414

Analysis of Electricity Demand in Ethiopia Using Partial Adjustment and ARIMA Modeling

By

Misganew Abebaw

Addis Ababa University

Office of Graduate Studies

Department of Statistics

May, 2011
Addis Ababa
Ethiopia

Analysis of Electricity Demand in Ethiopia Using Partial Adjustment and ARIMA Modeling

By

Misganew Abebaw

A Thesis Submitted to the Office of Graduate Programs of Addis Ababa University in Partial Fulfillment of the Requirement for the Degree of Master of Science in Statistics.

May 2011
Addis Ababa

Research Title: Analysis of Electricity Demand in Ethiopia Using Partial Adjustment and ARIMA modeling

Name of Candidate: Misganew Abebaw

College: Natural Sciences

Department: Statistics

Approved by the board of examiners

Chairman	Examiner	Examiner
Name:__________	Emmanuel G. yohannes (Dr.)	SOLOMON W. HARRAR
Sign:__________		
Date:__________	27/06/2011.	July 5, 2011

Acknowledgments

I wish to record my sincere thanks and appreciation to my advisor, Dr. Butte Gotu, for his unreserved hospitality and abiding patience. The value of his dedication, kind and untiring guidance and warmhearted advice throughout the execution of this thesis is not only difficult to estimate but also very hard to express adequately by the usual terms of acknowledgement. Without his assistance, the production of this paper would not have been possible.

I am indebted to Mesert Molla of Gondar University, Yabebal Ayalew and Birhan Fetene of Addis Ababa University, and Tewodros Tebekew for their helpful comments and suggestion, invaluable support and continuous encouragements during the writing of the thesis. I would also like to thank Tesfaye Mesele, monitoring and evaluation advisor, USAID contractor, for his priceless comments on the final draft of the thesis.

I am also grateful to my family especially Agegnew Abebaw and Habetamu Abebaw for their painstaking support through all the years of sacrifice I subjected you to.

Table of Content

LIST OF TABLES

Acronyms

AC	Autocorrelation
ACF	Autocorrelation Function
ADF	Augmented Dickey-Fuller
ARDL	Autoregressive Distributed Lag
ARIMA	Autoregressive Integrated Moving Average
AR	Autoregressive
ARMA	Autoregressive Moving Average
CLRM	Classical Linear Regression Model
CRDW	Co-integration Regression Durbin Watson
CSA	Central Statistic Agency
DF	Dickey-Fuller
ECM	Error Correction Model
EEPCo	Ethiopian Electric Power Corporation
ELPA	Ethiopian Light and Power Authority
GDP	Gross Domestic Product
GWh	Gigawatt-hour
ICS	Inter Connected System
LM	Lugerangian Multiplier
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MWh	Megawatt-hour
MOFED	Ministry Of Finance and Economic Development
OLS	Ordinary Least Squares
PP	Philips-Perron
PAC	Partial Autocorrelation
PACF	Partial Autocorrelation Function

RESET	Regression Specification Error Test
RMSE	Root Mean Square Error
SCS	Self Connected System
UEAP	Universal Electrification Access Program
UNIDO	United Nations Industrial Development Organization
VAR	Vector Autoregressive

Abstract

This paper examined demand for electricity in Ethiopia as a function of real gross domestic product per capita, real price of electricity and population growth rate between the period 1976 and 2010. Partial adjustment and ARIMA models were used to provide electricity demand estimation and forecast, and results were compared with government projections. In the partial adjustment model, we found that income and population growth are the main determinants of electricity demand in Ethiopia, while the effect of electricity price is insignificant. Moreover, elasticities of electricity demand in Ethiopia are low (inelastic demand) meaning that, the responsiveness of consumers' to price, income and population growth rate changes is limited. From the comparison of ARIMA (1, 2, 1) model forecasts and government projection, except reference scenario, we found that the government electricity demand projections are consistently higher for moderate and targeted scenario than forecasted values of ARIMA (1, 2, 1) model.

Key words: Partial adjustment model, Elasticity, Co-integration Analysis, ARIMA

1 Introduction

1.1 Background of the Study

According to Accram (2005), about 928 million people live in slums all over the world and most of these people are in developing nations. The challenges are copious, as the slum world brings with it mass poverty, unemployment and degradation. The figures according to him are expected to rise in the next few years. There is another problem facing city managers in Africa, it is providing electricity to the poor in order to aid development, but there seems to be no concrete arrangement in this direction. Odumosu (2005) states that for a long time now most role-players in the power industry have agreed that the best way to rapid development is through adequate provision of electricity. However, across the African continent, there are differences in levels of power provided such that classification on the basis of sufficiency is complex. Hence the continent can be categorized into four on the basis of population and land mass. There are small countries with sufficient power such as Lesotho, big countries with sufficient power such as Ethiopia and South Africa, small countries with insufficient power such as Benin Republic and big countries with insufficient power such as Nigeria. It is therefore imperative for Africa to generate adequate power for development to take place.

Ethiopia is a landlocked country situated in the Horn of Africa with a population estimated at over 73 million and its annual population growth rate of 2.5% ranks it 28th

fastest growing of 229 countries in the world. In terms of gross domestic product (GDP) per capita Ethiopia is rated 174th of 179 and in terms of human development index it is rated 169th of 177. These numbers indicate Ethiopia as one of the poorest countries in the world (Census, 2007).

Electricity has become an essential commodity in a modern society. Our daily lives depend on the use of electricity in various forms. Many countries in Africa are endowed with hydropower resources. In the rural areas of least developed countries, due to lack of adequate access to modern energy services, many social and development programmers cannot be implemented. Most role-players in the power industry have agreed that the best way to rapid development is through adequate provision of electricity. Ethiopia is endowed with abundant water resources distributed in many parts of the country. However, it has not made significant progress in the field of water resources development during the past four decades. In particular, the exploitation of hydropower potentials was not noticeably successful in spite of being given priority as in the national development plan (Solomon, 1997). The Ethiopian Electric Power Corporation (formerly Ethiopian Electric Light and Power Authority established by charter in 1956) is a statutory corporation established by the council of ministers Regulation No.118/97. It is mandated to engage in the business of electricity generation, transmission and distribution in Ethiopia and is owned by the government (Girma, 2000).

Ethiopian Electric Power Corporation (EEPCo) organizational reforms have been transformed during the last eight years. The targets set by the government for the organization in 1997 were simple; increase generating capacity, reach all regions of the country, increase sells and improve collection. In order to promote the achievement of these targets, the company was reorganized. The most significant measure was decentralization of operations and management, with the division of the company in to eight regions with 10-15 districts in each region. Each district is responsible for supply, maintenance and commercial matters in the district, including new connections, revenue collection and disconnection (Venkataravaman and Krishnaswamy, 2007).

Ethiopia is one of the few African countries with the potential to produce hydroelectric and geothermal power. As of mid-1991, however, no comprehensive assessment of this potential was available, although some estimates indicated that the total potential could be as much as 143 billion kilowatts. The main sources of this potential were thought to be the Abay (Blue Nile; 79.9 billion kilowatts), the Shebele (21.6 billion kilowatts), and the Omo (16.1 billion kilowatts). The remaining 25.9 billion kilowatts would come from rivers such as the Tekeze, Awash, Baro, Genale, and Mereb.

Ethiopia's first large hydroelectric generating facilities were constructed in the Awash river basin. The three plants-Awash I (Koka) with capacity, Awash II with 32,000 kilowatts capacity, and Awash III with 32,000 kilowatts capacity were finished between 1960 and 1972. In 1974 the Fincha River facility in central Welega opened with a generating capacity of 84,000 kilowatts. Other major power generating facilities included those at Bahir Dar (7,680 kilowatts) and Aba Samuel (6,560 kilowatts). The total installed capacity of thermal generating units amounted to 210,084 kilowatts in 1985/86.

Electric power production in 1985/86 totaled 998.7 million kilowatt, 83 percent of which was produced by hydroelectric power installations. Thermal generating units produced the remaining 17 percent. The thermal generating units in the public utility system, many of which were comparatively small, had a generating capacity of 95,635 kilowatts in 1985. Major units were located close to Asmera (31,900 kilowatts), Dire Dawa (4,500 kilowatts), Addis Ababa (3,100 kilowatts), and Aseb (3,100 kilowatts). In 1985/86 various business enterprises and local communities owned electrical generators of unspecified capacity.

The regional electrical distribution system included an interconnected system and a self-contained system. By 1988 most power generating sources, including all major hydroelectric power plants, were interconnected in a power grid. The interconnected system served more than 100 towns. Power from the Awash, Fincha, and Aba Samuel stations ran the central system, the largest component of the interconnected system. The Bahir Dar interconnected system, which served parts of Gojam and Gonder, and the Eritrean Region Electricity Supply Agency (ERESA) were two of the other major systems. A majority of the self-contained systems got their power from thermal power plants, with the power often being used for domestic purposes and to run small mills.

The Ethiopian Electric Light and Power Authority (ELPA), a government corporation, operated most of the country's power systems. Prior to the revolution, ELPA incorporated more than forty electric power stations and generated about 80 percent of the nation's total electrical output. Two Italian firms, Societa Elettrica dell Africa Orientale and Compagnia Nazionale Impresse Elettriche, chiefly serving Eritrea, produced another 16.5 percent of the country's electrical energy. Independent stations generated the remaining 3 to 4 percent. In 1975 the government nationalized all private utility companies and placed them under ELPA. Since then, utility services have been reserved exclusively to the state. In 1987 ELPA served about 170 towns and produced about 92 percent of the national electrical output. Mass organizations, sugar factories, and the Aseb refinery administered the remaining 8 percent.

In 1985/86, of the total 847.7 million kilowatt of power sold by ELPA, 59 percent was for industrial use, 29 percent for domestic use, 10 percent for commercial use, and the remaining 2 percent for other uses such as street lighting and agriculture. By 1987 about 9 percent of the total population (4.3 million people) were using electricity.

The Ethiopian Electric Power Corporation (EEPCO) is currently responsible for almost all of the electricity supply in Ethiopia. The organization has operating two systems:

the interconnected system (ICS) forming the national grid, and self contained system (SCS) consisting of isolated, individual supply center.

Table-1: EEPCo Grid (ICS) Hydro Power Generating Capacity

Power plant	Year	Installed capacity(MWh)	Average energy capability (GWh/year)	Percentage of installed capacity	Percentage of average installed capacity
Koka	1960	43.2	110	2.2	1.5
Tis Abbay I	1964	11.4	85	0.6	1.1
Awash II	1966	32	182	1.6	2.4
Awash III	1971	32	182	1.6	2.4
Finchaa	1973	100	660	5.2	8.8
Melka Wakena	1988	153	543	7.9	7.2
Tis Abbay II	2001	73	359	3.8	4.8
Finchaa 4th unit	2002	34	137	1.8	1.8
Gilgel Gibe I	2004	184	720	9.6	9.6
Tekeze	2009	300	1070	15.5	14.3
Gilgel Gibe II	2010	420	1886	21.7	25.2
Amerti Neshe	2010	97	198	5.0	2.6
Total		1939.6	7496		

Source: EEPCo website

Ethiopia is already over-dependent on hydropower in its energy portfolio. Besides its 25-year master plan, EEPCo is also undertaking an aggressive five-year plan called the

Universal Electrification Access Program (UEAP) to expand the domestic grid. The UEAP's goal is to increase access to electricity to 50% of the population, by extending the grid to a total of 6,000 rural towns and villages by 2010. At the beginning of the program in 2005, EEPCo had 700,000 customers in 470towns. EEPCo reports it has electrified an additional 758 towns and villages in 2007, and that access rates have increased from 15% in 2005 to 22% in 2007. (Terri, 2008).

There is a clear commitment from the Government of Ethiopia to expand electricity coverage in Ethiopia. Furthermore, the country has abundant, cheap, and environmentally friendly hydro resources, whose potential has been estimated at more than 30,000 MW. The government has emerged increasing electricity access (from about 13% in 2006 to about 50% by 2012) as an integral part of its strategy for promoting income generating activities and social services outside major urban centers to improve living standards and reduce poverty (UNIDO, 2006).

1.2 Statement of the Problem

The power generation in Ethiopia and the projected energy requirements from the year 1990 through 2040 indicate that the power generation needs to be increased by14 times by the year 2020 and about 25 times by 2040 (Solomon, 1997).

With only 6% of households connected to the national grid and 15% of the population having access to electricity from Ethiopian Electric Power Corporation (EEPCo), access to electricity in Ethiopia is one of the lowest by any standard. More importantly, despite the fact that over 80% of Ethiopia's population live in rural areas and agriculture is the prominent supporter of the economy, electricity supply from the national utility is concentrated almost entirely in urban areas (GTZ, 2010). This indicates that there is a high deficiency in electric power supply. In other words, despite low income in Ethiopia, the desire to use electricity is very strong, that is there is a wide gap between supply and demand for electricity in the country. Therefore, this study examines demand and supply of electricity in Ethiopia and point out the appropriate model for forecasting electricity demand by using Partial adjustment and ARIMA modeling.

1.3 Objective of the Study

The general objective of this study is to develop a model of electricity demand in Ethiopia with a view to obtaining a short-run and a long-run price, income and population growth rate elasticities. Moreover, this study has the following specific objectives,

- To compute price, income and population growth rate elasticities of electricity demand.
- To forecast the demand for electricity in Ethiopia.
- To compare the discrepancy between the official projections of electricity demand with ARIMA model forecasts.

1.4 Significance of the Study.

Electricity demand studies have important practical implications. A robust forecast of electricity demand is very important as overestimation of demand could result in unnecessary huge investments in electric generations and transmission assets. Such a result would create more financial burdens in a country like Ethiopia that is already struggling financially with one of the highest debt relative to its GDP. Alternatively, underestimation of the demand for electricity would lead to further shortage of supply of electricity, a situation from which the country has already been suffering. It is very important, therefore, to select an appropriate demand estimation model that will provide future demand of electricity in Ethiopia.

1.5 Limitation of the Study

There are some other variables that might affect electricity demand function, for instance, data on construction, industry etc., but not included because of the unavailability of data.

1.6 Organization of the Study

The organization of the rest of this paper is presented as follows. Chapter 2 deals with the review of literature on electricity demand analysis and Chapter 3 specifies the data and methodology. In Chapter 4 the analyses of data and interpretations are presented. Chapter 5 presents conclusion.

2 Litrature Review

Engle and Granger (1987) have developed a technique, popularly known as "co-integration and error correction method" (ECM), for analyzing time series properties and estimating elasticities based on this analysis, which enables full analysis of the properties of the relevant data before actual estimation. In their study, Engle and Granger have formulated a model estimation procedure and recommended a number of tests, among which the most notable and commonly used is the Augmented Dickey-Fuller (ADF) test. Subsequent improvements related to this approach have been in the form of inclusion of variables in the model and the development of new methods to identify co-integrating relationships.

Since the late 1980s, especially co-integration analysis has become the standard component of all studies of energy demand; and most scholars have done their data analysis based on co-integration. Since economic theory indicates that the demand for energy in general depends on price and income, most of the studies in this area have been concentrated on these two variables as the major determinants of energy demand. The popularity and widespread use of the co-integration originate from the fact that it justifies the use of data on non-stationary variables to estimate coefficients as long as the variables are co-integrated; that is, they have a long-run equilibrium relationship. Actually, this is also the basic reason for the use of co-integration technique in this study. The papers written in this area include that of Engle et al.

(1989); Hunt and Manning (1989), Hunt and Lynk (1992), Bentzen and Engsted (1994, 2001), Fouquet et al. (1993), Hunt and Witt (1995); and Beenstock and Goldin (1999).

Box and Jenkins (1970) there emerged an increased interest in analyzing time series data and using rather simple models for out-of sample forecasting. Indeed, the proposed ARIMA models turned out to deliver high quality forecasts, and in fact, in those days these time series forecasts were observed to be much better than those from large-scale macro-econometric models containing hundreds of equations and equally sized numbers of unknown parameters. Even though it can be argued that large-scale simultaneous equations models can be written as VAR models, which in turn can be written as ARIMA type models (Sims, 1980 and Zellner and Palm, 1974), the sheer infinite distance between large models and ARIMA forecast schemes created a need for "models in between". Possible candidates for this were the single-equation error correction models such as the well-known Davidson et al. (1978) consumption function. In fact, as discussed by Granger (2009), it was exactly the confrontation of such models with unit-root processes that led to the notion of co-integration. Second, initiated by Dickey and Fuller (1979, 1981)'s innovative work on testing for unit-roots in time series data and the application of their tools to US macroeconomic data in Nelson and Plosser (1982), there seemed to be an acute problem with analyzing such data. Before then, all economic time series data were supposed to be governed by deterministic trends, while suddenly the word was that they all had stochastic trends, also known as unit- roots. If that were true, then all previously constructed models were created using the wrong statistical tools, as Phillips (1986) showed that statistical theory for regressions with unit-root time series is markedly different than standard theory.

