FORECASTING OF PETROLEUM OIL IMPORT PRICES BY GARCH MODELS: THE CASE OF ETHIOPIA

LEYKUN GETANEH

A THESIS SUBMITTED TO

THE DEPARTMENT OF STATISTICS

PRESENTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS (APPLIED STATISTICS)

ADDIS ABABA UNIVERSITY

ADDIS ABABA, ETHIOPIA

NOVEMBER 2012

ADDIS ABABA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared by Leykun Getaneh, entitled: Forecasting of petroleum oil import prices by GARCH models: The case of Ethiopia and in partial fulfillment of the requirement for the degree of master of science in Statistics (Applied Statistics) complies with requirements of the university and meets the accepted standards with respect to originality and quality.

Signed by the examining committee:

Examiner ______________________________Signature_________________ Date________

Examiner ______________________________Signature_________________ Date________

Advisor ______________________________Signature_________________ Date________

____

Chair of the department or graduate program coordinator

Abstract

Forecasting of petroleum oil import prices by GARCH models: The case of Ethiopia

Leykun Getaneh

Addis Ababa University, 2012

The price of oil affects everyone, everyday. Petroleum oil (oil) is an important energy commodity to mankind. Several causes have made petroleum oil prices to be volatile. The fluctuation of petroleum oil import prices of Ethiopia has affected many related sectors, stock market indices and the country's economy. The main objectives of this study were to estimate and forecast petroleum oil import prices in Ethiopia. The data for this study were monthly petroleum oil import prices data obtained from the National Bank of Ethiopia for the period from July 1999 to June 2011. The Generalized Autoregressive Conditional Heteroscedasticity (GARCH) approach was used to forecast monthly petroleum oil import prices of Ethiopia. GARCH(1,1) was found to be a suitable model in forecasting monthly petroleum oil import prices of Ethiopia due to its ability to capture the non constant conditional variance. Based on the fitted GARCH(1,1) model monthly petroleum oil import prices were forecasted from July 2011 to June 2013. The forecasted results obtained from GARCH (1,1) model showed that import petroleum oil prices are expected to increase from July 2011 to June 2013 with some fluctuations as a result of volatility.

ACKNOWLEDGMENT

I would like to express my special thanks to my adviser Dr. Butte Gotu for his guidance and constructive suggestions.

My appreciation also goes to my father Getaneh Gebeye, my mother Emebet Gezehagn, my sister Mahlet Getaneh, my brothers Fiseha and Tewolde. I would like to express my gratitude to all those who gave me support to complete this thesis.

Table of Contents

List of Tables

List of Figures and Tables in the Appendix

List of Acronyms

ACF	Autocorrelation functions
ADF	Augmented Dickey-Fuller
AIC	Akaike Information Criterion
AICc	Akaike Information Criterion, bias corrected
AR	Autoregression
ARCH	Autoregressive Conditional Heteroscedasticity
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CGARCH	Component GARCH
DW	Durbin-Watson
EGARCH	Exponential GARCH
EViews	Econometric Views
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GED	Generalized Exponential distribution
I(0)	Integrated of order zero
I(1)	Integrated of order one
IGARCH	Integrated GARCH
JB	Jarque-Bera
LLF	Log-likelihood function

LM	Lagrange Multiplier
MA	Moving Average
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MSFE	Mean Squared Forecast Error
NBE	National Bank of Ethiopia
OPEC	Organization of the Petroleum Exporting Countries
PACF	Partial Autocorrelation Functions
PP	Phillips-Perron
QMS	Quantitative Micro Software
RMSE	Root Mean Squared Error
RSS	Residual sum of squares
SIC	Schwarz Information Criterion
Theil-U	Theil Inequality Coefficient
VaR	Value at Risk

1. Introduction

1.1. Background of the study

Petroleum-oil (oil) is a naturally occurring and flammable liquid found in rock formations in the earth. It consists of a complex mixture of hydrocarbons of various molecular weights plus other organic compounds.

Oil is a commodity, which unlike any other, affects everyone's daily life in a plethora of ways. The modern world is as dependent on oil as the human body is dependent on blood. Oil prices and availability affects transportation be it every day driving or flights, as well as economical growth, since goods have to be transported and oil is used almost everywhere in the secondary industry. Machines have to be oiled; engines need fuel and some products are entirely oil-based. The insulation on the wires in computers and the circuit boards are made from crude oil. Heating and of course the military is heavily dependent on oil prices and availability. The uniqueness of oil and its use in the global economy makes it challenging to find another comparable resource. Furthermore, one must not forget that oil is a non-renewable fossil resource. Presently economic activity and oil consumption are fairly correlated. It is not only for the aforementioned reasons that the interest in oil prices and in particular, in the ability of being able to forecast petroleum oil import prices is crucial (Fabian, 2010).

One of the most prominent stylized facts of returns on petroleum oil prices is that their volatility changes over time. In particular, periods of large movements in petroleum oil prices alternate with periods during which prices hardly change. This characteristic feature commonly is referred to as volatility clustering (Brooks, 2008).

In statistics, a sequence of random variables is heteroscedastic if the random variables have different variances. The term heteroscedastic means "differing variance" and comes from the Greek "hetero" ('different') and "skedasis" ('dispersion'). In contrast, a sequence of random variables is called homoscedastic if it has constant variance (Brooks, 2008).

We consider crude oil price data as heteroscedastic time series models where the conditional variance given in the past is no longer constant (Palma, 2007). In financial analysis, forecast of future volatility of a series under consideration are often of interest to assess the risk associated with certain assets. In that case, variance forecasts are of direct interest, of course (Lütkepohl, 2005).

The Nobel Prize winner, Robert Engle (1982), developed the ARCH model asserting that although many financial time series, such as, stock returns and exchange rates are unpredictable, there is apparent clustering in the variability or volatility. This is often referred to as conditional heteroscedasticity since it is assumed that overall the series could be stationary but the conditional expected value of the variance may be time dependent.

Later, Bollerslev (1986) had modified Engle's ARCH model into a more generalized model called GARCH model which is a simplified model to ARCH model but more powerful. Currently, this model is being widely used in many financial time series data. The simple GARCH model is able to detect the financial volatility in a time trend.

1.2. Statement of the problem

The price of the energy commodity is highly volatile throughout time. Since petroleum oil import prices variability or fluctuation does affect other sectors and stock market, the forecasting of future petroleum oil import prices becomes crucial.

This study attempts to forecast the volatility of petroleum oil import prices of Ethiopia.

1.3. Objectives of the Study

General objective

The main objectives of this study are to estimate and forecast petroleum oil import prices in Ethiopia.

Specific Objectives

- To fit and select the appropriate ARIMA and GARCH models for data on petroleum oil import prices;
- To determine the appropriate modeling approach for petroleum oil import prices in Ethiopia.

1.4. Scope of the Study

This study focuses on the Box-Jenkins and GARCH models to forecast petroleum oil import prices of Ethiopia. Since the oil price volatility is the main concern, the study uses only monthly data. The data were obtained from NBE for the period between October 1999 and June 2011.

1.5. Significance of the Study

Since petroleum oil market is highly volatile, the estimation of the time series model must be able to detect volatility. The result of this study could help:

- In understanding petroleum oil import prices volatility.
- In managing, creating and revising effective administrative procedure, rules and regulation and strategies for organizations working on import sectors.
- To adjust the domestic retail prices of petroleum products.
- As a basis to other researchers for further study.

2. Literature review

Several world events have led to major oil disruptions in the past. Most of these disruptions were related to political or military upheavals, especially occurred in the Middle East. Since 1973, there were four crises which have caused oil prices to be volatile. These include the 1973 Arab–Israeli war, the 1978–89 Iranian revolution, the 1980 Iran–Iraq war and the 1990–91 Gulf war which have resulted in initial shortfalls of between 4.0 and 5.6 million barrels per day (Marimoutou *et al.*, 2009). In 1999, the increase in Iraq oil production coincided with the Asian financial crises which caused the oil price to drop due to reduced demand.

Crude oil prices change due to many reasons. These reasons include OPEC policy, war and political uncertainty in several places such as the Middle East, supply disruptions due to natural or other disasters, and changing demand and imbalances between physical supply and demand (Marimoutou *et al.*, 2009).

Some other unexpected issues have also had some effect on oil prices. For example, the post-9/11 war on terror, labor strikes, hurricane threats to oil platforms, fires and terrorist threats at refineries, and other short-lived problems are responsible for the higher prices. Such problems do push prices into a higher level temporarily, but have not historically been fundamental to long-term price increases.

In 2009/10, total merchandize import, which showed a 21 percent yearly average growth in the past five years, rose moderately by 7 percent to USD 8.3 billion. Hence, the proportion of import in GDP increased to 26.5% in 2008/09. Imports of petroleum products bill went up marginally by 4.3% to USD 1.3 billion as a result of recovery in international oil prices and increase in volume of oil imports driven by the rising domestic demand (NBE, 2009/2010).

The forecast of crude oil prices is essential because it affects other sectors and stock markets. Various forecasting techniques can be employed to analyze data on petroleum oil import prices.

The merits of ARIMA models are two-fold (Wang *et al.*, 2005). ARIMA models are a class of typical linear models which are designed for linear time series and capture linear characteristics in time series. ARIMA models are widely used in many practical applications. However, the disadvantage of the ARIMA is that it cannot capture nonlinear patterns of complex time series if nonlinearity exists.

Moshiri and Foroutan (2006) modeled and forecasted daily crude oil futures prices that are listed in MYMEX from 1983 to 2003, applying linear and nonlinear time series models. They used EViews to estimate and forecast crude oil futures prices. They discovered that linear ARMA (1, 3) and nonlinear GARCH (2, 1) models were the most suitable. However, the GARCH model was better than the ARMA model.

Models for volatility and forecasting volatility stock market have been the subject of vast empirical and theoretical investigations over the past decade by academics and practitioners. There are a number of motivations for this line of inquiry. Arguably, volatility is one of the most important concepts in finance. Volatility, as measured by the standard deviation or variance of returns, is often used as a basic measure of the total risk of financial assets. Many value-at-risk models require the estimation or forecast of a volatility parameter. Also, the volatility of stock market prices enters directly into the Black-Scholes formula for deriving the prices of traded options (Brooks, 2008)

In the literature, there were quite a number of relevant researches regarding the Box-Jenkins method and GARCH-type models in forecasting crude oil prices.

Sadorsky (1999) found that oil price volatility shocks have asymmetric effects on the economy. The changes in oil prices affect economic activity but the changes in economic activity have little impact on oil prices. Therefore, oil price fluctuations have large macroeconomic impacts.

Sadorsky (2006) also found that the out-of-sample forecasts of a single equation GARCH model are more superior to those of state space, vector auto regression and bivariate GARCH models in forecasting petroleum futures prices.

Agnolucci (2009) stated that GARCH-type models are able to have a better performance than the implied volatility (IV) models in terms of predictive accuracy. Moreover, some conclusions are drawn, which included the mean returns from oil futures can be assumed to be constant across time, shocks to the conditional varianceof the series have been found to be highly persistent, the parameters in the models are robust to the distribution assumed for the errors and no leverage effect can be observed in the oil future series.

Siti *et al*. (2011) stated that GARCH models are able to have a better performance than the ARIMA model in terms of forecasting accuracy in forecasting daily crude oil prices. Cabedo and Moya (2003) compared the Value at Risk (VaR) from historical simulation with ARMA forecasts (HSAF) approach to those from a standard GARCH model. They found that the VaRs from the HSAF method provides a most flexible and efficient risk qualification that outperforms those from the standard GARCH model. Sadeghi and Shavvalpour (2006) introduced the HSAF and comparison with GARCH model to estimate the VaR of OPEC oil price. In their findings HSAF was shown to be more efficient.

Morana (2001) showed how the GARCH properties of oil price changes can be employed to forecast the oil price distribution over short-term horizons using a semi-parametric approach to oil price forecasting that was based on bootstrap approach.

Kang *et al.* (2009) investigated the efficacy of a volatility model for three crude oil markets. They are Brent, Dubai, and WTI with regard to its ability to forecast and identify volatility stylized facts, in particular volatility persistence. They assessed persistence in the volatility of the three crude oil prices using conditional volatility models. One of the models is component-GARCH (CGARCH) model that was developed by Engle and Lee (1999). This model was able to distinguish between short-run and long-run persistence of volatility. Another model is the fractionally integrated

GARCH (FIGARCH) model that was introduced by Baillie *et al.* (1996) that allows for a fractional integrated process in conditional variance. This model is better equipped to capture persistence than are the GARCH and Integrated GARCH (IGARCH) models. The CGARCH and FIGARCH models also provide superior performances in out-of-sample volatility forecasts. Kang *et al.* (2009) concluded that the CGARCH and FIGARCH models are useful for modelling and forecasting persistence in the volatility of crude oil prices.

Cheong (2009) evaluated the volatility behaviour of the two major crude oil markets particularly in the empirical stylized facts such as the asymmetric news impact, long-persistence volatility and tail behaviour in the crude oil series. A very powerful and flexible ARCH model is used to evaluate the aforementioned stylized facts for both the Brent and WTI markets. Although both the estimation and diagnostic evaluations are in favour of the Fractional Integrated Asymmetric Power ARCH (FIAPARCH) model, from empirical out-of-sample forecast it appears that the simplest parsimonious GARCH fits the Brent crude oil data better than the other models. On the other hand, the FIAPARCH out-of-sample WTI forecasts provide superior performance. These findings suggest that energy economists and financial analysts should consider not only the complexity, but also the parsimonious principle and actual performance of out-of-sample forecasts, in choosing a crude oil volatility model.

Amare and Fentaw (2010) examine the GARCH model fitting and volatility forecasting of export prices. Their main findings suggest that export prices volatility is consistent and past volatility is important in predicting future volatility

Narayan and Narayan (2007) examine the volatility of crude oil price using daily data for the period 1991–2006 by Exponential GARCH (EGARCH) model. Their main findings summarized that across the various sub-samples, there is inconsistent evidence of asymmetry and persistence of shocks; and over the full sample period, evidence suggests that shocks have permanent effects and asymmetric effects, on volatility. These findings imply that the behaviour of oil prices tends to change over short periods of time.

