure of forecast accuracy we can compute the mean absolute percentage error. MAPE usually expresses accuracy as a percentage.

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|Y_t - \hat{Y}_t|}{Y_t} \tag{5.1}$$

where Y_t and $\hat{Y}_t$ denote the actual and predicted values.

For the VEC model, the MAPE in forecasting DMD, CON, POP and GDP are 2.32, 2.35, 0.80 and 3.29, 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%, which is good enough.

Another measure of forecast accuracy is Theil's U statistic (Montgomery and Johnson, 1976) which measures the degree to which one time series X_i differs from another Y_i. The index is computed as

$$U = \frac{\sqrt{\frac{1}{n}\sum(X_i - Y_i)^2}}{\sqrt{\frac{1}{n}\sum X_i^2} + \sqrt{\frac{1}{n}\sum Y_i^2}} \tag{5.2}$$

U varies from 0 to 1 with 1 meaning maximum disagreement.

Substituting the actual and predicted series into (5.2), the Theil's U statistics are found to be 0.0039, 0.0041, 0.0006 and 0.0137, respectively, for DMD, CON, POP and GDP. All these values are close to zero indicating that the actual and forecasted values agree strongly.

Both MAPE and Theil's U statistics show that the model is good enough to forecast the series accurately.

CHAPTER SIX

CONCLUSIONS

The objective of this paper was to apply multivariate time series methods to electricity demand and consumption, population size and economic development as measured by gross domestic product (GDP) in Ethiopia with forty one years (1958 – 1998 E.C.) yearly time series data. 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 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 of study as a function of its lag and the lag of other variables. Based on the given data GDP, demand of electricity and population size cannot be used to forecast demand of electricity, GDP and consumption of electricity, respectively. IRF and FEVD analysis based on VAR (1) model were 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 POP. DMD has a positive and a negative response for a one SD change in GDP and POP, respectively, while CON is not affected by a one SD change. One SD innovations change of DMD and POP have almost equal and opposite impact on GDP.

The decomposition of the variations into the component shocks was shown by FEVD analysis. Demand shocks provide the biggest percentage (more than 64%) of the forecast error variance of DMD and CON in the first 10 years though its relative importance declines from over time. GDP and POP shocks increase their percentage as the contribution of DMD shock decreases. The percentage contribution of CON shock for its forecast variance error decreases from 42 % to less than 1% over the 10 year time. CON and POP shocks have much lower percentage contribution for the forecast error variance of GDP which account for less than 6%. This variance is highly contributed by GDP shocks accounting more than 67.8 % and 25.64% respectively.

The Johansen cointegration test gave us two cointegration equations for all the series. Based on this result a corresponding vector error correction model was fitted with lag one and two ranks of cointegration. The result showed that the model is even better than the previous VAR model in light of different statistical criteria. It has a better coefficient of determination and forecasting accuracy. This model can be used for forecasting all the series simultaneously.

Therefore, in planning and making decisions related to the power sector in Ethiopia, it has a significant importance to make use of such models.

APPENDIX A: List of Tables

Table 1. Raw Data for each of the Series

Year	DMD	CON	POP	GDP
1958	43.3	166.33	23.14	2979.2
1959	51.76	201.22	23.67	3103.4
1960	53.19	221.06	24.14	3199.4
1961	56.83	227.11	24.32	3331.1
1962	67.06	245.85	25.83	3451.4
1963	73.19	285.92	26.46	3606.1
1964	79.55	299.46	27.12	3777.8
1965	82.09	319.54	27.8	3879.6
1966	85.46	278.95	27.1	3936
1967	89.37	342.72	27.91	5103.2
1968	90.83	348.49	28.64	5530
1969	93.24	370.53	29.33	6146.3
1970	94.96	385.56	30.25	6487.4
1971	105.51	433.16	31.07	7086.5
1972	110.77	468.66	31.9	7624.7
1973	110.36	488.73	32.78	9324.55
1974	123.56	573.22	34.62	9374.01
1975	129.83	598.28	41.04	10326.5
1976	144.29	594.16	42.18	9675.79
1977	161.53	665.13	43.35	8734.69
1978	167.79	731.73	44.65	9597.31
1979	180.87	808.74	45.96	10948.73
1980	188.46	840.26	47.31	10947.85
1981	212.46	878.83	48.7	10986.38
1982	225.09	916.42	50.13	11432.73
1983	232.44	947.16	51.62	10938.12
1984	216.39	991.33	53.15	10534.57
1985	260.08	1034.49	54.77	11798.74
1986	288.25	1133.94	56.37	11999.35
1987	308.41	1178.36	58.06	12644.31
1988	308.16	1267.85	56.37	13930.9
1989	316.97	1322.61	58.12	14640.21
1990	344.47	1358.5	59.88	14428.98
1991	331.36	1332.33	61.67	15294.03
1992	340.62	1375.75	63.5	16112.33
1993	363.25	1412.14	65.34	17354.36

1994	403.52	1621.44	67.22	17632.35
1995	418.06	1706.74	69.13	16941.47
1996	467.42	1846.8	71.07	18900.94
1997	545.63	2069.27	73.44	20702.11
1998	614.27	2361.29	75.07	22179.08

Table 2. Predicted Series for each Variable

Year	P-DMD	P-CON	P-POP	P-GDP
1958	43.89954	174.7469	23.6811	3335.37
1959	54.24739	211.1796	24.25393	3533.543
1960	57.16651	224.8399	24.87054	3670.499
1961	60.48244	236.5779	24.99511	3788.014
1962	68.35435	263.5441	26.13419	3851.517
1963	77.37395	298.3054	26.96102	4100.898
1964	82.97736	318.4536	27.48861	4248.571
1965	86.43282	333.7458	28.24387	4375.784
1966	84.87904	320.6245	27.03747	4278.464
1967	94.67541	376.867	28.75325	5507.16
1968	95.18058	385.8123	29.53669	5856.285
1969	98.61409	408.5516	30.51127	6438.964
1970	100.2539	421.4908	31.52802	6741.678
1971	113.284	474.3805	32.58446	7383.416
1972	120.5368	508.9713	33.69746	7938.136
1973	120.81	538.1862	35.30811	9453.125
1974	139.2365	606.5315	37.3788	9708.397
1975	135.9705	621.3422	43.23674	10256.96
1976	145.0162	635.1459	43.33821	9558.434
1977	166.3082	693.6021	44.15886	8918.705
1978	177.571	750.3771	46.12313	9858.778
1979	195.4194	832.9697	48.11605	11244.03
1980	203.2365	861.0251	49.3386	11275.43
1981	224.8316	926.212	50.02411	11311.82
1982	237.1443	973.2504	51.34046	11735.85
1983	244.5802	993.4952	52.50834	11327.67
1984	235.2849	975.7545	54.86129	11135.71
1985	272.2027	1097.761	55.42694	12175.25
1986	304.7837	1207.388	56.91597	12587.56
1987	323.4178	1274.931	58.29926	13172.49
1988	337.7103	1348.585	58.40714	14772.11
1989	348.662	1399.998	60.45332	15493.59