Most of the studies on determinants of electricity demand functions have focused on developed countries. There are studies for Canada, Bernard et al. (1996) developed unit-root and co-integration techniques to estimate the long-run and short-run income and price elasticities of electricity demand. The results indicate that in the long-run demand for electricity is price elastic and income inelastic.

In the United States; Houthakker et al. (1974) and Houthakker (1980) estimate electricity demand functions with particular attention paid to the demand stickiness that is imposed by the capital-intensive nature of electricity consumption and to regional, seasonal, and sectoral variation. The analysis uses a partial adjustment model of electricity demand that is estimated in a fixed-effects OLS framework. This model formulation allows for the price elasticity to be expressed in both its short-run and long-run forms and also Hsing (1994) estimated electricity demand elasticities in the United States using partial adjustment model approach for six southern states, correcting for cross-sectional heteroskedasticity and autocorrelation.

For Asia–Pacific countries, such as Taiwan Holtedahl and Joutz (2004) analyzed electricity demand. The main variables of the study were income, population growth, electricity price, urbanization and weather. The Error Correction Model was applied to separate both the short and long-run effects of these variables on demand for electricity. The study concluded that in the long-run demand for electricity increased with increase in income where the price effect was negative and inelastic. However, the effects of price and income were smaller in short-run as compared to long-run. Furthermore, weather and urbanization also influenced demand for electricity in both short and long-run and Filippini and Pachauri (2004) estimates electricity demand in urban Indian households. Three electricity demand functions have been econometrically estimated using monthly data for the winter, monsoon and summer season in order to understand the extent to which factors like income, prices, household size and other household specific characteristics, influence variations observed in individual households' electricity demand. The results show electricity demand is income and price inelastic in all three seasons, and that household, demographic and geographical variables are significant in determining electricity demand.

European countries, such as Greece, Hondroyiannis (2004) examined the factors which brought changes in aggregate demand for electricity for both short and long-run periods. The study concluded that in the long-run real income, price level and weather played an important role in the demand for electricity. However, in the short-run changes in demand for electricity was affected only by weather condition and Filippini (1999) investigates demand for electricity for Switzerland using data of 48 cities form 1987-1990. Using the variables consumption of electricity per city in Kwh, electricity price index, personal income, household size, number of households in city, heating degree days and dummy for all households who face two part time tariffs, the OLS model estimates income and price elasticities as 0.391 and -0.595 respectively.

However, there is a scarcity of research on energy demand in developing countries and only a few of the studies accounted for the time-series properties of the response of energy consumption to changes in income and relevant prices. Amusa et al. (2009) analyzed the determinants of aggregate demand for electricity in South Africa by using bounds testing approach in an autoregressive distributed lag framework during the period 1960 to 2007. The results showed that demand for electricity was greatly affected by changes in income. However, the study found that changes in price of electricity had no effect on the demand for electricity.

In a recent study, a diversity of approaches to the estimation of electricity demand can be found in the literature ranging from aggregative analysis of the relationship between electricity demand, income and prices (Narayan et al., 2007; Lin, 2003, Holtedahl and Joutz, 2004), to more detailed disaggregated analysis (Bose and Shukla, 1999;) based on simultaneous model structure. In the most basic model, the demand for electricity, has been modeled as a function of a single variable, such as real income (Dincer and Dost, 1997) or temperature (Al-Zayer and Al-Ibrahim, 1996); real income and prices (Houthakker, et al., 1974; Zachariadis and Pashourtidou, 2006, Ziramba 2008) real income, real electricity price and price of natural gas (Narayan et al., 2007); real income, electricity prices, population growth, structural changes in the economy and efficiency improvement (Lin, 2003); real income, price of electricity and diesel (used in for captive power generation to meet the shortages), and reliability of power supply from utilities (Bose and Shukla 1999); real income, the real price of electricity, and the variable that captures the seasonal component of the demand for electricity (Chang and Martinez-Combo, 1997).

Using the partial adjustment model approach and simultaneous equation approach Kamerschena, and Porter (2004), investigate the electricity demand for period 1973-1998 for U.S. Estimates of price elasticities vary from -0.85 to -0.94 in 3SLS. However, estimates of price elasticities by partial-adjustment model shows biased results since the problem of endogeneity is not taken into consideration. Narayan and Smyth (2005) uses two models to estimate electricity demand in Australia. In model 1 the natural logs of levels of the variables are taken, whereas in model 2 the natural log of the ratio of the real price of electricity to the real price of natural gas, per capita electricity consumption, real per capita income, and temperature are used. In model 1 income and own price elasticities in short-run are 0.0121 and -0.263 respectively where as for long-run they are 0.323 and -0.541 respectively. In model 2 income elasticities in short-run and long-run are 0.0415 and 0.408 respectively. The relative price variable, in both-short and long-run, is significant at 1 percent. Erdogdu (2007) can also be separated into two parts. In the first part, the Partial adjustment model is employed with quarterly data including real GDP per capita, real price of electricity and net total electricity consumption per capita between 1984 and 2004. Erdogdu found that the price elasticity is not significantly different from zero and that the short and long- run income elasticities are 0.01 and 1.09 respectively. In the second part, Erdogdu undertook ARIMA modelling with a simple regression model for his forecast and annual data for electricity consumption data for the period 1923 to 2004. He concluded that there will be an annual 3.3% percent increase until 2014 with electricity consumption gradually increasing to about 156 thousand GWh in 2010 and about 160 thousand GWh in 2014.

3 Data and Methodology

3.1 Data and Variables of the Study

Data for the study were obtained from Ethiopian Electric Power Corporation (EEPCo), Ministry of Finance & Economic Development (MOFED), Central Statistical Agency (CSA) and World Development Indicator. The annual time series data on real electricity price, real GDP per capita, population growth rate and net electricity consumption per capita for the period 1976-2010 (a total of 34 observations on each variable) were used for estimation process.

The annual net electricity consumption and real price of electricity are obtained from EEPCo. The Annual electricity consumption is divided by population figures to get net electricity consumption per capita. The annual population and population growth rate are obtained from CSA and World development indicators. The annual time series data on Ethiopia real Gross Domestic Product (GDP) per capita at current prices is obtained from MOFED and World development indicators.

The variables of interest in this study are net electricity consumption per capita (E_t) as dependant variable and real GDP per capita (Y_t), real electricity price (P_t) and population growth rate (PG_t) as independent variables.

3.2 Methodology

The methodology adopted in this study follows the co-integration approach. A preliminary step in testing co-integration is to incorporate the unit root tests such as the Augmented Dickey-Fuller (ADF) (Dickey and Fuller, 1981) and Phillips-Perron (1988) tests to determine the degree of integration of the relevant variables and Co-integration Regression Durbin-Watson test to establish variables that have long-run relationship, followed by Partial adjustment model to obtaining the short and long-run behavior of electricity demand in Ethiopia. Finally, electricity demand forecast can be conducted by using autoregressive integrated moving average (ARIMA) model.

3.2.1 Co-integration Analysis

3.2.1.1 Stationarity

A stochastic process is said to be stationary if its mean and variance are constant over time and the value of the covariance between the two time periods depends only on the distance or gap between the two time periods and not the actual time at which the covariance is computed. Stationarity may be strong (i.e. the whole distribution of the variable does not depend on time) or weak stationary. A time series is said to be weak stationary if, $E(X_t)$ is constant and independent of time, $Var(X_t)$ is a finite, positive constant and independent of time and $Cov(X_t, X_k)$ is a finite function of $t - k$ but not of t or k.

In practical terms, it is difficult to test strong stationarity, especially with the short-runs of data. A more useful concept is that of weak stationary. This only requires that the mean, variance and covariances are independent of time.

3.2.1.2 Unit-Root Problem

The assumption of stationarity is somewhat unrealistic situation in most macroeconomic variables. Trivially, a non-stationary process arises when one of the conditions for stationarity does not hold. Why do we care for non-stationarity? With non- stationary series the effect of a shock never dies away and it leads to spurious regressions; that is one can regress completely unrelated series and find high R^2 and the standard tests are not valid.

The first step for an appropriate analysis is to determine if the data series are stationary or not. Time series data generally tend to be non-stationary. Due to the non-stationarity, regressions with time series data are very likely to result in spurious results. The problem stemming from spurious regression has been described by Granger and Newbold (1974). In order to ensure the condition of stationarity, a series ought to be integrated to the order of 0 ($I(0)$).

A series which is stationary without making any difference is said to be integrated of order 0 and is denoted I(0) and a series which is stationary after being differenced once is said to be integrated of order 1 and is denoted I(1). In general, a series which is stationary after being differenced d times is said to be integrated of order d and is denoted I(d). A series which is I(1) is also said to have a unit-root whereas I(d) has d roots.

Unit-root problem can be solved or stationarity can be achieved. Differencing techniques are normally used to transform a time series from a non-stationary to stationary by subtracting each datum in a series from its predecessor. As such the set of observations that correspond to the initial time period (t) when the measurement was taken is described as the series level. Differencing a series using differencing operations produces other sets of observations such as the first-differenced values, the second-differenced values and so on.

There are a number of unit root tests, but the most notable and commonly used are Augmented Dickey-Fuller and Phillips-Perron unit-root tests which are discussed in the following sections.

3.2.1.2.1 The Augmented Dickey-Fuller (ADF) Unit-Root Test

Generally, many time series data have a more complicated dynamic structure than is captured by a simple AR (1) model. Said and Dickey (1984) augment the basic autoregressive unit-root test to accommodate general ARMA (p, q) models with unknown orders and their test is referred to as the Augmented Dickey- Fuller (ADF) test. In such tests the null hypothesis states that a time series is I(1) while the alternative is I(0), assuming that the dynamics in the data have an ARMA structure. The ADF test is based on estimating the test regression

$$X_t = a_0 + a_2 t + \phi X_{t-1} + \sum_{j=1}^{p} \psi_j \Delta X_{t-j} + \varepsilon_t \qquad (1)$$

where a_0 is constant, $a_2 t$ is a trend, p is a lagged difference terms, ΔX_{t-j} are used to approximate the ARMA structure of the errors, and the value of p is set so that the error ε_t is serially uncorrelated. The error term is also assumed to be homoskedastic. Under the null hypothesis, X_t is I(1) which implies that $\phi = 1$. The ADF t-statistic and normalized bias statistic are based on the least squares estimates of (1) and are given by

$$ADF_t = t_{\phi=1} = \frac{\hat{\phi} - 1}{SE(\phi)}$$

$$ADF_n = \frac{T(\hat{\phi} - 1)}{1 - \hat{\psi}_1 - \hat{\psi}_2 - \cdots - \hat{\psi}_p}$$

An alternative formulation of the ADF test regression is

$$\Delta X_t = a_0 + a_2 t + \pi X_{t-1} + \sum_{j=1}^{p} \psi_j \Delta X_{t-j} + \varepsilon_t \quad (2)$$

where $\pi = \phi - 1$, the data needs to be differenced to make it stationary (unit-root problem/ non- stationary). The ADF test statistic is then the usual t-statistic for testing $\pi = 0$ and the ADF normalized bias statistics is

$$ADF_n = \frac{T\hat{\pi}}{1 - \hat{\psi}_1 - \hat{\psi}_2 - \cdots - \hat{\psi}_p}$$

The test regression (2) is often used in practice because the ADF t-statistic is the usual t-statistic reported for testing the significance of the coefficient X_{t-1}. ADF_t and ADF_n have the same asymptotic distribution as $t_{\pi=0}$ and $T\hat{\pi}$ under white noise errors provided p is selected approximately.

3.2.1.2.2 Phillips-Perron (PP) Unit-Root Test

Phillips and Perron (1988) developed a number of unit-root tests that have become popular in the analysis of financial time series. 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

$$\Delta X_t = a_0 + a_2 t + \pi X_{t-1} + \varepsilon_t \quad (3)$$

Where ε_t is $I(0)$ and may be heteroskedastic. The PP tests correct for any serial correlation and heteroskedasticity in the errors ε_t of the test regression by directly modifying the test statistics $t_\pi = 0$ and $T\hat{\pi}$. These modified statistics, denoted Z_t and Z_π are given by

$$Z_t = \left(\frac{\hat{\sigma}^2}{\hat{\lambda}^2}\right)^{1/2} . t_{\pi=0} - \frac{1}{2}\left(\frac{\hat{\lambda}^2 - \hat{\sigma}^2}{\hat{\lambda}^2}\right) . \left(\frac{T.SE(\hat{\pi})}{\hat{\sigma}^2}\right)$$

$$Z_\pi = T\hat{\pi} - \frac{1}{2}\left(\frac{T^2.SE(\hat{\pi})}{\hat{\sigma}^2}\right)(\hat{\lambda}^2 - \hat{\sigma}^2)$$

The terms $\hat{\sigma}^2$ and $\hat{\lambda}^2$are consistent estimates of the variance parameters

$$\sigma^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E[\varepsilon_t{}^2]$$

$$\lambda^2 = \lim_{T\to\infty} T^{-1} \sum_{t=1}^{T} E[S_T{}^2]$$

Where $S_T = \sum_{t=1}^{T} \varepsilon_t$. The sample variance of the least square residual $\hat{\varepsilon}_t$ is a consistent estimate of σ^2, and the Newey-West long-run variance estimate of ε_t using $\hat{\varepsilon}_t$ is a consistent estimate of λ^2.

Under the null hypothesis that $\pi = 0$, the PP Z_t and Z_π statistics have the same asymptotic distributions as the ADF t-statistic and normalized bias statistics. One advantage of the PP tests over the ADF tests is that the PP tests are robust to general forms of heteroskedasticity in the error term ε_t. Another advantage is that the user does not have to specify a lag length for the test regression.

3.2.1.3 Co-integration

Two or more variables are said to be co-integrated if they share common trends i.e. they have long-run equilibrium relationships. The technique of co-integration involves three steps. The first step requires a determination of the order of integration of the variables of interest using Augmented Dickey-Fuller and Phillips-Perron tests. In the second step, the co-integration regression using variables having the same order of integration is estimated. In the third step, residuals from the co-integration are subjected to test. The presence of co-integration is an evidence of a long-run equilibrium relationship between the variables. In this study, Co-integration-Regression Durbin-Watson test is applied to check the presence or absence of long-run relationship between the variables.

One method of finding out whether the variables are co-integrated or not is the Co-integration Regression Durbin-Watson (CRDW) test, whose critical values were first provided by Sargan and Bhargava (1983). In CRDW we use the Durbin-Watson d obtained from the co-integrating regression. Here, the null hypothesis is that $d = 0$ rather than the standard $d = 2$. This is because we know that $d \approx 2(1 - \hat{\rho})$, so if there is to be a unit root, the estimated ρ will be about 1, which implies that d will be about zero. Gujarati (1995) notes that the 1, 5 and 10 percent critical values to test the hypothesis that the true $d = 0$ are 0.511, 0.386 and 0.322, respectively. Thus, if the computed d value is smaller than, say, 0.511, we reject the null hypothesis of co-integration at the one percent level. Otherwise, do not reject the null, meaning that the variables in the model are co-integrated and there is a long term or equilibrium relationship between the variables.

3.2.2 Partial Adjustment Model

Partial adjustment model is a dynamic model that the variable of interest may be influenced by the value of the same variable in the previous time period. In these cases the explanatory variable determines the dependant variable with a lag. It is used in models where adjustment does not occur immediately, but takes a number of time periods to fully adjust (Houthakker et al. 1974). According to Narayan and Smyth (2005) and others specify, the derivation of partial adjustment electricity demand begins with the static representation. The simplest model can be written as

$$logE_t = \alpha + \beta_1 logP_t + \beta_2 logY_t + \beta_3 logPG_t + \varepsilon_t \qquad (4)$$

Where $logE_t$ is the natural logarithm of electricity consumption per capita, $logP_t$ is the natural logarithm of real electricity price, $logY_t$ is the natural logarithm of real income per capita, $logPG_t$ is the natural logarithm of population growth rate, ε_t is the random error, α is intercept term, and β_1, β_2 and β_3 are parameters to be estimated.

Income is considered to be the most important determinant for electricity consumption in the literature. Economic growth and its impact on living standards is the main driving force of electricity consumption growth. Therefore, higher real per capita income will increase purchases of electrical equipment and hence increase electricity demand. Therefore, income variables are expected to have positive sign. Nevertheless, an increase in the price of electricity will cause the electricity demand to decrease. That is price variables are expected to have negative sign. Population is another important factor to determine electricity demand in Ethiopia. Higher population level is expected to increase electricity consumption. A positive correlation between population growth and electricity demand is therefore expected.