Fong and See (2002) studied a Markov switching model of the conditional volatility of crude oil futures prices, and show that the regimes identified by their model capture major oil-related events. Hence, the significant regime shifts in the conditional volatility of crude oil futures contracts, which tends to dominate the GARCH effects. Alizadeh *et al.* (2008) investigated the hedging effectiveness of the Markov Regime Switching (MRS) models for WTI crude oil futures contracts. They then introduced a MRS vector error correction model (VECM) with GARCH error structure. This specification links the concept of disequilibrium with that of high uncertainty across high and low volatility regimes. The results indicated that using MRS models, market agents may be able to obtain superior gains, measured in terms of both variance reduction and increase in utility.

Engle's (1982) original work on ARCH was concerned with the volatility inflation. However, it is the applications of the ARCH model to financial time series that established and consolidated the significance of his contribution. Financial time series have characteristics that are well represented by the models with dynamic variances (Hill et al., 2008).

From the literature review, the analysis of petroleum oil prices has been a topic of extensive research. In this study, monthly petroleum oil import prices will be forecasted using GARCH models.

3. Data and Methodology

3.1. Data sources

The data collected were monthly petroleum oil import prices from October 1999 to June 2011. Values are quoted in birr per metric ton (MT) and these data were obtained from NBE. As a result, 141 observations were collected. We divided the whole time-period into two parts. The first time-period which is called "in-sample period" ranged from October 1999 to June 2010; and the data from this period were used to estimate GARCH and ARIMA models. The second time-period, which is called "out-of-sample period" ranged from July 2010 until June 2011; the data from this period were used to test forecast accuracy.

3.2. Methodology

3.2.1. Stationarity and Unit root test

The foundation of time series analysis is stationarity. A time series $\{X_t\}$ is said to be strictly stationary if the joint distribution of $\{x_{t_1}, x_{t_2}, \cdots, x_{t_k}\}$ is identical to that of $\{x_{t_{1+h}}, x_{t_{2+h}}, \cdots, x_{t_{k+h}}\}$ where h and k are an arbitrary positive integers. In other words, strict stationarity requires that the joint distribution of $\{x_{t_1}, x_{t_2}, \cdots, x_{t_k}\}$ is invariant under time shift. This is a very strong condition that is hard to verify empirically. A weaker version of stationarity is often assumed. A time series $\{X_t\}$ is weakly stationary if both the mean of X_t and the covariance between X_t and X_{t-h} are time-invariant, where h is an arbitrary integer (Tsay, 2010).

More specifically, $\{X_t\}$ is weakly stationary if (a) $(E(X_t) = \mu$, which is a constant mean, and (b) $cov(x_{t-h}, x_t) = \gamma_h$, which only depends on h. In practice, suppose that we have observed T data points $\{x_t | t = 1, \cdots, T\}$. The weak stationarity implies that the time plot of the data would show that the T values fluctuate with constant variation around a fixed level. In applications, weak stationarity enables one to make inferences concerning future observations (e.g., prediction).

In the finance literature, it is common to assume that an asset return series is weakly stationary (Tsay, 2010). This assumption can be checked empirically provided that a sufficient number of historical returns are available. For example, one can divide the data into subsamples and check the consistency of the results obtained across the subsamples.

3.2.1.1. Unit root tests

Many economic and financial time series exhibit a trending behavior or non stationarity in the mean. Unit root tests can be used to determine if trending data should be first differenced or regressed on deterministic functions of time to render the data stationary.

A test of stationarity (or non stationarity) that has become widely popular over the past several years is the unit root test. For a number of reasons, it is important to know whether or not an economic time series has a unit root. As Nelson and Plosser (1982) pointed out, non stationarity often has important economic implications. It is therefore very important to be able to detect the presence of unit roots in time series, normally by the use of what are called unit root tests. If a variable is stationary in level, i.e. without running any differencing, then the variable is said to be integrated of order zero, denoted by I(0). Similarly, if it becomes stationary by differencing once, then the variable is said to be integrated of order 1, written as I(1). Unit-root test helps to detect whether a variable is stationary or not. For these tests, the null hypothesis is that the time series has a unit root and the alternative is that it is I(0). The widely used unit-root tests are the Augmented Dickey Fuller test (ADF) and Phillips Perron (PP) (Zivot and Wang, 2006).

3.2.1.1.1 The Augmented Dickey-Fuller (ADF) unit root tests

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. The ADF test tests the null hypothesis that a time series x_t is I(1) against the alternative that it is I(0), assuming that the dynamics in the data have an ARMA structure. The ADF test is based on estimating the test regression (Zivot and Wang, 2006)

$$x_t = \beta^{'} D_t + \emptyset x_{t-1} + \sum_{j=1}^{p} \psi_j \, \Delta x_{t-j} + \varepsilon_t \tag{1}$$

where D_t is a vector of deterministic terms (constant, trend etc.). The p lagged difference terms, ∇x_{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 $\emptyset = 1$. The ADF t-statistic and normalized bias statistic are based on the least squares estimates of (1) and are given by

$$ADF_t = t_{\emptyset=1} = \frac{\widehat{\emptyset} - 1}{S.E(\widehat{\emptyset})}$$

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

An alternative formulation of the ADF test regression is

$$\Delta x_t = \beta' D_t + \pi x_{t-1} + \sum_{j=1}^{p} \psi_j \, \Delta x_{t-j} + \varepsilon_t \tag{2}$$

where $\pi = \emptyset - 1$. Under the null hypothesis, Δx_t is I(0) which implies that $\pi = 0$. The ADF t-statistic is then the usual t-statistic for testing $\pi = 0$ and the ADF normalized bias statistic is $T\hat{\pi}/(1 - \hat{\psi}_1 - \cdots - \hat{\psi}_p)$. 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} (Zivot and Wang, 2006)

3.2.1.1.2 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

$$\Delta x_t = \alpha_0 + \alpha_2 t + \pi x_{t-1} + \varepsilon_t \tag{3}$$

where ε_t is $I(0)$ and may be heteroskedastic. The PP test 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

$$z_t = t_{\pi=0}(\frac{\hat{\sigma}^2}{\hat{\lambda}^2})^{\frac{1}{2}} - \frac{T(\hat{\lambda}^2 - \hat{\sigma}^2)(se(\hat{\pi}))}{2\hat{\lambda}^2\hat{\sigma}^2}$$

$$z_\pi = T\hat{\pi} - \frac{T^2(\hat{\lambda}^2 - \hat{\sigma}^2)(se(\hat{\pi}))}{2\hat{\sigma}^2}$$

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

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

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

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 statistic. One advantage of the PP test over the ADF test is that the PP test is 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 (Zivot and Wang, 2006).

3.2.2 Time domain approach and ARIMA model

Before looking more closely at the particular statistical methods, it is appropriate to mention that two separate, but not necessarily mutually exclusive, approaches to time series analysis exist, commonly identified as the time domain approach and the frequency domain approach. In this study we used the time domain approach.

The time domain approach is generally motivated by the presumption that correlation between adjacent points in time is best explained in terms of a dependence of the current value on past values. The time domain approach focuses on modeling some future value of a time series as a parametric function of the current and past values. In this scenario, we begin with linear regressions of the present value of a time series on its own past values and on the past values of other series. This modeling leads one to use the results of the time domain approach as a forecasting tool and is particularly popular with economists for this reason (Shumway and Stoffer, 2010).

One approach, advocated in the landmark work of Box and Jenkins (1970) and Box et al. (1994), developed a systematic class of models called autoregressive integrated moving average (ARIMA) models to handle time correlated modeling and forecasting. The approach includes a provision for treating more than one input series through multivariate ARIMA or through transfer function modeling. The defining feature of these models is that they are multiplicative models, meaning that the observed data are assumed to result from products of factors involving differential or difference equation operators responding to a white noise input.

The ARIMA class of models is an important forecasting tool, and is the basis of many fundamental ideas in time-series analysis. The acronym ARIMA stands for 'autoregressive integrated moving average', and the different components of this general class of models will be introduced in turn. The original key reference is Box and Jenkins (1970), and ARIMA models are sometimes called Box-Jenkins models (Chatfield, 2000).

3.2.2.1. Autoregressive (AR) Process

Autoregressive models are based on the idea that the current value of the series, x_t, can be explained as a function of p past values, $x_{t-1}, x_{t-2}, \cdots, x_{t-p}$, where p determines the number of steps into the past needed to forecast the current value. An autoregressive model of order p, abbreviated AR(p), is the form

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + z_t \tag{4}$$

where x_t is stationary, and $\phi_1, \phi_2, \cdots, \phi_p$ are constant ($\phi_p \neq 0$). Although it is not necessary yet, we assume that z_t is a Gaussian white noise series with mean zero and variance σ_z^2. Using the back shift operator B, such that $Bx_t = x_{t-1}$, the AR (P) may be written more succinctly in the form

$$\phi(B)x_t = z_t \tag{5}$$

where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ is a polynomial in B of order p. The properties of AR processes defined by (4) can be examined by looking at the properties of the function ϕ.As B is an operator, the algebraic properties of ϕ have to be investigated by examining the properties of $\phi(x)$, say, where x denotes a complex variable, rather than by looking at $\phi(B)$. It can be shown that (5) has a unique causal stationary solution provided that the roots of $\phi(x) = 0$ lie outside the unit circle. This solution can be expressed in the form

$$x_t = \sum_{j \geq 0}^{\infty} \psi_j z_{t-j} \tag{6}$$

for some constants ψ_j such that $\sum |\psi_j| < \infty$. Typically an AR process is stationary provided that the roots of $\phi(x) = 0$ lie outside the unit circle (Chatfield, 2000). For higher-order stationary AR processes, the autocorrelation function (ACF) will typically be a mixture of terms which decrease exponentially or of damped sine or cosine waves. The ACF can be found by solving a set of difference equations called the Yule-Walker equations given by (Chatfield, 2000)

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \tag{7}$$

for $k = 1,2,\cdots$, where $\rho_0 = 0$.

A useful property of an AR(p) process is that it can be shown that the partial ACF is zero at all lags greater than p. This means that the sample partial ACF can be used to help determine the order of an AR process (assuming the order is unknown as is usually the case) by looking for the lag value at which the sample partial ACF 'cuts off' (meaning that it should be approximately zero, or at least not significantly different from zero, for higher lags) (Chatfield, 2000).

3.2.2.2. Moving average (MA) processes

A time series $\{x_t\}$ is said to be a moving average process of order q (abbreviated MA(q)) if it is a weighted linear sum of the last q random shocks so that

$$x_t = z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_q z_{t-q} \quad (8)$$

where $\{z_t\}$ denotes a purely random process with zero mean and constant variance σ_z^2. Equation (8) may alternatively be written in the form

$$x_t = \theta(B) z_t \quad (9)$$

where $\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$ is the polynomial in B of order q. Note that some authors (including Box et al., 1994) parameterize an MA process by replacing the plus signs in (8) with minus signs, presumably so that θ(B) has a similar form to ϕ(B) for AR processes, but this seems less natural in regard to MA processes. There is no difference in principle between the two notations but the signs of the θ values are reversed and this can cause confusion when comparing formulae from different sources or examining computer output.

It can be shown that a finite-order MA process is stationary for all parameter values. However, it is customary to impose a condition on the parameter values of an MA model, called the invertibility condition, in order to ensure that there is a unique MA model for a given ACF. This condition can be explained as follows. Suppose that $\{z_t\}$ and $\{z'_t\}$ are independent purely random processes and that $\theta \in (-1,1)$. Then it is straightforward to show that the two MA(1) processes defined by $x_t = z_t + \theta z_{t-1}$and $x_t = z'_t + \theta^{-1} z'_{t-1}$ have exactly the same autocorrelation function (ACF). Thus the polynomial θ(B) is not uniquely determined by the ACF. As a consequence, given a sample ACF, it is not possible to estimate a unique MA process from a given set of data without putting some constraint on what is allowed. To resolve this ambiguity, it is usually required that the polynomial $\theta(x)$ has all its roots outside the unit circle (Chatfield, 2000).

We can rewrite (8) in the form

$$x_t - \sum_{j \geq 1}^{\infty} \pi_j x_{t-j} = z_t \tag{10}$$

for some constants π_j such that $\sum |\pi_j| < \infty$. In other words, we can invert the function taking the z_t sequence to the x_t sequence and recover z_t from present and past values of x_t by a convergent sum. The negative sign of the π-coefficients in (10) is adopted by convention (chatfield, 2000) so that we are effectively rewriting an MA process of finite order as an AR(∞) process. The astute reader will notice that the invertibility condition (roots of $\theta(x)$ lie outside the unit circle) is the mirror image of the condition for stationarity of an AR process (roots of $\phi(x)$ lie outside the unit circle).

The ACF of an MA(q) process can readily be shown to be

$$\rho_k = \begin{cases} 1 & k = 0 \\ \sum_{i=0}^{q-k} \theta_i \theta_{i+k} \Big/ \sum_{i=0}^{q} {\theta_i}^2 & k = 1,2,\cdots,q \\ 0 & k > q \end{cases} \tag{11}$$

where $\theta_0 = 1$. Thus the ACF cuts off at lag q. This property may be used to try to assess the order of the process (i.e. what is the value of q?) by looking for the lag beyond which the sample ACF is not significantly different from zero (Chatfield, 2000).

3.2.2.3. Autoregressive Moving Average (ARMA) Processes

In some applications, the AR or MA models become cumbersome because one may need a higher-order model with many parameters to adequately describe the dynamic structure of the data. To overcome this difficulty, the autoregressive moving-average (ARMA) models are introduced (Box, *et al* (1994)). Basically, an ARMA model combines the ideas of AR and MA models into compact form so that the number of parameters used is kept small. For the return series in finance, the chance of using ARMA models is low (Tsay, 2010). However, the concept of ARMA models is highly relevant in volatility modeling.