1990	372.0868	1466.02	61.19062	15268.14
1991	354.8637	1429.252	63.24101	15851.52
1992	364.4821	1476.176	65.23812	16612.89
1993	383.5371	1551.171	66.76199	17665.09
1994	439.7849	1747.72	69.35361	18558.43
1995	459.4873	1809.407	71.17754	18165.08
1996	513.0151	2005.901	73.20396	20187.23
1997	599.5491	2303.437	75.36081	22300.7
1998	599.5491	2303.437	75.36081	22300.7

Table 3. One period Forecast of each Variable for 10 years

Year	DMD	CON	POP	GDP
1999	684.9742	2617.21	78.53824	23767.48
2000	766.6587	2910.804	82.45061	25569.9
2001	860.9252	3248.746	86.85408	27618.91
2002	969.7632	3638.234	91.82346	29953.71
2003	1095.501	4087.535	97.44874	32620.01
2004	1240.843	4606.224	103.8348	35670.86
2005	1408.927	5205.408	111.1034	39167.79
2006	1603.393	5897.969	119.3958	43182.22
2007	1828.463	6698.854	128.8758	47796.96
2008	2089.035	7625.399	139.7336	53108.1

Table4. IRF and FEVD for VAR(1)

Response of DMD:

Period	DMD	CON	POP	GDP
1	14.01	0.00	0.00	0.00
2	15.25	2.48	-1.94	1.99
3	18.36	2.69	-3.83	4.72
4	22.23	2.52	-5.97	7.95
5	26.75	2.24	-8.45	11.71
6	32.00	1.92	-11.31	16.07
7	38.10	1.54	-14.64	21.13

Variance Decomposition of DMD:

Period	S.E.	DMD	CON	POP	GDP
1	14.01	100.00	0.00	0.00	0.00
2	21.04	96.86	1.39	0.85	0.90
3	28.71	92.95	1.63	2.24	3.18
4	37.73	88.52	1.39	3.80	6.28
5	48.50	83.98	1.05	5.34	9.63
6	61.37	79.65	0.76	6.73	12.87
7	76.69	75.69	0.52	7.96	15.83

8	45.17	1.10	-18.50	26.99
9	53.38	0.59	-22.98	33.80
10	62.90	0.00	-28.17	41.70

Response of CON:

Period	DMD	CON	POP	GDP
1	30.41	26.33	0.00	0.00
2	47.53	14.67	-4.13	10.25
3	60.52	12.20	-10.40	20.40
4	74.51	11.07	-17.90	31.91
5	90.58	10.03	-26.64	45.23
6	109.18	8.86	-36.79	60.66
7	130.76	7.52	-48.57	78.57
8	155.80	5.96	-62.23	99.34
9	184.84	4.16	-78.07	123.43
10	218.52	2.06	-96.45	151.39

Response of POP:

Period	DMD	CON	POP	GDP
1	-0.06	0.10	0.89	0.00
2	-0.21	0.51	0.78	0.03
3	-0.12	0.56	0.70	0.13
4	0.03	0.57	0.61	0.25
5	0.20	0.56	0.52	0.40
6	0.41	0.54	0.41	0.57
7	0.64	0.53	0.28	0.76
8	0.92	0.51	0.13	0.99
9	1.24	0.49	-0.05	1.25
10	1.61	0.47	-0.25	1.56

Response of GDP:

Period	DMD	CON	POP	GDP
1	137.07	59.11	-21.28	624.72
2	176.04	84.48	-56.22	667.44
3	239.26	84.90	-92.85	721.52
4	315.43	80.84	-134.75	784.98
5	404.23	75.43	-183.25	858.70
6	507.31	69.03	-239.50	944.23
7	626.90	61.60	-304.75	1043.46
8	765.63	52.97	-380.45	1158.57

8	94.84	72.18	0.36	9.01	18.45
9	116.25	69.12	0.24	9.90	20.74
10	141.43	66.48	0.16	10.66	22.70

Variance Decomposition of CON:

Period	S.E.	DMD	CON	POP	GDP
1	40.23	57.16	42.84	0.00	0.00
2	64.92	75.56	21.55	0.40	2.49
3	92.47	80.08	12.36	1.46	6.10
4	124.76	79.67	7.58	2.86	9.89
5	163.17	77.39	4.81	4.34	13.47
6	208.94	74.50	3.11	5.75	16.64
7	263.33	71.56	2.04	7.02	19.38
8	327.71	68.81	1.35	8.14	21.70
9	403.61	66.33	0.90	9.11	23.66
10	492.83	64.15	0.61	9.94	25.30

Variance Decomposition of POP:

Period	S.E.	DMD	CON	POP	GDP
1	0.90	0.49	1.28	98.23	0.00
2	1.31	2.72	15.64	81.59	0.05
3	1.60	2.36	23.03	73.94	0.67
4	1.82	1.85	27.41	68.33	2.42
5	2.02	2.51	29.78	61.90	5.82
6	2.24	5.34	30.05	53.52	11.09
7	2.53	10.70	28.07	43.38	17.84
8	2.91	17.99	24.21	32.83	24.97
9	3.44	25.82	19.41	23.56	31.20
10	4.14	32.90	14.70	16.64	35.77

Variance Decomposition of GDP:

Period	S.E.	DMD	CON	POP	GDP
1	642.66	4.55	0.85	0.11	94.50
2	948.56	5.53	1.18	0.40	92.88
3	1222.07	7.17	1.19	0.82	90.82
4	1494.60	9.25	1.09	1.36	88.30
5	1781.53	11.66	0.95	2.02	85.38
6	2094.02	14.31	0.79	2.77	82.13
7	2442.00	17.11	0.65	3.59	78.65
8	2835.39	19.98	0.52	4.46	75.04

9	926.57	42.96	-468.26	1292.10		9	3284.60	22.85	0.40	5.36	71.39
10	1113.27	31.35	-570.12	1447.01		10	3801.03	25.64	0.31	6.25	67.80

