ADDIS ABABA UNIVERSITY
GRADUATE STUDIES PROGRAMME
DEPARTMENT OF STATISTICS

Modeling and Forecasting Cereal Price in Ethiopia: an Application of ARIMA and GARCH Models

KEDIR MAHAMEDHUSIEN

A Thesis submitted to the School of Graduate Studies of Addis Ababa University in partial fulfillment of the requirements for the Degree of Master of Science in Statistics.

JUNE 2013
ADDIS ABABA

Addis Ababa University
School of Graduate Studies

This is to certify that the thesis prepared by Kedir Mohamedhusien Abdelkadir, entitled: Modeling and Forecasting Cereal Price in Ethiopia: an Application of ARIMA and GARCH Models and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics (Applied Statistics) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

Examiner M. K. Sharma Signature M.K.Sh Date June-27-2013

Examiner Emmanuel G. Yohannes Signature ______ Date June-27-2013

Advisor Butte Gotu Signature ______ Date June-27-2013

____

Chair of Department or Graduate Program Coordinator

ACKNOWLEDGEMENT

First and foremost, I would like to extend my unshared thanks to the almighty God for providing me the opportunity for what I have achieved and for his mercy (Alhamdu Lilah).

My deepest gratitude goes to my thesis advisor Dr. Butte Gotu for his constructive advice and guidance in all phases of the study.

No words can suffice to express my feelings of gratitude to teachers and friends who were all time with me giving moralsupport and assistance throughout my research work.

Abstract

Models for Forecasting Cereal Price in Ethiopia: An Application of ARIMA and GARCH Models

Kedir Mohamedhusien

Addis Ababa University

Cereal production and marketing constitute the single largest sub-sector in Ethiopian economy. It accounts for roughly 60 percent of rural employment, 73% of total cultivated land and 68.3% of total output, 46 percent of a typical household's food expenditure more than 60% of caloric/ intake. According to available estimates, cereals production represents about 30 percent of gross domestic product (GDP). In this study we attempted to model cereal price and obtain forecasts at national level. The data used are monthly cereal price obtained from the Central Statistical Agency (CSA) for the periods from September 1996 to July 2012.

Seasonal ARIMA and GARCH were employed to analyze the monthly cereal price data. It was found that the Seasonal ARIMA(0,1,1)*(0,1,1) and ARMA(2,1)-GARCH(1,1) were adequate models for the data considered in this study. In the GARCH model, the value of the GARCH term for the return of cereal price is close to one indicating slow convergence of volatility to a steady state and high persistence in volatility. In addition, the constant term in the mean equation was significant and thus it follows an ARMA (2, 1) model. The point forecast results showed a very closer match with the pattern of the actual data and better forecasting accuracy in validation period. Almost all the in-sample forecast evaluations statistic indicated that the Seasonal ARIMA model is better in comparison to GARCH Model. However, almost all the out-sample forecast evaluation statistic shows the superiority of GARCH (1, 1) model over the Seasonal ARIMA model.

LIST OF ABBREVIATIONS

ACF	Autocorrelation Function
ADF-Test	Augmented Dickey-Fuller-Test
AICc	Akaike Information Criterion corrected for Bias
AIC	Akaike Information Criterion
ARCH	Autoregressive Conditional Heteroscedasticity
AR	Autoregressive
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CSA	Central Statistics Agency
DF	Dickey-Fuller test
EGARCH	Exponential Generalized Autoregressive Conditional Heteroscedasticity
EGTE	Ethiopian Grain Trade Enterprise
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GARCH-M	Generalized Autoregressive Conditional Heteroscedasticity mean
GDP	Gross Domestic Product
GED	Generalized Error Distribution
HDI	Human Development Index
IFPRI	International Food Policy Research Institute
JB	Jarque-Bera-Test
LM Test	Lagrange Multiplier Test
MA	Moving Average
ME	Mean Error
MAE	Mean Absolute Error
MAPE	Mean Absolute percentage Error
MFIs	Microfinance Institutions
ML	Maximum Likelihood
MoARD	Ministry of Agriculture and Rural Development
MoFED	Ministry of Finance and Economic Development
MSE	Mean Square Error
MSR	Mean Square Regression
PACF	Partial Autocorrelation Function
PP	Phillips-Peron
PSNP	Productive Safety Net Program PSNP
Q	Ljung-Box Q-test Statistic
Q-Q Plot	Quantile-Quantile plot
RMSE	Root mean Square Error
SIC	Schwarz Information Criterion
SV	Stochastic Volatility Models
UK	United Kingdom
VIX	Volatility Index

Table of Contents

List of Figures

List of Tables

CHAPTER ONE: Introduction

1.1. Background and Motivation of the Study

Official statistics of Ethiopia indicates an average real GDP growth of 11 percent over the last eight consecutive years between 2003/4 – 2010/11(Annual Report 2010/11 MoFED). Ethiopia's HDI score improved from .240 to 0.328 between the years 2000- 2010; that is a 2.73% average annual growth in HDI. This is the third fastest average annual HDI growth rate in the world. The top two are Rwanda and Serra Leone with an average annual HDI growth rate of 3.31% and 2.29%, respectively (Human Development Report 2010). In 2011, Ethiopia's HDI is 0.367, which makes Ethiopia a country which has the first fastest annual HDI growth rate in the world (HDR Report 2011). However, if you do much shopping, you have probably noticed that the cost of food has been rising at very brisk pace over the past years. So, why are food prices rising so fast? What will happen if food prices keep rise?

Nicolas Sarkozy urges Action against world food prices that rose 37 percent in a year, driving 44 million more people into poverty (Ruitenberg and Dreibus Jun 22, 2011. When food prices rise in the developed countries it may be painful for millions of these countries, but around the world, a rise in food prices may mean the difference between surviving and not surviving.

High food prices means buying less food, or having less money to spend on other things. Higher food prices are causing more people to go hungry. Because high food prices means spend a lot of money on food or not getting enough to eat. In general massive food price increase could result in:

- Increase hunger as food becomes unaffordable.
- Eat less fruit, vegetables, dairy and meat in order to afford staple foods such as wheat. As a result of less or no intake of protein and vitamins, causing malnutrition particularly for children, pregnant women and unborn children.
- Households are going into debt or selling off assets vital to future income such as cattle to pay for food.
- Less attention to health care, family planning and education
- Women who tend to manage the food budget may earn money by taking up risky employment such as sex work or domestic work.

Despite clear understanding on the impacts of high food price variability, much has not been done in modeling food price data. Therefore, there is need for understanding and modeling the rate and size of food price variabilities in developing countries like Ethiopia.

1.2. Statement of the Problem

The knowledge about domestic price and its volatility pattern influences decisions of policy makers to bring price and its volatility on staple food crops to satisfactory levels, as price and its volatility markets affect production levels, investment and income stability of consumers, whole sellers, producers and as well as the country's economic development. Moreover, it is generally accepted that profound understanding of the price and its volatility generating process, in particular, the speed of price volatility adjustment in response to shocks due to different factors is of crucial importance for a developing economy especially for Ethiopia whose policy is oriented towards price stability. Therefore, this study attempts to address the following problems: (i) Is there domestic price volatility on the cereal price? (ii) Can it be modeled and predicted by GARCH or Seasonal ARIMA Model?

1.3. Objective of the Study

The general objective of this study is to model cereal prices as well as volatility of cereal prices in Ethiopia.

The specific objectives are:

- Fit appropriate ARIMA and GARCH models for data on cereal prices
- Assess the volatility of producer price of cereals
- Make forecast of the producer price and price volatility for cereals
- Compare the result of ARIMA and GARCH models on cereal prices

1.4. Significance of the Study

As many studies indicated price volatility on the agricultural commodities has negative impact on the economy of the country through making income instability for producers, consumers, whole sellers, and leads to a major decline in future output (Gebremedhin et al. 2006). Thus, the findings of this study

- Could help to closely monitor the changes in the market supplies and price behavior such as the surpluses or scarcities that could require adjustments in response to the current food shortages (optimal control of the system).
- Contributes to identify the pattern of producer price volatility for the purpose of being able to make more informed decisions and to regulate its prices.

- Contributes to knowledge on agricultural commodities pricing.
- Would be used as a basis to other researchers for further investigations.

1.5. Organization of the Study

This work is organized into the following chapters. Chapter one is introduction of the thesis. This chapter briefly addresses the thesis objectives, significance, and background. Chapter 2 reviews the literature with emphasis on the statistical tools relevant to modeling prices. Chapter 3 explains the methodology applied in building ARIMA and GARCH models and estimating their parameters in various time lags for plausible description of the data. Chapter 4 presents the results of analysis. Chapter 5 concludes this paper.

CHAPTER TWO: Literature Review

There is now considerable empirical evidence that the volatility in agricultural prices is a concern for agricultural producers and for other agents along the food chain. Pindyck (2004) pointed out that changes in commodity prices can influence the total cost of production as well as the opportunity cost of producing commodities currently rather than later. It has also been argued that price volatility reduces welfare and competition by increasing consumer search costs (Zhang et al., 2008). In the same line, Apergis and Rezitis (2003) noted down that price volatility leads both producers and consumers to uncertainty and risk.

Price volatility on agricultural crop in developing countries like Ethiopia leads to a variety of risks like income instability on consumers, producers, as well as on overall economic growth. For poor consumers, consequences of price instability are severe. Since a large share of their income is spent on food, an unusual price increase forces them to cut down food intake, take their children out of school, or, in extreme cases, simply to starve. The recent hike in relative prices has increased the urban cost of living by 8-12 percent in urban areas and worsened income inequality significantly (Klugman, 2007).

The literature on inflation in Ethiopia is limited. Nevertheless, few studies have emerged in the light of Ethiopia's food price crisis. Most of these studies take a general approach, identifying and discussing various possible factors contributing to inflation. Several factors have been mentioned as causes of the recent global food price inflation. For example: rising population; rapid economic growth in emerging economies which resulted in increased food demand; high energy and fertilizer prices; increased use of food crops for bio-fuels; depreciation of the US dollar; and declining global stocks of food grains due to changes to buffer stock policies in the US and European Union (Ahmed, 2007). More recently, Gilbert

(2009) argued that the world food price hikes in 2006-2008 are mainly explained by depreciation of dollar and future market investments.

There is no consensus on why Ethiopia is experiencing such rapid price rises. Inflation growth has recently coincided with relatively high economic growth rates, whereas in the past inflation was traditionally associated with large agricultural supply shocks due to drought. World food price increases are traditionally believed to have rather small effects in Ethiopia because of the limited size of food imports, which amount to about 5 percent of agricultural GDP. Prices for major staple crops have been above import parity since early 2008, and though there has been an incentive to import ordinary cereals, estimates suggest that little informal or formal trade actually took place (IFPRI, 2008). Instead, the chief explanations have focused on high domestic demand, expansionary monetary policy, a shift from food aid to cash transfers, and structural factors due to reforms and investments in infrastructure (Ahmed, 2008; Dorosh and Subran, 2007; WB, 2007; IMF, 2008a; IMF, 2008b).

There is some evidence indicating that world food prices have been driven by higher grain prices. For instance the international price of wheat increased more than triple between 2002 and March 2008. The price has since then come down, but as of August 2008 it remained 70 percent higher than the average price in 2006. Similar trends have been exhibited for other cereals and food items (Ahmed, 2008; Ivanic and Martin, 2008).

According to the empirical findings of Mulat et al (2007), the role of cooperatives has been significantly increasing both in the input and output market through improved access to storage facilities and market information. Along with strong emergence of these farmers cooperatives, access to credit channeled through Microfinance Institutions (MFIs) and the

Ministry of Agriculture and Rural Development (MoARD) has also greatly contributed to the change in production and marketing behavior of farmers.

Moreover, since 2005 Ethiopian Government managed to secure food aid in terms of cash transfer from donors through Productive Safety Net Program (PSNP). This had eliminated imported food aid and pumping more cash for local. According to the WB 2007 study, of the total aid beneficiaries' cash recipients account for 45% to 64%, this has been an important source of food expenditure. In fact, in 2006 alone the PSNP transferred about ETB 800 million in cash and ETB 800 million equivalents in cereals (WB, 2007).

Over the past three decades, Ethiopia has experimented with a whole spectrum of agricultural price policies and then following the overthrow of the Derg regime in May 1991, various economic reform programs were launched, including major reforms in cereal markets. As part of the reorganization and re-structuring of government parastatals that began in 1992, Agricultural Market Corporation reorganized as a public enterprise and allowed to operate in the open market in competition with the private sector. The name of the agency was also changed to the Ethiopian Grain Trade Enterprise (EGTE) and one of its mandate is to stabilize prices with an objective to encourage production and protect consumers from price shocks and maintain a strategic food reserves for disaster response and emergency food security operations. However, the EGTE encountered at least three major problems in the subsequent years. First, there was a constant tension between fulfilling its mandate of price stabilization and that of competitiveness and profitability (Bekele, 2002). Second, EGTE was not effective in stabilizing grain prices due to its limited grain purchases and sales network and shortage of working capital. The closure of branch offices and purchase and/or sales centers in regions with less potential for grain production, and in remote areas reduced procurement and led to under utilization of EGTE resources (Lirenso, 1994). Finally, the EGTE was often not able to guarantee purchases at pre-announced prices due to logistic and capital constraints, which had led to shaken farmers confidence and loss of policy credibility (Meron et al., 2006). The most recent and important attempt towards market development in Ethiopia has been the establishment of the Ethiopian commodity exchange which imposed the ban on cereal export in February 2008 to stabilize domestic price on the agricultural food crops. This was based on the assumption that prices had increased because of exports. However, the IFPRI study concluded that the cross border trade of cereal was too small to influence the domestic market prices (Alemu et al., 2007).