A mixed autoregressive moving average model with *p* autoregressive terms and *q* moving average terms is abbreviated ARMA(*p*, *q*) and can be written as

$$\phi(B)x_t = \theta(B)z_t \tag{12}$$

where $\phi(B)$ and $\theta(B)$ are polynomials in B of finite order p and q, respectively. Equation (12) has a unique causal stationary solution provided that the roots of $\phi(x) = 0$ lie outside the unit circle. The process is invertible provided that the roots of $\theta(x) = 0$ lie outside the unit circle. In the stationary case, the ACF will generally be a mixture of damped exponentials or sinusoids.

The importance of ARMA process is that many real data sets may be approximated in a more parsimonious way (meaning fewer parameters are needed) by a mixed ARMA model rather than by a pure AR or pure MA process.

3.2.2.4. Autoregressive Integrated Moving Average (ARIMA) Processes

We now reach the more general class of models which is ARIMA models. In practice many time series are non-stationary and we cannot apply stationary AR, MA or ARMA processes directly. One possible way of handling non-stationary series is to apply differencing so as to make them stationary. The first differences, namely $(x_t - x_{t-1}) = (1 - B)x_t$, may themselves be differenced to give second differences, and so on. The dth differences may be written as $(1 - B)^d x_t$. If the origin data series is differenced d times before fitting an ARMA(p,q) process, then the model for the original undifferenced series is said to be an ARIMA(p,d,q) process where the letter 'I' in the acronym stands for integrated and d denotes the number of differences taken.

Mathematically, (12) is generalized to give:

$$\phi(B)(1-B)^d x_t = \theta(B) z_t \qquad (13)$$

The combined AR operator is now $\phi(B)(1-B)^d$. If we replace the operator B in this expression with a variable x, it can be seen straight away that the function $\phi(x)(1-x)^d$ has d roots on the unit circle indicating that the process is non-stationary – which is why differencing is needed (Chatfield, 2000).

3.2.3. Building ARIMA Models

There are a few basic steps to fitting ARIMA models to time series data. These steps involve plotting the data, possibly transforming the data, identifying the dependence orders of the model, parameter estimation, diagnostic checking and forecasting.

3.2.3.1. Model identification

The first step in ARIMA modeling is identification of the appropriate model. ARIMA procedures are mainly based on the analysis of the autocorrelation function (ACF) and the partial autocorrelation function (PACF), which are essential diagnostic instruments to identify the dependence structure of the series. The autocorrelation is the correlation of the 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:

$$\hat{\rho}(k) = \frac{\sum_{t=k+1}^{T}(X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^{T}(X_t - \bar{X})^2} \tag{14}$$

where $\bar{X}$ the sample mean of X. This is the correlation coefficient for values of the series k periods apart.

The partial autocorrelation at lag k is estimated by:

$$\hat{\phi}_k = \begin{cases} \rho_1 & for\ k = 1 \\ \dfrac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1}, j\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_{k-1}, j\rho_{k-j}} & for\ k > 1 \end{cases} \tag{15}$$

where $\hat{\rho}_k$ is the estimated autocorrelation at lag k and $\hat{\phi}_{k,j} = \phi_{k-1}, j - \phi_k \phi_{k-1,k-j}$

The first step in determining the AR and MA terms 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 from the current one. For a first-order difference, the current value is then considered, on average, as equal to the preceding value.

Taking difference may eliminate the trend and make the series stationary. 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.

Table 3.1. Behavior of the ACF and PACF for ARMA Models

Model	**ACF**	**PACF**
AR(p) $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + z_t$	Tails off	Cuts off after lag p
MA(q) $x_t = z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2} + \cdots + \theta_q z_{t-q}$	Cuts of after lag q	Tails off
ARMA(p ,q) $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + z_t + \theta_1 z_{t-1} + \theta_2 z_{t-2}$ $+ \cdots + \theta_q z_{t-q}$	Tails off	Tails off

Model selection criteria

The idea is to fit all ARMA(p,q) models with orders $p \leq p_{max}$ and $q \leq q_{max}$ and choose the values of p and q which minimize some model selection criteria. Model selection criteria for ARMA(p,q) models have the form:

Akaike's Information Criterion (AIC)

Akaike (1973) suggested measuring the goodness of fit for ARMA(p,q) by balancing the error of the fit against the number of parameters in the model.

$$AIC = \log \hat{\sigma}_k^2 + \frac{n + 2k}{n} \tag{16}$$

where $\hat{\sigma}_k^2$ is the maximum likelihood estimator for the variance and k is the number of parameters in the model. The value of k yielding the minimum AIC specifies the best model. The idea is roughly that minimizing $\hat{\sigma}_k^2$ would be a reasonable objective, except that it decreases monotonically as k increases. Therefore, we ought to penalize the error variance by a term proportional to the number of parameters.

AIC, Bias Corrected (AICc)

A corrected form, suggested by Sugiura (1978), and expanded by Hurvich and Tsai (1989), can be based on small-sample distributional results for the linear regression model. The corrected form is defined as follows.

$$\text{AICc} = \log \hat{\sigma}_k^2 + \frac{n + k}{n - k - 2} \tag{17}$$

where $\hat{\sigma}_k^2$ is the maximum likelihood estimator for the variance, k is the number of parameters in the model, and n is the sample size.

Bayesian Information Criterion (BIC)

We may also derive a correction term based on Bayesian arguments, as in Schwarz (1978), which leads to defined as

$$\text{BIC} = \log \hat{\sigma}_k^2 + \frac{k \log n}{n} \tag{18}$$

BIC is also called the Schwarz Information Criterion (SIC); Rissanen (1978) for an approach yielding the same statistic based on a minimum description length argument. Various simulation studies have tended to verify that BIC does well at getting the correct order in large samples, whereas AICc tends to be superior in smaller samples where the relative number of parameters is large (McQuarrie and Tsai, 1998) as pointed out Shumway and Stoffer (2010).

3.2.3.2. Model estimation

Having identified the appropriate p, d and q values, and the next step 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

Once a model has been identified and the parameters estimated, *diagnostic checks* are then applied to the fitted model. We next see whether the selected 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 respecification of ARIMA model.

Ljung-Box test

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

$$Q = T(T+2)\sum_{j=1}^{k}\frac{\hat{\rho}_j^{\,2}}{T-j} \tag{19}$$

where $\hat{\rho}_j$ is the *j-th* autocorrelation and T is the number of observations. If the series is not based upon results of ARIMA estimation, then under the null hypothesis, Q is asymptotically distributed as chi-square with degrees of freedom equal to the number of autocorrelations. If the series represents the residuals from ARIMA estimation, the appropriate degrees of freedom should be adjusted to represent the number of autocorrelations less the number of AR and MA terms previously estimated.

Heteroskedasticity tests

This displays the autocorrelations and partial autocorrelations of the squared residuals up to any specified number of lags and computes the Ljung-Box Q-statistic for the corresponding lags. The correlogram of the squared residuals can be used to check autoregressive conditional heteroskedasticity (ARCH) in the residuals. If there is no ARCH in the residuals, the autocorrelations and partial autocorrelations should be zero at all lags and the *Q*-statistic should not be significant.

Normality

This displays the frequency distribution of a series in a histogram. The histogram divides the series range (the distance between the maximum and minimum values) into a number of equal length intervals or bins and displays a count of the number of observations that fall into each bin. Complements of standard descriptive statistics are displayed along with the histogram. All of the statistics are calculated using the observations in the current sample.

Jarque-Bera

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

Bera and Jarque (1981) formalize these ideas by testing whether the coefficients of excess kurtosis are jointly zero. Denoting the errors by u and their variance by σ^2, the coefficients of skewness and kurtosis can be expressed, respectively, as

$$b_1 = \frac{E(u^3)}{(\sigma^2)^{3/2}} \quad \text{and} \quad b_2 = \frac{E(u^4)}{(\sigma^2)^2}$$

The kurtosis of the normal distribution is 3 so its excess kurtosis $(b_2 - 3)$ is zero.

The Jarque-Bera test statistic is given by

$$JB = T\left[\frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24}\right] \tag{20}$$

where T is the sample size. Under the null hypothesis this test statistic asymptotically follows a chi-square distribution with two degree of freedom (Brooks, 2008).

3.2.3.4. Forecasting

The last step in ARIMA modeling is forecasting. Once a model has been selected, estimated and checked, it is usually a straightforward task to compute forecasts.

3.2.4 Forecasting evaluation and accuracy criteria

The accuracy of a forecasting method is determined by analyzing forecast error experiences. The forecasting performance of the estimators is judged on the basis of the differences between predictions and realizations. The smaller the difference between the predictions and the actual values of the dependent variable, the better is the forecasting performance of the estimator. The forecasting performance of the system should be assessed using standard statistical tools such as Root Mean Square Error, Mean Absolute Error, Mean Absolute Percentage Error and Theil's inequality Coefficient. The first two forecast error statistics depend on the scale of the dependent variables, and the remaining two statistics are scale invariant (i.e. unit free). In most instances unit free measures are preferable (Challen and Hagger, 1983). As a result Theil's Inequality coefficient (Theil-U) and Mean Absolute Percentage Error (MAPE) are preferable. If the forecast is good, the Mean Absolute Percentage Error and Theil's Inequality coefficient should be as small as possible. Theil's Inequality coefficient (Theil-U) suggested by Theil (1996) is a measure of the fit of a forecast. It ranges between zero and one. When it is equal to zero it indicates that a forecast has a perfect fit. Theil-U =1 indicates that a forecast just as accurate as one of "no change'' ($\Delta y_t = 0$). Theil's Inequality coefficient can be decomposed into bias, variance, and covariance proportions each showing a different source of forecast error.

Suppose the forecast sample is $j = n + 1, n + 2, \cdots, n + h$. We will denote the actual and forecasted value in period t as y_t and $\hat{y}_t$ respectively.

3.2.4.1 Mean Absolute Error

Suppose that one analyst thinks the loss function is linear, rather than quadratic. Then the preferred estimate of accuracy is given by the mean Absolute error (abbreviated MAE) given by

$$MAE = \frac{\sum_{t=n+1}^{n+h} |y_t - \hat{y}_t|}{h} \tag{21}$$

Minimizing this measure of accuracy, rather than the MSE, will in general lead to different parameter estimates and could even lead to a different model structure.

3.2.4.2 Root Mean Squared Error

Root Mean Squared Error (RMSE) is calculated from

$$RMSE = \sqrt{\frac{\sum_{t=n+1}^{n+h} (\hat{y}_t - y_t)^2}{h}} \tag{22}$$

The RMSE is similar to MAE. The MAE and RMSE 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.

3.2.4.3 Mean Absolute Percentage Error (MAPE)

Another type of scale independent statistic involves percentage errors rather than raw errors. For example, the mean absolute percentage error is the mean 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. MAPE is given by

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

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

where PE is percentage error, x_t is actual observation for time period t, F_t is forecast for the the same period and MAPE is mean absolute percentage error. Statistics involving percentage errors only make sense for non negative variables having a meaningful zero. Fortunately, is covers most variables measured in practice.

3.2.4.4 Theil Inequality Coefficient

The U-statistic developed by Theil (1966) is an accuracy measure that emphasizes the importance of large errors (as in MSE) as well as providing a relative basis for comparison with the so-called Theil coefficients, named after Henri Theil, are also scale independent.

The measure of Theil Inequality Coefficient (Thiel-U) is calculated as follows:

$$Thiel-U=\frac{\sqrt{\frac{\sum_{t=n+1}^{n+1}(\hat{y}_t-y_t)^2}{h}}}{\sqrt{\frac{\sum_{t=n+1}^{n+h}\hat{y}_t{}^2}{h}}+\sqrt{\frac{\sum_{t=n+1}^{n+h}y_t{}^2}{h}}} \tag{24}$$

While MAPE and Theil-U are scale invariant, the Theil-U always lies between zero and one, where zero indicates a perfect fit.

3.2.4.5 Mean Squared Forecast Error

The Mean Squared Forecast Error (MSFE) can be decomposed as:

$$\frac{\sum(\hat{y}_t-y_t)^2}{h}=\left(\left(\frac{\sum\hat{y}_t}{h}\right)-\bar{y}\right)^2+(s_{\hat{y}}-s_y)^2+2(1-r)s_{\hat{y}}s_y$$

where $\frac{\sum \hat{y}_t}{h}, \bar{y}, s_{\hat{y}}, s_y$ are the means and (biased) standard deviations of $\hat{y}_t$ and y, and r is the correlation between $\hat{y}_t$ and y. The proportions are defined as:

Table 3.2. Bias, Variance, and Covariance proportions

Bias Proportion	$\frac{((\sum \frac{\hat{y}_t}{h} - \bar{y})^2}{\sum(\hat{y}_t - y_t)^2/h}$
Variance proportion	$\frac{(s_{\hat{y}} - s_y)^2}{\sum(\hat{y}_t - y_t)^2/h}$
Covariance proportion	$\frac{2(1-r)s_{\hat{y}}s_y}{\sum(\hat{y}_t - y_t)^2/h}$

- The bias proportion measures the mean of the forecast from the mean of the actual series.
- The variance proportion measures the variation of the forecast from the variation of the actual series.
- The covariance proportion measures the remaining unsystematic forecasting errors.

If the forecast is "good", the bias and variance proportions should be as small as possible so that most of the bias should be concentrated on the covariance proportion.

The bias, variance and covariance proportions add up to one. If the forecast is good, the bias and variance proportions should be small so that most of the bias should be concentrated on the covariance proportions as pointed out in Quantitative Micro Software (QMS, 2008).

3.2.5. Introduction to ARCH and GARCH models

The great workhorse of applied econometrics is the least squares method. This is a natural choice, because applied econometricians are typically called upon to determine how much one variable will change in response to a change in some other variable. Increasingly however, econometricians are being asked to forecast and analyze the size of the errors of the model. In this case, the questions are about volatility, and the standard tools have become the ARCH/ GARCH models (Engle, 2001).

Recent problems in finance have motivated the study of the volatility, or variability, of a time series. Although ARMA models assume a constant variance, models such as the autoregressive conditionally heteroscedastic or ARCH model, first introduced by Engle (1982), and were developed to model changes in volatility. These models were later extended to generalized ARCH, or GARCH models by Bollerslev (1986).