Appendix B: List of Figures

Fig. 1 Time Plots of the Original Series

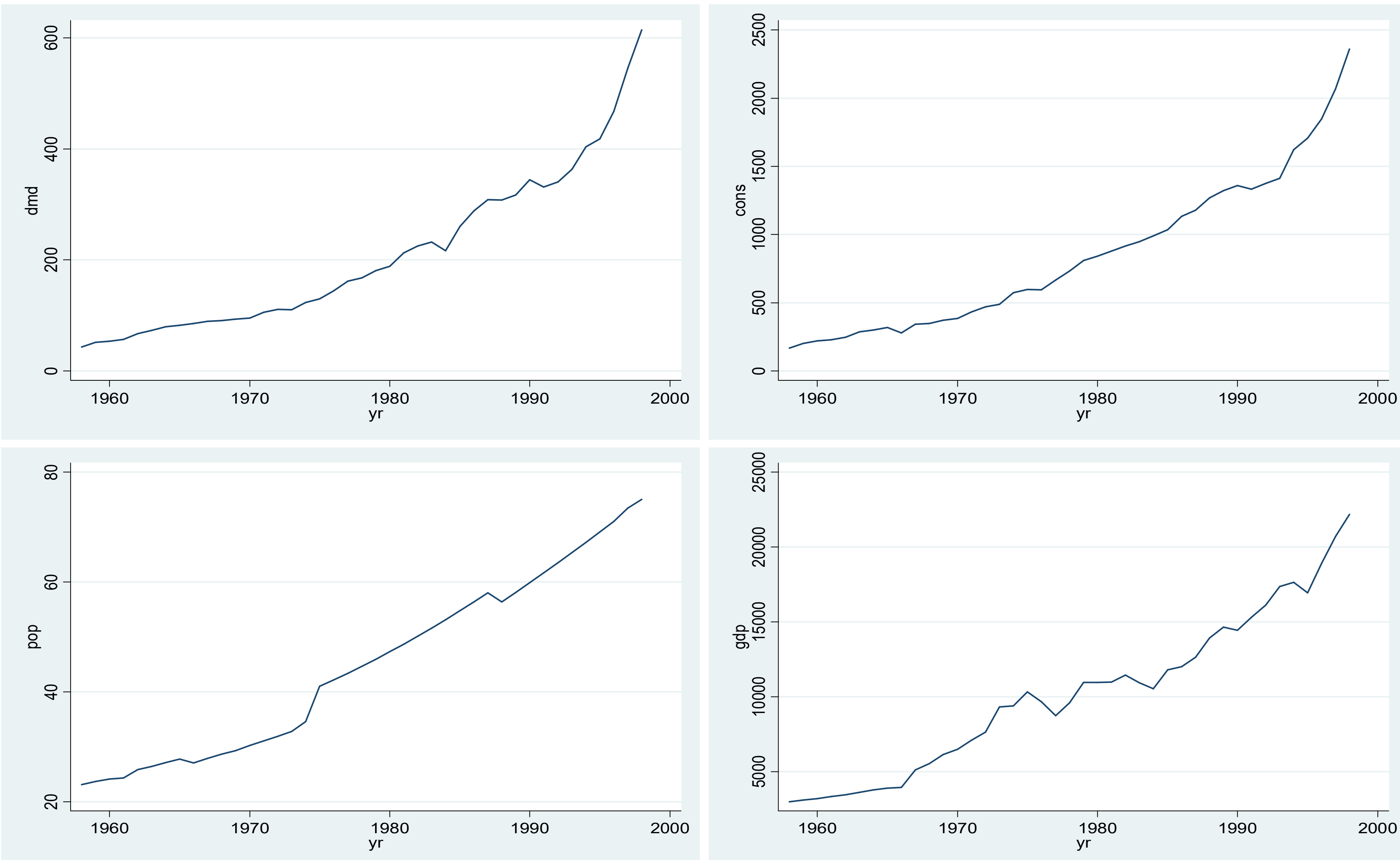
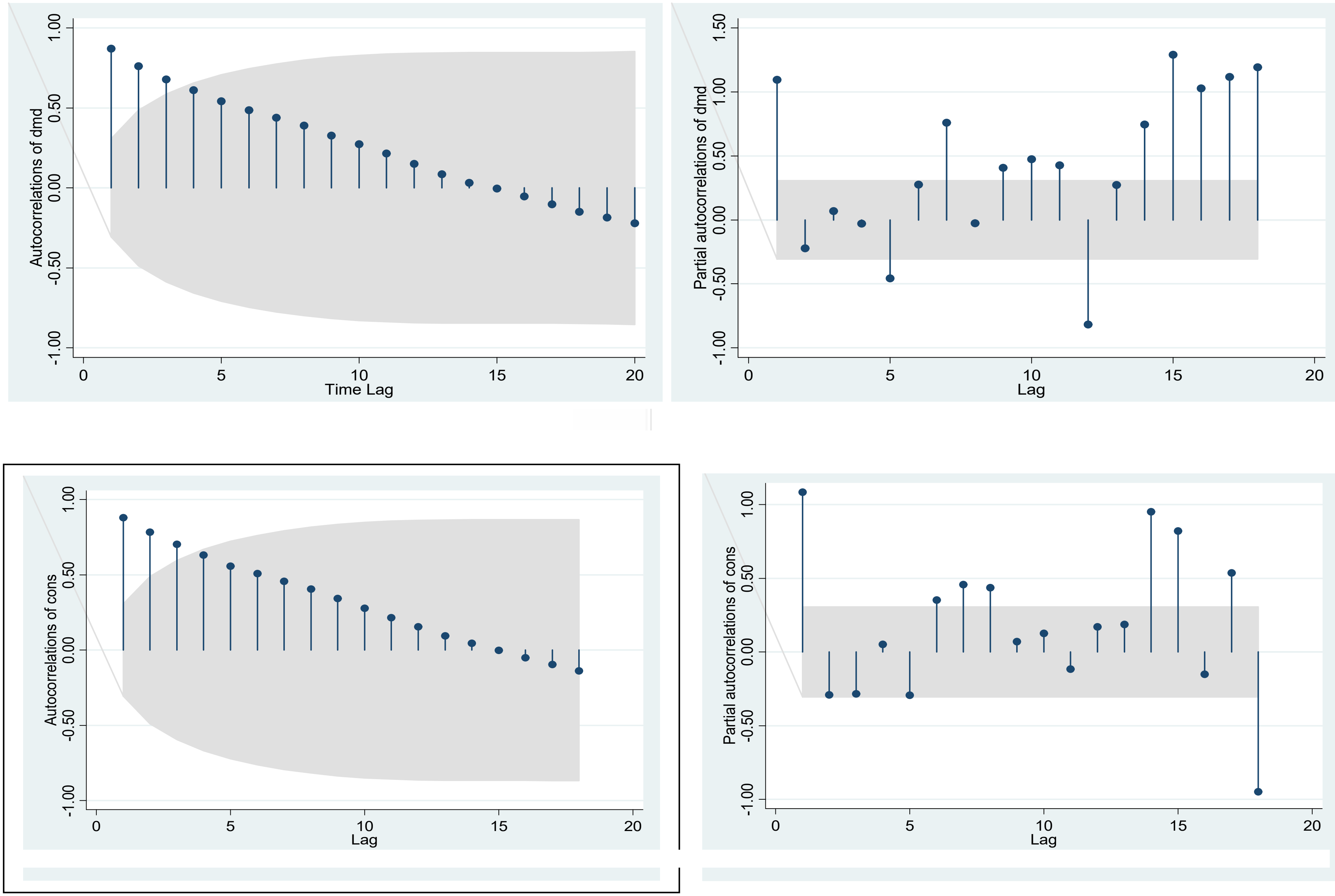
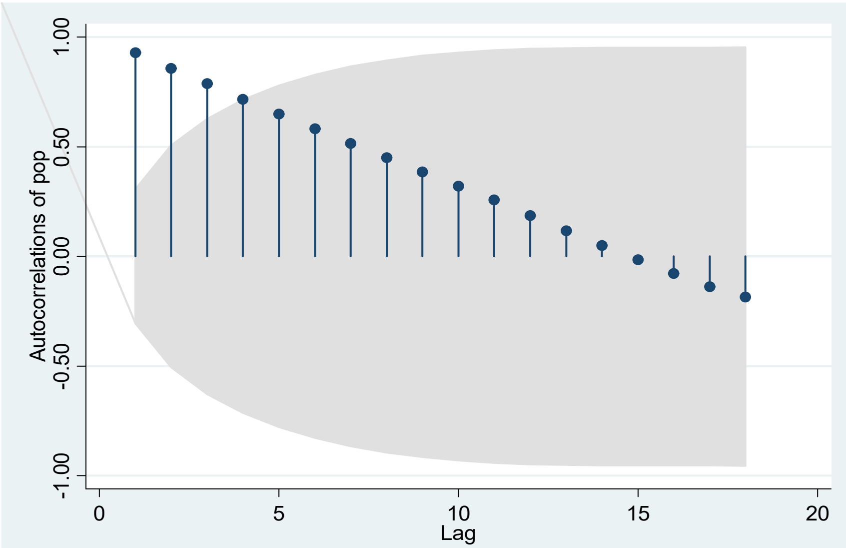
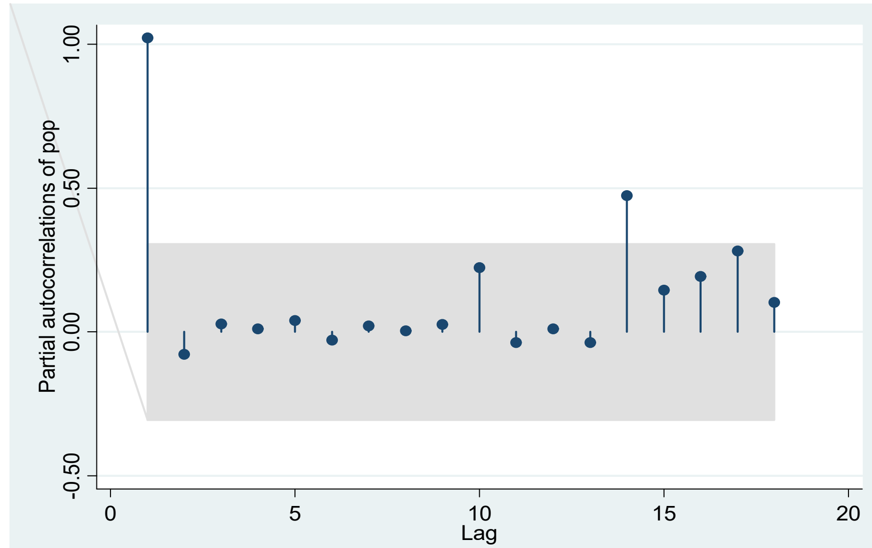