When it compare to statistical modeling of cereal price the following studies could be cited:

Stephan et al (2009) studies the development of cereal prices during the 20^{th} century using yearly data from 1900-2002. He found both, the ARIMA (1,1,2) and ARIMA (0,1,4) models selected for rice and wheat respectively, indicate the presence of unit moving average roots. As far as wheat and rice are concerned then, the conclusions obtained for the 1900-1992 sub-sample continue to support the inference of a significant drift coefficient in the difference stationary model selected by AIC while there appears to be no support for the inclusion of dummy variables. (Stephan et al 2009).

Shiferaw (2012) employed ARCH/GARCH models to capture the log-return price volatility under the study of price returns of crops. He found that GARCH (1,1), GARCH(1,2) and GARCH (2,1) models were the most appropriate fitted models to use one has to evaluate the volatility of the log-returns of price of Cereal, pulse and oil crops respectively. Prices volatility was persistent in all three categories of selected agricultural crops. The following studies employed GARCH and/or ARIMA models in analyzing various types of inflation or price volatility

There are a number of different volatility estimating and forecasting models such as GARCH-type and SV models. Among those models, the family of GARCH-type models introduced by Engle (1982), Bollerslev (1986) and Taylor (1986) and then developed by many other researchers is the most common approach and proved to be sufficient to capture the properties of return series. We narrow the discussion of empirical studies to the GARCH and ARIMA models.

In his research, Engle (1982) uses ARCH (4) models to estimate the mean and variance of inflation in the UK. He found that the ARCH effect is significant and the estimated variance considerably increased during crisis period. As mentioned in previous section, some problems with ARCH (q) models led to the more general framework GARCH (p, q) proposed by Bollerslev (1986) and Taylor (1986). Bollerslev (1986) focused on the paper of Engle and Kraft (1983) who attempted to explain the rate of growth of GNP deflator in US by its own lagged values. Bollerslev employed GARCH (1,1) and ARCH (8) in his study and found that GARCH (1,1) outperform the ARCH (8) in terms of best fit and reasonable lag structure.

Three years later, Akgiray (1989) employed GARCH (1, 1) model to explain the daily stock return series obtained from Centre for Research in security prices from 1963 to 1986. The empirical results showed the significant dependence in stock return. The conditional heteroscedasticity process allowed the improvement of the volatility forecast. Akgiray (1989) also argued that GARCH (1, 1) fit the data very satisfactorily and outperform usual historical estimates in forecasting variance.

Bollerslev (1986) provided the overview of some extension of ARCH type models and a survey of some empirical applications. They argued that GARCH (1, 1) is sufficiently enough in almost financial application without the need of more complicated models.

The forecasting ability of ARCH-type models is comprehensively researched in the study of Hansen and Lunde (2005). Their analysis consisted of 330 different ARCH-type models to estimate the volatility of DM exchange rate and IBM stock returns. The main findings are that the GARCH (1, 1) is superior in analyzing DM exchange rate rather than any other models. Whereas, GARCH (1, 1) underperforms in modeling IBM stock return compared to other models. The study suggested the good forecasting models require specifications, which are able to capture leverage effects.

Another prominent piece of paper is performed by Song et al. (1998). After the economic reform in China, Song et al. (1998) employed GARCH models to analyze the volatility on the two main official stock markets including Shanghai and Shenzhen Stock Exchanges in mainland China. The sample data was observed from May 1992 to February 1996. It is believed that the Chinese stock markets which have fewer listed companies and smaller capitalization than mature markets and have been fluctuated after the economic reform are considerably more volatile. There are three main outcomes, which can be concluded from the paper. Firstly, although various GARCH (p,q) specifications are fitted the data, GARCH-M (1,1) model shows to best model the stock return series of both Shanghai and Shenzhen markets. GARCH-M model indicates the higher returns for higher risks in the stock prices. Secondly, the evidence of volatility transmission between the two markets was found. Moreover, the one month forecast suggests the similar pattern of the conditional variances of the two markets' return.

The result gained from different studies varies between markets and depends on the specification of data and the choice of evaluation measures. In this study, we have used Seasonal ARIMA and GARCH models.

CHAPTER THREE: Data and Methodology

3.1. Data Sources

The purpose of this study is to model and forecast cereal price and volatility of cereal price of Ethiopian market. In order to propose suitable specifications, it is reasonable to investigate the nature of the data set. The characteristics of the data and their descriptive statistics partly indicate the appropriate models, which should be employed.

In this study, the data are secondary data on the monthly cereal price of Ethiopian markets obtained from the Central Statistical Agency (CSA). The focus on cereals is motivated by the place of cereals on production and consumption in Ethiopia. Cereals are by far the largest group in terms of their share in area cultivated, output, and consumption.

The data is limited to the four regions of Ethiopia, namely, Tigray, Amhara, Oromya and SNNP Region. My choice of these regions was not only due to availability of relatively long series of prices but also due to their large coverage cereal production, which is about 90% (Taffese, 2008). The choice of monthly data is obligatory, because daily or weekly cereal price data are not available.

The sample means for cereal price series starts on September 1996 and ends on July 2012 for the following considerations. Prior to September 1996 the data are not available. That means monthly cereal price data are only available starting from September 1996. Therefore, the research period is intentionally chosen from September 1996 to July 2012. The series from 1996 to 2012 constitutes 192 monthly observations. The time period is divided into two. The first period includes the mean cereal price data from September 1996 to august 2011. These are used for the estimation of the models' parameters. The remaining

12 observations, which are constructed from September 2011 to July 2012, are employed for out-of-sample forecasts.

3.2. Methodology

3.2.1. Preliminaries

Definitions

(i) Strictly stationary

A strictly stationary time series is one for which the probabilistic behavior of collection of values$\{x_{t_1}, x_{t_2}, \dots, x_{t_k}\}$ is identical to that of the time-shifted set

$$\{x_{t_1+h}, x_{t_2+h}, \dots, x_{t_k+h}\}$$

(ii) Weakly Stationary

A weakly stationary time series, x_t, is a finite variance process. Its mean function, μ_t, is constant and does not depend on time t i.e $\mu_t = \mu$ for all t. And the covariance function,$\gamma(s, t)$ depends on s and t only through their difference $|t - s|$ which we call lags; for all s and t and is denoted by $\gamma(h)$ where $h = t - s$. In this study, we will use stationary to mean weakly stationary.

Most of the probability theory of time series is concerned with stationary time series, and for this reason, time series analysis often requires one to turn a non-stationary series into a stationary one so that one can use this theory (Brock well and Davis, 1996). To apply this theory we need to test for stationarity, if not difference and test the differenced series.

3.2.2 Tests for Stationarity

3.2.2.1. Time Plot:

Regardless of which technique is used, the first step in any time series analysis is to construct a time plot of the data, and inspect the graph for any anomalies. A number of qualitative aspects are noticeable as you visually inspect the graph. A time plot of the data will typically suggest whether any differencing is needed to make the time series stationary. If differencing is called for, then difference the data once, d = 1, and inspect the time plot of ∇x_1. If Additional differencing is necessary, then try differencing again and inspect a time plot of $\nabla^2 x_1$. Be careful not to over difference because this may introduce dependence where none exists. For example, $x_t = w_t$ is serially uncorrelated; but $\nabla x_t = w_t - w_{t-1}$ is an invertible MA (1).

3.2.2.2. The Correlogram Test

The autocorrelation function, ACF, defined as

$$\rho(s,t) = \frac{\gamma(s,t)}{\sqrt{\gamma(t,t)\gamma(s,s)}} \qquad (1)$$

measures the linear predictability of the series at time t, say x_t, using only the value x_s.

One way to characterize a series with respect to its dependence over time is to plot its sample autocorrelation function. This function is abbreviated as ACF and its plot is usually referred to as a correlogram. The partial autocorrelation function, denoted by PACF, is similar to the ACF and can be described as the correlation between x_t and x_{t+h} after controlling for the common linear effects of the intermediate lags. Both functions are used in Box-Jenkins modeling as correlogram to reveal important information regarding the order of the autoregressive (AR) and moving average (MA) orders as mentioned in table 1 and 2 below. In addition to this time plots, the sample ACF can help in indicating whether differencing is needed to achieve stationarity. Because the polynomial $\phi(z) = (1 - z)^d$ has a unit root, the sample ACF, $\hat{\rho}(h)$, will not decay to zero fast as h increases. Thus, as Wei (1990) states, a slow decay in $\hat{\rho}(h)$ is an indication that differencing may be needed.

3.2.2.3. The Unit Root Test

While the stationary tests described in the above sections make use of subjective visual inspection of data plots and correlogram, other tests were developed to help with determining stationary. These tests are also known as unit root tests. Stationary tests are based for the most part on formal statistical tests and the difference between them lies in the strictness of the assumptions they use as well as in the form of the null and alternative hypotheses they adopt. The standard Dickey-Fuller test (DF) is based on iid. errors and has as a null hypothesis that there is a unit root. On the other hand, the Phillips-Petron test is nonparametric and allows for some heterogeneity and serial correlation in the innovations. There exist many other unit root test other than the ones mentioned above. However, in this study, we use the most popular test called the Dickey-Fuller test (DF).

Consider a causal AR(1) process (we assume throughout this section of Unit Root that the noise is Gaussian);

$$x_t = \phi x_{t-1} + w_t \qquad (2)$$

A unit root test provides a way to test whether eq. 6 is a random walk (the null case) as opposed to a causal process (the alternative). That is, it provides a procedure for testing

$H_0: \phi = 1$ the data has unit root and is non-stationary versus

$H_1 < 1$ the data does not have unit root and it is stationary

Toward a more general model, we note the DF test was established by noting that if $x_t = \phi x_{t-1} + w_t$, then subtracting x_{t-1} from both sides we obtain $\nabla x_t = (1-\phi)x_{t-1} + w_t$ is same as $\nabla x_t = \gamma x_{t-1} + w_t$, and one could test $H_0: \gamma = 0$ by regressing Δx_t on x_{t-1}. They formed Wald statistic and derived its limiting distribution. The test was extended to accommodate AR(p) models, $x_t = \sum_{j=1}^{p} \left(\phi_j x_{t-j}\right) + w_t$, as follows. From the model

Subtract x_{t-1} to obtain

$$\nabla x_t = \gamma x_{t-1} + \sum_{j=1}^{p-1} \Psi_j \Delta x_{t-j} + w_t \qquad (3)$$

Where, $\gamma = (\phi_1 + \phi_2 + \cdots + \phi_p - 1)$ and $\Psi_j = -\sum_{j=i}^{p} \phi_j \quad j = 2,3,\ldots,p$

To test the hypothesis that the process has a unit root at 1 (i.e., the AR polynomial $\phi(z) = 0\ when\ z = 1$), we can test

$H_0: \gamma = 0$ - (the data has unit root and is non-stationary)

Versus

$H_1: \gamma < 0$ - (the data is stationary and does not need differencing)

by estimating γ in the regression of ∇x_t on x_{t-1}, ∇x_{t-1} ...., ∇x_{t-p} and forming a Wald test based on $t_\gamma = \frac{\hat{\gamma}}{se(\hat{\gamma})}$. This test leads to the so-called augmented Dickey-Fuller test (ADF). While the calculations for obtaining the asymptotic null distribution change, the basic ideas and machinery remain the same as in the simple case. For ARMA (p, q) models, the ADF test can be used assuming p is large enough to capture the essential correlation structure; The general model for ADF that includes a constant or even non-stochastic trend is

$$\Delta x_t = \alpha + \beta t + \gamma x_t - \sum_{j=1}^{p-1} \Psi_j \nabla x_{t-j} + w_t \qquad (4)$$

Where, $\gamma = (\phi_1 + \phi_2 + \cdots + \phi_p - 1)$ and $\varphi_j = -\sum_{j=i}^{p} \phi_j \quad j = 2,3,\dots,p$

Example $x_t = \alpha + \beta t + \phi x_{t-1} + w_t$

If we assume $\beta = 0$, then under the null hypothesis, $\phi = 1$ *or equivalently* $\gamma = 0$, the process is a random walk with drift α. Under the alternative hypothesis, the process is a causal AR (1) with mean $\mu = \alpha(1-\phi)$. If we cannot assume $\beta = 0$, then the interest here is testing

H_0: $(\beta, \phi) = (0, 1)$ simultaneously, (The process is random walk with drift)

Versus

H_1: not H_0 (The process is stationary around a global trend)

the null that $(\beta, \phi) = (0, 1)$ simultaneously, versus the alternative that $\beta \neq 0$ *and* $|\phi| < 1$
In this case the null hypothesis will be the process is a random walk with drift, versus the alternative hypothesis that the process is stationary around a global trend.

The name, unit root, comes from the fact that the coefficient of x_{t-1} is unity, if the time series is non-stationary, and the Unit Root tests, as the name suggests, tests if ϕ is unity or not. Greene (2000) stated that a non stationary time series could be converted to a stationary time series by taking first or higher order difference.

3.2.3. Time Domain Approach and ARIMA Models

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

In this study a univariate Box-Jenkins Methods (Box et al 1994), in particular, seasonal Autoregressive Integrated Moving Average (Seasonal ARIMA) and Generalized Autoregressive Conditional Heteroscedastic (GARCH) methods are used.