Typically, for financial time series, the return y_t, does not have a constant conditional variance, and highly volatile periods tend to be clustered together. In other words, there is a strong dependence of sudden bursts of variability in a return on the series own past.

3.2.5.1. GARCH (*m*, *r*) model

Bollerslev (1986) extended Engle's original work by developing a technique that allows the conditional variance to be an ARMA process. Let the error process be

$$\varepsilon_t = v_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^{m} \alpha_j \, \varepsilon_{t-j}^2 + \sum_{j=1}^{r} \beta_j \, \sigma_{t-j}^2 \qquad (25)$$

where $\sigma_v^2 = 1$ and $\{v_t\}$ is a white noise process that is independent of past realization of ε_{t-j}, the conditional and unconditional means of ε_t are equal to zero. That is, $E(\varepsilon_t) = E(v_t \sigma_t) = 0$.

The generalized ARCH (m, r) model-called GARCH (m, r) allows both autoregressive and moving average components in the heteroscedastic variance. If we set $m = 0$ and $r = 1$, it is clear that the first order ARCH model is simply a GARCH (0,1) mode. If all values of β_j equal to zero, the GARCH(m, r) model is equivalent to an ARCH(r) model. The benefits of the GARCH model should be clear; the GARCH model is a more parsimonious model of the conditional variance than a high-order ARCH model that is much easier to identify and estimate. This is particularly true since all coefficients in (25) must be positive. Moreover, to ensure that the conditional variance is finite, all characteristic roots of (25), in the variance equation must lie inside the unit circle. Clearly, the more parsimonious model will entail fewer coefficient restrictions.

The key feature of GARCH models is that the conditional variance of the disturbances of the $\{y_t\}$ sequence constitutes an ARMA process. Hence, it is to be expected that the residuals from a fitted ARMA model should display this characteristic pattern. To explain, suppose that the estimate of $\{y_t\}$ is an ARMA process. If our model of $\{y_t\}$ is adequate, the ACF and PACF of the residuals should be indicative of a white noise process. However, the ACF of the squared residuals can help identify the order of the GARCH process (Enders, 1995).

GARCH (1,1) Model

The GARCH(1,1) is the simplest and most robust of the family of volatility models. The GARCH model that has been described is typically called the GARCH(1,1) model. The (1,1) in parentheses is a standard notation in which the first number refers to how many autoregressive lags, or ARCH terms, appear in the equation, while the second number refers to how many moving average lags are specified, which here is often called the number of GARCH terms. Sometimes models with more than one lag are needed to find good variance forecasts.

Although this model is directly set up to forecast for just one period, it turns out that based on the one-period forecast, a two-period forecast can be made. Ultimately, by repeating this step, long-horizon forecasts can be constructed. For the GARCH(1,1), the two-step forecast is a little closer to the long-run average variance than is the one-step forecast, and, ultimately, the distant-horizon forecast is the same for all time periods as long as $\alpha + \beta < 1$. This is just the unconditional variance. Thus, the GARCH models are mean reverting and conditionally heteroskedastic, but have a constant unconditional variance (Engle, 2001).

The GARCH(1,1) model often fits the data fairly well when the number of parameters is small. A high value of β_1 means that volatility is persistent and it takes a long time to change. A high value of α_1 means volatility is spiky and quick to react to market movements. The common estimates of β, are over 0.7, but α is usually less than 0.25 (Alexander, 1998). The GARCH(1,1) with positive intercept α_0 also has the attraction that it allows us to model the volatility as mean-reverting. If the volatility is high, it will tend to fall over time. If the volatility is low, it will tend to rise over time (Dowd, 2002).

The GARCH(1,1) model specification is:

$$y_t = \mu + \varepsilon_t$$

$$\varepsilon_t = v_t \sigma_t$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{26}$$

where v_t's are iid random variables with $E[v_t] = 0$, $var[v_t] = 1$ (It is often assumed that $v_t \sim N(0,1)$), $\alpha_0 > 0$, $\alpha_1 \geq 0$ and $\beta_1 \geq 0$ and y_t represents the dependent variable over period t and μ is a constant mean. Since σ_t^2 is the one-period ahead forecast variance based on past information, it is called conditional variance. The conditional variance equation specified contains a constant term, α_0; news about volatility from the previous period, measured as the lag of the squared residual from the mean equation, ε_{t-1}^2; last period forecast variance, σ_{t-1}^2. The last condition is to ensure nonnegative conditional variances, while the condition $0 < (\alpha_1 + \beta_1) < 1$ ensures stationarity and finite variance of the unconditional returns.

This specification is often interpreted in a financial context, where an agent or trader predicts this period's variance by forming a weighted average of a long term average (α_0), the forecasted variance from last period (σ_{t-1}^2), and information about volatility observed in the previous period (ε_{t-1}^2). If the asset return was unexpectedly large in either the upward or the downward direction, then the trader will increase the estimate of the variance for the next period. This model is also consistent with the volatility clustering often seen in financial returns data, where large changes in returns are likely to be followed by further large changes.

Furthermore, the GARCH(1,1) model is a very popular specification because it fits many data series well. It shows that the volatility changes with lagged shocks, ε_{t-1}^2, but there is also momentum in the system working via σ_{t-1}^2. This model became popular because it can capture long lags in the shocks with only a few parameters. A GARCH(1,1) model with three parameters ($\alpha_0 < \alpha_1 < \beta_1$) can capture similar effects to an ARCH(q) model requiring the estimation of $(q + 1)$ parameters, where q is large ($q \geq 6$) (Hill et al., 2008).

3.2.5.2. Testing for ARCH effects

Before estimating a full ARCH model for financial time series, it is usually good practice to test for the presence of ARCH effect in the residuals. If there are no ARCH effects in the residuals, then the ARCH model is unnecessary and misspecified.

The Box-Jenkins (1976) approach is based on the assumption that the residuals are homoskedastic (remain constant over time) for ARMA or ARIMA models. Thus, the presence of ARCH effect (whether or not volatility varies over time) has to be tested in series through the squared residuals of the series (Tsay, 2010). According to Tsay there are two available methods to test for ARCH effects. The methods to test for ARCH effect and there details are discussed below.

Ljung-Box Test

It was developed by Box and Pierce (1970) and modified by Ljung and Box (1978) and tests the joint significances of serial correlation in the standardized and squared standardized residuals for the first *k* lags instead of testing individual significance. They suggested that the hypothesis test:

$H_0 = \rho_1 = \rho_2 = \cdots = \rho_k = 0$ (ACF of the first *k* lags of the residuals series is zero) against H_1: not all $\rho_j = 0$, where ρ_j is the ACF at lag $j = 1,2, \dots, k$

And the statistic:

$$Q = T(T+2)\sum_{j=1}^{k}\frac{\hat{\rho}_j^2}{T-j}$$

where $\hat{\rho}_j$ is the *j-th* autocorrelation and *T* is the number of observations.

Lagrange multiplier (LM) test

This test was suggested by Engle (1982) and used to test significance of serial correlation in the squared residuals for the first *q* lags. This particular heteroskedasticity specification was motivated by the observation that in many financial time series, the magnitude of residuals appeared to be related to the magnitude of recent residuals. ARCH in itself does not invalidate standard LS inference. However, ignoring ARCH effects may result in loss of efficiency.

The null hypothesis of the test is that there is no serial correlation in the residuals up to the specified order. The test statistic is $LM = n * R^2$, where R^2 is R-square from the regression of $\hat{\varepsilon}_t^2$ on $\hat{\varepsilon}_{t-i}^2$, $i = 1, \dots, q$ and n is number of observations. This test statistic under the null hypothesis follows the chi-square distribution with q degree of freedom.

3.2.5.3. Parameter estimation

Since the model is no longer of the usual linear form, OLS cannot be used for GARCH model estimation. There are a variety of reasons for this, but the simplest and most fundamental is that OLS minimizes the residual sum of squares. The RSS depends only on the parameters in the conditional mean equation, and not longer an appropriate objective.

In order to estimate models from the GARCH family, another technique known as maximum likelihood is employed. Essentially the method works by finding the most likely values of the parameters given the actual data. More specifically, a log-likelihood function is formed and the values of the parameters that maximize it are sought.

Maximum likelihood estimation can be employed to find parameter values for both linear and non-linear models (Brooks, 2008).

The steps involved in actually estimating an ARCH or GARCH models are:

1. Specify the appropriate equation for the mean and the variance.
2. Specify the log-likelihood function (LLF) to maximize under a normality assumption for the disturbances.

Maximum likelihood methods are required to understand and write a program ARCH-type models. Suppose that values of $\{y_t\}$ drawn from a normal distribution having a mean of zero and conditional variance σ_t^2 . From a standard distribution theory, the log likelihood function using T independent observations is

$$\log L = -\frac{\mathrm{T}}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\frac{\varepsilon_t^2}{\sigma_t^2}$$

where logL is log of the likelihood of ε_t . The likelihood of ε_t is given by:

$$l_t = \left(\frac{1}{\sqrt{2\pi\sigma_t^2}}\right)e^{-\frac{\varepsilon_t^2}{2\sigma_t^2}}$$

Since the realization of ε_t are independent, the likelihood of the joint realizations of $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ is the product in the individual likelihoods. Hence, if all have the same variance, the likelihood of the joint realization is

$$L = \prod_{t=1}^{n}\left(\frac{1}{\sqrt{2\pi\sigma_t^2}}\right)e^{-\frac{\varepsilon_t^2}{2\sigma_t^2}}$$

So that the likelihood function is

$$\ln L = -\frac{\mathrm{n}}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\ln(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{n}\frac{\varepsilon_t^2}{\sigma_t^2}$$

Now, suppose that $\varepsilon_t = y_t - \mu$ and the conditional variance is $\sigma_t^2 = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2$. Substituting for σ_t^2 and y_t yields

$$\ln L = \frac{\mathrm{n}-1}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=2}^{n}\ln\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2 - \frac{1}{2}\sum_{t=2}^{n}\frac{(y_t - \mu)^2}{\alpha_0 + \alpha_1\varepsilon_{t-1}^2 + \beta_1\sigma_{t-1}^2}$$

Note that the initial observation is lost since ε_0 is outside the sample. Once it substitutes $(y_{t-1} - \mu)^2$ for ε_{t-1}^2, it is possible to maximize $\ln L$ with respect to α_0, α_1 *and* β_1. There are no simple solutions to the first order conditions for a maximum.

3.2.5.4 Diagnostic Checking

For a properly specified ARCH model, the standardized residuals $\hat{\mathrm{s}}_\mathrm{t} = \hat{\varepsilon}_\mathrm{t} / \hat{\sigma}_\mathrm{t}$ form a sequence of iid random variables. Therefore, one can check the adequacy of a fitted ARCH model by examining the series $\{y_t\}$. In particular, the Ljung–Box statistic of y_t can be used to check the adequacy of the mean equation and that y_t^2 can be used to test the validity of the volatility equation. The skewness, kurtosis, and quantile-to-quantile plot (i.e., QQ-plot) of $\{y_t\}$ can be used to check the validity of the distribution assumption (Tsay, 2010).

When a model has been fitted to a time series, it is advisable to check that the model really does provide an adequate description of the data. As with most statistical models, the goodness of fit of the ARCH-GARCH model is based on residuals and more specifically on the standardized residuals.

The following are the model adequacy checking methods used for this study

I. The ACF and PACF of the standardized residuals: If the model is adequate the ACF and PACF of squared standardized residuals should be indicative of a white noise process.

II. The standardized residuals should be identically and independently distributed standard normal even if student-t and generalized exponential distribution (GED) are assumed (Tsay, 2010). This can be checked through the Jarque-Bera test.

III. The Ljung-Box test is one of the most widely used lack-of-fit tests, that is, a test for the appropriateness of the fitted model. It was developed by Box and Pierce (1970) and modified by McLeod and Li (1983). The Ljung-Box test statistic uses the Q-statistic to test whether there is a group of significant k autocorrelations, to test whether the mean model is appropriately specified and to test for the remaining GARCH effects under the null hypothesis that there is no autocorrelation among k lags of standardized residuals and squared standardized residuals for mean and GARCH specification, respectively. Thus, if the statistic Q at all lags is non-significant, it indicates the absence of autocorrelation in the residuals, and this is evidence that the model selected fits the data well.

3.2.5.5. Forecasting

To forecast conditional variance one can use estimated GARCH model. There are two forecasting options in EViews: the dynamic and the static forecasting. Dynamic forecasting used to calculate multi-step forecasts starting from the first period in the forecast sample. Static forecasting calculates a sequence of one-step a head forecasts using the actual data.

4. Results and Analysis

4.1. Petroleum Oil Import Price Time Series

First, as with any data analysis, we should construct a time plot of the petroleum oil import prices. The in-sample period of petroleum oil prices data from October 1999 to June 2010 were plotted with the aid of statistical software in order to determine the type of the trend of the series: constant, linear and etc (Figure A1in the Appendix). The petroleum oil import prices series has a trend (exponentially trending).

4.2. Unit root test for stationarity

A stationary series must be obtained before the series can be used to estimate and develop a model. The unit roots test will help us to determine whether the series is stationary or not. The unit root tests first impose the null hypothesis that the time series has a unit root problem, versus the alternative hypothesis that the series is stationary. Augmented Dickey-Fuller test (ADF) and Phillip Perron (PP) tests are used to check the stationarity of the monthly petroleum oil import prices series. Initially, we check the stationarity of the original oil prices series (at level form). The result is displayed in Table 4.1.

Table 4.1. ADF test for petroleum oil import prices series (at level form)

Series	ADF t-statistic	1% crit. Value	5% crit. Value	10% crit. Value	p-value
POIL	-2.744851	-3.483312	-2.884665	-2.579180	0.0694

According to the ADF test for monthly petroleum oil import prices time series in Table 4.1, the computed ADF test statistic is greater than critical values "tau statistic" at 1% and 5% levels of significance, respectively. We don't reject the null hypothesis at 5% level of significance. That is, there is a unit root problem. Thus, the series is not stationary at level form (Table A1 in Appendix).