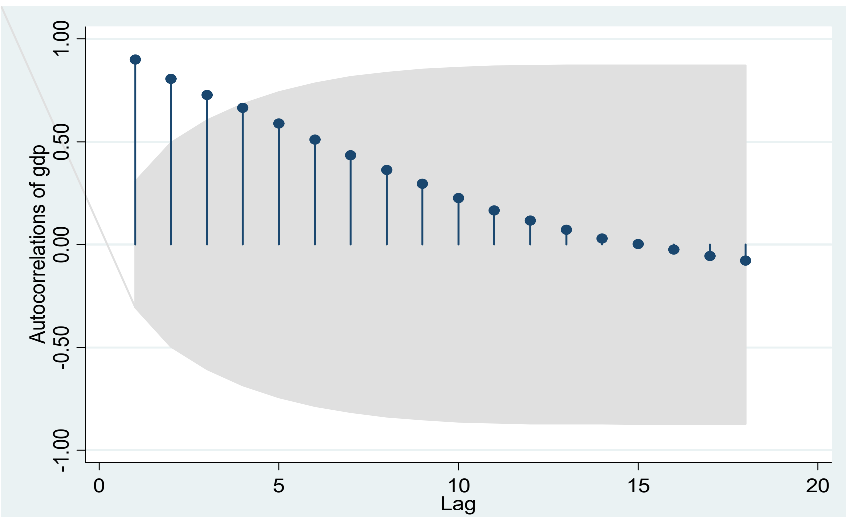
Fig. 2 Time Plots of the Differenced Series.

Fig.3 ACF and PACF for the Series




Autocorrelations of pop
Lag


Partial autocorrelations of pop
Lag


Autocorrelations of gdp
Lag

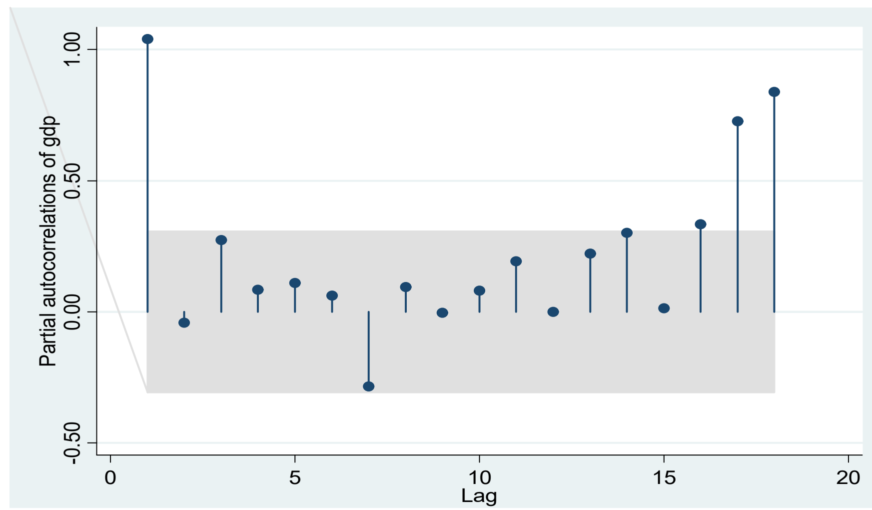

Partial autocorrelations of gdp
Lag

Fig. 4 ACF of the Residual Series after VAR (1)