The differences between the types of models tested here relate to the specification of the mean and variance of the series. ARIMA models have both a constant mean and a constant variance; GARCH models have a constant mean, but time-varying variance.

Assumptions of ARIMA Model

1. Expected value of residuals ω_t is zero, i.e. $E(\omega_t)=0$
2. Variance of the residuals ω_t is constant, i.e. var $(\omega_t) = \sigma_t^2 = \sigma^2$, for all t
3. Residuals ω_t are independently and normally distributed
4. No serial autocorrelations among each successive standardized residual

3.2.3.1. ARIMA Models

ARIMA Modeling takes in to account historical data and decomposes it into an autoregressive(AR) process, where there is a memory of past events(e.g., the cereal price of this month is related to the price of last month, and so forth, with a decreasing memory lag); an Integrated (I) process, which accounts for stabilizing or making the data stationary and making it easier to forecast; and a Moving Average (MA) of the forecast errors, such that the longer the historical data the more accurate the forecasts will be, as it learns over time.

There are many reasons why an ARIMA model is superior to common time-series analysis and multivariate regressions. The common finding in time series analysis and multivariate regression is that the error residuals are correlated with their own lagged values. This serial correlation violates the standard assumption of regression theory that disturbances are not correlated with other disturbances (Shumway and Stoffer 2006).

An ARIMA (p, d, q) model has p autoregressive terms (lagged values of the variable of interest), q moving average terms (lagged values of the error term), and d differencing operations (the number of differences needed to make a series stationary). Integrated refers to the differencing process. This describes a rich set of models for which numerical methods exist to fix models for any given p, d, and q. This suggests an automated approach to try all models in some subset of parameter space and select the best one.

3.2.3.2. AR Models

Autoregressive Integrated Moving Average ARIMA (p, d, q) models are the extension of the AR model that uses three components for modeling the serial correlation in the time series data. The first component is the autoregressive (AR) term. AR 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, $x_{t-1}, x_{t-2}, \ldots, x_{t-p}$ where p determines the number of steps into the past needed to forecast the current value (lagged values of the variable of interest). An autoregressive model of order p, abbreviated AR (p) (Box et al., 1994), has the form

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

$$x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + \omega_t \qquad (5)$$

Where x_t is stationary, $\phi_1, \phi_2, \dots, \phi_p$ are constants ($\phi_p \neq 0$). We assume that ω_t is a Gaussian white noise series with mean zero and variance δ_w^2. The order of such a model can be determined by analysis of the Partial autocorrelation function, PACF, which will cuts off after lag p. It is convenient to write the model in lag operators,

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

$$\rightarrow \left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) x_t = \omega_t$$

$$\rightarrow \phi(B) x_t = \omega_t \qquad (6)$$

The values of ϕ which make the process stationary are such that the roots of $\phi(B) = 0$ lie outside the unit circle in the complex plane where B is the backward shift operator such that

$B^j x_t = x_{t-j}$ for all j= 0 1, 2, 3, ...p (Chatfield, 1996). If all zero roots of $\phi(B)$ are larger than one in absolute value, there is a stationary process $x_{t,}$, which satisfies the autoregressive equation given as

$$x_t = \sum_{j=0}^{\infty} \Psi_j w_{t-j} = \Psi(B) w_t \qquad (7)$$

Where $\Psi(B) = \sum_{j=0}^{\infty} \Psi_j B^j$, $\sum_{j=0}^{\infty} |\Psi_j| < \infty, \Psi_0 = 1$ and B^j is the backward shift operator *i.e.* $B^j w_t = w_{t-j}$ *for all* $j = 1, 2, 3, \dots$

Recall that for a causal model, all of the roots are outside the unit circle, $|z_j| > 1, for\ j = 1; \dots; r$. Where z_j are the j^{th} root. If all the roots are real, then $\hat{\rho}_x(h)$ dampens exponentially fast to zero as $h \rightarrow \infty$. If some of the roots are complex, then they will be in conjugate pairs and $\hat{\rho}_x(h)$ will dampen, in a sinusoidal fashion, exponentially fast to zero as $h \rightarrow \infty$. In the case of complex roots, the time series will appear to be cyclic in nature. This, of course, is also true for ARMA models in which the AR part has complex roots.

3.2.3.3. Differencing Models

The Second component of ARIMA model is the integration (d) order term. Each integration order corresponds to differencing the time series. I(1) means differencing the data once. I(d) means differencing the data d times. Differencing transforms a time series X into another series Y where $y_t = x_t - x_{t-1}$. Differencing does not require estimating a parameter, although it costs us one series point per difference. Differencing is a better way to remove locally varying trends to make it stationary than explicitly subtracting a fitted trend. The first difference accounts for a trend that influences the change of the mean of the time series, the second for a change in the slope and so on.

One advantage of differencing over de-trending to remove trend is that no parameters are estimated in the differencing operation. One disadvantage, however, is that differencing does not yield an estimate of the stationary process y_t. If an estimate of y_t is essential, then de-trending may be more appropriate. If the goal is to make the data stationary, then differencing may be more appropriate. Differencing is also a viable tool if the trend is fixed

3.2.3.4. MA Models

The third component of the ARIMA model is the moving average (MA) term. The MA(q) model uses the q lags of the forecast errors to improve the forecast. A time series is said to be a moving average process of order q if it is a weighted linear sum of the last q random shocks /errors. In general, the moving average model of order q or MA (q) is defined to be

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \ldots + \theta_q w_{t-q} = \sum_{i=1}^{q} (\theta_i w_{t-i}) + w_t \quad (8)$$

Where there is q lags in the moving average and $\theta_1, \theta_2, \ldots \theta_q$, and $(\theta_q \neq 0)$ are parameters. The noise w_t is assumed to be Gaussian white noise. The history of the model dictates how long the effects of the random shocks last.

The order of such a model can be determined by analysis of the autocorrelation function, ACF, which cuts of after q lags and partial ACF decays exponentially fast. Unlike the autoregressive process, the moving average process is stationary for any values of the parameters $\theta_1, \theta_2, \ldots \theta_q$; (Diebold et al, 2006)

By mimicking the criterion of causality for AR models, we will choose the model with an infinite AR representation. Such a process is called an invertible process. If all roots of $\theta(B) = 0$ lie outside the unit circle, the MA process has an autoregressive representation of generally infinite order $w_t = \sum_{j=0}^{\infty} \pi_j (x_{t-j}) = \pi(B) x_t \; with \; \sum_{j=0}^{\infty} |\pi_j| < \infty$. A property required on occasion in the analysis of such time series is that of invertibility.

3.2.3.5. ARMA Models

Finally, an ARMA (p, q) model has the combined form:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q}$$

Or

$$x_t = \sum_{j=1}^{p} (\phi_j x_{t-j}) + \sum_{i=1}^{q} (\theta_1 w_{t-i}) + w_t \quad (9)$$

With $\phi_p \neq 0$, and $\theta_q \neq 0$, and $\delta_w^2 > 0$. The parameters, p and q are called the autoregressive and the moving average orders respectively.

Backward shift operator B will simplify eq. 9 of the time series model as follows

$$\phi(B)x_t = \theta(B)w_t \qquad (10)$$

Where $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

In the definition of ARMA (p, q) we have three main problems that we need to address; these are parameter redundancy; stationary AR models that depend on the future, and MA models that are not unique. (Shum-way and Stoffer, 2006)

To address the first problem, we will henceforth refer to an ARMA (p, q) model to mean that it is in its simplest form. That is, in addition to the original definition given in eq.9, we will also require that $\phi(B)$ and $\theta(B)$ have no common factors.

An ARMA (p, q) model, $\phi(B)x_t = \theta(B)w_t$ is said to be causal only when all the roots of $\phi(B)$ are larger than one in absolute value. Equivalently, if all the roots of $\phi(B)$ are larger than one in absolute value, then it will have an infinite order of MA representation. Again, an ARMA (p, q) is said to be invertible only if all the roots of $\theta(B)$ are larger than one in absolute value. Equivalently, it will have an infinite AR order representation.

Parameter Redundancy of ARMA Model

For instance a white noise $x_t = w_t \Rightarrow x_t = 0.5x_{t-1} - 0.5w_{t-1} + w_t$ may have significant parameter estimates and we may claim the data are correlated when in fact they are not. In addition, we might fit an ARMA (1, 1) model to white noise data and find that the parameter estimates are significant. If we were unaware of parameter redundancy, we might claim the data are correlated when in fact they are not. Such problems can be avoided by writing them in exponential form using back shift operators $B^j w_{t-j}$ and avoid common terms. As the example $(1 - B)x_t = (1 - B)w_t \Rightarrow x_t = w_t$ this works for any ARMA(p, q) Model.

Causality of ARMA Model: Stationary AR models that do not depend on the future

An ARMA (p, q) model, $\phi(B)x_t = \theta(B)w_t$ is said to be causal, if the time series $\{x_t;\ t = 0, \pm 1,\ \pm 2, \ldots\}$ can be written as a one-sided linear process:

$$x_t = \sum_{j=0}^{\infty} \Psi_j\left(w_{t-j}\right) = \Psi(B)w_t \qquad (11)$$

Where $\Psi(B) = \sum_{j=0}^{\infty} \Psi_j B^j$ *and* $\sum_{j=0}^{\infty} |\Psi_j| < \infty$ and $\Psi_0 = 1$ and B^j is the back shift operator i.e. $B^j w_t = w_{t-j}$ for all $j = 0,\ 1,\ 2, \ldots$

We might wonder whether there is a stationary AR (1) process with $|\phi| > 1$. i.e.

$$x_t = \sum_{j=0}^{\infty} \phi^{-j}\left(w_{t+j}\right)$$

Such processes are called explosive (non-causal) because the values of the time series quickly become large in magnitude and if $|\phi| = 1$ of course this is not stationary. We have to check if ARMA (p, q) can be written as a one sided linear process as given in eq.(11).

Invertibility of ARMA Models: Non-uniqueness of MA Models

By mimicking from the criterion of causality for AR model, we will choose the model with an infinite AR representation for the MA model. Such a process is called an invertible process. An ARMA (p, q) model, $\phi(B)x_t = \theta(B)w_t$, is said to be invertible, if the time series $\{x_t; t = 0, \pm1, \pm2, ...\}$ can be written as a one-sided linear process:

$$w_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = \pi(B)x_t \qquad (12)$$

Where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ *and* $\sum_{j=0}^{\infty} |\pi_j| < \infty$ *absolute sumability of* π_j and $\pi_0 = 1$and B^j is the back shift operator i.e. $B^j x_t = x_{t-j}$ for all $j = 0, 1, 2, ...$

3.2.3.6. Seasonal ARIMA Models:

In this section, we introduce several modifications made to the ARIMA model to account for seasonal and non-stationary behavior. Often, the dependence on the past tends to occur most strongly at multiples of some underlying seasonal lag s. For example, with monthly economic data, there is a strong yearly component occurring at lags that are multiples of s = 12, because of the strong connections of all activity to the calendar year. Data taken quarterly will exhibit the yearly repetitive period at s = 4 quarters. Natural phenomena such as temperature also have strong components corresponding to seasons. Hence, the natural variability of many physical, biological, and economic processes tends to match with seasonal fluctuations.

It is appropriate to introduce autoregressive and moving average polynomials that identify with the seasonal lags. The resulting pure seasonal autoregressive moving average model, say, seasonal $ARMA(P,Q)_s$ then takes the form

$$\Phi_P(B^s)x_t = \Theta_Q(B^s)w_t$$

With the following definition,

The operators

$$\Phi_P(B^s) = 1 - \Phi_1(B^s) - \Phi_2(B^{2s}) - \cdots - \Phi_P(B^{Ps}) \quad (13)$$

And

$$\Theta_Q(B^s) = 1 - \Theta_1 B^s - \Theta_2 B^{2s} - \cdots - \Theta_Q B^{Qs} \quad (14)$$

are the seasonal autoregressive operator and the seasonal moving average operator of orders P and Q, respectively, with seasonal period s.

Analogous to the properties of non-seasonal ARMA models, the pure seasonal $ARMA(P,Q)_s$ is causal only when the roots of $\Phi_P(B^s)$ lie outside the unit circle, and it is invertible only when the roots of $\Theta_Q(B^S)$ lie outside the unit circle.

3.2.3.7. Building ARIMA Models

There are a few basic steps to fitting ARIMA models to time series data. These steps involve plotting the data; possibly transform the data, identifying the dependence order of the model, parameter estimation, diagnostics and model choice. The Box-Jenkins approach allows one to decide whether to go back to the identification stage or not, according to the fitting level that the model presents. The original Box-Jenkins modeling procedure involves an iterative three-stage process of model selection, parameter estimation and diagnostic checking. But, further explanations of the process by Makridakis et al. (1998) often add a preliminary stage of data preparation and a final stage of model application (or forecasting including calculation and evaluation of the forecast).

i. Data preparation involves transformations and differencing. Transformations of a data, (for instance a Box-Cox class of power transformations to stabilize the variance when the variance of the data grows with time, or logarithms for a process that involve as a fairly small and stable percent-change). The data are differenced until there are no obvious patterns such as trend or seasonality left in the data.

ii. Model selection in the Box-Jenkins framework uses various graphs based on the transformed and differenced data to try to identify potential ARIMA processes which might provide a good fit to the data. Later developments have led to other additional model selection tools such as Akaike's Information Criterion.

iii. Parameter estimation means finding the values of the model coefficients, which provide the best fit to the data. There are sophisticated computational algorithms designed to do this.

iv. Model checking involves testing the assumptions of the model to identify any area where the model is inadequate. If we find the model to be inadequate, it is necessary to go back to Step (ii) and try to identify a better model.