In the Phillip-Perron (PP) unit root test the null hypothesis of the test again states that there is a unit root problem in the series. Table 4.2 below summarizes the result of Philips-Perron unit root test for petroleum oil import prices.

Table 4.2: Phillips-Perron (PP) test for the petroleum oil import prices (at level form)

Series	PP t-statistic	1% crit. Value	5% crit. Value	10% crit. Value	p-value
POIL	-9.901176	-3.482035	-2.884109	-2.578884	0.0000

Table 4.2 shows that the Phillips-Perron t-statistic for the petroleum oil import price series is smaller than the critical values - "tau-statistic" at 1%, 5% and 10% significance

indicating the PP t-statistic is significant (Table A2 in the Appendix). But ADF test indicates the series is not stationary or a unit-root problem.

The petroleum oil import prices series show some nonstationarity that can be improved by transforming to logarithms and some additional nonstationarity that can be corrected by differencing the logarithms. We also checked the stationarity or unit-root problem for the first order difference of transformed petroleum oil import prices series. Similarly, we used ADF and PP unit-root tests to determine the series' stationarity. Table 4.3 below shows the result of ADF unit-root test for the first order differenced log petroleum oil import prices series.

Table 4.3. ADF unit-root test after first difference of log series

Series	ADF t-statistic	1% crit. Value	5% crit. Value	10% crit. Value	p-value
$\Delta \log \text{POIL}$	-9.9767	-3.483751	-2.884856	-2.579282	0.0000

The computed ADF t-statistic for the first order logarithm difference from the original petroleum oil import prices denoted as $\Delta \log \text{POIL}$ sometimes named as growth rate (or return) of oil prices series is smaller than the critical values at 1%, 5% and 10% significance level respectively, which leads to the rejection of the null hypothesis of a unit root problem. That is, the first difference of the transformed petroleum oil import prices series does not have a unit-root problem and stationary (Table A3 in the Appendix).

Table 4.4. PP unit-root test after first difference of log series

Series	PP t-statistic	1% crit. Value	5% crit. Value	10% crit. Value	p-value
$\Delta \log POIL$	-70.80035	-3.482453	-2.884291	-2.578981	0.0001

The PP test also gives a similar inference to ADF test. The computed PP t-statistic for $\Delta \log POIL$ series is very small compared to test critical values at 1%, 5% and 10% significance levels. We reject the null hypothesis that there is a unit-root problem in series. Thus, we make the same inference as in the ADF test where the series has not unit root problem in first differenced series. That is, the monthly petroleum oil import prices series is stationary at first difference of the transformed series (Table A4 in the Appendix 4).

Therefore, based on the ADF and PP test we can conclude that the first order difference of the log transformed petroleum oil import prices series ($\Delta \log POIL$) is stationary.

4.3. Normality test

The histogram and normality test statistic including mean, median, maximum and minimum values, standard deviation, skewness, kurtosis, and Jarque-Bera test of first order difference of log petroleum oil import prices series ($\Delta \log POIL$) (Figure A2 in the Appendix). The histogram is centered and peaked at zero. The mean and standard deviation values approximated a standard normal distribution of N(0,1). In order to formally examine the normality assumption, we run some classes of tests for monthly petroleum oil import prices series. The first test used is the Jarque-Bera test. This test uses information on the third and fourth moments of a distribution and is based upon the assumption of a Gaussian distribution. The Jarque-Bera test indicated that the first order difference log series is not normally distributed. That is, we reject the hypothesis of normal distribution at 5% level of significance. But we can say $\Delta \log \mathrm{POIL}$ series is fairly normal from the graph of histogram.

According to the information obtained above, we have the stationary series after differencing the logarithm series from the original petroleum oil prices series. In the next step, we use the first order logarithm difference for petroleum oil prices series ($\Delta \log \mathrm{POIL}$) to find our models using Box-Jenkins and GARCH approaches.

4.4. ARIMA Model

Box-Jenkins modeling of a stationary time series involves the following four steps: model identification, model estimation, diagnostic checking and forecasting. We have the stationary time series after first order differencing. Now, the model that we are looking at is ARIMA$(p, 1, q)$. We have to identify the model, estimate suitable parameters, diagnostic checking for residuals and finally achieve our objective of forecasting the future petroleum oil prices.

4.4.1. Model Identification

After suitably transforming of the data, the next step is to identify preliminary values of the autoregressive order, p, the order of differencing, d, and the moving average order, q. But we have already settled the problem of selecting d. When preliminary values of d have been settled, the next step is to look at the sample ACF and PACF of the transformed data, denoted as ∇logpoil.

Firstly, we compute sample ACF and PACF the stationary series which consists of the sample ACF and PACF values (Figure A3 in the Appendix). We also calculate the Ljung-Box Q-statistic. We observe the patterns of the ACF and PACF, and then determine the parameter values p and q for ARIMA model. Using Table 3.1 as a guide, preliminary values of p and q are chosen. Inspecting the sample ACF and PACF, we might feel that the ACF is cutting off at lag 1 and the PACF is tailing off. This would suggest the petroleum oil growth rate ($\Delta \log \mathrm{POIL}$) follows an MA(1) process, or log poil follows an ARIMA(0,1,1) model (or ∇logpoil follows MA(1)). Rather than focus on one model, we will also suggest that it appears that the ACF is tailing off after lag 1 and PACF is tailing off after lag 2. This suggests an ARMA(2,1) model for the growth rate, or ARIMA(2,1,1) for log poil. The others ACF tailing off after lag 1 and PACF tailing off slowly after lag 1. Thus, the p and q values for the ARIMA(p,1,q) model are set at 1 respectively. So, we temporarily set our ARIMA model to be ARIMA(1,1,1). As preliminary analysis, we will fit ARIMA(0,1,1), ARIMA(1,1,1), ARIMA(2,1,1) and ARIMA(3,1,1) models.

4.4.2. Model Estimation

After identifying the competing ARIMA models, the next step is estimating the parameter coefficients. We assume we have n observations, $x_1, x_2, \cdots, x_n$, from a causal and invertible Gaussian ARMA(p,q) process in which, initially, the order parameters, p and q, are known. The goal is to estimate the parameters, $\emptyset_1, \emptyset_2, \cdots, \emptyset_p, \theta_1, \theta_2, \cdots, \theta_q$, and σ_w^2.

Table 4.5. Parameter estimates summary statistics and selection criteria for the four candidate models

Model	Parameter	Coefficients	Standard error	t-statistic	Prob.	Information criteria
ARIMA(0,1,1)	μ	0.0103	0.0074	1.3965	0.165	AIC=0.0024
	θ_1	-0.87	0.044	-19.732	0.000	AICc=0.0195
						BIC=-0.9530
						$\hat{\sigma}_W^2 = 0.3574$
ARIMA(1,1,1)	μ	0.0104	0.0069	1.5578	0.1218	AIC=0.0083
	$\emptyset_1$	0.1132	0.1035	1.106	0.2709	AICc=0.0265
	θ_1	-0.897	0.0519	-19.718	0.000	BIC=-0.9248
						$\hat{\sigma}_W^2 = 0.354$
ARIMA(2,1,1)	μ	0.009	0.0079	1.2415	0.2168	AIC=0.0146

	$\emptyset_1$	0.0856	0.1085	0.819	0.4141	AICc=0.0341
	$\emptyset_2$	-0.1186	0.1020	-1.1828	0.2392	BIC=-0.8962
	θ_1	-0.8603	0.0720	-14.489	0.000	$\hat{\sigma}_W^2 = 0.3507$
ARIMA(3,1,1)	μ	0.0089	0.0085	1.033	0.3036	AIC=0.0289
	$\emptyset_1$	0.0407	0.1336	0.343	0.7321	AICc=0.0499
	$\emptyset_2$	-0.1432	0.1103	-1.382	0.1695	BIC=-0.8597
	$\emptyset_3$	-0.060	0.1141	-0.564	0.5737	$\hat{\sigma}_W^2 = 0.3502$
	θ_1	0.8225	0.1052	-10.265	0.000	

To choose the final model, we compare the AIC, the AICc and the BIC for the candidate models. Out of a class of candidate models, the best-fitted model is the one with the smallest AIC, AICc and BIC statistic as given by equation (16,17 and 18), respectively. According to the model selection criterion, ARIMA (0,1,1) model for the logarithms of the monthly petroleum import prices series had been chosen to be the most appropriate. From the t-statistic for the coefficient variable ARIMA (0,1,1), the null hypothesis that the coefficients are equal to zero is rejected. But the constant is insignificant, so it can be omitted from the equation of the model. The estimated parameter coefficients of ARIMA (0,1,1) model for the logarithms of petroleum oil import prices are $\mu = 0.0103$ and $\theta = -0.87$ (Table A5 in the Appendix).

So the equation of the fitted ARIMA model is given by

$$x_t = \varepsilon_t - 0.87\varepsilon_{t-1}$$

4.4.3. Diagnostic checking

The next step in model fitting is diagnostics. This investigation includes the analysis of the residuals as well as model comparisons. Again, the first step involves a time plot of the innovations (or residuals), $x_t - \hat{x}_t^{t-1}$, or of standardized innovations.

$$e_t = (x_t - \hat{x}_t^{t-1}) \Big/ \sqrt{\hat{p}_t^{t-1}}$$

where $\hat{x}_t^{t-1}$ is the one-step-ahead prediction of x_t based on the fitted model and $\hat{\mathrm{p}}_{\mathrm{t}}^{\mathrm{t}-1}$ is the estimated one-step-ahead error variance. If the model fits well, the standardized residuals should behave as an iid sequence with mean zero and variance one. The time plot should be inspected for any obvious departures from this assumption (Figure A4 in the Appendix). We see some spiky residuals that indicate high volatile period. Since the residuals are changing with time, a volatile series is obtained.

We could also inspect the sample autocorrelations of the residuals for the ARIMA(0,1,1) model (Figure A5 in the Appendix), say, $\hat{\rho}_e(h)$, for any patterns or large values. For a white noise process sequence, the sample autocorrelations are approximately independently and normally distributed with zero means and variances 1/n. hence, a good check on the correlation structure of the residuals is to plot $\hat{\rho}_e(h)$ versus h along with the error bounds of $\pm 2/\sqrt{n}$. As we see from sample ACF and PACF of residuals of

ARIMA (0,1,1) model the ACF and PACF are both relatively small or approximately zero.

In addition to plotting a correlogram of the residuals ($\hat{\rho}_e(h)$), we can perform a general test that takes into consideration the magnitude of $\hat{\rho}_e(h)$ as a group is called Q-statistic. The Q-statistic in Figure A5 in the Appendix shows that the model is adequate. That means the ACF patterns are white noise.

4.4.3.1. Test for serial correlation

An alternative test to the Q-statistic for testing serial correlation is Breusch-Godfrey LM test. Unlike the Durbin-Watson statistic for AR (1) errors, the Breusch-Godfrey LM test used to test for higher order ARMA errors. The null hypothesis of the LM test claims that there is no serial correlation up to lag order p (Table A6 in the Appendix). The F-statistic and Breusch-Godfrey LM test statistic are 1.214684 and 1.223918, respectively indicating that the tests do not reject the hypothesis of no serial correlation up to order 1. The p-value of 0.2686 indicates that we have no evidence to reject the null hypothesis of no serial correlation up to lag 1. The Q-statistic and the LM test both indicate that the residuals are not serially correlated and the equation is appropriate for forecasting. Once again, we justify the model is adequate.

4.4.3.2. Test of normality

Investigation of marginal normality can be accomplished visually by looking at a histogram of the residuals. The result shows that the mean 0.004728 and the standard deviation is 0.599. The skewness and kurtosis are 0.462 and 5.57 respectively. These indicate that the residuals have a little leptokurtic (i.e., one which has fatter tails and is more peaked at the mean) and slightly skewed. The Jarque-Bera test indicated that the residuals are not normally distributed at 5% level of significance (Figure A6 in the Appendix). This is one of the shortcomings of using an ARIMA model to forecast a volatile monthly petroleum oil import prices series. The GARCH model provides a framework for modeling and analyzing time series that display some of these characteristics.

4.5. GARCH Models

4.5.1. Testing for ARCH effect

Before estimating a GARCH-type model, we must examine the existence of heteroscedasticity in monthly petroleum oil import prices series for ARCH effects to make sure that this class of models is appropriate for the data. The monthly petroleum oil import prices data that are used in this study show volatility periods. Thus, it is suitable to assume heteroscedasticity where conditional variance is not constant throughout the time trend.

4.5.1.1. LM ARCH Test

There is a test for heteroscedasticity developed by Engle (1982) called Lagrange Multiplier (LM) ARCH test. This test is used to determine the occurrence of ARCH effect in the residuals. The results of LM ARCH test for ARIMA(0,1,1) model is displayed in Table A7 in the Appendix. The top part of the output displays F-statistic and ARCH-LM test statistic. The F-statistic value of 7.2 is taken from the test equation for residuals squared. The p-value indicates the F-statistic is significant. This suggests the presence of ARCH effect in the model. The ARCH-LM test statistic of 6.9 also gives the same result for F-statistic as the one under chi-square with one degree of freedom.

4.5.1.2.Diagnostic checking for residuals squared

Another important criterion to determine whether a series contains heteroscedastic (ARCH effect) is by checking the sample ACF and PACF of the residuals squared. At this point, we also need to observe the patterns in the ACF and PACF of residuals squared for ARIMA(0,1,1) model. It shows that there are spikes at the second lag for both ACF and PACF of residuals. This indicates that the ARCH effect does occur in the residuals for the ARIMA(0,1,1) model (Figure A7 in the Appendix).

4.5.2. Parameter estimation

The method to estimate the parameters is done by using EViews software. The maximum likelihood estimator will find the parameter coefficients for conditional mean and conditional variance equations. Using EViews, the parameter coefficients of the first order difference for the logarithm monthly petroleum oil import prices are obtained (Table A8 in the Appendix). From the conditional mean equation, the parameter found is $\mu=0.039$. The standard normal distribution z-test has rejected the parameter coefficients equal to zero, while the conditional variance equation gives α_0=0.145, α_1=0.385 and β_1=0.393.