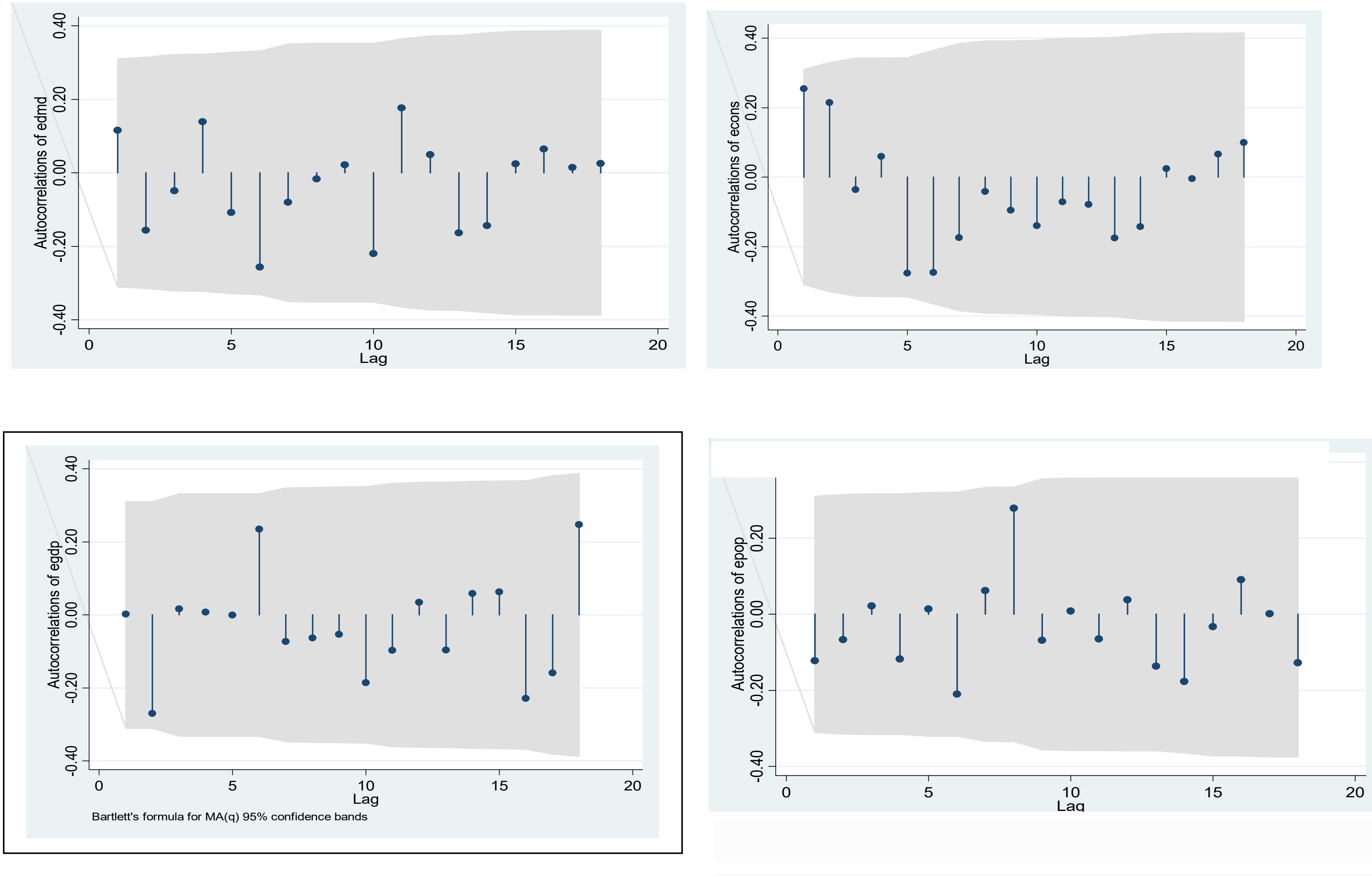


Fig. 5 Scatter Plots of the Error Terms after VEC Model is used

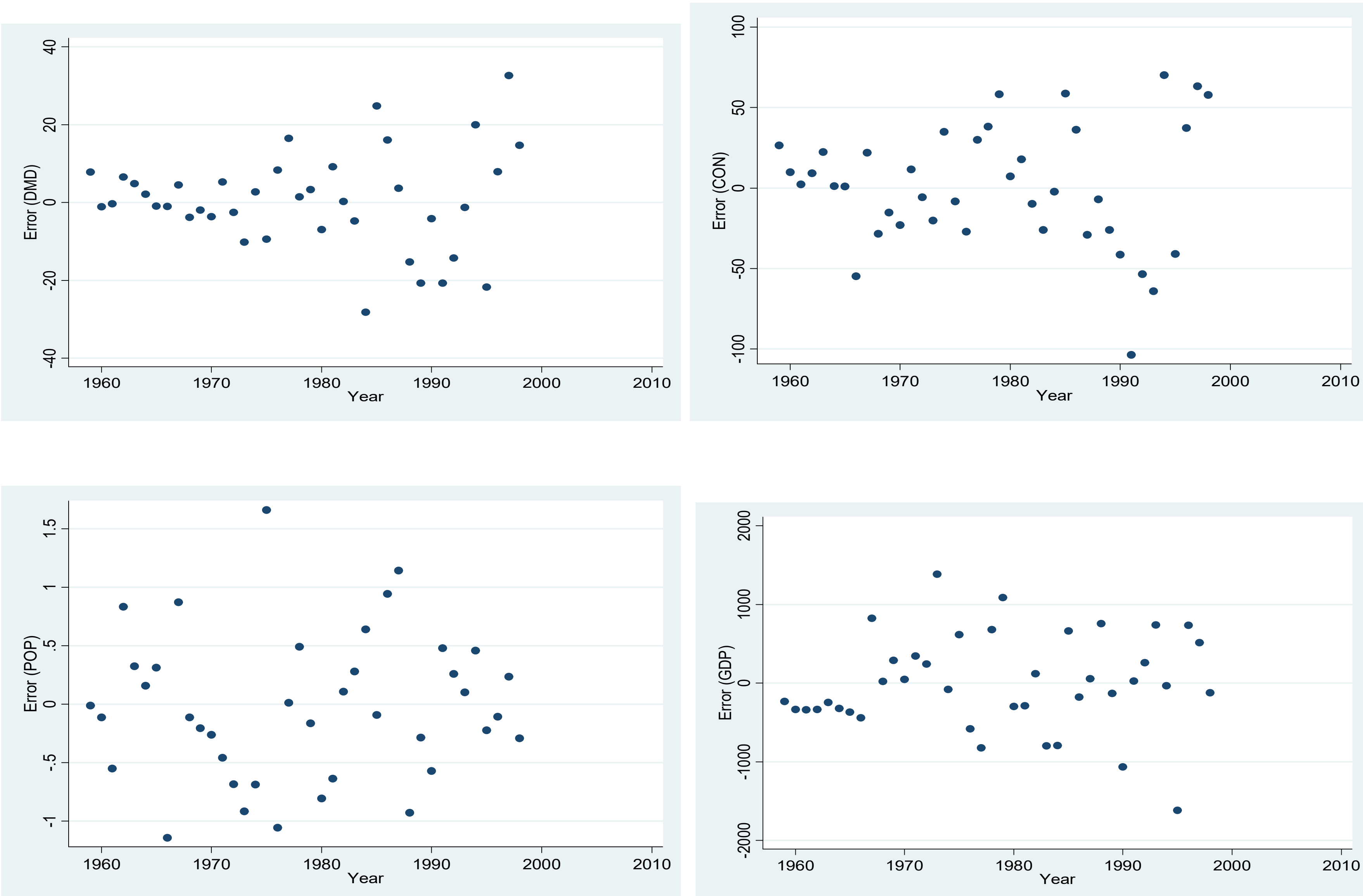