In order to capture short and long-run elasticities of electricity demand separately a dynamic version of reduced form model, called "partial adjustment model", is used. Suppose that "desired" (or targeted) electricity demand E_t^* is determined by the model.

$$logE_t^* = \alpha + \beta_1 logP_t + \beta_2 logY_t + \beta_3 logPG_t + \varepsilon_t \tag{5}$$

And the actual electricity consumption is subjected to adjustment process

$$logE_t - logE_{t-1} = \vartheta(logE_t^* - logE_{t-1}) \tag{6}$$

Where $logE_t$ is the natural logarithm of electricity consumption per capita in period t, $logE_{t-1}$ is the natural logarithm of electricity consumption per capita in period t-1, $logE_t^*$ is the natural logarithm of targeted(desired) level of electricity consumption per capita and ϑ is the speed of adjustment coefficient ($0 < \vartheta < 1$). If ϑ is 0, it means there is no adjustment from one time period to the next and if ϑ is 1, it means adjustment from one time period to the next is instantaneous. The smaller the value of ϑ the greater the adjustment lags. The left side of equation (6) is the actual change and also the right side is the desired change. Equation (6) postulates that the current level, $logE_t$ will move only partially from the previous position, $logE_{t-1}$ to the desired level, $logE_t^*$. The amount of the adjustment between time t and t-1 is equal to $\vartheta(logE_t^* - logE_{t-1})$.

Substituting (5) in (6) to yield our partial adjustment model

$$logE_t - logE_{t-1} = \vartheta(logE_t^* - logE_{t-1})$$

$$logE_t = \vartheta\alpha + \vartheta\beta_1 logP_t + \vartheta\beta_2 logY_t + \vartheta\beta_3 logPG_t + \vartheta\varepsilon_t - \vartheta logE_{t-1} + logE_{t-1}$$

$$logE_t = \vartheta\alpha + \vartheta\beta_1 logP_t + \vartheta\beta_2 logY_t + \vartheta\beta_3 logPG_t + (1-\vartheta)logE_{t-1} + \vartheta\varepsilon_t \qquad (7)$$

Where $\vartheta\beta_1$, $\vartheta\beta_2$ and $\vartheta\beta_3$ are the short-run price, income and population growth rate elasticities of electricity demand respectively. The parameters β_1, β_2 and β_3 are long-run price, income and population elasticities of electricity demand. $\vartheta\varepsilon_t$ is random error $\alpha, \beta_1, \beta_2, \beta_3$ and ϑ can be estimated by using ordinary least squares (OLS) method.

3.2.3 Autoregressive Integrated Moving Average (ARIMA) Modeling

ARIMA stands for autoregressive-integrated-moving average model. In statistics, ARIMA models are typically applied to time series data for forecasting. An autoregressive integrated moving average (ARIMA) model is a univariate time series model. Univariate means that the model involves only one variable. Different ARIMA models are identified by the number of autoregressive parameters (p), the degree of differencing (d) and the number of moving average parameters (q). In short form, we write as ARIMA(p, d, q). If there is no differencing term, i.e $d = 0$, we call it ARMA (p, q). The ARMA (p, q) process is a combination of the p^{th} order autoregressive and q^{th} order moving average processes. An ARMA (p, q) process can often describe stationary time series with fewer parameters than either a MA or AR process by itself. The ARMA (p, q) is given by

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \quad (8)$$

Or equivalently by

$$\emptyset(B) X_t = \theta(B) \varepsilon_t \quad (9)$$

Where $\emptyset(B) = 1 - \emptyset_1 B - \emptyset_2 B^2 - \cdots - \emptyset_p B^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ and B is the backward shift operator.

The general forms of the ARIMA (p, d, q) can be written as described by Judge et al. (1988).

$$w_t = \emptyset_1 w_{t-1} + \emptyset_2 w_{t-2} \ldots + \emptyset_p w_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \ldots - \theta_q \varepsilon_{t-q} \quad (10)$$

That is

$$\Delta^d X_t = \emptyset_1 \Delta^d X_{t-1} + \emptyset_2 \Delta^d X_{t-2} + \cdots + \emptyset_p \Delta^d X_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \cdots - \theta_q \varepsilon_{t-q} \quad (11)$$

equivalently,

$$\emptyset(B)\, W_t = \theta(B) \varepsilon_t \quad (12)$$

where X_t= the actual values of the series at time t

X_{t-i}=the value of the series at time t-i

ε_t =the error term at time t

ε_{t-i} =the error term at time t-i

$\emptyset_i$ = the Autoregressive parameter

θ_i =The moving average parameter

$w_t = \Delta^d X_t$, the X_t series differenced d times

Where $\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}$ ... are uncorrelated each other and

There are some basic steps to construct ARIMA model. These are: identification of the model, estimation of the unknown parameters, diagnostic checking and forecasting.

3.2.3.1 Identification of the model

The first step in ARIMA modelling is identification of the appropriate model. ARIMA procedures are mainly based on the analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which are essential diagnostic instruments to identify the dependence structure of the series. Autocorrelation is the correlation of a series with itself, lagged by a particular number of observations. The partial autocorrelation is the partial correlation of a series with itself, lagged by a particular number of observations and controlling for all correlations for lags of lower order.

The autocorrelation of a series X at lag k is estimated by

$$\rho_k = \frac{cov\{X_t, X_{t-k}\}}{var\{X_t\}} = \frac{\sum_{t=k+1}^{n}(X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^{n}(X_t - \bar{X})^2} \tag{13}$$

The estimates of partial autocorrelation at lag k by

$$\phi_k = \begin{cases} \rho_1 & for\ k = 1 \\ \frac{\rho_k - \sum_{j=1}^{k-1}\phi_{k-1,j}\,\rho_{k-j}}{1-\sum_{j=1}^{k-1}\phi_{k-1,j}\,\rho_{k-j}} & for\ k > 1 \end{cases} \tag{14}$$

Where ρ_k is the estimated autocorrelation at lag k and where,

$$\phi_{k,j} = \phi_{k-1,j} - \phi_k \phi_{k-1,k-j}$$

The first step in determining the AR and MA is to ensure that the series is stationary. Mathematically, the stationarity of a time series refers to the time invariance of the data generating process, revealed by the stability over time of mean and variance.

Stationarity with respect to the mean implies that time series fluctuate around a constant value over time. Thus, the autocorrelation plot of a stationary variable will usually decay into noise and/or hit negative values within three or four lags. The general shape of a non-stationary series presents a trend over time. The presence of positive and persistent autocorrelations implies the need to introduce at least one difference term into the model. The differencing operation is realized by subtracting the previous value to the current one.

Taking difference may eliminate the trend and make the series stationery. To determine whether stationarity has been achieved, one may examine the ACF of the differenced series that should converge quite rapidly to zero as the value of the lag increases. Moreover, to determine whether taking difference is appropriate, one may refer to the examination of the Box-Ljung statistics (Q-statistics). Ljung-Box Q-statistics can be computed as

$$Q_{LB} = n(n+2)\sum_{j=1}^{k} \frac{\hat{\rho}_j^{\,2}}{n-J} \tag{15}$$

Where ρ_j is the j^{th} autocorrelation and n is is the number of observation. If the series is not based upon the results of ARIMA estimation, then under the null hypothesis Q is asymptotically distributed as a X^2 with degrees of freedom equal to the number of autocorrelations. If the series is represents the residuals from ARIMA estimation, the appropriate degrees of freedom should be adjusted to represent the number of autocorrelations less the number of AR and MA terms previously estimated. The Q-statistic at lag k is a test statistic for the null hypothesis that there is no autocorrelation up to order k.

3.2.3.2 Estimation of ARIMA model

Having identified the appropriate p and q values, the next stage is to estimate the parameters of the autoregressive and moving average terms included in the model. We can estimate the unknown parameters by ordinary least squares or by maximum likelihood methods.

3.2.3.3 Diagnostic checking

Having chosen a particular ARIMA model, and having estimated its parameters, we next see whether the chosen model fits the data reasonably well, for it is possible that another ARIMA model might do the job as well. One simple test of the chosen model is to see if the residuals estimated from this model are white noise; if they are, we can accept the particular fit; if not, we must start over or check whether any of the autocorrelations and partial correlations of the residuals is individually significant or not. If they are not statistically significant, then it means that the residuals are purely random and there is no need to respecify of ARIMA model.

3.2.3.4 Forecasting

The last step in ARIMA modeling is forecasting. One of the reasons for the popularity of the ARIMA modeling is its success in forecasting.

3.2.3.4.1 Minimum mean square error forecasts

The general ARIMA(p, d, q) model can be written as

$$\varphi(B)X_t = \theta(B)\varepsilon_t \tag{16}$$

Where $\varphi(B)$=$\emptyset(B)\Delta^{\mathrm{d}}$. Forecasting a value $X_{t+l}, l \geq 1$, when we are currently standing at time t is said to be made at origin t for lead time l. The three explicit forms of the model that an observation X_{t+l} generated by the process (16) may be expressed as follows (Box and Jenkins, 1976)

(1) Directly in terms of the difference equation by

$$X_{t+l} = \emptyset_1 X_{t+l-1} + \emptyset_2 X_{t+l-2} + \cdots + \emptyset_{p+d} X_{t+l-p-d} - \theta_1 \varepsilon_{t+l-1} - \cdots - \theta_q \varepsilon_{t+l-q} + \varepsilon_{t+1} \quad (17)$$

(2) Infinite weighted sum of current and previous shocks ε_j

$$X_{t+l} = \sum_{j=\infty}^{t+l} \psi_{t+l-j} \varepsilon_j = \sum_{j=o}^{\infty} \psi_j \varepsilon_{t+l-j} \quad (18)$$

Under stability condition and $\sum_{j=0}^{\infty} {\psi_j}^2 < \infty$

Where $\psi_0 = 1$ and ψ weight may be obtained by equating coefficient in

$$\emptyset(B)(1 + \psi_1 B + \psi_2 B^2 + \cdots) = \theta(B)$$

(3) Infinite weighted sum of previous observations, plus a random shock

$$X_{t+l} = \sum_{j=1}^{\infty} \pi_j X_{t+l-j} + \varepsilon_{t+l} \quad (19)$$

$\sum_{j=1}^{\infty} \mathrm{I}\pi_j\mathrm{I} = 1$, π weights may be obtained from

$$\emptyset(B) = (1 - \pi_1 B - \pi_2 B^2 - \cdots)\theta(B)$$

Standing at origin t, we are to take a forecast $\hat{X}_t(l)$ of X_{t+l} which is to be a linear function of current and previous observations of $X_t, X_{t-1}, \ldots.$ then it will also be a linear function of current and previous shocks $\varepsilon_t, \varepsilon_{t-1}, \ldots.$

The minimum mean square error forecast $\hat{X}_t(l)$ for lead time l is the conditional expectation of $E[\,X_{t+l}]$, of X_{t+l}, at the origin t. For simplicity in notation the square brackets imply that the conditional expectation is taken at time t. Thus,

$[\varepsilon_{t+l}] = E[\varepsilon_{t+l}]$, $[X_{t+l}] = E[X_{t+l}]$, for $l > 0$,the three different ways of expressing the forecasts are:

Forecasts from difference equation form: taking conditional expectations at time t in (17), we obtain

$$[X_{t+l}] = \hat{X}_t(l) = \emptyset_1[X_{t+l-1}] + \emptyset_2[X_{t+l-2}] + \cdots + \emptyset_{p+d}[X_{t+l-p-d}] - \theta_1[\varepsilon_{t+l-1}] - \cdots - \theta_q[\varepsilon_{t+l-q}] + [\varepsilon_{t+1}] \quad (20)$$

Forecast from integrated form: using (18) we get

$$[X_{t+l}] = \hat{X}_t(l) = \sum_{j=o}^{\infty} \psi_j[\varepsilon_{t+l-j}] \quad (21)$$

Forecasts as a weighted average of previous observations and forecasts made at previous lead times from the same origin: taking conditional expectation in (19) we obtain

$$[X_{t+l}] = \hat{X}_t(l) = \sum_{j=1}^{\infty} \pi_j[X_{t+l-j}] + [\varepsilon_{t+l}] \quad (22)$$

To calculate the conditional expectations we have to use the following expressions:

$[X_{t-j}] = E[X_{t-j}] = X_{t-j}$ $\quad j = 0, 1, 2 \ldots$ which have already happened at the origin t, are left unchanged.

$[X_{t+j}] = E[X_{t+j}] = \hat{X}_t(j)$ $\quad j = 1, 2, \ldots$ which have not yet happened, are replaced by their forecast.

$[\varepsilon_{t-j}] = E[\varepsilon_{t-j}] = \varepsilon_{t-j}$ $\quad j = 0, 1, 2, \ldots$ which have happened, are available from the actual vales minus the fitted. That is $\varepsilon_{t-j} = X_{t-j} - \hat{X}_{t-j}$

$[\varepsilon_{t+j}] = E[\varepsilon_{t+j}] = 0$ $\quad$ j=1, 2, 3… which have not yet happened, are replaced by zeros.

3.2.3.4.2 Forecasting accuracy measures

Once forecasts are made they can be evaluated if the actual values of the series to be forecasted are observed. There are some measurements of the accuracy of forecasts. The criteria used are standard measures which summarise the errors generated by a particular forecasting technique. These are root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and Theil's inequality coefficient (Theil-U).

The first two forecast error statistics depend on the scale of the dependent variable. These should be used as relative measures to compare forecasts for the same series across different models; the smaller the error, the better the forecasting ability of that model according to that criterion. The remaining two statistics are scale invariant. Therefore, this study used mean absolute percentage error and Theil's inequality coefficient.

3.2.3.4.2.1 Mean Absolute Percentage Errors (MAPE)

The mean absolute percentage error is the mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign so as to avoid the problem of positive and negative values cancelling one another (Makridakis and Wheelwright 1998). MAPE can be calculated as

$$PE_t = \left(\frac{X_t - F_t}{X_t}\right) * 100$$

$$MAPE = \frac{1}{\mathrm{n}}\sum_{t=1}^{\mathrm{n}} |PE_t| \tag{23}$$

Where PE is percentage error, X_t is actual observation for time period t, F_t is forecast for the same period and MAPE is mean absolute percentage error. MAPE less than 4% indicates good forecasting ability.

3.2.3.4.2.2 Theil's inequality coefficient (U-Statistics)

Theil's U statistic according to Greene (2000) is related to the coefficient of determination (R^2) but it isn't bounded between 0 and 1. Like MAPE statistics, high values suggest poor performance in the forecast; however, and unlike MAPE, the Theil-U corrected the performance scale that MAPE had. Theil's U can be estimated as

$$U = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(X_t - F_t)^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n}F_t^{\,2}} + \sqrt{\frac{1}{n}\sum_{t=1}^{n}X_t^{\,2}}} \tag{24}$$

The scaling of U is such that it will always lie between 0 and 1. If $U = 0$, $X_t = F_t$ for all forecasts and there is a perfect fit; if U = 1 the predictive performance is as bad as it possibly could be.

The Theil's inequality coefficient can be decomposed in to bias, variance and covariance proportions.

$$\text{Bias Proportion} = \frac{(\bar{F} - \bar{X})^2}{\sum_{t=1}^{n}(X_t - F_t)^2 / n} \tag{25}$$

$$\text{Variance Proportion} = \frac{(S_F - S_X)^2}{\sum_{t=1}^{n}(X_t - F_t) / n} \tag{26}$$

$$\text{Covariance Proportion} = \frac{2(1 - \mathrm{r})S_X S_F}{\sum_{t=1}^{n}(X_t - F_t)^2 / n} \tag{27}$$

Where $\bar{F}$, $\bar{X}$, S_F and S_X are means and standard deviations of F and X respectively, and r is the correlation between F and X.

The bias and the variance proportions show how far the mean of the forecast series is from the mean of the actual series and how far the variation of the forecast is from the variation of the actual series, respectively. The covariance proportion measures the remaining asymmetric forecast errors. The sum of these three measures would be one. If the forecast is good, the bias and the variance proportions should be small.

4 Data Analysis

4.1 Descriptive Analysis

EViews 5.0 the windows-based forecasting and econometric analysis package was used to estimate the demand for electricity in Ethiopia. The analysis used time series yearly data.