The GARCH(1,1) model can be written into conditional mean and conditional variance Equations as:

$$y_t \ = 0.039 + \varepsilon_t$$

$$\sigma_t^2 = 0.145 + 0.385\varepsilon_{t-1}^2 + 0.393\sigma_{t-1}^2$$

A graphical plot for conditional standard deviation and conditional variance is presented in Figure A8 and Figure A9 in the Appendix. The values for conditional standard deviation are obtained by taking the square root from the conditional variances. From the conditional standard deviation plot in Figure A8 the long spikes indicates high volatile periods of the series. With GARCH(1,1) model, the volatility clustering will be detected. In case of the conditional variance there are lesser spikes compared to conditional standard deviation graph. But it does point out some of the high volatile clusters in the series. However, the extraordinary long spikes are the high volatile periods in the data series.

4.5.3. Diagnostic Checking GARCH(1,1) Model

After we have estimated the parameters, the next step will be diagnostic checking on the adequacy for GARCH(1,1) model. It can be done by inspecting the correlogram of standardized residuals squared which consists of autocorrelation and partial autocorrelation (Figure A10 in the Appendix). The sample ACF and PACF of standardized residuals squared are approximately zero. The insignificance Ljung-Box Q-statistics also provides the same evidence with p-value that GARCH(1,1) model is adequate. We can conclude that the model is adequate.

In diagnostic checking stage, a test for assessing of conditional hetroscedasticity in the data conducted with ARCH-LM test on the residuals (Table A9 in the Appendix).

The ARCH-LM for one lag difference of residuals squared is 0.0127 under $\chi^2(1)$. But, the null hypothesis that stated that there is no autoregressive conditional hetroscedasticity up to order q is not rejected since the p-value 0.91 is greater than significance level of 5%. On the other hand, the computed F-statistic is 0.0125 which does not lead to the rejection of the null hypothesis at 5% level of significance. The ARCH-LM test on the residuals of this model indicates that the conditional hetroscedasticity is no longer present in the data. This means that we do not need higher order GARCH models and that the GARCH (1,1) model is able to capture appropriately the GARCH effects.

The distribution of the standardized residuals will be summarized in the histogram chart and normality test (Figure A11 in the Appendix). The histogram chart shows that the standardized residuals are evenly distributed. The mean and standard deviation are -0.0069 and 1.0026, respectively implying that the standardized residuals are normally distributed $N(0,1)$. The skewness and kurtosis are 0.387 and 5.458, respectively. The distribution is a bit positively skewed and peaked (leptokurtic). The Jarque-Bera test indicated that the standardized residuals are normally distributed implying that the conditional distribution of $\Delta \log \mathrm{POIL}$ is fairly normal. We can therefore conclude that the GARCH (1,1) model is a parsimonious model and there are no remaining ARCH effects that need to be modeled by higher-order GARCH models.

4.5.4. Forecasting

Having estimated the models, our next step is forecasting. Apart from forecasting the conditional variance, we also forecast the conditional mean at the same time. Here, our monthly forecast petroleum oil prices are the conditional mean from the original series. To forecast from GARCH(1,1) model, we use a sequence of one-step ahead static forecast (Figure A12 in the Appendix). The solid line represents the forecasted import prices whereas the dotted lines are forecast import prices with ± 2 standard errors. The forecast petroleum oil import prices fluctuate between 10287 birr and 17837 birr in one-year out-of-sample period (Figure A13 in the Appendix). The forecast of conditional variance are not constant. Since conditional heteroscedasticity means finding the non-constant variance that may exist in time series data its trend is non-linear. The actual and forecast monthly petroleum oil import prices by GARCH(1,1) model are plotted in Figure A14 in the Appendix. It can be concluded that the trend of forecast petroleum oil import prices follows the actual petroleum oil import prices for one year (12 months) out-of-sample period.

4.5.4.1 Forecasting Evaluation of GARCH(1,1) Model

In the forecasting stage, we calculate RMSE, MAE, MAPE, Theil-U and MSFE values from GARCH(1,1) model. These are presented in Table 4.6. If the actual values and forecast values are closer to each other, a small forecast error will be obtained (Table A10 in the Appendix).

Table 4.6. Forecasting evaluation of GARCH(1,1) model

Forecast performance	**GARCH(1,1)**
RMSE	1722.848
MAE	1144.322
MAPE	8.277502
Theil-U	0.066243
Bias proportion	0.000901
Variance proportion	0.002590
Covariance proportion	0.996509

Table A11 in the Appendix shows the out-sample forecasts of GARCH model compared with the actual series over the period July 2010 to June 2011.

In this study petroleum oil import price forecasts have been made using GARCH model. The GARCH(1,1) model gives a better estimate when there are volatility clustering in the data series. This is due to the GARCH model's ability to capture the volatility by the conditional variance of being non-constant throughout the time.

Based on the fitted GARCH(1,1) model monthly petroleum oil import prices were forecasted for the period from July 2011to June 2013 (Table A12 in the Appendix). The forecasted results obtained from GARCH (1,1) model show that the import petroleum oil prices are expected to increase from July 2011 to June 2013 with some fluctuations (or volatility) due to its ability to capture volatility.

5. Conclusions and recommendations

5.1. Conclusions

The objective of this study was to develop appropriate time series models that could be useful for forecasting petroleum oil import prices. ARIMA is a popular forecasting method. It is a general class of Box-Jenkins model for stationary time series. The Box-Jenkins approach involves four stages. These stages are model identification, model estimation, diagnostic checking and forecasting. In this study, the model that has been selected for forecasting petroleum oil import prices is ARIMA(0,1,1). But this model was not adequate.

Despite the fact that this approach has been used extensively in various fields such as economics, agriculture and business, it does not perform very well when there exists volatility in the data series. To handle volatility, this study used the GARCH model. Most of the time, GARCH models can accommodate volatility clustering and leptokurtosis very easily. The GARCH approach involves model identification, model estimation, diagnostic checking and forecasting. In this study, the model that has been selected for forecasting petroleum oil import prices was GARCH(1,1). This model performed well because of its ability to capture the volatility by the conditional variance of being non-constant throughout the time. This study concluded that the GARCH(1,1) model is suitable to forecast petroleum oil import prices because the values for RMSE, MAE, MAPE, Theil-U and MSFE calculated using this model were small.

Based on the fitted GARCH(1,1) model monthly petroleum oil import prices were forecasted from July 2011 to June 2013. The forecasted results obtained from GARCH (1,1) model showed that import petroleum oil prices are expected to increase from July 2011 to June 2013 with some fluctuations as a result of volatility.

5.2. Recommendations

The focus of this study was estimating and forecasting petroleum oil import prices by using GARCH model. Future studies in this area can also use a hybrid method, specifically ARMA/GARCH model. The hybrid model combines the Box-Jenkins with GARCH. The hybrid model is an alternative to forecast petroleum oil import prices because it contains both qualities of the Box-Jenkins and GARCH methods. In addition to these other GARCH-type models that could be investigated to forecast petroleum oil import prices data include Integrated GARCH (IGARCH) and Exponential GARCH (EGARCH).

Another suggestion for future work is to compare the model with other models that are able to capture volatility in the petroleum oil import prices such as implied volatility and stochastic volatility and the like.

References

- Agnolucci, P. (2009). Volatility in Crude Oil Futures: A Comparison of the Predictive Ability of GARCH and Implied Volatility Models. *Energy Economics,* **31**, 316-321.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principal. In *2nd Int. Symp. Inform. Theory*, 267-281. B.N. Petrov and F. Csake, eds. Budapest: Akademia Kiado.
- Amare, T. and Fentaw, A. (2010). Application of GARCH models in forecasting volatility of export prices. *Journal of the Ethiopian Statistical Association*, **19**, 77-91
- Alexander, C. (1998). Volatility and Correlation: Measurement, Models and Applications. *Risk Measurement and Analysis.* ***Vol. 1****: Measuring and Modelling Financial Risk*, 125-171.
- Alizadeh, A. H., Nomikos, N. K. and Pouliasis, P. K. (2008). A Markov Regime Switching Approach for Hedging Energy Commodities. *Journal of banking & Finance* **32**, 1970-1983.
- Baillie, R.T., Bollerslev, T. and Mikkelsen, H.O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* **74**, 3–30.
- Bera, A. K. and Jarque, C. M. (1981). An Efficient Large-Sample Test for Normality of Observations and Regression Residuals, *Economics letters*, **7**(4): 313-318.
- Bollerslev, T. (1986). Generalised Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* **31**, 307-27.
- Box, G.E.P. and D.A. Pierce (1970). Distributions of residual autocorrelations in autoregressive integrated moving average models. *J. Am. Stat. Assoc.*, **72**, 397-402.

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

- Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control. San Francisco: Holden Day.
- Box, G.E.P., G.M. Jenkins and G.C. Reinsel (1994). *Time Series Analysis, Forecasting, and Control, 3rd ed.* Englewood Cliffs, NJ: Prentice Hall.
- Brooks, C. (2008). *Introductory Econometrics for Finance*. (2nd ed.) Cambridge: Cambridge University Press.
- Cabedo, J. D. and Moya, I. (2003). Estimating Oil Price "Value at Risk" Using the Historical Simulation Approach. *Energy Economics,* **25**, 239-253.
- Challen, J. and Hagger, S. (1983). Economics: Mathematical models. St. Martin's Press, New York.
- Chatfield, C. (2000). Time-Series Forecasting. London: Chapman & Hall/CRC
- Cheng, W. H. (2008). Overestimation in the Traditional GARCH Model during Jump Periods. *Economics Bulletin*, **3**(68), 1-20.
- Cheong, C. W. (2009). Modeling and Forecasting Crude Oil Markets Using ARCH-type Models. *Energy Policy,* **37**, 2346-2355.
- Cotter, J. & Stevenson, S. (2006). Multivariate Modeling of Daily REIT Volatility, *Journal of Real Estate Finance and Economics*, forthcoming.
- Dowd, K. (2002). *An Introduction to Market Risk Measurement*. Chichester: John Wiley & Sons.
- Enders, W. (1995). *Applied Econometric Time Series*. Giles and Mizra, New York.
- Engle, R. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica,* **50**, 987–1007.
- Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Econometric perspectives,* **15**, 157–168.
- Engle, R. and Lee, G. (1999). A permanent and Transitory Component Model of Stock Return Volatility. In: Engle, R. F. and White, H. (Eds.), Cointegration,

Casuality and Forecasting: A Festschrift in Honour of Clive W. J. Granger. Oxford University Press, Oxford, 475–497.

- Fabian, T. B. (2010). Models for Oil price prediction and forecasting. *Unpublished thesis*. San Diego State University.
- Fong, W. M. and See, K. H. (2002). A Markov Switching Model of the Conditional Volatility of Crude Oil Price. *Energy Economics,* **24**, 71-95.
- Hill, R. C., Griffiths, W.E. and Lim, G. C. (2008). Principles of Econometrics. (3rd ed.). Danvers, John Wiley & Sons, Inc.
- Hurvich, C.M and C.-L. Tsai (1989). Regression and time series model selection in small samples. *Biometrika*, **76**, 297-307.
- Kang, S. H., Kang, S. M. and Yoon, S. M. (2009). Forecasting Volatility of Crude Oil Markets. *Energy Economics,* **31**, 119-125.
- Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models, *Biometrika,* **65**: 297-303
- Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin Heidelberg, Springer-Verlag.
- Marimoutou, V., Raggad, B. and Trabelsi, A. (2009). Extreme Value Theory and Value at Risk: Application to Oil Market. *Energy Economics,* **31**, 519-530
- McLeod, A. I. and Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis,* **4**, 269–273.
- McQuarrie, A.D.R. and C-L. Tsai (1998). Regression and Time Series Model Selection, Singapore: World Scientific.
- Morana, C. (2001). A Semi parametric Approach to Short-term Oil Price Forecasting. *Energy Economics,* **23**, 325-338.
- Moshiri, S. and Foroutan, F. (2006). Forecasting Nonlinear Crude Oil Futures Prices. *The Energy Journal,* **27**(4), 81-95.

- Narayan, P. K. and Narayan, S. (2007). Modelling Oil Price Volatility. *Energy Policy* **35**, 6549-6553
- National bank of Ethiopia (2009/10). Annual Report. Documentation and Publication Service, Addis Ababa.
- Nelson, C.R. and Plosser, C.I. (1982). Trends and Random walks in Microeconomic Time Series. *Journal of Monetary Economics*, **10**, 139-162
- OPEC Press Release (2007, September 11). 145th Meeting of the OPEC Conference. *OPEC News*. Retrieved No 10/2007, from http:// www.opec.org/
- Palma, W. (2007). *Long-Memory Time Series: Theory and Methods*. New Jersey, John Wiley & Sons, Inc.
- Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika*, **75**, 335-346.
- QMS (2008). *EViews 6: User's Guide*. Irvine CA: Quantitative Micro Software.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, **14**, 465-471.
- Sadeghi, M. and Shavvalpour, S. (2006). Energy Risk Management and Value at Risk Modeling. *Energy Policy*, **34**, 3367-3373.
- Sadorsky, P. (1999). Oil Price Shocks and Stock Market Activity. *Energy Economics,* **21**(5), 449–469
- Sadorsky, P. (2006). Modeling and Forecasting Petroleum Futures Volatility. *Energy Economics,* **28**(4), 467–488.
- Said, S.E. and D. Dickey (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, **71**, 599-607
- Schwarz, F. (1978). Estimating the dimension of a model. *Ann. Stat.*, **6**, 461-464
- Shumway, R.H. and D.S. Stoffer (2010). Time Series Analysis and Its Applications with R Examples, Second Edition. New York, Springer

- Siti, R. Y., Maizah, H. A., Lee, C. N. and Noryanti, M., 2011. A Comparative Study on Box-Jenkins and Garch Models in Forecasting Crude Oil Prices. *Journal of Applied Sciences,* **11***:* 1129-1135.
- Sugiura, N. (1978). Further analysis of the data by Akaike's information criterion and the finite corrections, *Commun. Statist, A, Theory Methods,* **7**, 13-26.
- Theil, H. (1966). Applied economic Forecasting (North-Holland, Amsterdam).
- Theil, H. (1996). Economic Forecasts and Policy. University of Florida, Florida.
- Tsay, R.S. (2010). Analysis of Financial Time Series, Third Edition. New Jersey, John Wiley & Sons, Inc.
- Wang, S., Yu, L. and Lai, K. K. (2005). Crude Oil Price Forecasting With TEI@I Methodology. *Journal of Systems and Complexity*, **18**(2), 145-166.
- Zivot, E. and Wang, J. (2006). Modeling Financial Time Series with S-PLUS, Second Edition. Springer, New York.