Fig. 6 Normal Probability Plot of Residuals after VECM

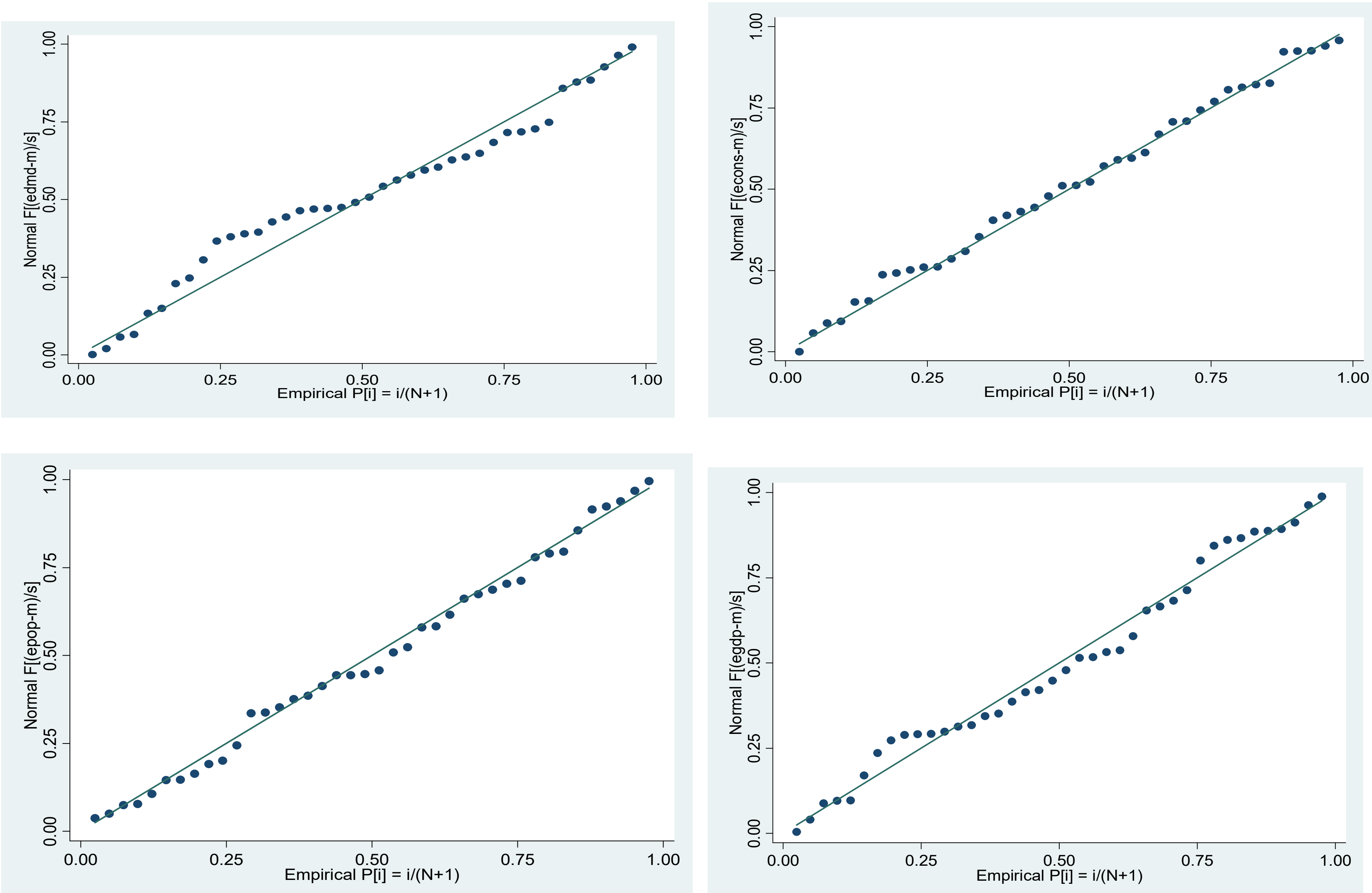


Fig. 7 Actual versus Predicted values of the Series

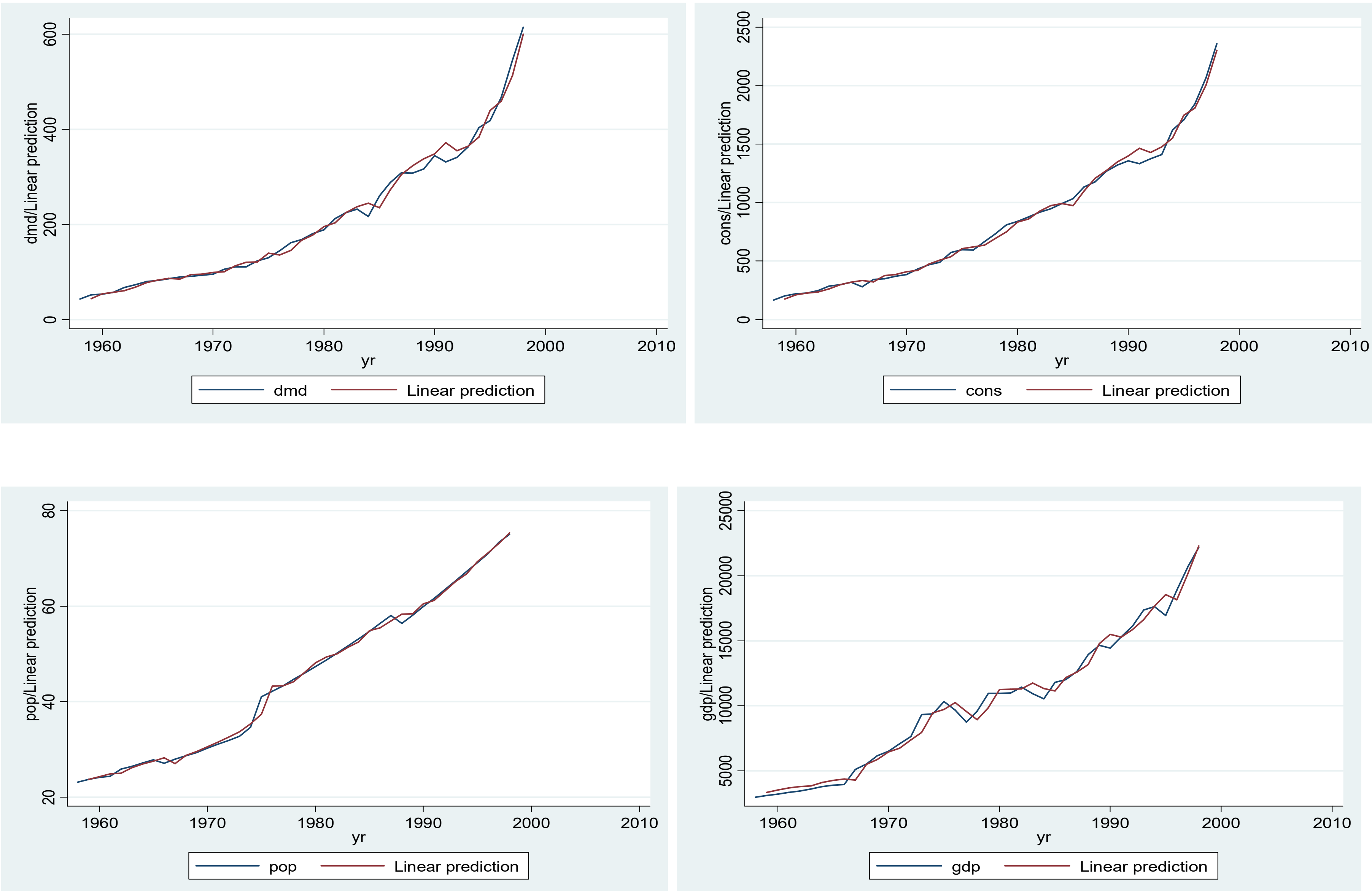