Forecasting is what the whole procedure is designed to accomplish once the model has been selected, estimated and checked.

3.2.3.8. Model Identification

Graphic Inspection:

Regardless of which technique is used, the first step in any time series analysis is to construct a time plot of the data, and inspect the graph for any anomalies. A number of qualitative aspects are noticeable as you visually inspect the graph. If, for example, the variability in the data grows with time, it will be necessary to transform the data to stabilize the variance. A time plot of the data will typically suggest whether any differencing is needed. If differencing is called for, then difference the data once, $d = 1$, and inspect the time plot of ∇x_t. If additional differencing is necessary, then try differencing again and inspect a time plot of $\nabla^2 x_t$.

Parameter Estimation:

When preliminary values of d have been settled, the next step is to look at the sample ACF and PACF of $\nabla^d x_t$ for whatever values of d have been chosen. Once the degree of differencing has been determined, we precede to select the autoregressive and moving-average orders by examining the sample autocorrelations and sample partial autocorrelations. The PACF for MA models behaves much like the ACF for AR models. Also, the PACF for AR models behaves much like the ACF for MA models. Because an invertible ARMA model has an infinite AR representation, the PACF will not cut off. We may summarize these results in Table 1. In addition, Table 2 summarizes for the seasonal part.

Table 1. Behavior of the ACF and PACF for ARMA Models

	AR(p)	**MA(q)**	**ARMA(p, q)**
ACF	Tails off	Cuts of after lag q	Tails off
PACF	Cuts of after lag p	Tails off	Tails off

Table 2. Behavior of the ACF and PACF for Pure Seasonal ARMA Models

	$AR(P)_S$	$MA(Q)_S$	$ARMA(P,Q)_S$
ACF	Tails off at lags ks, k = 1, 2, ...	Cuts of after lag Qs	Tails off at lags ks
PACF	Cuts of after lag Ps	Tails off at lags ks k = 1, 2, ...	Tails off at lags ks

*The values at non-seasonal lags h $\neq$ ks, for $k = 1, 2, ...$ are zero

Using Table 1 and 2 as a guide, preliminary values of P, Q, p and q are chosen. Recall that, if $p = 0$ and $q > 0$, the ACF cuts of after lag q, and the PACF tails of. If $q = 0$ and $p > 0$, the PACF cuts of after lag p, and the ACF tails of. If $p > 0$ and $q > 0$, both the ACF and PACF will tail of. The same applies to the seasonal orders at the seasonal values. Because we are dealing with estimates, it will not always be clear whether the sample ACF or PACF is tailing-of or cutting-of. Also, two models that are seemingly different can actually be very similar. With this in mind, we should not worry about being so precise at this stage of the model fitting. At this stage, a few preliminary values of p, d, q, P, D, and Q should be at hand, and we can start estimating the parameters.

After choosing the most appropriate model, the model parameters are estimated by using several estimation procedures .In general, nonlinear estimation method is used to estimate the above identified ARIMA model parameters to maximize the likelihood function of the observed series given the parameter values.

3.2.3.9. Model Selection Criterion

There are different techniques that can be used to test models against one another using the F test given by $F_{q-r,n-q} = \frac{MSR}{MSE}$ and ANOVA table. These tests have been used in the past in a stepwise manner, where variable are added or deleted when the values from the F-test either exceeds or fail to exceed some predetermined levels. An alternative is to focus on a procedure for model selection that does not proceed sequentially, but simply evaluates each model on its own merits. Suppose we consider a normal regression model with k coefficients and denote the maximum likelihood estimator for the variance as

$$\hat{\delta}_k^2 = \frac{SSE_k}{n} \qquad (15)$$

Where SSE_k denotes the residuals sum of squares under the model with k regression coefficients. Then, Akaike (1969, 1973, 1974) suggested measuring the goodness of fit for this particular model by balancing the error of the fit against the number of parameters in the model.

Akaike's Information Criterion

We introduce Akaike's Information Criterion (AIC)

$$\text{AIC} = -2\log L_k + 2k \tag{16}$$

Where L_k is the maximized log-likelihood; and k is the number of parameters in the model. For the normal regression problem, AIC can be reduced to the form given by (21). AIC is an estimate of the Kullback-Leibler discrepancy between a true model and a candidate model;

$$\text{AIC} = \text{Log}\hat{\delta}_k^2 + \frac{n+2k}{n} \tag{17}$$

$\hat{\delta}_k^2$ is given by (19), k is the number of parameters in the model and n is the sample size. (Shumway and Stoffer 2006 and Tsay 2005)

The value of k yielding the minimum AIC specifies the best model. The idea is roughly that minimizing $\hat{\delta}_k^2$ 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. The choice for penalty term given by (21) is not the only one, and a considerable literature is available advocating different penalty terms. 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.

Bias corrected (AICc)

AIC bias corrected is defined as

$$\mathrm{AICc} = \mathrm{Log}\hat{\delta}_k^2 + \frac{n+k}{n-k-2} \tag{18}$$

Where $\hat{\delta}_k^2$ is given by (19), k is the number of parameters in the model and n is the sample size..

We may also derive a correction term based on Bayesian arguments, as in Schwarz (1978), which leads to an alternative measure known as Bayesian Information Criteria (BIC).

Bayesian Information Criterion

BIC is defined as

$$\mathrm{BIC} = \mathrm{Log}\hat{\delta}_k^2 + \frac{k \log n}{n} \tag{19}$$

Using the same notation as in the above definition

BIC is also called the Schwarz Information Criterion (SIC). 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; see McQuarrie and Tsai (1998) for detailed comparisons.

3.2.3.10. Diagnostic Test

The diagnostic checking of model adequacy is the last stage of model building. The tests related to independence, normality and homoscedasticity should be performed at this stage as the residuals (e_t) from an ARIMA model are assumed to be independent, homoscedastic, and normally distributed. Several diagnostic statistics and plots of the residuals can be used to examine the goodness of fit of the tentative model to the historical data. if the model is inadequate, 3-step model building process is typically repeated several time until a satisfactory model is finally obtained. The final selected model can then be used for prediction purposes (Wei, 1990).

3.2.3.10.1. Testing for Independence:

Testing for independence (randomness) against serial dependence is a fundamental problem in time series analysis. To determine whether a time series, (e_t), is independent, the function (ACF) of the series is examined. If the ACF is significantly different from zero, this implies that there is dependence between observations. Therefore, ACF is a powerful complementary tool for testing independence (Janacek and Swift, 1993; Ferguson et al., 2000). The residual autocorrelation function (RACF) should be obtained to determine whether the residuals are white noise. There are different applications related to the Residual ACF for the independence of residuals. The first one is the correlogram drawn by plotting $\rho_e(h)$ against lag h.

$$\rho_e(h) = \frac{\gamma(t+h,t)}{\sqrt{\gamma(t+h,t+h)\gamma(t,t)}} = \frac{\gamma(h)}{\gamma(0)} \qquad (20)$$

Under the assumption that,(e_t), follows a white noise process the standard errors of these $\rho_e(h)$ are approximately equal to $\frac{1}{\sqrt{n}}$. Thus, under the null hypothesis that, (e_t) follows a white noise process, roughly 95% of the ($\rho_e(h)$) should fall within the range of $\pm\frac{1.96}{\sqrt{n}}$. If more than 5% of the $\{\rho_e(h)\}$ fall outside of this range, then, (e_t) most likely does not follow a white noise process (Lehmann and Rode, 2001).

There are many statistical tests used for diagnostic checking of randomness. In this study, the Ljung-box Q-statistic is used as alternative approaches for the diagnostic checking of residuals for independence.

Ljung-Box Q-test Statistic: These residuals from a model fit will not quite have the properties of a white noise sequence and the variance of $\hat{\rho}_e(h)$ can be much less than $\frac{1}{\sqrt{n}}$. For example, it may be the case that, individually, each $\hat{\rho}_e(h)$ is small in magnitude, say each one is just slightly less that $\frac{1.96}{\sqrt{n}}$ in magnitude, but, collectively, the values are large (Box and Pierce (1970) and McLeod (1978). The Q(H) statistic is calculated by the following equation (Ljung and Box, 1978):

$$Q = n(n-1)\sum_{h=1}^{H} \frac{\hat{\rho}_e^2(h)}{n-h} \qquad (21)$$

The value H in eq. 21 is chosen somewhat arbitrarily, typically, H = 20.

The null hypothesis for this test is H_0: the model is adequate. Under the null hypothesis of model adequacy, asymptotically for large n, Q follows a chi square distribution with H-p-q degrees of freedom, i.e. $Q \sim \chi^2_{H-p-q}$. Thus, we would reject the null hypothesis at level α if the value of Q exceeds the (1 − α) quantile of the χ^2_{H-p-q} distribution. The basic idea is that if w_t is white noise, then $n\hat{\rho}_w^2(h)$ for $h = 1, \ldots, H$, are asymptotically independent χ_1^2 random variables. This means that $n\sum_{h=1}^{H} \hat{\rho}_w^2(h)$ is approximately a χ^2_H random variable. Because the test involves the ACF of residuals from a model fit, there is a loss of p + q degrees of freedom; the other values in the above equation are used to adjust the statistic to better match the asymptotic chi-squared distribution.

3.2.3.10.2. Testing for Normality

Investigation of marginal normality can be accomplished visually by looking at a histogram of the residuals. The empirical distribution of the data (the histogram) should be bell-shaped and resemble the normal distribution. This might be difficult to see if the sample is small. In this case we might proceed by regressing the data against the quantiles of a normal distribution with the same mean and variance as the residual. Lack of fit to the regression line suggests a departure from normality.

In addition to this, a normal probability plot or a Quantile-Quantile plot can help in identifying departures from normality. Q-Q plot of the standardized residuals against the standard normal distribution is a graphical tool for assessing normality. We have used Jarque-Bera test (eq. 30) and Shapiro test (eq. 33) that we can apply to the residuals.

Jarque-Bera Test: The Jarque–Bera test is a goodness-of-fit test of whether sample data have the skewness and Kurtosis matching a normal distribution. The test is named after Carlos Jarque and Anil K. Bera. The test statistic *JB* is defined as

$$\mathrm{JB} = \frac{\mathrm{n}}{6}\{\mathrm{s}^2 + \frac{1}{4}[\mathrm{K} - 3]^2\} \tag{22}$$

Where n is the number of observations; S is the sample skewness, and K is the sample kurtosis:

$$\mathrm{s} = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(x_i - \bar{\mathrm{x}})^3}{(\frac{1}{\mathrm{n}}\sum_{\mathrm{i}=1}^{\mathrm{n}}(x_i - \bar{\mathrm{x}})^2)^{\frac{3}{2}}} \tag{23}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\}^2} \tag{24}$$

Where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of third and fourth central moments, respectively, $\bar{x}$ is the sample mean, and $\hat{\sigma}^2$ is the estimate of the second central moment, the variance.

If the data come from a normal distribution, the JB statistic asymptotically has a chi-square distribution with two degree of freedom, so we can use the statistic to test the hypothesis that the data are from a normal distribution. The null hypothesis is a joint hypothesis of the skewness being zero and an expected excess kurtosis of 0 (which is the same as a kurtosis of 3). As the definition of JB shows, any deviation from this increases the JB statistic (Jarque and Bera,. (1981)).

For small samples, the chi-squared approximation is overly sensitive, often rejecting the null hypothesis when it is in fact true. Furthermore, the distribution of p-values departs from a uniform distribution and becomes a right-skewed unimodal distribution, especially for small p-values. This leads to a large Type I error rate.

Shapiro-Wilk Test: The Shapiro-Wilk test tests the null hypothesis that a sample $x_1, \dots, x_n$ came from a normally distributed population. The test statistic is:

$$W = \frac{\sum_{i=1}^{n}(a_i x_{(i)})^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{25}$$

Where $x_{(i)}$ is the i^{th} order statistic and $\bar{x}$ is the sample mean; the constants a_i are given by

$$(a_1, \dots, a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{\frac{1}{2}}} \tag{26}$$

$$m = (m_1, \dots, m_n)' \tag{27}$$

and $m_1, \ldots, m_n$ are the expected values of the order statistics of independent and identically distributed random variables sampled from the standard normal distribution, and V is the covariance matrix of those order statistics (Shapiro and Wilk, (1965).

Interpretation of the test: If the p-value is less than the chosen alpha level, reject the null hypothesis and conclude that data are not from a normally distributed population. If the p-value is greater than the chosen alpha level, then we do not reject the null hypothesis that the data came from a normally distributed population.

3.2.3.10.3. Testing for ARCH Effects:

To recognize the presence of conditional heteroscedasticity is the same as to indentify whether an ARCH process appears in the innovation term sequence. The squared residual series are conducted to test the conditional heteroscedasticity which is known as ARCH effect (Tsay, 2005).

According to Tsay (2005), there are two available methods to test for the ARCH effects. The first is Ljung-Box Q-test statistic proposed by McLeod and Li (1983) and the second is Lagrange multiplier test.

Lagrange Multiplier Test: LM test, suggested by Engle (1982), is used to test significance of serial correlation in the squared residuals for the first p lags. This particular heteroscedasticity specification was motivated by the observation that in many financial time series, the magnitude of the residuals appeared to be related to the magnitude of recent residuals. ARCH in itself does not invalidate standard least square inference. Nevertheless, ignoring ARCH effects may result in loss of efficiency.