Appendix

Figure A1: The time plot of monthly petroleum oil import prices data.

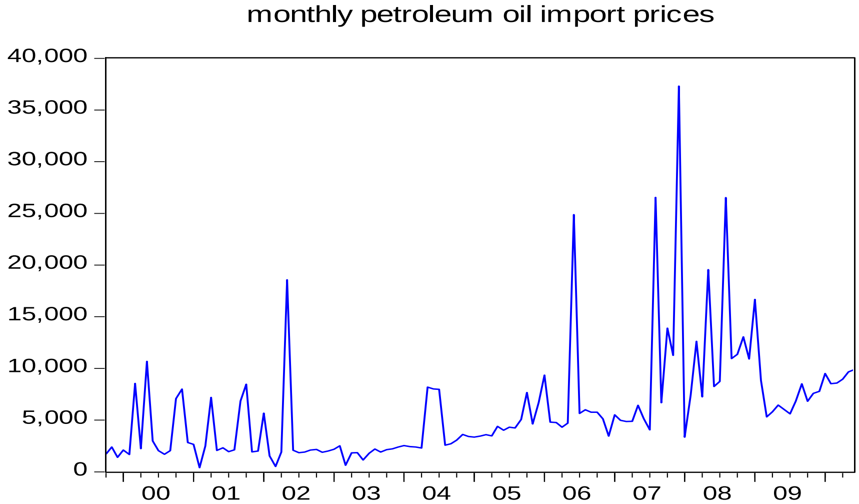


Figure A2: Histogram and normality test for first order difference of log series (Δ **log POIL**)

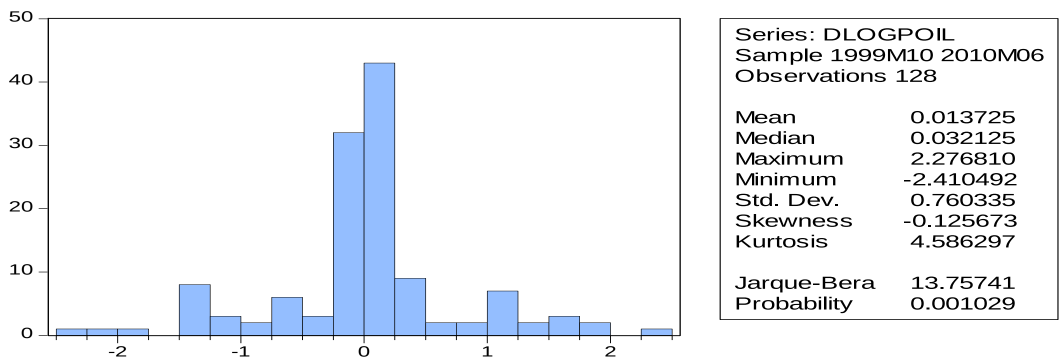


Figure A3: Sample ACF and PACF of the differenced log monthly petroleum oil import prices series (**∇logpetoil**)

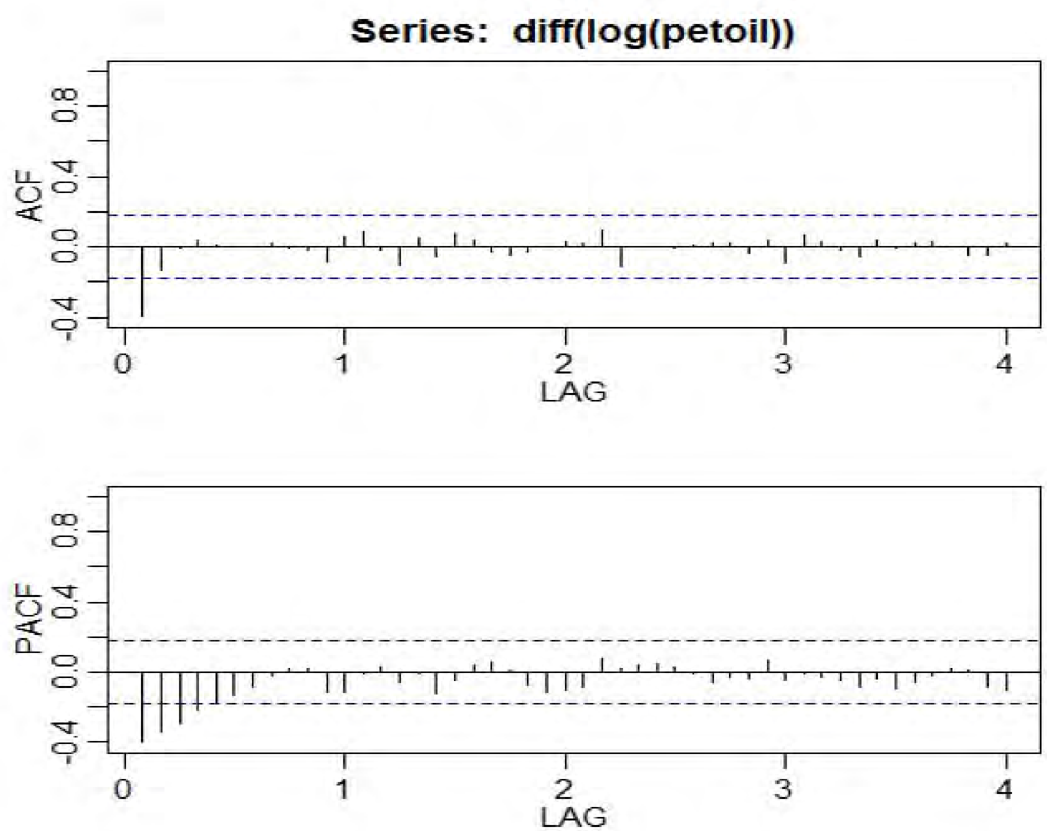


Figure A4: Residual plot of **∇logpoil**

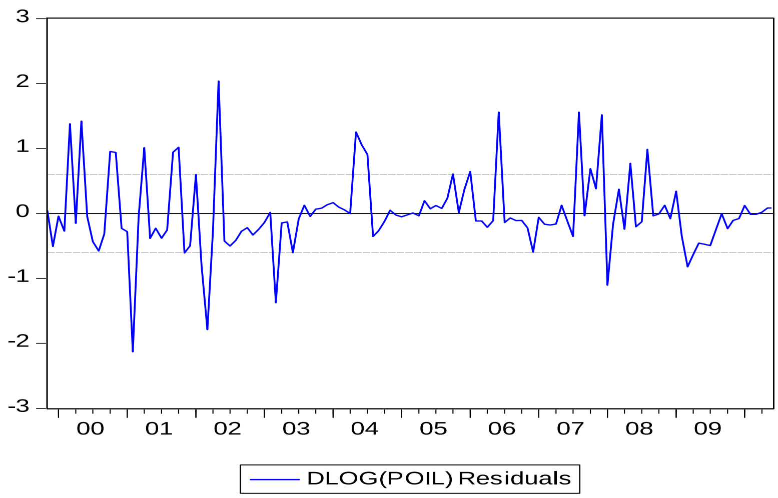


Figure A5: Sample ACF and PACF of residuals for ARIMA(0,1,1)

Date: 10/13/12 Time: 19:10
Sample: 1999M11 2010M06
Included observations: 128
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.085	0.085	0.9514	
		2	-0.104	-0.112	2.3703	0.124
		3	-0.042	-0.023	2.6018	0.272
		4	0.038	0.033	2.7929	0.425
		5	0.056	0.043	3.2107	0.523
		6	0.042	0.040	3.4467	0.631
		7	0.042	0.048	3.6879	0.719
		8	0.024	0.027	3.7652	0.806
		9	-0.036	-0.033	3.9449	0.862
		10	-0.090	-0.084	5.0982	0.826
		11	-0.105	-0.107	6.6804	0.755
		12	0.046	0.037	6.9797	0.801
		13	0.098	0.064	8.3584	0.757
		14	-0.022	-0.026	8.4281	0.815

Figure A6: Histogram and Jarque-Bera normality test for ARIMA(0,1,1) model

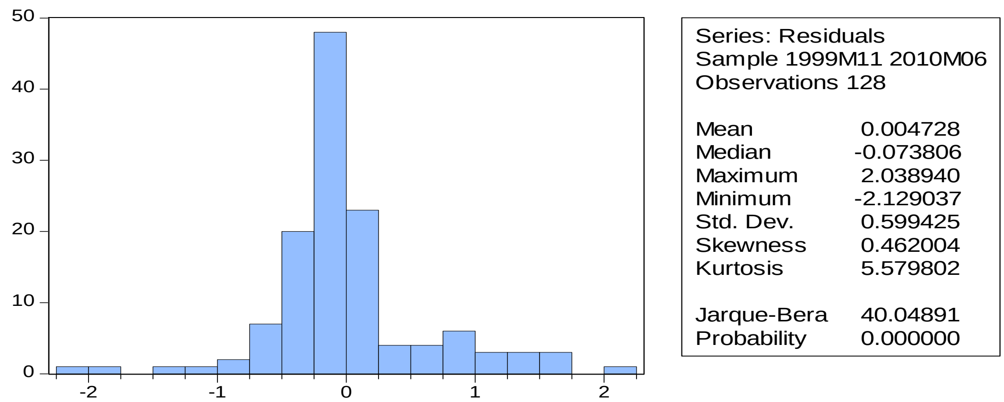


Figure A7: Sample ACF and PACF of the squared residuals by ARIMA(0,1,1) model

Date: 11/04/12 Time: 22:13
Sample: 1999M11 2010M06
Included observations: 128
Q-statistic probabilities adjusted for 1 ARMA term(s)

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.029	-0.029	0.1086	
		2	0.201	0.201	5.4564	0.019
		3	-0.002	0.009	5.4569	0.065
		4	0.017	-0.024	5.4951	0.139
		5	0.036	0.035	5.6663	0.225
		6	-0.032	-0.031	5.8089	0.325
		7	0.005	-0.012	5.8123	0.445
		8	0.031	0.046	5.9452	0.546
		9	0.004	0.008	5.9480	0.653
		10	-0.014	-0.032	5.9767	0.742
		11	0.059	0.061	6.4639	0.775
		12	0.012	0.024	6.4851	0.839
		13	0.159	0.139	10.133	0.604
		14	0.002	0.006	10.134	0.683
		15	0.165	0.116	14.139	0.439
		16	-0.005	-0.006	14.142	0.515

Figure A8: Conditional standard deviation for GARCH(1,1) model

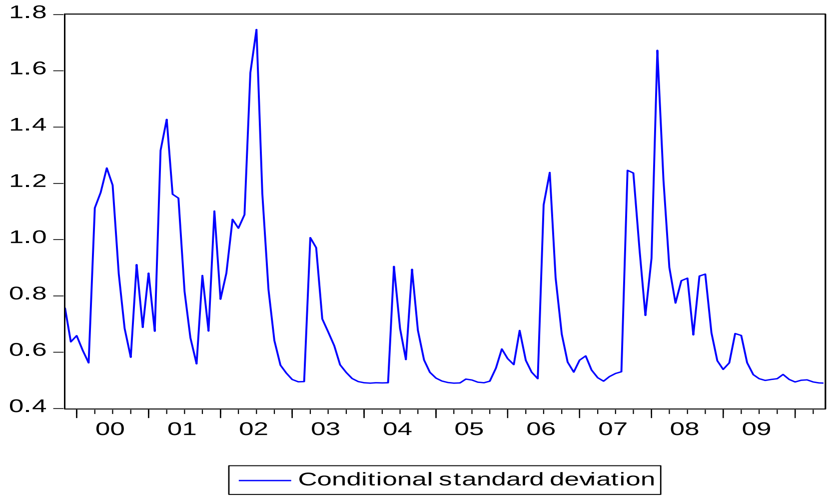


Figure A9: Conditional variance for GARCH(1,1) model

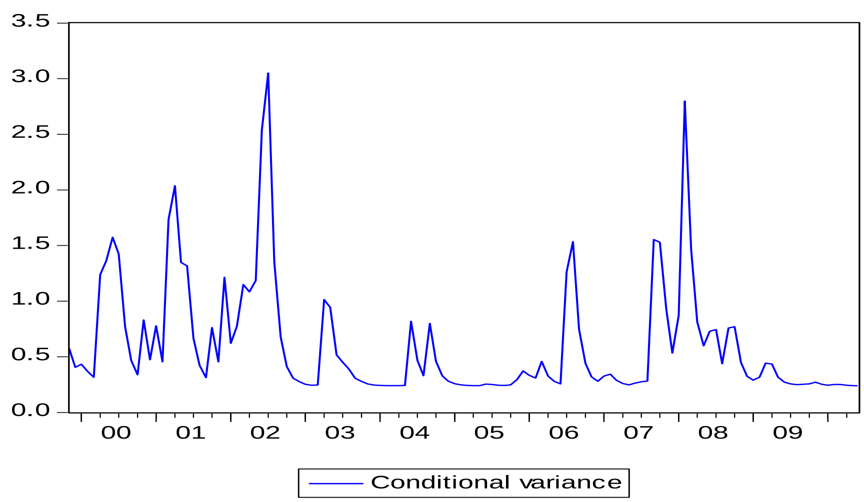
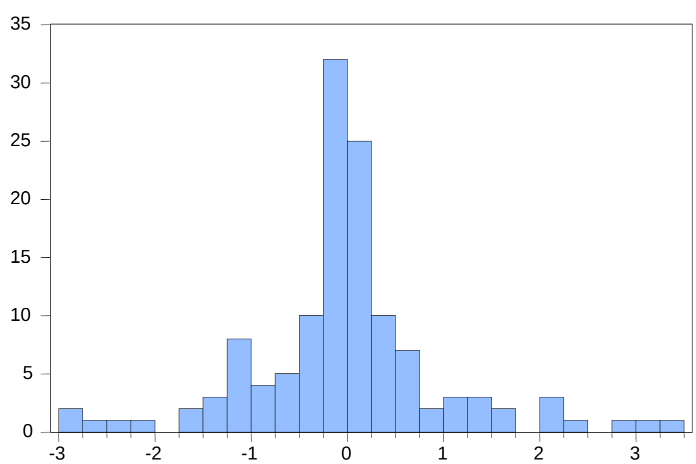