Before going to present inferential statistics it is better to start with descriptive analysis. In our study, the data consist of yearly net electricity consumption per capita, real price of electricity, gross domestic product per capita and population growth rate from the period 1976-2010. From Appendix 1, the time series plots of net electricity consumption per capita, real price of electricity and gross domestic product per capita show an increasing trend but population growth rate increased from 1979- 1992, after that to some extent decreased. As shown in table 9, the average electricity consumption per capita in Ethiopia is 21.03 KW, the average real electricity price is 0.27 birr/KW, the average population growth rate is 2.80 and the average real income per capita is 985.22 birr in the last three decades. As can be seen in the same table, there is the highest variability in real price of electricity as compared to the other which is about 54.52% and also smallest variability in the population growth rate which is 7.95%.

4.2 Integration and Co-integration Analysis

In this section we discuss first, unit-root tests and the order of integration of variables by using Augmented Dickey- Fuller and Phillips-Perron tests. Then, test for co-integration of the variables were conducted by using Co-integration Regression Durbin-Watson test with a view to identify the existence of a long-run relationship.

4.2.1 Tests for Unit-Root

The time series plots of real electricity prices (P), real GDP per capita (Y), net electricity consumption per capita (E) and population growth rate (PG) at level form reveal that there is a trend in all variables. The plots also indicate that real electricity prices, real GDP per capita, net electricity consumption per capita and population growth rate have non-constant means and non-constant variances; that is, they seem to be non-stationary. But they seem to be stationary at the first difference (Appendix 1). The formal tests of unit-root have also been used to check whether the variables are stationary at level or after differencing, and for identifying the order of integration.

4.2.1.1 Augmented Dickey-Fuller Test (ADF)

The estimated standard procedure for co-integration analysis is to start with unit-root tests on the time series data that are being analyzed. The null hypothesis of the test is that there is a unit-root problem in the variables. Since equation (7) implies that the electricity demand in time t is affected by the first lags of the variables; one lags have been used in ADF unit root tests.

Table-2 below summarizes the results of ADF unit-root test for the variables at level form (without differencing).

Table-2: Summary of ADF unit-roots test in the variables (in level form with a trend and intercept)

Variable	ADF test statistic (t-statistic)	1% crit. Value	5% crit. Value	10% crit. Value	Results
LogE	-2.175469	-4.252879	-3.548490	-3.207094	do not reject
LogP	-3.017138	-4.262735	-3.552973	-3.209642	do not reject
LogY	-1.287734	-4.252879	-3.548490	-3.207094	do not reject
LogPG	0.055646	-4.296729	-3.568379	-3.218382	do not reject

Since the computed ADF test statistics (t-statistic) for the natural logarithms of net electricity consumption per capita, real electricity prices, income per capita and population growth rate are greater than the critical values-" tau statistic" at 1% significance level, we cannot reject the null hypothesis. That means there is a unit-root problem in the variables. Based on ADF test, the variables are non-stationary at level form (Appendix 2-A to Appendix 2-D). However, we can transform the time series data from non-stationary to stationary by taking difference of the variables. Table 3 below summarizes the result of ADF unit-root test for the first differenced variables.

Table-3: Summary of ADF tests for unit-roots in the variables (after first difference with a trend and intercept)

Variable	ADF test statistic (t-statistic)	1% crit. Value	5% crit. Value	10% crit. Value	Results
ΔLogE	-5.831010	-4.262735	-3.552973	-3.209642	reject the null
ΔLogP	-4.407153	-4.262735	-3.552973	-3.209642	reject the null
ΔLogY	-5.441581	-4.252879	-3.548490	-3.209642	reject the null
ΔLogPG	-9.957219	-4.296729	-3.568379	-3.218382	reject the null

The computed ADF test statistics (t-statistic) for the natural logarithm of all first differenced variables are smaller than the critical value-"tau statistics" at 1% significance level, which leads to the rejection of the null hypothesis that there is a unit-root problem in the variables. This means that the variables do not have a unit-root problem and all variables are stationary at first difference (Appendix 2- E to Appendix 2-H).

4.2.1.2 Phillips-Perron Test (PP)

To make sure that the result of ADF test for unit-root is viable, we need also see other unit-root test namely Phillips-Perron unit-root test. The null hypothesis of the test again states that there is a unit-root problem in the variables. Table 4 below summarizes the results of Phillips-Perron unit-root test at level form (without differencing).

Table-4: Summary of Phillips-Perron tests for unit-roots in the variables (in level form with a trend and intercept)

Variable	PP test statistic (Adj. t-Stat)	1% crit. Value	5% crit. Value	10% crit. Value	Results
LogE	-2.175469	-4.252879	-3.548490	-3.207094	do not reject
LogP	-2.206202	-4.252879	-3.548490	-3.207094	do not reject
LogY	-1.353747	-4.252879	-3.548490	-3.207090	do not reject
LogPG	-1.705377	-4.273277	-3.557759	-3.212361	do not reject

The Phillips-Perron statistics (Adj. t-Stat) are all insignificant at 1% level of significance, which leads to the non-rejection of the null hypothesis that there is a unit-root problem in the variables. Based on the PP test it is obvious that the variables are non-stationary at level form (Appendix 2-I to Appendix 2-L). Here also, differencing has the effect of making non-stationary variables stationary. Table 5 below summarizes the results of Phillips-Perron unit-root test for the first differenced variables.

Table-5: Summary of Phillips-Perron tests for unit-roots in the variables (after first difference with a trend and intercept)

Variable	PP test statistic (Adj. t-Stat)	1% crit. Value	5% crit. Value	10% crit. Value	Results
ΔLogE	-5.991758	-4.262735	-3.552973	-3.209642	reject the null
ΔLogP	-4.428630	-4.262735	-3.552973	-3.209642	reject the null
ΔLogY	-5.212105	-4.262735	-3.552973	-3.209642	reject the null
ΔLogPG	-9.709399	-4.296729	-3.568379	-3.218382	reject the null

Since the computed PP statistics for the natural logarithms of all first differences of net electricity consumption per capita, real electricity prices, income per capita and population growth rate are smaller than the critical value- "tau statistics" at 1% significance level, we reject the null hypothesis that there is a unit-root problem in the variables. This means that unit-root does not exist in the first differenced variables, i.e. the variables are stationary at first difference (Appendix 2-M to Appendix 2-P).

Therefore, based on the ADF and PP unit-root test results which indicates that first differenced variables are stationary, it can be stated that all variables are integrated of order one, I(1).

4.2.2 Test for Co-integration

If an OLS regression is estimated with non-stationary data, then there is a possibility that the regression is spurious (nonsense), meaning that one can regress completely unrelated series and find high R^2, the standard tests are not valid (t-ratios does not have t distribution, F- statistics doesn't have an F distribution). In general, standard statistical inference does not make any sense. But still the regression is meaningful when the variables are co-integrated. Therefore, Co-integration Regression Durbin-Watson test is carried out to make sure whether the variables are co integrated or not.

From the partial adjustment model, equation (7) the calculated Durbin-Watson statistics is 1.646202, which is above the 1% critical value of 0.511. Therefore, do not reject the null hypothesis of co-integration at 1% level of significance. Based on the CRDW test, the variables of $logE, logP, logY$ and $logPG$ are co-integrated, that is there exists a stable long-run relationship between them (Appendix 2-R).

4.3 Partial Adjustment Model

The fitted model for equation (4) is given by (Appendix 2-Q):

$$logE_t = -1.5924 + 0.2780logP_t + 0.7082logY_t + 0.0744logPG_t \tag{28}$$

The last column of estimation output gives the probability of drawing a t-statistic. It is also known as the "p-value". The estimated coefficients of constant term and population growth rate are not significant (the p-value of constant term is 0.1938 and the p-value of population growth rate is 0.6838) but the coefficients of price and income are significant. Price has a positive sign which is not supported by the demand theory, but income has the correct sign. The economic theory states that there is a positive relationship between demand and income and an inverse relationship between demand and price.

"R-squared" measures the success of the regression in predicting the values of the dependant variable within the sample. "R-squared" equals one if the regression fits perfectly. In our model, it is 0.92 and therefore it can be concluded that the model fits the data well. Since "R-squared" never decreases as more regressors are added; "Adjusted R-squared", which penalizes for addition of irrelevant regressors, is a better measure of "goodness-of-fit". In our model, "Adjusted R-squared" is almost 0.917 and, therefore, it is also above the expected value, which is above around 0.80.

Durbin-Watson statistic measures the serial correlation (or, autocorrelation), AR (1), in the residuals. As a rule, if it is less than 2, there is an evidence of positive serial correlation. Durbin-Watson statistic in our estimation output is about 0.747036. The low values of the Durbin-Watson statistic is indicative of the presence of serial correlation left in the residuals of the estimated equation. If uncorrected, serial correlation in the residuals will lead to incorrect estimates of standard errors and invalid statistical inference for the coefficients of the equation. F-statistic and its p-values, Prob(F-statistic), come from a joint test of the null hypothesis that all slope coefficients in the regression are zero, we can reject the null hypothesis (Appendix 2-Q).

The above model does not show short and long-run parameters separately. In order to capture short and long-run elasticities of electricity demand separately a dynamic version of reduced form model, called "partial adjustment model" are used. The fitted model for equation (7) is given by (Appendix 2-R)

$$logE_t = -0.6513 + 0.0857logP_t + 0.2517logY_t + 0.1724logPG_t + 0.6724logE_{t-1} \tag{29}$$

The demand model estimated noticeably shows that the price variable still has not expected sign and also its coefficient are not significant even at 10% level of significance (the p-value of price is 0.1199) and also the constant term are not significant (the p-value of the constant term is 0.4800) but expectedly the income variable has a positive sign and the coefficient of income is significant at 10% level of significance, population growth rate is significant at 10% level of significance and previous period electricity consumption is significant at 1% level of significance. With regard to the fit, the regression as measured by R^2 and adjusted R^2, which are 0.96 and 0.95 respectively, a very good fit is attained.

4.3.1 Short and Long-Run Elasticity

Based on the estimated coefficients of the explanatory variables, we can measure the elasticity of the dependent variables. The consumption of electricity is used as a dependent variable. The estimated short-run and long-run elasticties of electricity demand are as follow

From equation (7) we can have the speed of adjustment (ϑ), short-run parameters ($\vartheta\beta_1$, $\vartheta\beta_2$, and $\vartheta\beta_3$) and long-run parameters (β_1, β_2 and β_3) of electricity demand in Ethiopia.

$\widehat{\vartheta\alpha}= -0.6513$, $\widehat{\vartheta\beta_1}=0.0857$ $\widehat{\vartheta\beta_2} = 0.2517$, $\widehat{\vartheta\beta_3} = 0.1724$, $(1-\hat{\vartheta}) = 0.6724$. From this $\hat{\vartheta} = 0.3276$ and $\hat{\beta}_1 = 0.2616$, $\hat{\beta}_2 = 0.7683$ and $\hat{\beta}_3=0.5263$

The coefficient of lagged electricity consumption $(1-\vartheta)$ which reflects as an adjustment parameter is equal to 0.6724. That is the speed of adjustment to the long-run equilibrium demand level equal to 0.3276, which is considerable high. This means that demand takes less time to reach long-run equilibrium. In other words, if any exogenous or external shock to the demand is going to happen, it gets adjusted fairly quickly towards the long-run equilibrium values.

4.3.1.1 Income Elasticity of Demand

Income variable is significant at 10% level of significance. The long-run elasticity is 0.7683 in the partial adjustment model while it is 0.2517 in the short-run. As expected, the long-run income elasticity is greater than the short-run one. The result indicates that electricity consumption is a normal good as it increases with income.

In keeping with the model, if real income increases by 100%, electricity demand increases by 25.17% and 76.83% in the short-run and long-run, respectively,

4.3.1.2 Price Elasticity of Demand

Against the priori expectation, price elasticity is positive and is not significant even at 10% level of significance. The long-run price elasticity is 0.2616, while the short-run value is 0.0857. According to the theory regarding price elasticity of electricity demand, an increasing price of electricity should lead to the reduction of the rate of utilization in the short-run. In the long-run people are expected to changes the composition of electricity appliance in such a way that demand for electricity further reduces. But our results do not support this expectation. This could be simply because the price of electricity is usually fixed by the government and reviewed infrequently and also the inadequate supply of electricity in Ethiopia.

4.3.1.3 Population Growth Rate Elasticity of Demand

The population variable is significant at 10% significance level. The long-run population growth rate elasticity is 0.5263 while the short-run population growth rate elasticity is 0.1724.

In keeping with the model, if population growth rate increases by 100%, electricity demand increases by 17.24% and 52.63% in the short-run and long-run, respectively, in Ethiopia.

In general, the effects of price, income and population growth rate are smaller in the short-run compared to the long-run. Elasticities of electricity demand in Ethiopia are relatively low meaning that the reaction or sensitivity of consumers' to price, income and population growth rate changes is low.

4.3.2 Diagnostic Checking

4.3.2.1 Test for Heteroskedasticity in the Error Term

Heteroskedasticity is violations of the assumption of the Classical Linear Regression Model (CLRM) that the error term has constant variance. In this study, White heteroskedasticity test (without cross terms) are carried out (Appendix 2-S). The top part of the output displays F-statistic and Obs* R-squared statistic. Obs* R-squared statistic is a White's test statistic. It is computed as the number of observations times R^2 from the test regression. White's test statistic is asymptotically distributed as a X^2 with degrees of freedom equal to the number of slop coefficients, excluding the constant. Under the null there is no heteroskedasticity in the error term since the probability value of 0.314398 is greater than 1% significance level. Therefore, we do not reject the null hypothesis of no heteroskedasticity in the error term.

4.3.2.2 Test for Serial Correlation

The Durbin-Watson statistic is not appropriate as a test for serial correlation; in this case, since there is a lagged dependant variable on the right hand side of the equation (7). Therefore, another test, namely Bruesch-Godfrey Serial correlation LM was applied (Appendix 2-T). The top part of the output represents the test statistics and associated probability values. The statistic labeled "Obs∗R-squared" is the LM test statistic for the null hypothesis of no serial correlation up to lag p. The LM statistics is computed as the number of observation times the R^2 from the test regression. The LM test statistic is asymptotically distributed as X^2 with p degrees of freedom (p is the number of lags, in our case p = 1). The probability value of 0.456720 indicates that we cannot reject the null hypothesis of no serial correlation.

4.3.2.3 Specification and Stability test

4.3.2.3.1 Ramsey's RESET Test

We carried out Ramsey RESET test. RESET stands for Regression Specification Error Test proposed by Ramsey (1969). Further assumption that we make about the model is that the functional form is correct. It is important to include all the relevant variables in the model, if we exclude an important explanatory variable; the regression has omitted variable bias. This means the estimates are unreliable and the t and F statistics cannot be relied on. Equally there can be a problem if we include variables that are not relevant, as this can reduce the efficiency of the regression; however, this is not as serious as the omitted variable bias. Therefore, Ramsey's RESET test is carried out to make sure that our functional form of the model is acceptable or not. This test does not indicate any specification problem in our model since the probability value of 0.981692 is so high. These means that there is no important variable excluded from the model and also there is no irrelevant variable included in the model (Appendix 2-U).

4.3.2.3.2 Chow breakpoint test

Chow test is a parameter stability test. A series of data can often contain a structural break, due to a change in policy or sudden shock to the economy. In order to test for a structural break, we often use the Chow test. The model in effect uses an F-test to determine whether a single regression is more efficient than two separate regressions involving splitting the data into two sub-samples. The Chow test basically tests whether the single regression line or the two separate regression lines fit the data best.

From Appendix 2-V, represents the F-statistic and Log likelihood ratio statistic and associated probability values. The F-statistic is based on the comparison of the restricted and unrestricted sum of squared residuals. The F statistic has an exact finite sample F- distribution if the errors are independent and identically distributed normal random variables. The log likelihood ratio statistic is based on the comparison of the restricted and unrestricted maximum of the log likelihood function. The LR test statistic has an asymptotic $\chi 2$ distribution with degrees of freedom equal to (m-1)k under the null hypothesis of no structural change, where m is the number of subsamples and k is the number of estimated coefficients. The p-value of F statistic is high which 0.739247 is, therefore, we do not reject the null hypothesis of no structural change in the electricity consumption function before and after 1999.

In summary, our variables are non-stationary, integrated of order one (I(1)), and there exists a long-run relationship between them. Therefore, the results obtained from partial adjustment model are meaningful. In addition, there is no serial correlation and heteroskedasticity in the error term. There is no important variable excluded from the model and also there is no irrelevant variable included in the model that means there is no specification problem in our model.

Thus, we do not need to look another adjusted partial adjustment model by fitting a certain functional form. In the partial adjustment model, price can be used as an explanatory variable. Therefore, based on those results, we may conclude that equation (7) is the appropriate model for electricity demand in Ethiopia excluding price variable.