In testing the significance of serial correlation first select the time series model of interest for x_t and perform the regression and calculate the residuals; e_t. Then perform the second regression, of the following form, for a selected value of p.

$$e_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i e_{t-i}^2 + v_t\,, \; for \;\; t = q+1, q+2, \ldots, T \qquad (28)$$

The null hypothesis is

H_0: In the absence of ARCH components, we have $\alpha_i = 0$, for all $i = 1, 2, \ldots, q$

Against the alternative hypothesis.

H_1: In the presence of ARCH components, at least one of the estimated coefficients must be significant

Under the null hypothesis the test statistic,

$$LM = TR^2 \qquad (29)$$

follows the Chi-square distribution with q degree of freedom. Where T is the number of observations and R^2 is R-squared obtained from the regression of $e_t^2 \; on \; e_{t-i}^2, i = 1, 2, \ldots, q$.

Ljung-Box Test for the squared residuals: The Ljung-Box Q-statistic follows chi-square distribution with H degrees of freedom if the squared residuals are uncorrelated. The value H is chosen somewhat arbitrarily, typically, H=20. It is also recommended to consider up to n/4 values of H. The null hypothesis of Q (H) test is that the first H lags of ACF of the squared residuals series are equal to zero, meaning, there are no ARCH or GARCH errors. Rejecting the null thus means that there exists such errors in the conditional variance.

3.2.3.11. Forecasting

In forecasting, the goal is to predict future values of a time series, $x_{n+m}, m = 1, 2, ...,$ based on the data collected to the present, $\boldsymbol{x} = \{x_n, x_{n-1}, ..., x_1\}$. Throughout the forecasting section, we will assume x_t is stationary and the model parameters are known. The assumption includes specific values for all the parameters are known. Although this is never true in practice, the use of estimated parameters for large sample sizes does not seriously affect the results.

The minimum mean square error predictor of x_{n+m} is

$$x_{n+m}^{n} = E(x_{m+n}|\boldsymbol{x}) \tag{30}$$

Where for example $x^{n}{}_{n+1}$ denotes the forecast at the next period $n + 1$ based on the observed data, $x_1, x_2, ..., x_n$

The conditional expectation minimizes the mean square error

$$MSE = E[x_{m+n} - g(\boldsymbol{x})]^2 \tag{31}$$

Where $g(x)$ is a function of the observations $\boldsymbol{x}$;

For ARMA models in general, the prediction equations will not be as simple as the pure AR case. R-package uses the recursive solution due to Levinson (1947) and Durbin (1960). This is because, for large n, the use of the forecasting formula is prohibitive because it requires the inversion of large matrix. Using the Durbin-Levinson Algorithm forecasting equations can be solved iteratively. The other method is the innovations Algorithm in which the one-step-ahead predictors, x_{t+1}^{t} , and their mean-square errors,p_{t+1}^{t} , can be calculated iteratively (Shumway and Stoffer, 2006).

Prediction Limits: As in all statistical endeavors, in addition to forecasting or predicting the unknown x_{n+m} we would like to assess the precision of our predictions. To assess the precision of the forecasts, typically calculate prediction intervals along with the forecasts. In general, $(1-\alpha)$ prediction intervals are of the form

$$x_{n+m}^{n} \mp c_{\frac{\alpha}{2}}\sqrt{P_{n+m}^{n}} \tag{32}$$

where $c_{\frac{\alpha}{2}}$ is chosen to get the desired degree of confidence. For example, if the process is Gaussian, then choosing $c_{\frac{\alpha}{2}} = 1.96$ will yield an approximate 95% prediction interval for x_{n+m}. If we are interested in establishing prediction intervals over more than one time-period, then $c_{\frac{\alpha}{2}}$ should be adjusted appropriately, for example, by using Bonferroni's inequality (Johnson and Wichern, 1992).

Forecasting evaluation and Accuracy Criteria

Evaluating the performance of different forecasting models plays a very important role in choosing the most accurate models. In particular, the researchers and investors need to decide the evaluating criteria on which to base. Although the vast number of papers has studied the construction of modeling and forecasting volatility, a few of them focus on the volatility forecasting evaluation.

Previous empirical studies have employed various error statistics to compare the forecast performance of GARCH type models. The most widely used evaluation measures used are Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), and Theil's U statistics. Theil's U-statistic is presented in both of its specifications, these being labeled UI and UII respectively.

Denoting a series of interest as x_t and a forecast of it as f_t the resulting forecast error is given as $e_t = x_t - f_t$ for $t = 1,...,n$. Using this notation, the (fairly standard) set of forecast evaluation statistics considered can be presented as below:

Root mean Square Error

$$RMSE = \sqrt{\frac{\sum(x_t - f_t)^2}{n}} = \sqrt{\frac{\sum_{t=1}^{n} e_t^2}{n}} \tag{33}$$

Mean Absolute Error

$$MAE = \frac{\sum|x_t - f_t|}{n} = \frac{\sum_{t=1}^{n} |e_t|}{n} \tag{34}$$

Mean Absolute Percentage Error

$$\mathrm{MAPE} = \frac{\sum \frac{|x_t - f_t|}{x_t} * 100\%}{\mathrm{n}} \tag{35}$$

Theil's Inequality Coefficient UI

$$UI = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n}(x_t - f_t)^2}}{\sqrt{\frac{1}{n}\sum {f_t}^2} + \sqrt{\frac{1}{n}\sum {x_t}^2}} \tag{36}$$

Theil's Inequality Coefficient UII

$$UII = \frac{\sqrt{\frac{1}{n}\sum_{t=1}^{n-1}(\frac{f_t - x_t}{x_t})^2}}{\sqrt{\frac{1}{n}\sum_{t=1}^{n-1}(\frac{x_{t+1} - x_t}{x_t})^2}} \tag{37}$$

where n is the number of forecast errors.

The first two forecast error statistics depend on the scale of the dependent variable. These should be used as relative measures to compare forecasts for the same series across different models; the smaller the error, the better the forecasting ability of that model according to that criterion. The remaining three statistics are scale invariant.

Both Theil's UI and UII statistic are a relative accuracy measures that compare the forecasted results with a naïve forecast. They also square the deviations to give more weight to large errors and to exaggerate errors, which can help eliminate methods with large errors.

The Theil's UI inequality coefficient always lies between zero and one. However, the Theil's UII inequality coefficient lies always above zero. In both cases zero indicates a perfect fit. Theil's UII has easer interpretation as compared to UI.

Table.3. Interpreting Theil's UII

If Theil's UII statistic is	Interpretation
$UII < 1$	The forecasting technique is better than guessing.
$UII = 1$	The forecasting technique is almost as good as guessing.
$UII > 1$	The forecasting technique is worse than guessing.

The mean square forecast error can be decomposed as:

$$\frac{1}{n}\Sigma(f_t - x_t)^2 = ((\Sigma\frac{f_t}{n}) - \bar{x})^2 + (s_f - s_x)^2 + 2(1-r)s_f s_x \qquad (38)$$

Where $\Sigma\frac{f_t}{n}$, $\bar{x}$, s_f, s_x are the means and standard deviations of f_t and x, and r is the correlation between f and x. the proportions are defined as:

Bias Proportion $$\frac{((\Sigma\frac{f_t}{n}) - \bar{x})^2}{\frac{1}{n}\Sigma(f_t - x_t)^2} \qquad (39)$$

Variance Proportion $$\frac{(s_f - s_x)^2}{\frac{1}{n}\Sigma(f_t - x_t)^2} \quad (40)$$

Covariance Proportion $$\frac{2(1-r)s_f s_x}{\frac{1}{n}\Sigma(f_t - x_t)^2} \quad (41)$$

Where x_t and f_t denote the actual and forecasted value in period t, respectively.

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

Note that the bias, variance, and covariance proportions add up to one (Eviews Manual II).

In practice, when comparing the different models, it is rarely the case that one model dominate the other with respect to all evaluation measures. The common way to solve the problem is to carry out the average figures of some statistical measures and then compare the forecast models based on the parameter obtained.

3.2.4. Introduction to ARCH and GARCH Models

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); were developed to model changes in volatility. These models were later extended to generalized ARCH, or GARCH models by Bollerslev (1986). In GARCH models, the moments of a time series are considered as variant (i.e., the error term: real value minus forecasted value does not have zero mean and constant variance as with an ARIMA process). Here the error term is assumed to be serially correlated and can be modeled by an Auto Regressive (AR) process. Thus, a GARCH process can measure the implied volatility of a time series due to price spikes. In traditional ARMA estimation, the basic assumptions on the error terms are white-noise which include zero mean and constant variance, or specifically

$$i. E(\varepsilon_t) = 0, \quad ii. var(\varepsilon_t) = \sigma^2 \quad and \quad iii. E(\varepsilon_t \varepsilon_s) = 0 \quad for\ all\ t \neq s \qquad (42)$$

In particular, the homoscedastic assumption (ii) of constant variance does not necessarily need to hold. The class of models where the constant variance assumption does not hold is named heteroscedastic. If x_t is the price of cereals at time t, then the return or relative gain,$r_{t,}$ of the price of cereals at time t is

$$r_t = \frac{x_t - x_{t-1}}{x_{t-1}} => x_t = (1 + r_t)x_{t-1} \qquad (43)$$

Thus, based on this if the return represents a small (in magnitude,$r_t < 0.2$) percentage change then taking logarithm

$$\ln(x_t) = \ln[(1 + r_t)x_{t-1}]$$

$$=> \ln(x_t) = ln(1 + r_t) + \ln(x_{t-1})$$

$$=> \ln(x_t) - \ln(x_{t-1}) = ln(1 + r_t) \approx r_t$$

$$=> \nabla\ln(x_t) \approx r_t \qquad (44)$$

Either value $\nabla\ln(x_t)\ or\ \frac{x_t - x_{t-1}}{x_{t-1}}, where\ \nabla\ln(x_t) = \ln(x_t) - \ln(x_{t-1})$, will be called the return, and will be donated by r_t.

The basic idea behind volatility study is that the series $\{r_t\}$ is either serially uncorrelated or with minor lower order serial correlations, but it is a dependent series It is the study of r_t that is the focus of ARCH, GARCH, and other volatility models (Shumway, and Stoffer

2006). Recently there has been interest in stochastic volatility models. Typically, for financial series, the return r_t does not have a constant variance, and highly volatile periods tend to be clustered together.

To put the volatility models in proper perspective, it is informative to consider the conditional mean and variance of r_t given F_{t-1}; that is,

$$\mu_t = E(\ r_t|F_{t-1}\ ), \qquad \sigma^2_{t|t-1} = Var(r_t|F_{t-1}) = [(r_t - \mu_t)^2|F_{t-1}] \qquad (45\ )$$

where F_{t-1} denotes the information set available at $t-1$. Typically, F_{t-1}consists of all linear functions of the past returns.

The equation for μ_t in (eq. 45) should be simple, and we assume that r_t follows a simple time series model such as a stationary ARMA(*p, q*) model with some explanatory variables. In other words, we entertain the model

$$r_t = \mu_t + a_t, \qquad \mu_t = \phi_0 + \sum_{i=1}^{k} \beta_i x_{it} + \sum_{i=1}^{p} \phi_i r_{t-i} - \sum_{i=1}^{q} \theta_i a_{t-i} \qquad (46)$$

for r_t, where k, p, and q are non-negative integers, and x_{it} are explanatory variables. (Tsay, 2005). Combining Eqs. (45) and (46), we have

$$\sigma^2_{t|t-1} = Var(r_t|F_{t-1}) = \ Var(a_t|F_{t-1}) \qquad (47)$$

The conditional heteroscedastic models of this study are concerned with the evolution of $\sigma^2_{t|t-1}$. The manner under which $\sigma^2_{t|t-1}$ evolves over time distinguishes one volatility model from another.

Throughout this study, a_t is referred to as the *shock* or *Error term* at time *t* and $\sigma_{t|t-1}$ is the positive square root of $\sigma^2_{t|t-1}$. The model for μ_t in Eq. (46) is referred to as the *mean* equation for r_t and the model for $\sigma^2_{t|t-1}$ is the *volatility* equation for r_t . Therefore, modeling conditional heteroscedasticity amounts to augmenting a dynamic equation, which governs the time evolution of the conditional variance of the asset return, to a time series model. (Tsay, 2005)

3.2.4.1. ARCH Model

A more sophisticated volatility model is the Autoregressive Conditional Heteroscedasticity ARCH (q) model suggested by Engle (1982). ARCH model is the first attempt to capture the above characteristics without the assumption of constant variances which commonly exists in many conventional financial econometrics models. It is unlikely in financial time series that the error terms will be constant overtime, therefore allowing for conditional heteroscedasticity in stock return analysis is reasonable.

The conditional variance, also referred to as the *conditional volatility*, of r_t will be denoted by $\sigma^2_{t|t-1}$, with the subscript $t - 1$ signifying that the conditioning is upon returns through time $t - 1$. When r_t is available, the squared error a_t^2 provides an unbiased estimator of $\sigma^2_{t|t-1}$.