Fig. 8 Impulse Response Functions for DMD, CON, POP and GDP

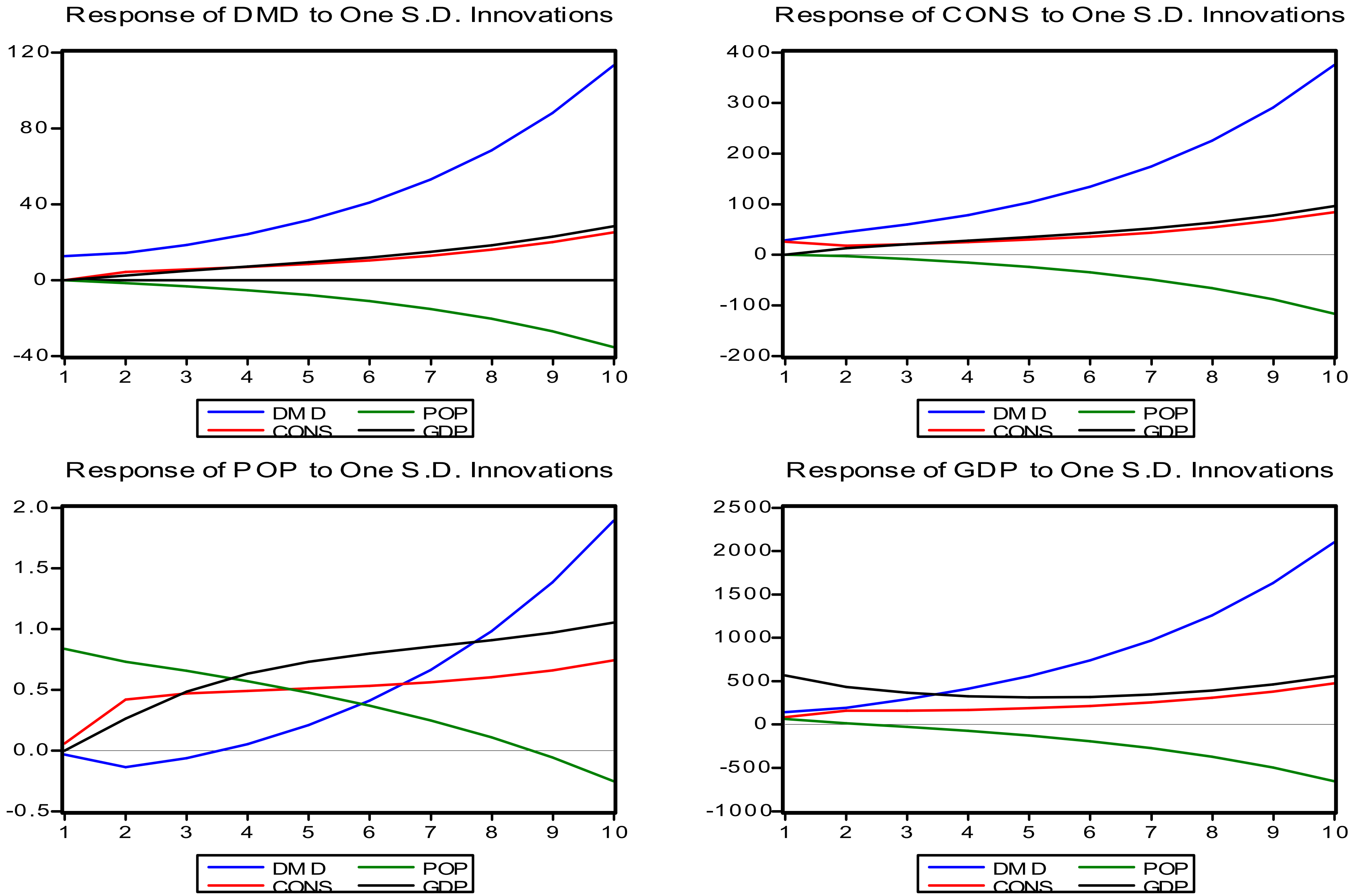
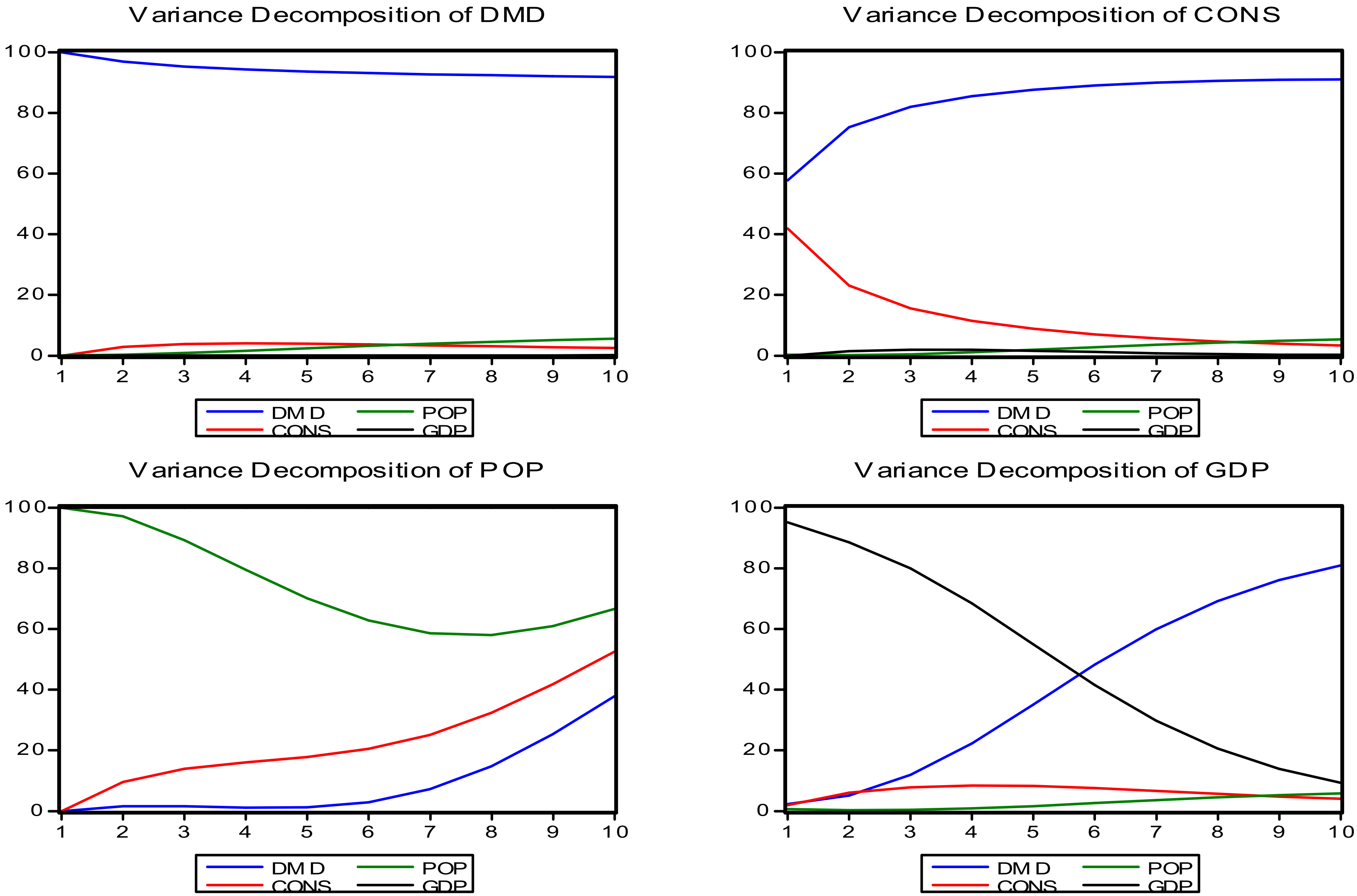