4.4 Electricity Demand Forecast for Ethiopia: 2011-2015

Electricity supply planning requires efficient management of existing power systems and optimization of the decisions concerning additional capacity. Demand prediction is an important aspect in the development of any model for electricity planning. The form of the demand depends on the type of planning and accuracy that is required. There are various methods used for forecasts. An ARIMA model predicts a value in a response time series as a linear combination of its own past values or past errors (shocks) or the combination of the two.

ARIMA models are well known for being simple, robust, and parsimonious, and for providing good results. On the other hand, Vector Autoregressive (VAR) models are used to build multivariate models in which the independent variables are taken into account in the forecast.

The absolute value of elasticity measures the relative change in the electricity consumption due to a relative change in the price, income and population growth rate; inelastic demand implies that the reaction or responsiveness of consumers' to price, income and population growth rate changes is limited. As a result, forecasting electricity demand linking with the independent variables may not produce good results. In this study, therefore, ARIMA model is used to forecast electricity demand in Ethiopia. Unlike previous section, the data here is not converted into natural logarithms and, therefore, the unit is GWh.

4.4.1 Building ARIMA Model.

In order to create ARIMA models, the steps of the Box and Jenkins' method are followed: that is, identifying the dependence order of the model; p, q and d, parameter estimation, diagnostic checking and forecasting.

The first step in the analysis using ARIMA model is to identify the model. Identification consists of specifying the appropriate structure of AR, MA or ARMA and order of integration of the model to be constructed. Autocorrelation function (ACF) and partial autocorrelation function (PACF) of the series together are the most powerful tool usually applied to reveal the correct values of the parameter. Appendix 3-A shows the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of the annual electricity consumption data. The last two columns reported in the correlogram are the Ljung-Box Q- statistics and their p-values. The null hypothesis is that there is no autocorrelation in all lags.

Inspecting the sample ACF and PACF we might feel that, the autocorrelation function dies off after lag three and the partial autocorrelation functions cut at the first lag. All p-values of Q-statistics have zero values indicates that there should be autocorrelation left in all residuals. The ACF indicates the existence of a trend, which implies that electricity consumption data is non-stationary. Since the data is non-stationary, we have to make it stationary before making further analysis.

The correlogram of the first difference (Appendix 3-B) still indicates trend, and also the presence of persistent autocorrelation in the residuals implies the need to introduce the second difference. The correlogram presented in Appendix 3-C shows that there is no trend after second difference. The autocorrelation and partial autocorrelation functions except the first lag are zero and also the Q-statistics are insignificant with large p-values that there is no autocorrelation except the first lag. These observed facts suggest that ARMA (1, 2, 1) model can be fitted. The second step in the analysis of ARIMA model is estimation of parameter. Once the model is identified, then the unknown parameters can be estimated by ordinary least squares methods as follows. Let E_t^* be the second difference electricity consumption data. Then, the model can be expressed as

$$E_t^* = \delta + \emptyset_1 E^*_{t-1} + \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{30}$$

From appendix 3-D, we have

$$E_t^* = 6.84 + 0.244E^*_{t-1} - 4.396\varepsilon_{t-1} + \varepsilon_t \tag{31}$$

The third step is diagnostic checking (white noise test). In order to ensure that the best forecasting model has been built, diagnostic checking must be performed. If our model is correctly specified, the residual from the model should be nearly white noise. This means that there should be no serial correlation left in the residuals. In other words, none of the autocorrelations and partial autocorrelations of the series are significantly different from zero. This is true for all lags.

The correlogram for equation (31) has not significant spike at all lags and all subsequent Q-statistics are not significant. From the normality assumption of correlogram of residual square given in Appendix 3-H indicates that there is no autocorrelation left in the residuals. This result clearly indicates that there is no need for respecification of the model (Appendix3-E).

The last step is forecasting. To produce the forecast of electricity consumption, we have to start from the general forms of ARIMA (p, d, q) model. As before, E_t is electricity consumption data and l =1, 2, 3,... is a lead time forecast. Then, we have

$$E_{t+l} = \emptyset_1 E_{t+l-1} + \cdots + \emptyset_{p+d} E_{t+l-p-d} - \theta_1 \varepsilon_{t+l-1} - \cdots - \theta_q \varepsilon_{t+l-q} + \varepsilon_{t+l} \tag{32}$$

Taking a conditional expectation of the above equation at a time t, we obtain

$$[E_{t+l}] = \hat{E}_t(l) = \emptyset_1[E_{t+l-1}] + \cdots + \emptyset_{p+d}\left[E_{t+l-p-d}\right] - \theta_1[\varepsilon_{t+l-1}] - \cdots - \theta_q\left[\varepsilon_{t+l-q}\right] + [\varepsilon_{t+l}]$$

As pointed out in the methodology part there are three ways of expressing forecasts. These are, forecast from difference equation form, forecasts in integrated form and forecasts as a weighted average of previous observations plus a random shock. In line with these, we can write the model in the difference equation form

$$\emptyset(B)\Delta^2 E_{t+l} = \delta + \theta(B)\varepsilon_{t+l} \tag{33}$$

Since our model is ARIMA (1, 2, 1). Now the model becomes

$$(1-\emptyset_1 B)(1-B)^2 E_{t+l} = \delta + (1-\theta_1 B)\varepsilon_{t+l}$$

$$\Leftrightarrow (1-\emptyset_1 B)(1-2B+B^2)E_{t+l} = \delta + (1-\theta_1 B)\varepsilon_{t+l}$$

$$\Leftrightarrow (1-2B+B^2-\emptyset_1 B+2\emptyset_1 B^2-\emptyset_1 B^3)E_{t+l} = \delta + (1-\theta_1 B)\varepsilon_{t+l}$$

$$\Leftrightarrow (1-(2+\emptyset_1)B+(1+2\emptyset_1)B^2-\emptyset_1 B^3)E_{t+l} = \delta + (1-\theta_1 B)\varepsilon_{t+l}$$

$$\Leftrightarrow E_{t+l}-(2+\emptyset_1)E_{t+l-1}+(1+2\emptyset_1)E_{t+l-2}-\emptyset_1 E_{t+l-3} = \delta + \varepsilon_{t+l}-\theta_1\varepsilon_{t+l-1}$$

$$\Leftrightarrow E_{t+l} = \delta + (2+\emptyset_1)E_{t+l-1}-(1+2\emptyset_1)E_{t+l-2}+\emptyset_1 E_{t+l-3}-\theta_1\varepsilon_{t+l-1}+\varepsilon_{t+l}$$

We have the values of δ , $\emptyset_1$ and $\theta_1$from equation (31). Therefore, the final model used for forecast is

$$E_{t+l} = 6.84 + (2+0.244)E_{t+l-1} - (1+2*0.244)E_{t+l-2} + 0.244E_{t+l-3} - 4.396\varepsilon_{t+l-1} + \varepsilon_{t+l}$$

$$E_{t+l} = 6.84 + 2.244E_{t+l-1} - 1.488E_{t+l-2} + 0.244E_{t+l-3} - 4.396\varepsilon_{t+l-1} + \varepsilon_{t+l} \tag{34}$$

Forecasts are to be made at the origin $t = 2010$ for lead times $l = 1, 2, 3, 4, 5$ for the electricity consumption data series in the coming five years. Therefore, the forecast equations at the origin 2010 for equation (34) are given by

$$\hat{E}_{2010}(1) = E_{2011} = 6.84 + 2.244E_{2010} - 1.488E_{2009} + 0.244E_{2008} - 4.396\varepsilon_{2010}$$

$$\hat{E}_{2010}(2) = E_{2012} = 6.84 + 2.244\hat{E}_{2010}(1) - 1.488E_{2010} + 0.244E_{2009}$$

$$\hat{E}_{2010}(3) = E_{2013} = 6.84 + 2.244\hat{E}_{2010}(2) - 1.488\hat{E}_{2010}(1) + 0.244E_{2010}$$

For general case

$$\hat{E}_{2010}(l) = 6.84 + 2.244\hat{E}_{2010}(l-1) - 1.488\hat{E}_{2010}(l-2) + 0.244\hat{E}_{2010}(l-3), \quad l = 4,4,5,\dots. \tag{35}$$

In general the moving average operator $\theta(\text{B})$, the forecast equation will depend directly on the $\varepsilon's$ but forecast at a longer lead times will not. In our forecast model only $\hat{E}_{2010}(1)$ which depend on ε_{2010}, the others become zero. Because, the conditional expectation of $[\varepsilon_{2010+j}]$ where $j = 1,2$ ... is zero.

The residual can be obtained as the actual minus the fitted values of electricity consumption data. From appendix 3-F, we have the values of the residual. So, we can calculate the forecast values by using the above formula.

ARIMA models are developed basically to forecast the corresponding variable. There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former are used to develop confidence in the model and the latter to generate genuine desired forecasts.

4.4.2 Sample Period Forecasts

The sample period forecasts are obtained simply by plugging the actual values of the variables in the estimated equation (34). We obtained values for $\hat{E}_t$ together with the actual values of E_t. To judge the forecasting ability of the fitted ARIMA model, important measures of the sample period forecasts accuracy were computed. The Mean Absolute Percentage Error (MAPE) turn out to be 2.80% which is less than 4% and Theil's inequality coefficient (U-statistic) turn out to be 0.015 which is close to zero. Besides that the bias and the variance proportion are also very small which are 0.037 and 0.001 respectively. There is no significant difference between actual versus fitted graph. Thus, those measures indicate that the forecasting inaccuracy is low (Appendix 3-G).

4.4.3 Post Sample Forecasts

Once a satisfactory model is obtained, generation of forecasts is an easy process. The principal objective of developing an ARIMA model for a variable is to generate post sample period forecasts for that variable. This is done through using equation (34). The forecasted values for electricity consumption data by considering 2010 as origin and lead time $l = 1,2,3,4,5$ are given below.

Table-6: Electricity demand forecasts (origin 2010, lead time $l = 1, 2, 3, 4, 5$).

Year	Forecasted electricity consumption(GWh)	Annual % change
$\hat{E}_{2010}(1)=E_{2011}$	3319.488	12.21
$\hat{E}_{2010}(2)= E_{2012}$	3761.446	13.31
$\hat{E}_{2010}(3)= E_{2013}$	4229.976	12.46
$\hat{E}_{2010}(4)= E_{2014}$	4711.828	11.39
$\hat{E}_{2010}(5)= E_{2015}$	5203.772	10.44

The average annual percentage change for the electricity consumption data is 11.92%. Based on our model the forecasted result of net electricity consumption in Ethiopia will increase an average 11.92% during the coming five years.

4.4.4 Government Projections of Net Electricity Consumption Data in Ethiopia from 2011-2015

The government of Ethiopia gave more emphasis to the power generation and has planned to increase it from about 2000MW in 2010 to 10,000MW in 2015, end of the growth and transformation plan of the country. The following table shows the government projection of electricity consumption in three scenarios.

Table-7: Government projections of electricity demand in Ethiopia (2011-2015)

Year	Government Projections for Electricity Consumption (GWh)			Annual % Change		
	Reference Scenario	Moderate Scenario	Targeted scenario	Reference Scenario	Moderate Scenario	Targeted Scenario
2011	3572.16	4184.79	5101.07	4.4	11.9	5.4
2012	3931.27	4634.37	5844.40	10.1	10.7	14.6
2013	4308.37	5104.38	6663.17	9.6	10.1	14.0
2014	4703.55	5594.39	7563.63	9.2	9.6	13.5
2015	5117.13	6104.38	8553.16	8.8	9.1	13.1

Source: Ethiopian Electric Power Corporation (EEPCo).

The average annual percentage change for the government projection of electricity consumption is 8.1% for reference scenario, 10.2% for moderate scenario and 11.4% for targeted scenario.

4.4.5 Comparison of ARIMA (1, 2, 1) Model Forecast Result with the Government Projection.

From table (6) we have seen that ARIMA (1, 2, 1) model forecasted results gives as indication that there exists increasing electricity consumption in Ethiopia in the coming five years an average 11.92%. Table 8 below shows the comparison of government projection with ARIMA (1, 2, 1) model forecasts.

Table-8: Comparisons of government projection with ARIMA (1, 2, 1) model

Year	Government projection (GWh)			ARIMA (1,2,1) model forecast	Difference between government projection and ARIMA(1,2,1) model forecast			Difference between government projection and ARIMA(1,2,1) model forecast as % of ARIMA model forecast		
	Reference scenario	Moderate scenario	Targeted Scenario		Reference Scenario	Moderate Scenario	Targeted scenario	Reference Scenario	Moderate scenario	Targeted scenario
2011	3572.16	4184.79	5101.07	3319.488	252.672	865.302	1781.582	7.6	26.1	53.7
2012	3931.27	4634.37	5844.40	3761.446	169.824	872.924	2082.954	4.5	23.2	55.4
2013	4308.37	5104.38	6663.17	4229.976	78.394	874.404	2433.194	1.9	20.7	57.5
2014	4703.55	5594.39	7563.63	4711.828	-8.278	882.562	2851.802	-0.2	18.7	60.5
2015	5117.13	6104.38	8553.16	5203.772	-86.642	900.608	3349.388	-1.7	17.3	64.4

The results show relatively small differences between forecasted values and reference scenario projection, large differences among moderate scenario, targeted scenario and ARIMA forecasts in actual term.

5 Conclusion

The paper examined demand for electricity in Ethiopia employing annual data over the period 1976-2010. Partial adjustment and ARIMA modeling were applied to estimate and forecast electricity demand in Ethiopia.

In this study, short-run and long-run elasticities of electricity demand are obtained. The income elasticity, as expected has a positive sign and is significant. This result indicates that electricity consumption is a normal good as it increases with income. The short-run and long-run income elasticities were found to be 0.2517 and 0.7683 respectively. The population growth rate was found to be an important determinant of electricity consumption in Ethiopia. Its elasticities are 0.1724 and 0.5263 in the short and long-run respectively. ARIMA model offers a good technique for predicting the magnitude of any variable. In our study the developed model for electricity demand was found to be ARIMA (1, 2, 1). The forecasted results obtained indicate that electricity demand increases at an average rate of 11.92% during the coming five years. The comparison of government projected and forecasted results of this study shows that the government projected values, except reference scenario, are consistently higher than the forecasted values of this study.

Future Direction of the Study

Due to the nature of ARIMA models and low elasticities obtained from the partial adjustment model, the present study has only employed net total consumption data for forecasting. There is an apparent need for further work with more variables that will examine the demand for electricity in Ethiopia taking in to account exogenous variables. In this regard Autoregressive Moving Average with eXogenous (ARMAX) models and transfer function model may have merits in allowing a more detailed and accurate understanding of the trend of electricity in Ethiopia. See Shumwy and Stoffer (2006) and references therein for more on these two models.

References

Accram, A., 2005. Dying for electricity in energy in Africa, 2(1), 22-32.

Al-Zayer, J., Al-Ibrahim, A., 1996. Modelling the impact of temperature on electricity consumption in the eastern provinces of Saudi Arabia. Journal of Forecasting, 15 (2), 97–106.

Amusa, H., Amusa, K., Mabugu, R., 2009. Aggregate demand for electricity in South Africa: An analysis using the bounds testing approach to co- integration. Energy Policy, 37, 4167-4175.

Beenstock, M., Goldin, E., Nabot, D., 1999. The demand for electricity in Israel. Energy Economics, 21, 168–183.

Bentzen, J., Engsted, T., 1994. Short and long-run elasticities in energy demand: A co-integration approach. Energy Economics, 15(1), 9-16.

Bentzen, J., Engsted, T., 2001. A revival of the autoregressive distributed lag model in estimating energy demand relationships. Energy, 26(1), 45-55.

Bernard, J.,T., Denis B., Donald B., 1996. Quebec residential electricity demand: A micro econometric approach. Canadian Journal of Economics, 29 (1), 92-113.

Bose, R.K., Shukla, M., 1999. Elasticities of electricity demand in India. Energy Policy, 27, 137–146.

Box, G.E.P., Jenkins, G.M., 1970. Time series analysis: forecasting and control. San Francisco, CA . Holden-Day.

Box, G.E.P., Jenkins, G.M., 1976, Time series analysis: forecasting and control. San Francisco, Holden Day.

Chang, Y.,E. Martinez. C., 1997. Electricity demand analysis using co-integration and error-correction models with time varying parameters: The Mexican case. Department of Economics, Rice University, United States.

Davidson, J.E.H., D.F. Hendry, F. Srba, J.S. Yeo, 1978. Econometric modeling of the aggregate time-series relationship between consumers' expenditure and income in the United Kingdom. Economic Journal, 88, 661-692.

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

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

Dincer, I., Dost, S., 1997. Energy and GDP. International Journal of Energy Research, 21, 153–167.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction representation, estimation and testing. Econometrica, 55(2), 251-276.