Unlike historical estimation process using sample standard deviations, the ARCH model constructs the conditional variance $\sigma^2_{t|t-1}$ of asset returns with maximum likelihood method. The ARCH (q) model proposed by Engle (1982) formulates volatility as follows:

$$a_t = \sigma_{t|t-1}\varepsilon_t \qquad (48)$$

$$\sigma^2_{t|t-1} = \alpha_0 + \sum_{i=1}^{q} \alpha_i a^2_{t-i} \qquad (49)$$

The time varying volatility is captured by allowing volatility to depend on the lagged values of the innovation terms r_t and q is chosen such that the residuals of the variance equation are white noise. That is, $\varepsilon_t \sim iid\ N(0,1)$. Here $\{\varepsilon_t\}$ is a sequence of independently and identically distributed random variables each with zero mean and unit variance (also known as the innovations). All of the coefficients in the conditional variance equation are required to be non-negative. For the case $q=1$, $\alpha_0 \; and \; \alpha_1 > 0$ is the condition to assure that $\sigma^2_{t|t-1} > 0$. (Shumway and Stoffer, 2006). The ARCH effect is exhibited by α_1 to capture the short-run persistence.

$$\sigma^2_{t|t-1} = \alpha_0 + \alpha_1 a^2_{t-1} \qquad (50)$$

In addition, ε_t is independent of $r_{t-j}, j = 1, 2, 3, \ldots$ .The innovation ε_t is presumed to have unit variance so that the conditional variance of a_t equals $\sigma^2_{t|t-1}$. This implies the conditional distribution of a_t given F_{t-1} is Gaussian (Cryer and Chan, 2008)

$$a_t|F_{t-1} \sim N(0, \alpha_0 + \alpha_1 r^2_{t-1} = \sigma^2_{t|t-1}) \qquad (51)$$

While the ARCH model resembles a regression model, the fact that the conditional variance is not directly observable (and hence is called a latent variable) introduces some subtlety in the use of ARCH models in data analysis. For example, it is not obvious how to explore the regression relationship graphically. To do so, it is pertinent to replace the conditional variance by some observable in Eq. 50. Let

$$\eta_t = a^2_t - \sigma^2_{t|t-1} \qquad (52)$$

It can be verified that $\{\eta_t\}$ is a serially uncorrelated series with zero mean. Moreover, η_t is uncorrelated with past returns. Now consider eq. 50

$\sigma^2_{t|t-1} = \alpha_0 + \alpha_1 a^2_{t-1}$ and substituting by $\sigma^2_{t|t-1} = a^2_t - \eta_t$

$=> a^2_t - \eta_t = \alpha_0 + \alpha_1 a^2_{t-1}$

$=> a^2_t = \alpha_0 + \alpha_1 a^2_{t-1} + \eta_t$ (53)

Thus, the squared return series satisfies an AR(1) model under the assumption of an ARCH(1) model for the return series. Based on this useful observation, an ARCH (1) model may be specified as an AR (1) specification for the squared returns is warranted by techniques of ARMA models. (Cryer and Chan, 2008)

Besides its value in terms of data analysis, the deduced AR (1) model for the squared returns can be exploited to gain theoretical insights on the parameterization of the ARCH model. For example, because the squared returns must be nonnegative, it makes sense to always restrict the parameters α_0 and α_1 to be nonnegative. In addition, if the return series is stationary with variance σ^2, then taking expectation on both sides of Eq. 53 as follows yields

$$\begin{aligned} E(a_t^2) &= E(\alpha_0 + \alpha_1 a_{t-1}^2 + \eta_t) \\ \sigma^2 &= \alpha_0 + \alpha_1 \sigma^2 \\ \sigma^2 &= \frac{\alpha_0}{1-\alpha_1} \end{aligned} \tag{54}$$

That is, $\sigma^2 = \frac{\alpha_0}{1-\alpha_1}$ and hence $0 \leq \alpha_1 < 1$. Indeed, it can be shown (Ling and McAleer, 2002) that the condition $0 \leq \alpha_1 < 1$ is necessary and sufficient for the (weak) stationarity of the ARCH (1) model.

Recall that weak stationarity of a process requires that the mean of the process is constant and the covariance of the process at any two epochs is finite and identical whenever the lags of the two epochs are the same. In particular, the variance is constant for a weakly stationary process. The condition $0 \leq \alpha_0 < 1$ implies that there exists an initial distribution for a_0 such that a_1 defined by Eq. 47 and eq. 50 for $t \geq 1$ is weakly stationary in the sense above. It is interesting to observe that weak stationarity does not preclude the possibility of a non-constant conditional variance process, as is the case for the ARCH(1) model. It can be checked that the ARCH (1) process is white noise. Hence, it is an example of a white noise that admits a non-constant conditional variance process as defined by Eq. (50) that varies with the lag one of the squared process.

A satisfying feature of the ARCH(1) model is that, even if the innovation η_t has a normal distribution, the stationary distribution of an ARCH(1) model with $0 \leq \alpha_1 < 1$ has fat tails; that is, its kurtosis, $E(a_t^4)/\delta^4 - 3$, is greater than zero. Recall that, the excess kurtosis of a normal distribution is always equal to 0. In addition, a distribution with positive kurtosis is said to be fat-tailed, while one with a negative kurtosis is called a light-tailed distribution.

A stationary ARCH(1) model need not have finite fourth moments. The existence of finite higher moments will further restrict the parameter range—a feature also shared by higher-order analogues of the ARCH model and its variants. In addition the kurtosis of a stationary ARCH (1) process is greater than zero. This verifies our earlier statement that an ARCH(1) process has fat tails even with normal innovations. In other words, the fat tail is a result of the volatility clustering as specified by Eq. (50).

The ARCH model is simple. However, many parameters are required to estimate the volatility of price returns. The problem of parsimony among the other problems of ARCH model such as how to specify the value of p and the violation of non-negativity constraints led to more general framework GARCH (p, q) proposed by Bollerslev (1986) and Taylor (1986).

3.2.4.2. GARCH Model

Extending the framework of Engle (1982), Bollerslev (1986) and Taylor (1986) generalized the ARCH (q) model to ***GARCH (p, q)*** in which they added the q lags of past conditional variance into the equation. GARCH (p, q) model allows for both autoregressive and moving average components in the heteroscedastic variance. Empirical findings suggest that GARCH model is more parsimonious than ARCH model. GARCH (p, q) is specified as follows:

$$\sigma_{t|t-1}^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \cdots + \alpha_q a_{t-q}^2 + \beta_1 \sigma_{t-1|t-2}^2 + \cdots + \beta_p \sigma_{t-p|t-p-1}^2$$

$$a_t = \sigma_{t|t-1}\varepsilon_t, \qquad \sigma_{t|t-1}^2 = \alpha_0 + \textstyle\sum_{i=1}^{q} \alpha_i a_{t-i}^2 + \sum_{j=1}^{p} \beta_j \, \sigma_{t-j|t-1-j}^2 \quad (55)$$

The restrictions $\alpha_0 > 0, \alpha_i \geq 0, \beta_i \geq 0$, ensure that the variance is always greater than zero and the restriction ($\sum_{i=1}^{p} \alpha_i + \sum_{j=1}^{q} \beta_j < 1$) is necessary and sufficient condition for the stability of the conditional variance equation (Cryer and Chan, 2008).

The above-specified term α_0 is generally interpreted as long-term volatility to which the system converges. On the other hand, the ARCH term $\alpha_i a_{t-i}^2$ reflects the effect of lagged shocks or surprises on the volatility at time t. Moreover, the GARCH term $\beta_i \sigma_{t|t-i}^2$ measures the effect of past-expected variance on the current volatility and thus high probability of volatility clustering. The term $\sum_{i=1}^{p} \alpha_i + \sum_{i=1}^{q} \beta_i$ indicates the speed of convergence of the forecast of variance to a steady state. The closer to one it is, the slower the convergence.

The unconditional Variance under the GARCH (1, 1) specification is given by

$$var(a_1) = \frac{\alpha_0}{1 - \alpha_1 - \beta}$$

When $\alpha_1 + \beta < 1$

- $\alpha_1 + \beta \geq 1$ is termed "non-stationary in variance
- $\alpha_1 + \beta = 1$ is termed the integrated GARCH
- For non-stationary in variance, the conditional variance forecasts will not converge on their unconditional value as the horizon increases. (Brooks, 2002)

According Tsay (2005) GARCH model can be written as an infinite order of ARCH model. A GARCH (0, 1) model is simply the first-order ARCH model. GARCH (1, 1) model is better than ARCH, because it's more parsimonious, avoids over fitting. Second is less likely to violate non-negativity constraints. Although GARCH model has been the most popular volatility model, it has three main problems. Firstly, the estimated models may violate the non-negativity constraint. Secondly, GARCH model does not take into account the leverage effect and finally it does not allow direct feedback between the conditional variance and conditional mean (Brooks, 2002).

3.2.4.3. Parameter Estimation GARCH (p, q):

Under the presence of ARCH effects, the OLS estimation is not efficient since variance of residuals is not constant and volatility models used in financial econometrics are non-linear in conditional variance. In addition, OLS involves minimizing the sum of squares of residuals, sum of squares of residuals depends on the coefficient of the mean not on conditional variance. Therefore, as many studies indicated, the commonly used method known as the maximum-likelihood estimation was employed to estimate parameters of GARCH family model. In maximum likelihood estimation the distributional assumption on residual is the core point. Thus, in this study, normal, student-t and the GED were considered to estimate parameters as financial time series data possess volatility clustering and leptokurtosis characteristics that leads to the use different distributional assumption for residuals (Bollerslev, 1986). However, appropriate distribution for the residual was identified based on in-sample forecast error statistics to check predictive ability of the model under specified error distributions and final analysis was done based on selected distribution for residuals in the mean equation.

Moreover, in ML estimation method the conditional maximum likelihood estimates for the parameters are obtained by maximizing the conditional log-likelihood function. However, maximization of the log-likelihood function of the model analytically in terms of its parameter is impossible because of nonlinearity of GARCH model. As a result, maximization of the log-likelihood function was done through numerical iteration method using statistical package fGARCH of R-software version 2.15.1, which is uniquely developed for financial time series data.

3.2.4.4. Maximum Likelihood Parameter Estimation Approach for GARCH Family Model

Building a volatility model for an asset return series consists of four steps: (Tsay, 2005)

1. Specify the mean equation by testing for serial dependence in the data and, if necessary, building an econometric model (e.g., an ARMA model) for the return series to remove any linear dependence.

2. Use the residuals of the mean equation to test for ARCH effects.

3. Specify a volatility model if ARCH effects are statistically significant and perform a joint estimation of the mean and volatility equations.

4. Check the fitted model carefully and refine it if necessary.

For most asset return series, the serial correlations are weak, if any. Thus, building a mean equation amounts to removing the sample mean from the data if the sample mean is significantly different from zero.

Prediction of Future Conditional Variance with ARCH Model

A main use of the ARCH model is to predict the future conditional variances. For example, one might be interested in forecasting the h-step-ahead conditional variance

$$\sigma^2_{t+h|t} = E(a^2_{t+h}|F_{t-1})$$

For $h = 1$, the ARCH (1) model implies that

$$\begin{aligned} \sigma^2_{t+1/t} &= E(a^2_{t+1}|F_{t-1}) \\ &= \alpha_0 + \alpha_1 a^2_t \end{aligned}$$

Substituting $\alpha_0 = \sigma^2(1-\alpha_1)$ in the above equation from eq. 54 we get

$$\sigma^2_{t+1|t} = \sigma^2(1-\alpha_1) + \alpha_1 a^2_t \quad (56)$$

which is a weighted average of the long-run variance and the current squared return. Similarly, using the iterated expectation formula, we have (Cryer and Chan, 2008):

$$\begin{aligned} \sigma^2_{t+\mathrm{h}|t} &= E\left(a^2_{t+\mathrm{h}}\middle|F_{t-1}\right) \\ &= E[E(\sigma^2_{t+h|t+h-1}\varepsilon^2_{t+h}|F_{t-1})|F_{t-1}] \\ &= E[\sigma^2_{t+h|t+h-1}E(\varepsilon^2_{t+h}|F_{t-1})|F_{t-1}] \\ &= E[\sigma^2_{t+h|t+h-1}E(\varepsilon^2_{t+h})|F_{t-1}] \\ &= E[\sigma^2_{t+h|t+h-1}|F_{t-1}] \\ &= E[\alpha_0 + \alpha_1 a^2_{t+h-1}|F_{t-1}] \\ &= \alpha_0 + \alpha_1 E[a^2_{t+h-1}|F_{t-1}] \\ &= \alpha_0 + \alpha_1 \sigma^2_{t+h-1|t} \end{aligned} \quad (57)$$

where we adopt the convention that $\sigma^2_{t+h|t} = a^2_{t+h}$ for h < 0. The formula above provides a recursive recipe for computing the h-step-ahead conditional variance.

3.2.4.5. Diagnostic test (Model Adequacy)

For a properly specified ARCH/GARCH model, the standardized residuals

$$\hat{a}_t = \frac{a_t}{\sigma_{t|t-1}}$$

form a sequence of iid random variables. Therefore, one can check the adequacy of a fitted ARCH model by examining the series $\{\hat{a}_t\}$. In particular, the Ljung–Box statistics of $\tilde{}$*at* can be used to check the adequacy of the mean equation and that of $\hat{a}_t^2$ can be used to test the validity of the volatility equation (Tsay).

The followings are the model adequacy checking method that was used for this study

A) The ACF and the PACF of the standardized residuals, $\hat{a}_t$: If the model is adequate, the ACF's of standardized residuals should be indicative of a white noise process.