Figure A10: Sample ACF and PACF of standardized residuals square for GARCH (1,1) model

Date: 10/13/12 Time: 19:34
Sample: 1999M11 2010M06
Included observations: 128

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.010	0.010	0.0131	0.909
		2	-0.024	-0.024	0.0877	0.957
		3	-0.034	-0.033	0.2373	0.971
		4	0.038	0.039	0.4347	0.980
		5	0.054	0.052	0.8274	0.975
		6	-0.071	-0.071	1.5088	0.959
		7	0.047	0.054	1.8184	0.969
		8	-0.059	-0.062	2.2999	0.970
		9	0.046	0.042	2.5969	0.978
		10	-0.066	-0.065	3.2065	0.976
		11	0.053	0.059	3.6069	0.980
		12	-0.018	-0.028	3.6522	0.989
		13	0.037	0.050	3.8548	0.993

Figure A11: Histogram and normality test for standardized residuals



Series: Standardized Residuals
Sample 1999M11 2010M06
Observations 128

Mean	-0.006967
Median	-0.010911
Maximum	3.460722
Minimum	-2.858298
Std. Dev.	1.002607
Skewness	0.387544
Kurtosis	5.458811
Jarque-Bera	35.44807
Probability	0.053000

Figure A12: Forecast petroleum oil import prices by GARCH(1,1) model

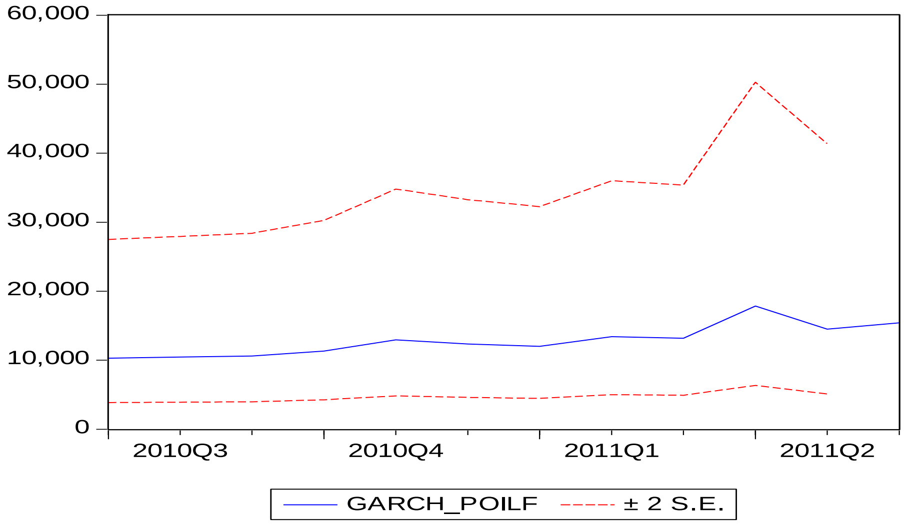


Figure A13: Conditional variance forecast by GARCH(1,1) model

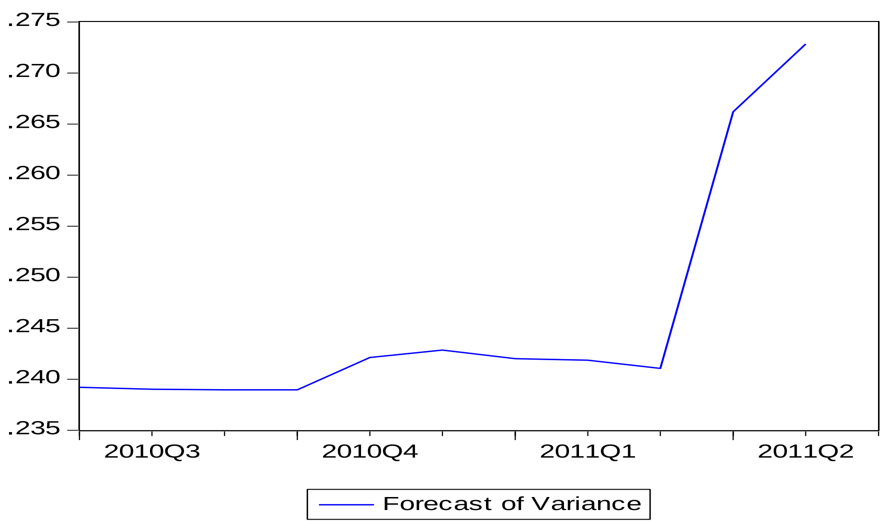


Figure A14: The plot of actual import prices against forecast import prices by GARCH(1,1) model

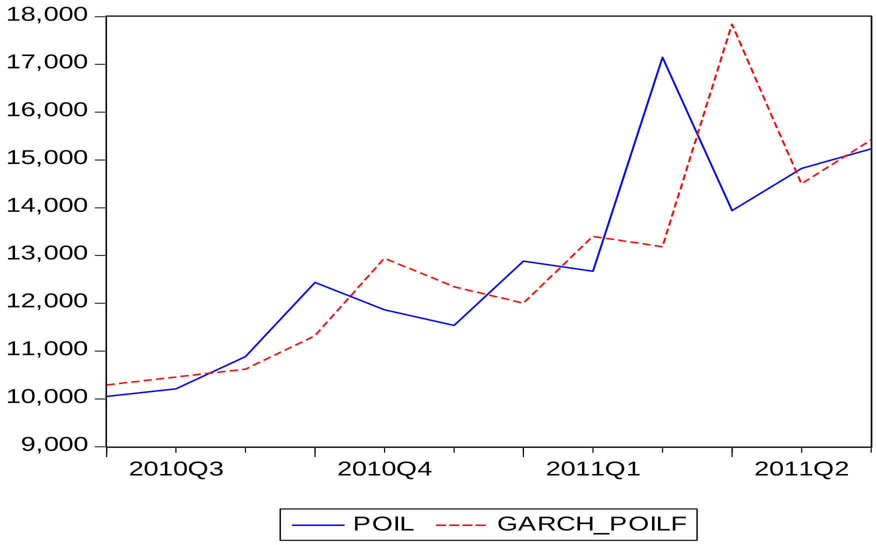


Table A1: ADF test for petroleum oil import prices (at level form)

Null Hypothesis: POIL has a unit root

Exogenous: Constant

Lag Length: 3 (Automatic based on SIC, MAXLAG=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-2.744851	0.0694
Test critical values:	1% level	-3.483312	
	5% level	-2.884665	
	10% level	-2.579180	

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

Table A2: PP test for the petroleum oil prices (at level form)

Null Hypothesis: POIL has a unit root

Exogenous: Constant

Bandwidth: 8 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-9.901176	0.0000
Test critical values:	1% level	-3.482035	
	5% level	-2.884109	
	10% level	-2.578884	

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

Table A3: ADF unit-root test after first difference of log series

Null Hypothesis: DLOGPOIL has a unit root

Exogenous: Constant

Lag Length: 3 (Automatic based on SIC, MAXLAG=12)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-9.767656	0.0000
Test critical values:	1% level	-3.483751	
	5% level	-2.884856	
	10% level	-2.579282	

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

Table A4: PP unit-root test after first difference of log series

Null Hypothesis: DLOGPOIL has a unit root

Exogenous: Constant

Bandwidth: 76 (Newey-West using Bartlett kernel)

		Adj. t-Stat	Prob.*
Phillips-Perron test statistic		-70.80035	0.0001
Test critical values:	1% level	-3.482453	
	5% level	-2.884291	
	10% level	-2.578981	

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

Table A5: Estimation equation of ARIMA(0,1,1)

Dependent Variable: DLOG(POIL)

Method: Least Squares

Date: 08/06/12 Time: 12:33

Sample (adjusted): 1999M11 2010M06

Included observations: 128 after adjustments

Convergence achieved after 7 iterations

MA Backcast: 1999M10

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.010334	0.007400	1.396505	0.1650
MA(1)	-0.868780	0.044029	-19.73202	0.0000

R-squared	0.378433	Mean dependent var	0.013725
Adjusted R-squared	0.373500	S.D. dependent var	0.760335
S.E. of regression	0.601818	Akaike info criterion	1.837779
Sum squared resid	45.63533	Schwarz criterion	1.882342
Log likelihood	-115.6179	Hannan-Quinn criter.	1.855885
F-statistic	76.71346	Durbin-Watson stat	1.829280

Prob(F-statistic)	0.000000

Inverted MA Roots	.87

Table A6: Serial correlation Breush-Godfrey LM test for ARIMA(0,1,1) model

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	1.214684	Prob. F(1,125)	0.2725
Obs*R-squared	1.223918	Prob. Chi-Square(1)	0.2686

Table A7: ARCH-LM test for ARIMA(0,1,1) model

Heteroskedasticity Test: ARCH

F-statistic	7.212494	Prob. F(1,125)	0.0082
Obs*R-squared	6.928141	Prob. Chi-Square(1)	0.0085

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 11/04/12 Time: 23:37
Sample (adjusted): 1999M12 2010M06
Included observations: 127 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.515440	0.374300	-6.720393	0.0000
RESID^2(-1)	0.233011	0.086763	2.685609	0.0082

R-squared	0.054552	Mean dependent var	-3.282181
Adjusted R-squared	0.046989	S.D. dependent var	2.794246
S.E. of regression	2.727807	Akaike info criterion	4.860496
Sum squared resid	930.1161	Schwarz criterion	4.905286
Log likelihood	-306.6415	Hannan-Quinn criter.	4.878693
F-statistic	7.212494	Durbin-Watson stat	2.069218
Prob(F-statistic)	0.008223		

Table A8: Parameter estimation of GARCH(1,1) model

Dependent Variable: DLOG(POIL)

Method: ML – ARCH

Date: 08/06/12 Time: 12:34

Sample (adjusted): 1999M11 2010M06

Included observations: 128 after adjustments

Convergence achieved after 14 iterations

Presample variance: backcast (parameter = 0.7)

GARCH = C(2) + C(3)*RESID(-1)^2 + C(4)*GARCH(-1)

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	0.039427	0.049256	0.800468	0.4234
		Variance Equation		
C	0.144785	0.048315	2.996719	0.0027
RESID(-1)^2	0.384998	0.166781	2.308410	0.0210
GARCH(-1)	0.393093	0.159409	2.465937	0.0137

R-squared	-0.001152	Mean dependent var	0.013725
Adjusted R-squared	-0.025373	S.D. dependent var	0.760335
S.E. of regression	0.769920	Akaike info criterion	2.139983
Sum squared resid	73.50437	Schwarz criterion	2.229109
Log likelihood	-132.9589	Hannan-Quinn criter.	2.176196
Durbin-Watson stat	2.781728		

Table A9: ARCH-LM test for GARCH (1,1) model

Heteroskedasticity Test: ARCH

F-statistic	0.012490	Prob. F(1,125)	0.9112
Obs*R-squared	0.012688	Prob. Chi-Square(1)	0.9103

Test Equation:

Dependent Variable: WGT_RESID^2

Method: Least Squares

Date: 08/09/12 Time: 23:51

Sample (adjusted): 1999M12 2010M06

Included observations: 127 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.993972	0.209085	4.753907	0.0000
WGT_RESID^2(-1)	0.009998	0.089461	0.111757	0.9112

R-squared	0.000100	Mean dependent var	1.004023
Adjusted R-squared	-0.007899	S.D. dependent var	2.118823
S.E. of regression	2.127175	Akaike info criterion	4.363089
Sum squared resid	565.6092	Schwarz criterion	4.407880
Log likelihood	-275.0562	Hannan-Quinn criter.	4.381287
F-statistic	0.012490	Durbin-Watson stat	1.998231
Prob(F-statistic)	0.911195		

Table A10: Forecast evaluation for GARCH(1,1) model

Forecast: GARCH_POILF

Actual: POIL

Forecast sample: 2010M07 2011M06

Included observations: 12

Root Mean Squared Error	1722.848
Mean Absolute Error	1144.322
Mean Absolute Percentage Error	8.277502
Theil Inequality Coefficient	0.066243
Bias Proportion	0.000901
Variance Proportion	0.002590
Covariance Proportion	0.996509

Table A11: Out-sample forecasting of petroleum oil import prices series for GARCH model

Months	Poil(obs)	GARCH_f
Jul-10	10051	10287.73
Aug-10	10209	10455.20
Sep-10	10884	10619.56
Oct-10	12439	11321.70
Nov-10	11866	12939.24
Dec-10	11536	12343.19
Jan-11	12882	11999.92

Feb-11	12672	13400.05
Mar-11	17148	13181.61
Apr-11	13939	17837.61
May-11	14825	14499.56
Jun-11	15235	15421.19

Table A12: Forecasts from GARCH(1,1) model

Months	GARCH(1,1)_POILF (Birr/Mt)
Jul-11	15847.7
Aug-11	16060.9
Sep-11	16385.5
Oct-11	17936.4
Nov-11	16664.2
Dec-11	16956.5
Jan-12	17668.1
Feb-12	19789.1
Mar-12	17927.1
Apr-12	18309.9

May-12	19077.5
Jun-12	18735.3
Jul-12	19474.9
Aug-12	20909.4
Sep-12	20783.5
Oct-12	20128.2
Nov-12	20658.7
Dec-12	21849.7
Jan-13	21174.6
Feb-13	21482.5
Mar-13	22284.5
Apr-13	22662.1
May-13	22946.1
Jun-13	23568.2