Engle, R.F., Granger, C.W.J., Hallman, J.J., 1989. Merging short and long-run forecasts: An application of seasonal co-integration to monthly electricity sales forecasting. Journal of Econometrics, 40(1), 45-62.

Erdogdu, E., 2007. Electricity demand analysis using co-integration and ARIMA modeling: A case study of Turkey. Energy Policy, 35, 1129-1146.

Federal democratic republic of Ethiopia population census commission, 2007. Population and housing census.

Filippini, M., 1999. Swiss residential demand for electricity. Applied Economics, 6, 533–538.

Filippini, M., Pachauri, S., 2004. Elasticities of electricity demand in urban Indian households. Energy Policy, 32(3), 429-436.

Fouquet, R., Hawdon, D., Pearson, P., Robinson, C., Stevens, P., 1993. The SEEC United Kingdom energy demand forecast. Surrey Energy Economics Centre (SEEC): Occasional Paper 1, Department of Economics, University of Surrey, Guildford, UK.

Girma, H., 2000. Encyclopedia international of laws. Addis Ababa. www.akababi.org/energylaw.[accessed 23 November 2010]

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

Granger, C.W.J., Newbold, P., 1974. Spurious regressions in econometrics. Journal of Econometrics, 2(2), 111-120.

Greene, W., 2000. Econometric Analysis, 4th edition. New Jersey, Prentice-Hall.

GTZ. 2010. Eastern Africa energy resource base. www.regionalenergy.net.com [accessed 12 August 2010].

Gujarati, D., 1995. Basic Econometrics. 3rd Edition. New York, McGraw Hill.

Holtedahl, P., & Joutz, F. L., 2004. Residential electricity demand in Taiwan. Energy Economics, 26, 201-224.

Hondroyiannis, G. 2004. Estimating residential demand for electricity in Greece. Energy Policy, 26, 319- 334.

Houthakker, H. S., Verleger, P. K., Sheehan, D. P., 1974. Dynamic demand analysis for gasoline and residential electricity: American Journal of Agricultural Economics, 56(2), 412-418.

Houthakker, H.S., 1980. Residential electricity revisited. Energy Journal, 1, 29–41.

Hsing, Y., 1994. Estimation of residential demand for electricity with the cross sectionally correlated and time-wise autoregressive model. Resource and Energy Economics, 16, 255–263.

Hunt L.C., Manning, D.N., 1989. Energy price and income elasticities of demand: Some estimates for the UK using the co-integration procedure. Scottish Journal of Political Economy, 36(2), 183-193.

Hunt L.C., Witt, R., 1995. An analysis of UK energy demand using multivariate co-integration. Surrey Energy Economics Centre (SEEC): Discussion Paper No: 86; Department of Economics, University of Surrey, Guildford, UK.

Hunt L.C., Lynk, E.L., 1992. Industrial energy demand in the UK: A co-integration approach. David Hawdon (Ed.), Energy demand: Evidence and expectations. Surrey University Press.London.

Judge, G.G., R. Carter Hill, E.G. William, I. Helmut, 1988. Introduction to the theory and practice of econometrics. 2nd ed., John Wiley and Son, INC. New York, Toronto, Singapore.

Kamerschena, D. R., D. V. Porter, 2004. The demand for residential, industrial and total electricity. Energy Economics, 26, 87–100.

Lin, B.Q., 2003. Electricity demand in the people's republic of China: investment requirement and environmental impact. Economics and research department working paper series number No.37. Asian Development Bank.

Makridakis, S., S. C. Wheelwright. 1998. Forecasting methods and applications. New York, John Wiley & Sons.

Narayan, P.K., Smyth, R., 2005. The residential demand for electricity in Australia: An application of the bounds testing approach to co-integration. Energy Policy, 33, 457–464.

Narayan, P.K., Smyth, R., Prasad, A., 2007. Electricity consumption in G7 countries: A panel co-integration analysis of residential demand elasticities. Energy Policy, 35, 4485–4494.

Nelson, C.R., C.I. Plosser, 1982. Trends and random walks in macroeconomic time series. Journal of Monetary Economics, 10, 139-162.

Odumosu, T., 2005. Africa needs power for NEPAD to work in energy in Africa, 2(1), 34-42.

Phillips, P.C.B., 1986. Understanding spurious regressions in econometrics. Journal of Econometrics, 33, 311-340.

Phillips, P. C. B., Perron, P., 1988. Testing for a unit-root in time series regression. Biometrika, 75, 335–346.

Ramsey, J. B., 1969. Tests for specification errors in classical linear least squares regression analysis. Journal of the Royal Statistical Society B, 32, 350–371.

Said, S.E., D. Dickey, 1984. Testing for unit-roots in autoregressive moving-average models with unknown order. Biometrika, 71, 599-607.

Sargan, J. D., Bhargava, A. S., 1983. Testing residuals from least squares regression for being generated by the Gaussian random walk. Econometrica, 51, 213–248.

Sims, C. A., 1980. Macroeconomics and reality. Econometrica, 48, 1–48.

Shumwy,R.H, Stoffer,D.S., 2006. Time series Analysis and its applications with R examples. 2nd edition. Spring.

Solomon, S., 2007. Hydropower of Ethiopia: Status, potential and prospect. www2.webng.com/bahirdarab/hydropower. [Accessed 3 September2010]

Terri, H., 2008. What cost Ethiopia's dam boom. www.internationalrivers.org [accessed 13 August 2010].

UNIDO, 2006. Analysis of financing models for small hydropower plants on the basis of case studies in Ethiopia. www.unido.org/fileadmin/media/documents/pdf/financial [16 October 2010]

Venkataravaman, Krishnaswamy, Gray Stuggins, eds., 2007.Closing the electricity supply- demand gap. *siteresources.worldbank.org/.../EnergyandMiningSectorBoardPaperNo20[accessed 14 July 2010]*

______ Interconnected system (ICS) hydropower generating capacity. www.eepco.gov.et [19 December 2010]

Zachariadis, T., N. Pashourtidou, 2006. An empirical analysis of electricity consumption in cyprus. Department Of Economics, University Of Cyprus.

Zellner, A., Palm, F.C., 1974. Time series analysis and simultaneous equation econometric models. Journal of Econometrics, 2, 17-54.

Ziramba, E., 2008. The demand for residential electricity in South Africa. Energy Policy, 36, 3460– 3466.

Appendix I Descriptive Analysis

Table-9: Summary statistics for the variables

	PG	E	P	Y
Mean	2.875106	21.03125	0.268500	1078.006
Median	2.801700	18.18094	0.205300	985.2223
Maximum	3.320800	35.67290	0.514800	1909.900
Minimum	2.430000	12.48744	0.102800	720.5400
Std.Dev.	0.228694	6.814168	0.146389	281.6196
CV	7.9543%	32.4002%	54.5221%	26.1241%
Skewness	0.093093	0.926222	0.440899	1.471613
Kurtosis	2.152022	2.886259	1.588254	4.503705

Appendix 1: Time Series Plots of Net Electricity Consumption Per capita, Real Price of Electricity, Real Income Per capita and population growth rate.

Figure-1: Net electricity consumption per capita at level form

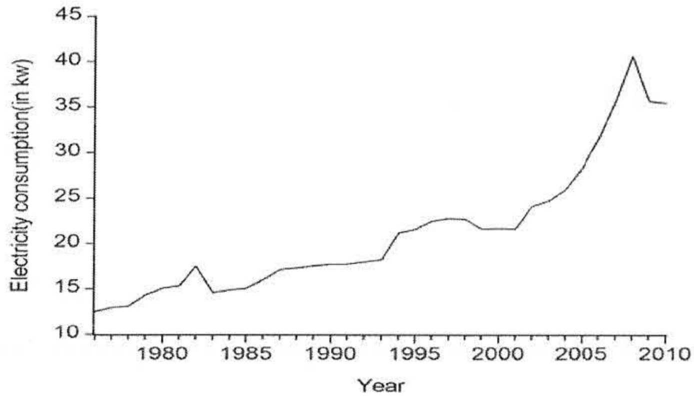


Figure-2: Net electricity consumption per capita at first difference

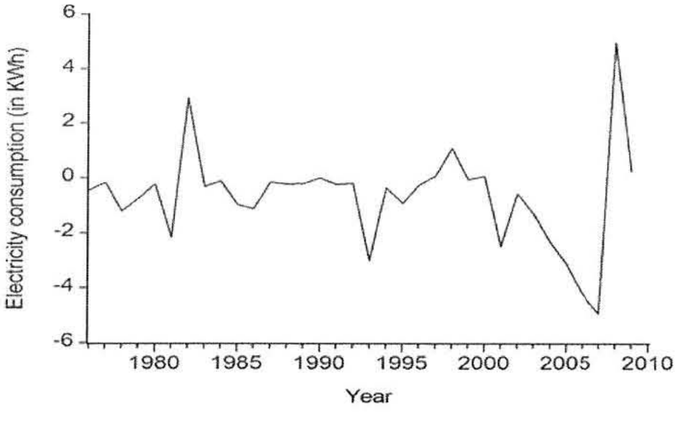


Figure-3: Real price of electricity at level form

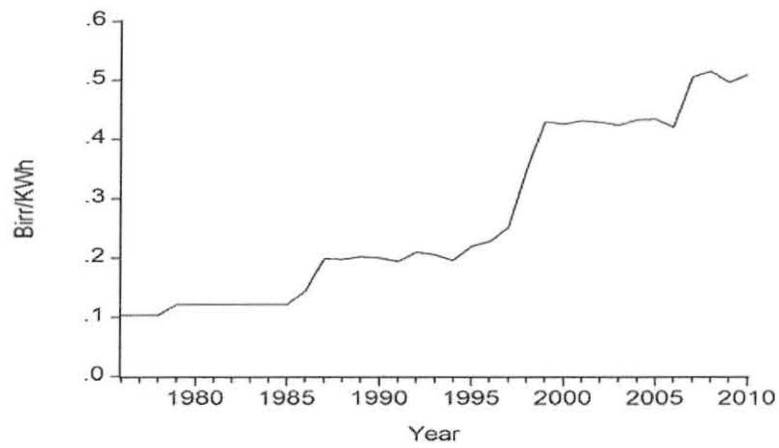


Figure-4: Real price of electricity at first difference

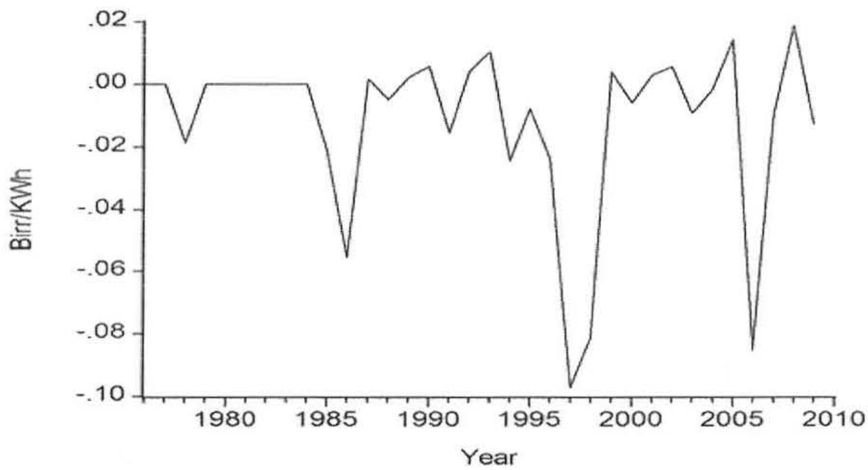


Figure-5: Real income per capita at level form

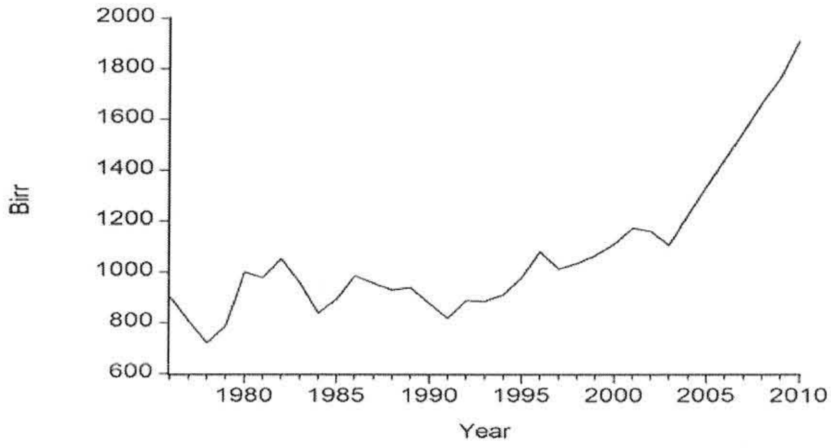


Figure-6: Real income per capita at first difference

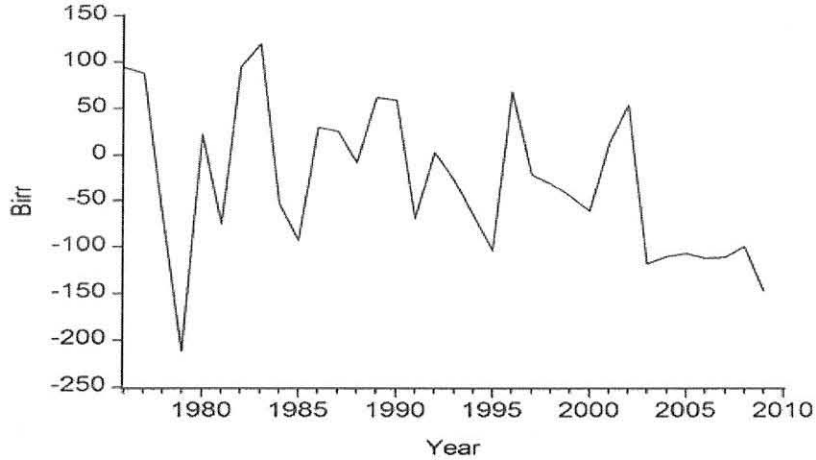


Figure-7: Population growth rate at level form

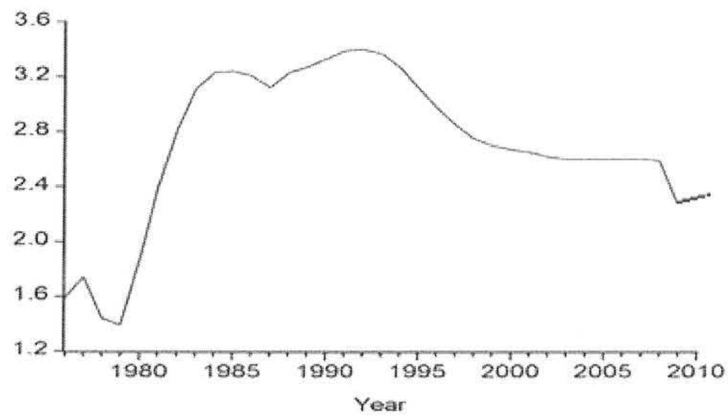
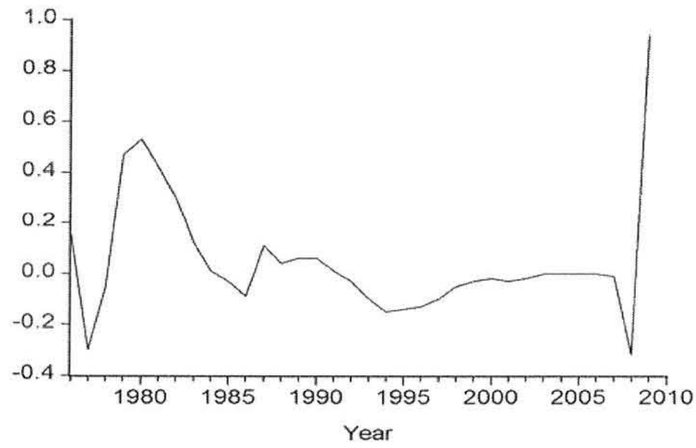


Figure-8: Population growth rate at first difference



Appendix 2: Estimation Outputs.