B) The standardized residuals are assumed to follow identical and independently distributed standard normal distributions (Tsay, 2002 and Gourieroux, 1997). This was checked through Histogram, Jarque-Bera test, and Shapiro test.

C) Testing the presence of ARCH Effect:

The purpose of performing the ARCH test on the residuals is to test for the presence of heteroscedastic errors in the model. In the presence of heteroscedastic errors, the standard errors associated with the parameters in the model will be biased downward. In addition, Weiss (1986) concluded that ignoring the ARCH effect will inevitably result in identifying

ARMA models containing too many parameters, that is over-parameterization. Thus, the presence of ARCH effect has to be tested in series through the squared residuals, $\hat{a}_t^2$, of the series (Tsay, 2010). According to Tsay there are two available methods to test for ARCH effects. These are Lagrange Multiplier (LM) test and Ljung-Box test applied to the squared residuals.

3.2.4.6. Forecasting evaluation and Accuracy Criteria

Evaluating the performance of different forecasting models plays a very important role in choosing the most accurate models. In particular, the researchers and investors need to decide the evaluating criteria on which to base. Although the vast number of papers has studied the construction of modeling and forecasting volatility, a few of them focus on the volatility forecasting evaluation.

The most widely used evaluation measures are Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percent Error (MAPE), and Theil's U statistics. Theil's U-statistic is presented in both of its specifications, these being labeled UI and UII respectively.

3.2.4.7. Forecasting by GARCH Family Model

Conditional variance forecasts from GARCH family models are obtained with similar approach to forecasts from ARMA models by iterating with the conditional expectations operator. In other words, when the estimation of the unknown parameters is done, estimates of the standard deviation series can be calculated recursively via the definition of the Conditional variance for the GARCH (P, Q) family process which helps to examine past behavior of domestic price volatility for cereals under consideration that possesses volatility clustering.

3.2.4.8. Handling Missing Values:

Similar to other statistical analysis, missing values are problems frequently encountered in time series analysis. We need to address missing data that we come across than simply omit. We need to replace them with appropriately estimated values so that the arrangement of data between time-periods will be balanced appropriately. In order to replace those observations, there are several options available in the literature. As discussed in Liu and William (2001), missing data in a time series may be estimated using one of the following methods. First, replace with the mean of the series. Alternatively, replace with the naïve forecast. Naive model is the simplest form of a univariate forecast model. It uses the current time value for the next time, that is $\hat{x}_{t+1} = x_t$. Another method could be, replace with a simple trend forecast. This is accomplished by estimating the regression equation of the form, $x_t = \alpha + \beta t$, where t is the time for the periods prior to the missing value. Then use the equation to estimate and replace the missing observation. As an alternative method one can replace a missing value, replace with an average of the last two known observations that bound the missing observation.

CHAPTER FOUR: Data Analysis and Results

4.1. Descriptive Data Analysis

In this section we will describe cereal prices of Ethiopia focusing on monthly cereal price (Teff, Barley, Sorghum, Maize and Wheat) of Ethiopia recorded by Ethiopian Statistical Agency, CSA.

Table 4. Summary of monthly cereal price

	(1996, 09) to (2006)	(2007,01) to (2007,12)	(2008,01) to (2012, 07)	Over all
Mean	1.488303	3.612489	4.909066	2.618787
Minimum	0.38087	2.167467	2.906365	0.38087
Quartile (1)	1.003888	3.074827	4.141044	1.321283
Median	1.4607	3.606022	5.096429	1.925415
Quartile(3)	1.909618	4.06411	5.61062	3.956769
Max	2.9409	5.229783	6.931824	6.931824

Table 4 above and Figure-3 in the Appendix show the summary of monthly cereal price. From the table we can see that the mean cereal price is 2.62 Birr/kg, minimum is 0.38 Birr/kg occurred in January 2002, and maximum is 6.93 Birr/kg, which occurred in Jul-2009. The table is constructed by dividing the data into three parts, that is, before world food crisis (before January 2007), during the crisis (from Jan 2007 - Dec 2007) and after the crisis (from Jan 2008 - Jul 2012). As clearly seen form Figure 3, the mean price of cereals after the crisis is three times higher than the mean price of cereals before crisis.

Following depreciation of Birr against US Dollar, Ethiopia experienced 200% increase in cereal price, which is a record rise in prices as compared to their respective prices during the period of 2006-2007. The 2011 price is about 200% higher than the previous three years prices, which were already very high. Possible reasons for the high prices include high-level prices helping farmers to hold more grain in stock; increase in informal cross-border trade; injection of cash into the economy via the safety net payments; higher livestock prices for good quality animals.

The time plot of the data given in Figure 3 shows there is some Seasonal pattern with an upward trend i.e. increasing toward the end of the data. Therefore, the first step was differencing the data once, d =1, and inspecting the time plot of ∇x_t to see if additional differencing is necessary. Inspecting the graph of differenced series, Figure 4, does not indicate any trend. But the differenced series is not stationary as shown by the ADF test given in Table 5 below. The series, therefore, may have not only an upward trend but also seasonal behavior due to high production in the harvest season implying less price and high price in summer season because of low production. Finally inspecting the plot of the seasonal difference of the differenced data, Figure 6, indicates that the series is stationary as it will be confirmed by the ADF test (Table 5).

In addition to time plots, the sample ACF can help in indicating whether there is a need for differencing. Figure 7 exhibits the sample autocorrelation function (ACF) plot of the data. We note that the trend in the data, the slow decay in the sample ACF, $\hat{\rho}_h$, and the fact that the PACF at the first lag is nearly 1, all indicate non-stationary behavior. Following the recommended procedure, first difference was taken and then the sample ACF and PACF of the differenced $\nabla x_t = x_t - x_{t-1}$ were inspected (see Figure 8 Appendix).

The peaks at seasonal lags, $h = 1s, 2s, 3s, 4s$ where s = 12 (i.e., $h = 12,\ 24,\ 36,\ 48$) with relatively slow decay suggest a seasonal difference.

Therefore, it can be considered that the series has both trend and seasonal variation, and we need both simple and seasonal differencing to make the series stationary- a conclusion that is also consistent with the result of the time series plot.

Next, we perform some formal tests of stationarity to ascertain the visual inspection of non-stationary behavior.

4.2. Tests

4.2.1. Augmented Dickey-Fuller Unit Root Test:

For ARMA (p, q) models, the ADF test can be used assuming p is large enough to capture the essential correlation structure. We have chosen the lag order p for the ADF test, for each series, based on the AR Model that has smallest AIC (Tsay 2005). The summary of the results of the Augmented Dickey-Fuller (ADF) test is given in Table 5 below.

Table 5: Summary of ADF unit-root test

(in level, after first differencing, after Seasonal differencing and after Seasonal difference of the difference Series)

Series	ADF test statistic	1%crit.Value	5%crit. Value	10% crit. Value
Original series	-1.7572	-4.04	-3.45	-3.15
First regular differencing	-3.4306	-4.04	-3.45	-3.15
First Seasonally differenced	-2.8551	-4.04	-3.45	-3.15
Seasonal difference of the difference Series	-4.3928	-4.04	-3.45	-3.15

From Table 5 the calculated ADF test statistic for the cereal price (original), seasonal differenced cereal price and first regular differenced cereal price are -1.7572, -2.8551 and -

3.4306 respectively. All the calculated ADF test statistic values mentioned above are greater than the tabulated (critical) values of -4.04 at 1% and -3.45 at 5% significance levels. According to these results, we do not have enough evidence to reject the null hypothesis that the data have unit root and are non-stationary at 1% and 5% significance levels. Hence, we conclude that the original series, as well as the series obtained after seasonal differencing and those obtained after regular differencing are not stationary. However, the computed ADF test statistic ,-4.3928, of the series obtained after seasonal differencing of the differenced series is smaller than the tabulated values of -4.04 at 1% and -3.45 at 5% significance. This leads to the rejection of the null hypothesis that there is a unit-root problem at first seasonal difference of the differenced series. Consequently, consistent with visual inspection we conclude that, the series obtained after seasonal difference of the differenced series is stationary.

Therefore the Patterns of monthly price of cereals series plot, the correlogram and the ADF-test all suggest the need of seasonal difference of the differenced series so as to have stationary data.

These tests for stationarity seem to agree and suggest that seasonal differencing of the differenced series is necessary to achieve stationarity around a constant mean, approximately 0.00 and with standard deviation of 0.89 Birr (Figure 7). Moreover, the ACF and PACF (Figures 10) also indicate that the monthly price series is stationary in both mean and variance after seasonal differencing of the differenced series. $\nabla\nabla_{12}x_t$.

4.2.2. Tests of Randomness

According to Harvey (1993), the simplest time series is a random model, in which the observations vary around a constant mean, have a constant variance, and are statistically independent. In other words, it is just white noise, which has no serial correlation, meaning that there is no point in attempting to fit a time series model to such type of data. Therefore, it is important to perform test of randomness before any attempt to modeling process to our series. Hence, we check the time series through the following test to investigate the hypothesis that the seasonal difference of the differenced series is serially uncorrelated.

Tests of Randomness using Graphic Inspection:

The visual inspection of the autocorrelation function plot provides useful information to identify the type of time series (Chatfield, 1996). Sample ACF has a sampling distribution that allows us to assess whether the data comes from a completely random or white noise series or whether correlations are statistically significant at some lags. Under general conditions, if x_t is white noise, then for large n, the sample ACF, $\hat{\rho}_x(\mathrm{h})$ for $\mathrm{h} = 1,2,\dots,\mathrm{H}$, where H is fixed but arbitrary, is approximately normally distributed with zero mean and standard deviation given by $\hat{\sigma}_\rho(h) = \frac{1}{\sqrt{n}}$.[1]

Based on the previous result, we obtain a rough method of assessing whether peaks in $\hat{\rho}_x(\mathrm{h})$ are significant by determining whether the observed peak is outside the interval $\frac{2}{\sqrt{\mathrm{n}}}$ (or minus/plus two standard errors $\mp 2\hat{\sigma}_\rho(h)$); for a white noise sequence, approximately 95% of the sample ACFs should be within these limits.

---

[1] See property 1.1. Large sample distribution of ACF for white noise p 29 [Shum-way and Stoffer]

Figure 10 exhibits the graph of sample ACF from which we can observe visually that the ACFs, for instance, at lag 1 and lag 12 are significantly different from zero. i.e. $\rho(1) = -0.48,\ and\ \rho(12) = -0.49$ are greater in absolute value than $\frac{2}{\sqrt{n}} = 0.15$.Where n = 180 with p-values <0.001. This indicates that there is some sort of dependence between values of $\nabla\nabla_{12}x_t$ series.

4.3. Building ARIMA Model for Monthly Cereal Price

Fitting a model to time series data involves plotting the data, transforming the data when appropriate, identifying the dependence orders of the model, parameter estimation, diagnostics tests, and model choice. In this section, a univariate Seasonal ARIMA methodology is used to model monthly cereal price in Ethiopia.

4.3.1. Model Identification

Once the degree of differencing has been determined, we precede to select the autoregressive and moving-average orders by examining the sample autocorrelations and sample partial autocorrelations. To use the sample autocorrelation and sample partial autocorrelation functions for tentative model parameters identification, we consider the ACF and PACF shown in Figure-10 of the appendix. Using Table-1 and Table-2 as a guide, we choose preliminary values of p, q, P and Q. Because since we are dealing with estimates, it will not always be clear whether the sample ACF or PACF is tailing off or cutting off. In addition to this, two models that are seemingly different, they can actually be very similar (Shumway and Stoffer, 2010). With this in mind, we should not worry about being so precise at this stage of the model fitting. At this stage, a few preliminary values of p, q, P and Q should be at hand, and we can start estimating the parameters.

First, concentrating on the seasonal lags, the characteristics of the ACF and PACF of our data series in Figure-10 tend to show a strong peak at h = 12 in the autocorrelation function, combined with peaks at h = 12, 24 in the partial autocorrelation function. Hence, it appears that either:

(i) the ACF is cutting off after lag 1s and the PACF is tailing off in the seasonal lags,

(ii) the PACF is cutting off after lag 2s and the ACF is tailing off in the seasonal lags,

(iii) The ACF and PACF are both tailing off in the seasonal lags.

Table 2 suggests either (i) a seasonal moving average of order Q = 1, or (ii) a seasonal autoregressive of order P = 2 or (iii) due to the fact that both the ACF and PACF may be tailing off at the seasonal lags, perhaps both components, P = 1 and Q = 1, are needed.

To identify the between-season model, we focus on the lags $h = 1,2, \ldots,$ and identify the order based on Table-1.

First, we set the ACF to be tailing-off and the PACF to cut-off after lag 2, we identify p=2 and q=0. In addition, it is possible to think of the PACF to be tailing-off and the ACF to cut-off after lag 1, leading to identify p=0 and q=1.

Fitting the following models suggested by these observations, we obtain:
SARIMA (1, 1, 2) × $(0, 1, 1)_{12}$
SARIMA (0, 1, 1) × $(0, 1, 1)_{12}$
SARIMA (0, 1, 1) × $(2, 1, 0)_{12}$
SARIMA (2, 1, 0) × $(2, 1, 0)_{12}$
SARIMA (2, 1, 0) × $(0, 1, 1)_{12}$

In this section, we assume that we have n observations($x_1,\ldots,x_n$) from a causal and invertible Gaussian SARMA (p, q)*$(P, Q)_{12}$ process in which initial order of parameters, p, q, P and Q are known. Our goal is to estimate the value of the parameters: $\theta_1, \dots \theta_q, \phi_1, \dots \phi_p, \Theta_1, \dots \Theta_Q, \Phi_1, \dots, \Phi_P$. Specifically, we estimate $\theta_1, \theta_2, \phi_1, \phi_2, \Theta_1, \Phi_1$ *and* Φ_2.