Fig. 9 Decomposition of each FEVD as Line Graphs in VAR (1)



Bibliography

- Baraket, H., and M. A. M. Eissa (1989). "Forecasting Monthly Peak Demand in Fast Growing Electric Utility Using a Composite Multiregression Decomposition Model." International Energy Outlook, Washington, D. C., PP:35-41.
- Bo Q. Lin: Electricity Demand in the Republic of China: Investment requirement and Environmental Impact. Asian Development Bank, Philippines, March 2003.
- Box G. E. P., Jenkins G. M., Reinsel G.C.(1994): Time Series Analysis, Forecasting and Control (3rd Ed). Englewood Cliffs: Prentice Hall.
- Campbell, J. A. Lo and C. MacKinlay (1997). The Econometrics of Financial Markets. Princeton University Press, New Jersey.
- Casdagli M. (1992): Nonlinear Modelling and Forecasting, Reading MA, Addison-Welsley.
- Chatfield C. (1984): The Analysis of Time Series. An Introduction (2nd Edition), Chapman and Hall, London.
- Cuthbertson, K. (1996). Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange. John Wiley and Sons, Chichester.
- Drossu R. (1996): Rapid Design of Neural Networks for Time Series Prediction, Washington State University, IEEE Computational Science and Engineering.

- Emanuel Gebreyohannes (2006). A Non-linear Approach to Modelling and Forecasting the Residential Electricity Consumption in Ethiopia. Journal of the Ethiopian Statistical Association, Volume XV, PP: 49 – 58.
- Engle, R. F., and C. W. J. Granger, 1987. Cointegration and Error Correction: Representation, Estimation and Testing. *Econometrica* PP: 251-256.
- Ethiopian Electric Power Corporation (1998). Annual Bulletin, Addis Ababa.
- Gelfand J. and W.H.Smith (1990). Marcov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference. London: Chapman and Hall.
- Granger, C.W.J. (1969). "Investigating Causal Relations by Econometric Models and Cross Spectral Methods," Econometrica. Volume IV, PP:112-120.
- Greene, W. H. (2000). Econometric Analysis. New Jersey: Prentice Hall.
- Hamilton, J.D. (1994). Time Series Analysis. Princeton University Press, Princeton.
- Holtedahl P. and F.L. Joutz (2004): Residential electricity demand in Taiwan. Department of Economics, the George Washington University, Washington DC, USA.
- Kout J., J. Kléma1, M. Vejmelka (1989): Predictive System for Multivariate Time Series, Czech Republic.

- Johansen, S. (1988). Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford: Oxford University Press.
- Johansen, S. and K.Juselius (1990). Maximum Likelihood Estimation and Inferences on Cointegration with Applications to the Demand for Money. Oxford Bulletin of Economics and Statistics, PP169-210.
- Lapedes, A.S (1987). Nonlinear Signal Processing using Neural Networks: Prediction and Modeling, Los Alamos National Laboratory.
- Lee, B.-S. (1992). "Causal Relations among Stock Returns, Interest Rates, Real Activity, and Inflation," Journal of Finance. Volume X, PP:18 – 21.
- Lütkepohl, H. (1991): Introduction to Multiple Time Series Analysis. Berlin: Springer-Verlag.
- Mills, T.C. (1999). The Econometric Modeling of Financial Time Series, Second Edition. Cambridge University Press, Cambridge.
- Montgomery, D.C. and L.A Johnson (1976): Forecasting and Time Series Analysis, (2nd Ed). Mc Graw-Hill Book Company, New York.
- Reinsel, D. E. (1993). "Vector Autoregressions and Reality," Journal of Business and Economic Statistics, PP:23 – 30.
- Pole, A.M West, P. Harrison (1994). Applied Bayesian Forecasting and Time Series Analysis, Chapman-Hall., New York
- Samuel Faye Gamtessa: Household's Consumption Pattern and Demand for Energy in Urban Ethiopia, Department of Economics, Addis Ababa University.

- Sims, C.A. (1980). “Macroeconomics and Reality,” Econometrica, Volume I, PP: 12 - 40.
- Sims, C. A., and Zha, T. (1998). Bayesian Methods for Dynamic Multivariate Models. Journal of International Economics, Volume IX PP:14-20.
- Smith, R.L.(1999) : Time Series, University of North Carolina, Chapel Hill, Wilmington.
- Stock, J.H. and M.W. Watson (2001). “Vector Auto regressions,” Journal of Economic Perspectives, Volume III, PP: 4 - 11.
- Street N. (1998): A Neural Network Model for Prognostic Prediction, Oklahoma State University, San Francisco, PP: 540-546.
- Tong H. (1990): Non-linear Time Series: A Dynamical Approach, Oxford University Press, Oxford.
- Tsay, R. (2001). Analysis of Financial Time Series. John Wiley & Sons, New York.
- Thiesing, F., (1997): Forecasting Sale Using Neural Network, Computational Intelligence, University of Osnabrueck, Germany.
- Waggoner, D. F., and T. Zha (1999). “Conditional Forecasts in Dynamic Multivariate Models,” Review of Economics and Statistics.
- Watson, M. (1994). “Vector Autoregressions and Cointegration,” in Handbook of Econometrics, Volume IV, PP: 63.
- Weigend A., N. Gershenfeld (eds) (1993). *Time Series Prediction – Forecasting the Future and Understanding the Past*, Addison-Wesley, Reading, Massachussets.

- Zha, T. (1998). “Dynamic Multivariate Model for Use in Formulating Policy”, Economic Review, Federal Reserve Bank of Atlanta, First Quarter.
- Zhang Chun. (1987). Electric Power Economics and Management. Beijing: Hydroelectricity Press.

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for degrees in any other University and all sources of material used for the thesis have been duly acknowledged.

Samuel Abera

…………………………

November, 2007