Appendix 2-A: ADF test output for variable logE

Null Hypothesis: LOG(E) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.175469	0.4874
Test critical values:	1% level	-4.252879	
	5% level	-3.548490	
	10% level	-3.207094	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(E))
Method: Least Squares
Date: 12/21/10 Time: 03:34
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(E(-1))	-0.272710	0.125357	-2.175469	0.0373
C	0.699351	0.311815	2.242836	0.0322
@TREND(1976)	0.008349	0.003806	2.193607	0.0359

R-squared	0.136194	Mean dependent var	0.030678
Adjusted R-squared	0.080465	S.D. dependent var	0.067945
S.E. of regression	0.065154	Akaike info criterion	-2.540022
Sum squared resid	0.131597	Schwarz criterion	-2.405343
Log likelihood	46.18037	F-statistic	2.443846
Durbin-Watson stat	1.866944	Prob(F-statistic)	0.103383

Appendix 2-B: ADF test output for variable logP

Null Hypothesis: LOG(P) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.017138	0.1429
Test critical values:	1% level	-4.252879	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(P))
Method: Least Squares
Date: 12/21/10 Time: 03:38
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(P(-1))	-0.388220	0.128672	-3.017138	0.0053
D(LOG(P(-1)))	0.410758	0.170071	2.415218	0.0223
C	-0.919801	0.321445	-2.861455	0.0077
@TREND(1976)	0.021069	0.007286	2.891869	0.0072

R-squared	0.276073	Mean dependent var	0.048463
Adjusted R-squared	0.201185	S.D. dependent var	0.098613
S.E. of regression	0.088137	Akaike info criterion	-1.906639
Sum squared resid	0.225275	Schwarz criterion	-1.725244
Log likelihood	35.45954	F-statistic	3.686438
Durbin-Watson stat	2.030662	Prob(F-statistic)	0.023079

<u>Appendix 2-C</u>: ADF test output for variable logY

Null Hypothesis: LOG(Y) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.287734	0.8741
Test critical values:	1% level	-4.252879	
	5% level	-3.548490	
	10% level	-3.207094	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(Y))
Method: Least Squares
Date: 12/21/10 Time: 03:43
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(Y(-1))	-0.142938	0.111000	-1.287734	0.2074
C	0.928115	0.736805	1.259647	0.2172
@TREND(1976)	0.004886	0.002344	2.084592	0.0454

R-squared	0.136470	Mean dependent var	0.022061
Adjusted R-squared	0.080758	S.D. dependent var	0.079782
S.E. of regression	0.076493	Akaike info criterion	-2.219146
Sum squared resid	0.181385	Schwarz criterion	-2.084468
Log likelihood	40.72549	F-statistic	2.449576
Durbin-Watson stat	1.581447	Prob(F-statistic)	0.102873

Appendix 2-D: ADF test output for variable LogPG

Null Hypothesis: LOG(PG) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		0.055046	0.9952
Test critical values:	1% level	-4.296729	
	5% level	-3.568379	
	10% level	-3.218382	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(PG))
Method: Least Squares
Date: 04/12/11 Time: 15:46
Sample (adjusted): 1978 2010
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(PG(-1))	0.006933	0.125943	0.055046	0.9565
D(LOG(PG(-1)))	-0.870221	0.226086	-3.849069	0.0007
C	0.045330	0.137971	0.328548	0.7451
@TREND(1976)	-0.002952	0.001249	-2.362871	0.0259

R-squared	0.448768	Mean dependent var	0.002869
Adjusted R-squared	0.385164	S.D. dependent var	0.080209
S.E. of regression	0.062893	Akaike info criterion	-2.571193
Sum squared resid	0.102844	Schwarz criterion	-2.384367
Log likelihood	42.56790	F-statistic	7.055687
Durbin-Watson stat	1.708838	Prob(F-statistic)	0.001268

Appendix 2-E: ADF test output for variable ΔlogE

Null Hypothesis: D(LOG(E)) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=0)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.831060	0.0002
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(E),2)
Method: Least Squares
Date: 12/21/10 Time: 03:49
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(E(-1)))	-1.069325	0.183384	-5.831060	0.0000
C	0.022540	0.026628	0.846479	0.4040
@TREND(1976)	0.000566	0.001303	0.434878	0.6668

R-squared	0.531404	Mean dependent var	-0.001266
Adjusted R-squared	0.500165	S.D. dependent var	0.100283
S.E. of regression	0.070899	Akaike info criterion	-2.368605
Sum squared resid	0.150801	Schwarz criterion	-2.232559
Log likelihood	42.08199	F-statistic	17.01055
Durbin-Watson stat	1.973772	Prob(F-statistic)	0.000012

Appendix 2-F: ADF test output for variable ΔlogP

Null Hypothesis: D(LOG(P)) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-4.407153	0.0070
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(P),2)
Method: Least Squares
Date: 12/21/10 Time: 03:54
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(P(-1)))	-0.782506	0.177554	-4.407153	0.0001
C	0.044717	0.037914	1.179432	0.2475
@TREND(1976)	-0.000368	0.001816	-0.202827	0.8406

R-squared	0.393567	Mean dependent var	0.000760
Adjusted R-squared	0.353138	S.D. dependent var	0.123501
S.E. of regression	0.099329	Akaike info criterion	-1.694244
Sum squared resid	0.295989	Schwarz criterion	-1.558198
Log likelihood	30.95503	F-statistic	9.734805
Durbin-Watson stat	1.921693	Prob(F-statistic)	0.000552

Appendix 2-G: ADF test output for variable ΔlogY

Null Hypothesis: D(LOG(Y)) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 1 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.441581	0.0005
Test critical values:	1% level	-4.262735	
	5% level	-3.557759	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.
Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(Y),2)
Method: Least Squares
Date: 12/21/10 Time: 04:07
Sample (adjusted): 1979 2010
Included observations: 32 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(Y(-1)))	-1.232605	0.226516	-5.441581	0.0000
D(LOG(Y(-1)),2)	0.317907	0.167589	1.896944	0.0682
C	0.001246	0.029071	0.042876	0.9661
@TREND(1976)	0.001796	0.001467	1.224028	0.2311

R-squared	0.553113	Mean dependent var	0.006096
Adjusted R-squared	0.505233	S.D. dependent var	0.103409
S.E. of regression	0.072738	Akaike info criterion	-2.287449
Sum squared resid	0.148141	Schwarz criterion	-2.104232
Log likelihood	40.59918	F-statistic	11.55190
Durbin-Watson stat	1.960743	Prob(F-statistic)	0.000042

Appendix 2-H: ADF test for variable ΔLogPG

Null Hypothesis: D(LOG(PG)) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic based on SIC, MAXLAG=1)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.957219	0.0000
Test critical values:	1% level	-4.296729	
	5% level	-3.568379	
	10% level	-3.218382	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LOG(PG),2)
Method: Least Squares
Date: 04/12/11 Time: 15:50
Sample (adjusted): 1978 2010
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(PG(-1)))	-1.863537	0.187154	-9.957219	0.0000
C	0.052801	0.024316	2.171493	0.0389
@TREND(1976)	-0.002963	0.001211	-2.446056	0.0212

R-squared	0.788590	Mean dependent var	0.006806
Adjusted R-squared	0.772930	S.D. dependent var	0.129525
S.E. of regression	0.061721	Akaike info criterion	-2.637743
Sum squared resid	0.102856	Schwarz criterion	-2.497624
Log likelihood	42.56615	F-statistic	50.35707
Durbin-Watson stat	1.708814	Prob(F-statistic)	0.000000

Appendix 2-I: PP test output for variable logE

Null Hypothesis: LOG(E) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 0 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-2.175469	0.4874
Test critical values:	1% level	-4.252879	
	5% level	-3.548490	
	10% level	-3.207094	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.003871
HAC corrected variance (Bartlett kernel)	0.003871

Phillips-Perron Test Equation
Dependent Variable: D(LOG(E))
Method: Least Squares
Date: 12/21/10 Time: 04:14
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(E(-1))	-0.272710	0.125357	-2.175469	0.0373
C	0.699351	0.311815	2.242836	0.0322
@TREND(1976)	0.008349	0.003806	2.193607	0.0359

R-squared	0.136194	Mean dependent var	0.030678
Adjusted R-squared	0.080465	S.D. dependent var	0.067945
S.E. of regression	0.065154	Akaike info criterion	-2.540022
Sum squared resid	0.131597	Schwarz criterion	-2.405343
Log likelihood	46.18037	F-statistic	2.443846
Durbin-Watson stat	1.866944	Prob(F-statistic)	0.103383

Appendix 2-J: PP test output for variable logP

Null Hypothesis: LOG(P) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 0 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-2.206202	0.4712
Test critical values:	1% level	-4.252879	
	5% level	-3.548490	
	10% level	-3.207094	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.007968
HAC corrected variance (Bartlett kernel)	0.007968

Phillips-Perron Test Equation
Dependent Variable: D(LOG(P))
Method: Least Squares
Date: 12/21/10 Time: 04:21
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(P(-1))	-0.275126	0.124706	-2.206202	0.0349
C	-0.624300	0.306738	-2.035290	0.0504
@TREND(1976)	0.014929	0.006999	2.132890	0.0410

R-squared	0.135771	Mean dependent var	0.047037
Adjusted R-squared	0.080014	S.D. dependent var	0.097462
S.E. of regression	0.093482	Akaike info criterion	-1.817999
Sum squared resid	0.270905	Schwarz criterion	-1.683320
Log likelihood	33.90598	F-statistic	2.435061
Durbin-Watson stat	1.391755	Prob(F-statistic)	0.104171

Appendix 2-K: PP test output for variable logY

Null Hypothesis: LOG(Y) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 3 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.353747	0.8564
Test critical values:	1% level	-4.252879	
	5% level	-3.548490	
	10% level	-3.207094	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.005335
HAC corrected variance (Bartlett kernel)	0.005645

Phillips-Perron Test Equation
Dependent Variable: D(LOG(Y))
Method: Least Squares
Date: 12/21/10 Time: 04:23
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(Y(-1))	-0.142938	0.111000	-1.287734	0.2074
C	0.928115	0.736805	1.259647	0.2172
@TREND(1976)	0.004886	0.002344	2.084592	0.0454

R-squared	0.136470	Mean dependent var	0.022061
Adjusted R-squared	0.080758	S.D. dependent var	0.079782
S.E. of regression	0.076493	Akaike info criterion	-2.219146
Sum squared resid	0.181385	Schwarz criterion	-2.084468
Log likelihood	40.72549	F-statistic	2.449576
Durbin-Watson stat	1.581447	Prob(F-statistic)	0.102873

Appendix 2-L: PP test output for variable logPG

Null Hypothesis: LOG(PG) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 3 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-1.705377	0.7255
Test critical values:	1% level	-4.273277	
	5% level	-3.557759	
	10% level	-3.212361	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.005161
HAC corrected variance (Bartlett kernel)	0.004250

Phillips-Perron Test Equation
Dependent Variable: D(LOG(PG))
Method: Least Squares
Date: 04/12/11 Time: 15:57
Sample (adjusted): 1977 2010
Included observations: 32 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LOG(PG(-1))	-0.234124	0.124401	-1.882003	0.0699
C	0.276720	0.139357	1.985694	0.0566
@TREND(1976)	-0.001594	0.001375	-1.159723	0.2556

R-squared	0.120494	Mean dependent var	0.004137
Adjusted R-squared	0.059838	S.D. dependent var	0.077829
S.E. of regression	0.075465	Akaike info criterion	-2.241244
Sum squared resid	0.165153	Schwarz criterion	-2.103832
Log likelihood	38.85991	F-statistic	1.986522
Durbin-Watson stat	2.472123	Prob(F-statistic)	0.155405

Appendix 2-M: PP test output for variable ΔlogE

Null Hypothesis: D(LOG(E)) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 5 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-5.991758	0.0001
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.004570
HAC corrected variance (Bartlett kernel)	0.002920

Phillips-Perron Test Equation
Dependent Variable: D(LOG(E),2)
Method: Least Squares
Date: 12/24/10 Time: 01:11
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(E(-1)))	-1.069325	0.183384	-5.831060	0.0000
C	0.022540	0.026628	0.846479	0.4040
@TREND(1976)	0.000566	0.001303	0.434878	0.6668

R-squared	0.531404	Mean dependent var	-0.001266
Adjusted R-squared	0.500165	S.D. dependent var	0.100283
S.E. of regression	0.070899	Akaike info criterion	-2.368605
Sum squared resid	0.150801	Schwarz criterion	-2.232559
Log likelihood	42.08199	F-statistic	17.01055
Durbin-Watson stat	1.973772	Prob(F-statistic)	0.000012

Appendix 2-N: PP test output for variable ΔlogP

Null Hypothesis: D(LOG(P)) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 1 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-4.428630	0.0067
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.008969
HAC corrected variance (Bartlett kernel)	0.009291

Phillips-Perron Test Equation
Dependent Variable: D(LOG(P),2)
Method: Least Squares
Date: 12/21/10 Time: 04:30
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(P(-1)))	-0.782506	0.177554	-4.407153	0.0001
C	0.044717	0.037914	1.179432	0.2475
@TREND(1976)	-0.000368	0.001816	-0.202827	0.8406

R-squared	0.393567	Mean dependent var	0.000760
Adjusted R-squared	0.353138	S.D. dependent var	0.123501
S.E. of regression	0.099329	Akaike info criterion	-1.694244
Sum squared resid	0.295989	Schwarz criterion	-1.558198
Log likelihood	30.95503	F-statistic	9.734805
Durbin-Watson stat	1.921693	Prob(F-statistic)	0.000552

Appendix 2-O: PP test output for variable ΔlogY

Null Hypothesis: D(LOG(Y)) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 6 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-5.212105	0.0009
Test critical values:	1% level	-4.262735	
	5% level	-3.552973	
	10% level	-3.209642	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.005434
HAC corrected variance (Bartlett kernel)	0.002499

Phillips-Perron Test Equation
Dependent Variable: D(LOG(Y),2)
Method: Least Squares
Date: 12/21/10 Time: 04:33
Sample (adjusted): 1978 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(Y(-1)))	-0.892076	0.177041	-5.038810	0.0000
C	-0.005585	0.029026	-0.192424	0.8487
@TREND(1976)	0.001636	0.001471	1.111985	0.2750

R-squared	0.459241	Mean dependent var	0.005747
Adjusted R-squared	0.423190	S.D. dependent var	0.101800
S.E. of regression	0.077315	Akaike info criterion	-2.195344
Sum squared resid	0.179329	Schwarz criterion	-2.059298
Log likelihood	39.22318	F-statistic	12.73878
Durbin-Watson stat	1.983225	Prob(F-statistic)	0.000099

Appendix 2-P: PP test output for variable ΔlogPG

Null Hypothesis: D(LOG(PG)) has a unit root
Exogenous: Constant, Linear Trend
Bandwidth: 1 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-9.709399	0.0000
Test critical values:	1% level	-4.296729	
	5% level	-3.568379	
	10% level	-3.218382	

*MacKinnon (1996) one-sided p-values.

Residual variance (no correction)	0.003429
HAC corrected variance (Bartlett kernel)	0.003842

Phillips-Perron Test Equation
Dependent Variable: D(LOG(PG),2)
Method: Least Squares
Date: 04/12/11 Time: 16:03
Sample (adjusted): 1978 2010
Included observations: 30 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LOG(PG(-1)))	-1.863537	0.187154	-9.957219	0.0000
C	0.052801	0.024316	2.171493	0.0389
@TREND(1976)	-0.002963	0.001211	-2.446056	0.0212

R-squared	0.788590	Mean dependent var	0.006806
Adjusted R-squared	0.772930	S.D. dependent var	0.129525
S.E. of regression	0.061721	Akaike info criterion	-2.637743
Sum squared resid	0.102856	Schwarz criterion	-2.497624
Log likelihood	42.56615	F-statistic	50.35707
Durbin-Watson stat	1.708814	Prob(F-statistic)	0.000000

Appendix 2-Q: OLS Output for equation 4

Dependent Variable: LOG(E)
Method: Least Squares
Date: 04/12/11 Time: 17:51
Sample: 1976 2010
Included observations: 34

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.592372	1.197861	-1.329346	0.1938
LOG(P)	0.278013	0.050565	5.498141	0.0000
LOG(Y)	0.708186	0.145404	4.870467	0.0000
LOG(PG)	0.074359	0.180826	0.411220	0.6838

R-squared	0.924862	Mean dependent var	3.007369
Adjusted R-squared	0.917349	S.D. dependent var	0.317789
S.E. of regression	0.091362	Akaike info criterion	-1.837851
Sum squared resid	0.250409	Schwarz criterion	-1.658279
Log likelihood	35.24346	F-statistic	123.0892
Durbin-Watson stat	0.747036	Prob(F-statistic)	0.000000

Appendix 2-R: OLS Output for equation 7 (Partial adjustment model)

Dependent Variable: LOG(E)
Method: Least Squares
Date: 04/12/11 Time: 17:54
Sample (adjusted): 1977 2010
Included observations: 33 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.651265	0.909706	-0.715907	0.4800
LOG(P)	0.085656	0.053399	1.604080	0.1199
LOG(Y)	0.251695	0.146449	1.718656	0.0967
LOG(PG)	0.172430	0.134610	1.700516	0.0793
LOG(E(-1))	0.672429	0.143535	4.684778	0.0001

R-squared	0.958782	Mean dependent var	3.021995
Adjusted R-squared	0.952894	S.D. dependent var	0.310879
S.E. of regression	0.067473	Akaike info criterion	-2.415451
Sum squared resid	0.127473	Schwarz criterion	-2.188707
Log likelihood	44.85494	F-statistic	162.8292
Durbin-Watson stat	1.646204	Prob(F-statistic)	0.000000

<u>Appendix 2-S</u>: White heteroskedasticity test output for equation 7 (Partial adjustment model)

White Heteroskedasticity Test:

F-statistic	1.184349	Probability	0.348629
Obs*R-squared	9.340408	Probability	0.314398

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 04/12/11 Time: 17:57
Sample: 1977 2010
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.881380	1.763915	-0.499673	0.6219
LOG(P)	0.030933	0.032522	0.951133	0.3510
(LOG(P))^2	0.016091	0.011882	1.354211	0.1883
LOG(Y)	0.161450	0.562678	0.286931	0.7766
(LOG(Y))^2	-0.012214	0.040583	-0.300969	0.7660
LOG(PG)	-0.019477	0.155549	-0.125216	0.9014
(LOG(PG))^2	0.011135	0.072650	0.153274	0.8795
LOG(E(-1))	0.204752	0.212636	0.962921	0.3452
(LOG(E(-1)))^2	-0.027029	0.035586	-0.759549	0.4549

R-squared	0.283043	Mean dependent var	0.003863
Adjusted R-squared	0.044057	S.D. dependent var	0.006666
S.E. of regression	0.006517	Akaike info criterion	-7.001775
Sum squared resid	0.001019	Schwarz criterion	-6.593636
Log likelihood	124.5293	F-statistic	1.184349
Durbin-Watson stat	1.740729	Prob(F-statistic)	0.348629

Appendix 2-T: Breusch-Godfrey serial correlation test output for equation 7 (Partial adjustment model)

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	0.460946	Probability	0.502961
Obs*R-squared	0.553922	Probability	0.456720

Test Equation:
Dependent Variable: RESID
Method: Least Squares
Date: 04/12/11 Time: 17:58
Presample and interior missing value lagged residuals set to zero.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.056178	0.922311	-0.060910	0.9519
LOG(P)	0.018639	0.060507	0.308052	0.7604
LOG(Y)	0.037013	0.157608	0.234844	0.8161
LOG(PG)	0.014406	0.137571	0.104714	0.9174
LOG(E(-1))	-0.063600	0.172575	-0.368535	0.7153
RESID(-1)	0.160978	0.237106	0.678930	0.5030

R-squared	0.016786	Mean dependent var	1.05E-16
Adjusted R-squared	-0.165291	S.D. dependent var	0.063115
S.E. of regression	0.068132	Akaike info criterion	-2.371773
Sum squared resid	0.125333	Schwarz criterion	-2.099681
Log likelihood	45.13425	F-statistic	0.092189
Durbin-Watson stat	1.881991	Prob(F-statistic)	0.992771

Appendix 2-U: Ramsey RESET test output for equation 7 (Partial adjustment model)

Ramsey RESET Test:

F-statistic	0.000536	Probability	0.981692
Log likelihood ratio	0.000656	Probability	0.979572

Test Equation:
Dependent Variable: LOG(E)
Method: Least Squares
Date: 04/12/11 Time: 18:00
Sample: 1977 2010
Included observations: 33

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.591401	2.745684	-0.215393	0.8311
LOG(P)	0.084189	0.083486	1.008420	0.3222
LOG(Y)	0.244775	0.333938	0.732995	0.4699
LOG(PG)	0.031901	0.142152	0.224417	0.8241
LOG(E(-1))	0.654981	0.767377	0.853533	0.4009
FITTED^2	0.004065	0.175496	0.023161	0.9817

R-squared	0.958783	Mean dependent var	3.021995
Adjusted R-squared	0.951150	S.D. dependent var	0.310879
S.E. of regression	0.068710	Akaike info criterion	-2.354865
Sum squared resid	0.127470	Schwarz criterion	-2.082772
Log likelihood	44.85527	F-statistic	125.6137
Durbin-Watson stat	1.642538	Prob(F-statistic)	0.000000

Appendix 2-V : Chow breakpoint test

Chow Breakpoint Test: 1999

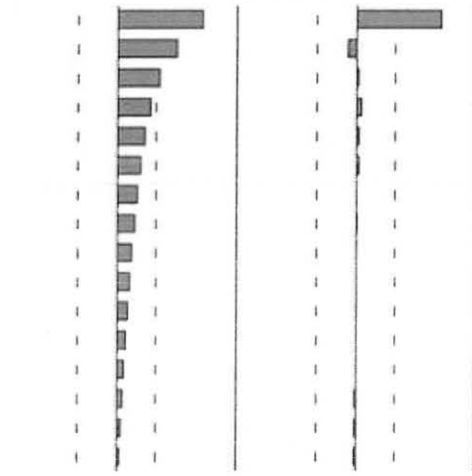
F-statistic	0.546516	Probability	0.739247
Log likelihood ratio	3.704705	Probability	0.592664

APPENDIX 3: Electricity Demand Forecast in Ethiopia for 2011-2015

Appendix 3-A: The correlogram of Ethiopian electricity consumption data up to 16 Lags

Correlogram of E

Date: 04/10/11 Time: 04:52
Sample: 1976 2010
Included observations: 35



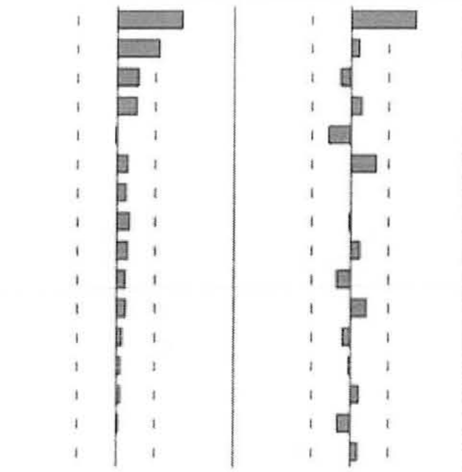
Autocorrelation / Partial Correlation (lag)	AC	PAC	Q-Stat	Prob
1	0.745	0.745	21.124	0.000
2	0.524	-0.069	31.884	0.000
3	0.376	0.025	37.612	0.000
4	0.291	0.043	41.152	0.000
5	0.240	0.030	43.644	0.000
6	0.207	0.025	45.560	0.000
7	0.177	0.007	47.009	0.000
8	0.152	0.011	48.115	0.000
9	0.129	0.005	48.948	0.000
10	0.109	0.003	49.568	0.000
11	0.092	0.001	50.022	0.000
12	0.076	0.000	50.352	0.000
13	0.064	0.000	50.591	0.000
14	0.052	-0.003	50.758	0.000
15	0.040	-0.007	50.860	0.000
16	0.025	-0.013	50.902	0.000

Appendix 3-B: The correlogram of the first-differenced data up to 16 Lags

Correlogram of D(E)

Date: 04/10/11 Time: 04:56
Sample: 1976 2010
Included observations: 34

Autocorrelation | Partial Correlation



	AC	PAC	Q-Stat	Prob
1	0.576	0.576	12.321	0.000
2	0.385	0.079	17.984	0.000
3	0.194	-0.083	19.476	0.000
4	0.173	0.106	20.703	0.000
5	-0.005	-0.196	20.704	0.001
6	0.099	0.227	21.135	0.002
7	0.089	0.001	21.490	0.003
8	0.115	-0.002	22.111	0.005
9	0.102	0.088	22.624	0.007
10	0.079	-0.121	22.941	0.011
11	0.083	0.140	23.309	0.016
12	0.055	-0.063	23.480	0.024
13	0.037	-0.016	23.558	0.035
14	0.034	0.079	23.630	0.051
15	0.024	-0.114	23.667	0.071
16	0.002	0.071	23.668	0.097

Appendix 3-C: The Correlogram of the second-differenced data up to 16 Lags

Correlogram of D(E,2)

Date: 04/10/11 Time: 04:58
Sample: 1976 2010
Included observations: 33

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.338	-0.338	4.1281	0.042
		2	0.006	-0.122	4.1296	0.127
		3	-0.108	-0.169	4.5821	0.205
		4	0.143	0.052	5.3978	0.249
		5	-0.047	0.010	5.4873	0.359
		6	0.068	0.074	5.6858	0.459
		7	-0.071	-0.001	5.9073	0.551
		8	0.053	0.029	6.0359	0.643
		9	0.019	0.061	6.0534	0.735
		10	-0.027	-0.015	6.0903	0.808
		11	0.026	0.036	6.1265	0.865
		12	0.017	0.035	6.1415	0.909
		13	-0.035	-0.024	6.2108	0.938
		14	0.018	0.005	6.2295	0.960
		15	0.016	0.015	6.2450	0.975
		16	-0.003	0.005	6.2455	0.985

Appendix 3-D: Estimation output of OLS for equation (30)

Dependent Variable: D(D(E))
Method: Least Squares
Date: 04/10/11 Time: 12:56
Sample (adjusted): 1979 2010
Included observations: 32 after adjustments
Convergence achieved after 52 iterations
Backcast: OFF (Roots of MA process too large)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.840329	2.683642	2.548897	0.0164
D(D(E(-1)))	0.244187	0.203965	1.197204	0.2409
MA(1)	4.396069	0.217449	-6.420225	0.0000

R-squared	0.575517	Mean dependent var	1.333378
Adjusted R-squared	0.546242	S.D. dependent var	71.98854
S.E. of regression	48.49260	Akaike info criterion	10.68976
Sum squared resid	68194.42	Schwarz criterion	10.82717
Log likelihood	-168.0361	F-statistic	19.65920
Durbin-Watson stat	2.045465	Prob(F-statistic)	0.000004

Inverted MA Roots 1.40
Estimated MA process is noninvertible

<u>Appendix 3-E:</u> The correlogram of the residual from equation (31)

Correlogram of Residuals

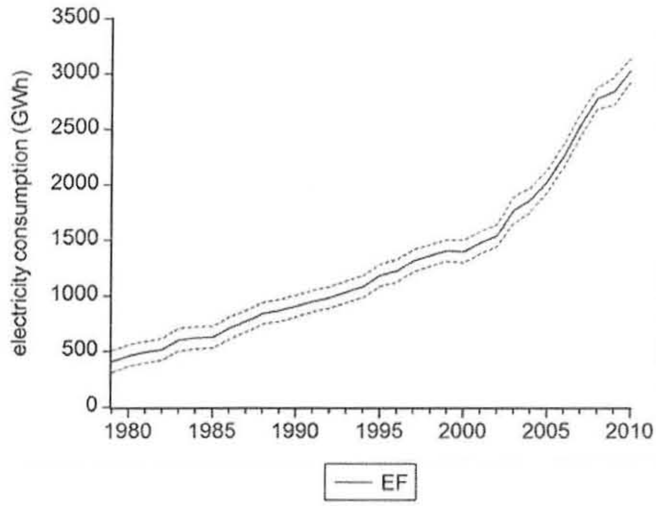
Date: 04/12/11 Time: 02:38
Sample: 1979 2010
Included observations: 32
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.053	-0.053	0.0982	
		2	0.001	-0.002	0.0983	0.754
		3	-0.134	-0.134	0.7695	0.681
		4	0.141	0.129	1.5420	0.673
		5	-0.043	-0.033	1.6169	0.806
		6	0.033	0.015	1.6619	0.894
		7	-0.057	-0.022	1.8028	0.937
		8	0.069	0.041	2.0191	0.959
		9	0.032	0.052	2.0672	0.979
		10	-0.021	-0.036	2.0890	0.990
		11	0.011	0.037	2.0956	0.996
		12	0.009	0.003	2.0999	0.998
		13	-0.025	-0.037	2.1346	0.999
		14	0.027	0.039	2.1777	1.000
		15	0.011	0.007	2.1854	1.000
		16	0.000	-0.004	2.1854	1.000

Appendix 3-F: Actual, fitted and residual table

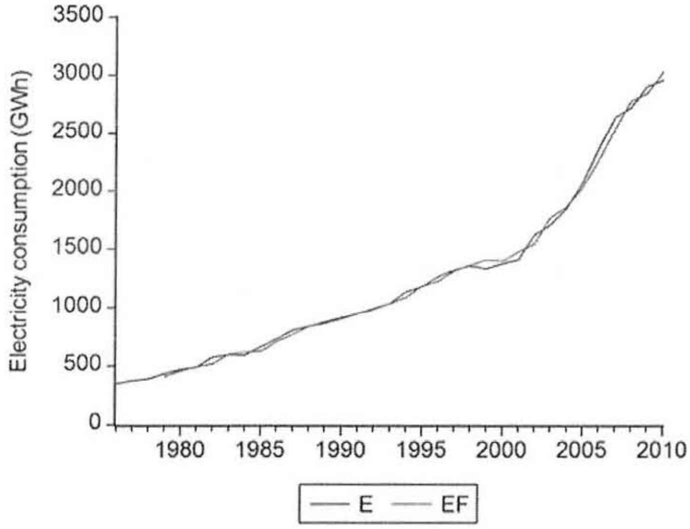
obs	Actual	Fitted	Residual
1979	32.5632	6.94122	25.6220
1980	-12.0959	-26.1214	14.0256
1981	-15.4355	-12.8673	-2.56817
1982	64.4225	9.96930	54.4532
1983	-59.4310	-64.1341	4.70311
1984	-29.1787	-2.94291	-26.2358
1985	75.0865	41.9672	33.1193
1986	-4.36649	-33.8181	29.4516
1987	10.4103	-33.8013	44.2117
1988	-45.4905	-53.3598	7.86931
1989	7.05100	-6.45006	13.5011
1990	-0.98370	-10.9150	9.93131
1991	-6.84340	-6.46986	-0.37354
1992	13.4213	7.46990	5.95140
1993	-1.00364	-0.02080	-0.98284
1994	56.2907	8.68950	47.6012
1995	-55.0352	-55.1359	0.10076
1996	45.0735	3.73107	41.3424
1997	-34.7332	-47.1588	12.4256
1998	-18.8631	-12.0916	-6.77152
1999	-62.0632	15.5890	-77.6522
2000	69.5875	111.287	-41.6995
2001	-7.03870	69.7597	-76.7984
2002	172.929	113.618	59.3110
2003	-124.003	-63.9473	-60.0560
2004	54.7504	82.8857	-28.1353
2005	82.4127	49.9698	32.4429
2006	69.5512	-32.4102	101.961
2007	-16.4212	-129.799	113.378
2008	-194.619	-151.152	-43.4669
2009	101.846	55.3290	46.5174
2010	-125.122	-50.7189	-74.4034

Appendix 3-G: Sample period forecasts and their evaluation.



Forecast: EF
Actual: E
Forecast sample: 1976 2010
Adjusted sample: 1979 2010
Included observations: 32

Root Mean Squared Error	46.16358
Mean Absolute Error	35.44266
Mean Abs. Percent Error	2.803694
Theil Inequality Coefficient	0.015377
Bias Proportion	0.037008
Variance Proportion	0.001089
Covariance Proportion	0.961903



Correlogram of Residuals Squared

Date: 06/29/11 Time: 08:45
Sample: 1979 2010
Included observations: 32
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.247	0.247	2.1471	
		2	-0.018	-0.085	2.1594	0.142
		3	-0.044	-0.019	2.2312	0.328
		4	-0.044	-0.031	2.3068	0.511
		5	-0.033	-0.019	2.3512	0.671
		6	-0.037	-0.030	2.4080	0.790
		7	0.009	0.023	2.4116	0.878
		8	0.009	-0.006	2.4150	0.933
		9	0.025	0.024	2.4451	0.964
		10	0.042	0.031	2.5339	0.980
		11	0.023	0.006	2.5609	0.990
		12	0.038	0.037	2.6383	0.995
		13	0.027	0.015	2.6786	0.997
		14	0.018	0.016	2.6987	0.999
		15	0.007	0.006	2.7017	0.999
		16	0.000	0.004	2.7017	1.000

Declaration

I, the undersigned, hereby declare that this thesis entitled, "Analysis of Electricity Demand in Ethiopia Using Partial Adjustment and ARIMA modeling" is my own work, and that all the source I have used or quoted have been indicated or acknowledge by means of completed references.

Misganaw Abebaw

Name

27/06/2011

Signature and date

This thesis has been submitted for examination with my approval as a university advisor.

Butte Gotu

Name

Signature and date