4.3.2. Parameter Estimation of ARIMA Models

Estimating the parameters for Box-Jenkins models follows a non-linear estimation and parameter estimates are usually obtained by maximum likelihood method, which is asymptotically correct for any time series (Brockwell and Davis, 1996). Hence, we use maximum likelihood estimation method for monthly mean price of cereal, to estimate the parameters. The results are summarized in Table- 6 below.

Table 6. Parameter Estimates for Suggested SARIMA Models

Model	Parameter	Estimate	Std. error	t-value	p-value	Criteria
i.	ϕ_1	-0.8937	0.1261	-7.087232	<0.001	AIC = 313.67
	θ_1					AICc = 314.2
	θ_2	-0.5337	0.0853	-6.256741	<0.001	BIC = 332.38
	Θ_1					$\hat{\sigma}_w^2 = 0.321$
	Θ_2	-0.0258	0.1079	-0.23911	0.811*	
	SARIMA(1, 1, 2)(0, 1, 2)*					
ii.	ϕ_1	-0.8897	0.1247	-7.134723	<0.001	AIC = 311.73
	θ_1					AICc = 312.1
	θ_2	-0.5273	0.0807	-6.534077	<0.001	BIC = 327.32
	Θ_1	-0.0807	0.0903	-9.459579	<0.001	$\hat{\sigma}_w^2 = 0.323$
	SARIMA (1, 1, 2)*(0, 1, 1)					
iii.	θ_1	-0.5385	0.0709	-7.595205	<0.001	AIC = **309.51**
	Θ_1	-0.8634	0.0891	-9.690236	<0.001	AICc = **309.65**
						BIC = **318.86**
	SRIMA(0, 1, 1)*(0, 1, 1)					$\hat{\sigma}_w^2 = 0.326$
iv.	θ_1	-0.5271	0.0709	-7.434415	<0.001	AIC = 318.68
	ϕ_1					AICc = 318.93
	ϕ_2	-0.4347	0.0746	-5.827078	<0.001	BIC = 331.15
	SARIMA(0,1,1)*(2,1,0)					$\hat{\sigma}_w^2 = 0.358$
v.	ϕ_1					AIC = 318.68

	ϕ_2	-0.2050	0.0758	-2.704485	0.0068	AICc = 318.93
	Φ_1					BIC = 331.15
	Φ_2	-0.4494	0.741	-6.061777	<0.001	$\hat{\sigma}_w^2 = 0.364$
	SARIMA(2, 1, 0)*(2, 1, 0)					
vi.	ϕ_1					AIC = 315.77
	ϕ_2	-0.1892	0.0758	-2.496042	0.01256	AICc = 316.02
	Θ_1	-0.8831	0.0971	-9.094784	<0.001	BIC = 328.25
	SARIMA(2, 1, 0)*(0, 1, 1)					$\hat{\sigma}_w^2 = 0.332$

*Stands for the parameter that is not significantly different from zero

Where

(i) stands for $SARIMA\ (1,1,2) * (0,1,2)_{12}$

(ii) stands for $SARIMA\ (1,1,2) * (0,1,1)_{12}$

(iii) stands for $SARIMA\ (0,1,1) * (0,1,1)_{12}$

(iv) stands for $SARIMA\ (0,1,1) * (2,1,0)_{12}$

(v) stands for $SARIMA\ (2,1,0) * (2,1,0)_{12}$

(vi) stands for $SARIMA\ (2,1,0) * (0,1,1)_{12}$

(vii) stands for $SARIMA\ (0,1,1) * (0,1,2)_{12}$

(viii) stands for $SARIMA\ (3,1,1) * (0,1,2)_{12}$

Table-6 above displays the list of the parameters for each temporally entertained model. For each model parameter, the table presents the estimated value, standard error, *t*- value, p-value, AIC, AICc, BIC and variance ($\hat{\sigma}_w^2$) for the estimate. As indicated by McDowall et al., (1980), parameters must differ significantly from zero and all significant parameters must be included in the model. The T-ratios (t_{cal}) related to the parameter estimate were compared with the critical value of ± 1.96 obtained from the t-distribution at the 0.05 significance level. The t-test for the parameters checks whether some lagged terms in the model that we may omit or not.

4.3.3. Model Selection

The AIC, AICc, BIC and variance of the estimate ($\hat{\sigma}_w^2$) in Table-6 are computed according to Eqs. (20, 21, 22, and 23 respectively) and are used to compare selected models fit best to the monthly cereal price series. The model with the smallest information criteria is said to fit the data better. Since model (iii) has lowest AIC=**309.51**, AICc=**309.65**, BIC=**318.86** and variance of estimate, $\hat{\sigma}_k^2 = 0.326$, compared to other models, it fits the monthly cereal price series best, and can be further analyzed. Therefore the model to be used is

$$(1-B)(1-B^{12})x_t = [1+0.5385B][1+0.8634B^{12}]\widehat{w}_t, \qquad \hat{\sigma}_w^2 = 0.3263$$

where standard errors of the two MA parameters are 0.0709 and 0.08934, respectively.

The estimated values of the parameters for Seasonal ARIMA(0,1,1)*(0,1,1) are $\hat{\theta}_1 = -0.5385 \neq 0$ *and* $\widehat{\Theta_1} = -0.8634 \neq 0$. However, the results being different from zero does not guarantee that both parameters are different from zero. Hence, we need to test the extent of co-linearity that might have influenced the result.

4.3.4. Test for Co-linearity

The correlations of the parameter estimates are shown in Table-7 below. We can use this to assess the extent to which co-linearity may have influenced the results. If two parameter estimates are highly correlated, you might consider dropping one of them from the model. However, in our case correlation between seasonal moving average and non-seasonal moving average parameter is, -0.140113, which is small. Therefore, we drop no parameter from the model. Another problem is that, even though the parameters are significant there could be a concern that we have not identified the model correctly. i.e. we need to test for model adequacy.

Table 7: Correlations of Parameter Estimates for the fitted model

parameter	$\hat{\theta}_1$	$\hat{\Theta}_1$
$\hat{\theta}_1$	1	**-0.140113**
$\hat{\Theta}_1$	-0.140113	**1**

4.3.5. Test of Model Adequacy

After we have built a model, we should allow for additional parameters in the fitted model and test whether the added parameters are significantly different from zero. If they are different from zero, then there is a reason for concern that we have not identified the model correctly. For example, we start with over fitting by including one more seasonal moving average parameter (Θ_2) to the Seasonal ARIMA model (iii) to examine whether this model with more parameters would adequately be fit to the monthly cereal price data. The inclusion of this parameter can be determined by testing its significance and the improvement in the measures of goodness of fit of the model. This is given in Table 22 appendix in which we have added one more parameter deliberately. The estimated parameter, $\Theta_2 = 0.0731(0.9326)$ is found to be not statistically significant. Where the value we see in parenthesis is the P-value of the parameter. We conclude that, there is no reason for concern that we have not identified the model correctly.

This is also further verified by plotting theoretical graph of autocorrelation function and partial autocorrelation function. The plots are presented in Figure 10 and Figure 11 (appendix). We found that the two graphs to be close to each other implying our model is good- a conclusion that is consistent with model adequacy, co-linearity and model selection criterion. The last step is to diagnose the residuals.

4.4. Diagnostic Checking for the selected ARIMA Model:

In this section, we will assess how well the selected model, $(\text{SARIMA}(0,1,1) * (0,1,1)_{12})$, fit the actual cereal price data. If the model fits the data well, the residuals of the fitted model are random (Chatfield, 1996).

4.4.1. Testing for Independence:

In this section, we will assess dependence of the residuals of the selected model, $\text{SARIMA}(0,1,1) * (0,1,1)_{12}$. Figure 14 displays a plot of the standardized residuals, the ACF of the residuals, the p-value associated with the Ljung-Box Q-statistics, (eq.25), at lags H=3 through H = 20 (with corresponding degrees of freedom H-2).

Inspection of the time plot of the standardized residuals in Figure 14 shows no obvious patterns. The ACF of the standardized residuals shows no apparent departure from the model assumptions, and the Ljung-Box Q-statistic is never significant at the lags shown. Hence, the residuals are random.

Unless the time series is Gaussian, it is not enough that the residuals are uncorrelated. For example, it is possible in the non-Gaussian case to have an uncorrelated process for which values contiguous (neighboring) in time are highly dependent. As an example, we mention the family of GARCH models. Therefore, we need to test for the presence of ARCH effect left.

4.4.2. Testing for ARCH Effect:

For the diagnostic checking of the presence of ARCH effect in residuals, LM ARCH test and the Ljung-Box Q-test for the squared residuals are used. The results are found in Tables 14 and 16, respectively. The LM ARCH test statistics is 0.1560(0.925) and Ljung box Q-statistics, for instance at 10 d.f., is 8.056(0.6263). Where the values in parentheses are their corresponding p-values. The p-values of the test statistic for both tests are greater than the 5% level of significance. Both tests fail to reject no ARCH left hypothesis in residuals. Therefore, we conclude that no autoregressive conditional heteroscedasticity left in the residuals- a conclusion consistent with the independence of the residuals given in (Testing for independence in 4.4.1.). The last but very important test for model fit is to check whether the residuals of the model are normally distributed.

4.4.3. Testing for Normality:

Finally, we check the normality assumption of the residuals. If the residuals are normally distributed, the histogram should be bell-shaped and the points in the QQ-plots should lie alongside a straight line. In addition to this, Both the Jarque-Bera statistic and Shapiro-test statistic should not be significant.

Figure 12 Appendix displays the histogram of the residuals. The histogram describes the distribution that is close to normality. Nevertheless, we cannot be certain. Let us look at the q-q plot of the residuals displayed in Figure 14 for closer comparison. The graph does not show a great deviation from normality, not even at the tail sides. Hence, we conclude that the residuals follow a normal distribution. Next, we conduct a normality test of hypothesis. We have used two different methods, Shapiro-Wilk and Jarque-Berra tests. The results for the calculated test statistic for Shapiro-Wilk test and Jarque-Berra test are 0.9884(0.1495) and 1.2548(0.534) respectively as displayed in Table 14. The values in parentheses are p-values.

The Jarque-Bera statistic tests the residuals of the fit for normality based on the observed skewness and kurtosis, and it appears that the residuals have normal skewness and kurtosis. The Shapiro-Wilk statistic tests the residuals of the fit for normality based on the empirical order statistics. The p-values of the test statistics for both tests are greater than the 5% level of significance. These tests imply that we do not have enough evidence to reject the null hypothesis that the residuals are normal. Therefore, we conclude that the residuals of the fitted seasonal $\text{ARIMA}(0,1,1) * (0,1,1)_{12}$ model are normally distributed. All the residual tests are consistent in showing the adequacy of the fitted model. Therefore, we conclude that the fitted model

$$(1-B)(1-B^{12})x_t = [1+0.5385B][1+0.8634B^{12}]\widehat{w}_t, \qquad \hat{\sigma}_w^2 = 0.3263$$

is adequate.

4.5. GARCH Models:

4.5.1. Unit Root Test for Log Return Cereal Price Series

We choose $p = 12$ because AIC selects an AR (11) model for the log return cereal price data, $logx_t - logx_{t-1}$. Other values of p are also used, but they do not alter the conclusion of the test. From Table 18 with p =12, the ADF test statistic is -4.8706, which is smaller than the tabulated values of -4.04 at 1% and -3.45 at 5% significance. Hence, we reject the null hypothesis that the data has unit root and is non-stationary.

4.5.2. Model Identification for the Log Return Cereal Price Data

Once the degree of differencing has been determined, we precede to select the autoregressive and moving-average orders by examining the sample autocorrelations and sample partial autocorrelations.

To use the sample autocorrelation and sample partial autocorrelations functions for tentative model parameters identification, we consider the ACF and PACF shown in Figure 15 of appendix. Using Table-1 as a guide,

First, we set the ACF to be tailing-off and the PACF to be cut-off after lag 3, we identify p=3 and q=0. In addition, it is possible to think of the PACF to be tailing-off and the ACF to cut-off after lag 3, leading to identify p=0 and q=3.

Fitting the following models suggested by these observations, we obtain:

i. ARMA (1, 0)
ii. ARMA (3, 2)
iii. ARMA (3, 3)

4.5.3. Parameter Estimation for the ARMA Model

We use maximum likelihood estimation method for monthly log return mean price of cereal price, to estimate the parameters. The results are summarized in Table- 8 below.

Table 8 Summary of parameter estimates and selection criterion

Model	Parameter	Estimate	Std.error	t-value	p-value	Information Criteria
a.	ϕ_1	-0.4505	0.0665	- 6.774	<0.001	AIC=123.77
						AICc=123.84
			$\hat{\sigma}_a^2$= 0.1142			BIC=130.15
b.	ϕ_3	-0.3317	0.0870	- 3.813	<0.001	AIC=114.65
	θ_2	0.8886	0.1761	5.046	<0.001	AICc=115.14
			$\hat{\sigma}_a^2$= 0.1033			BIC=133.78
c.	ϕ_3	0.3332	0.1140	2.923	0.003	AIC=116.82
	θ_3	-0.6387	0.1242	- 5.143	<0.001	AICc=117.48
			$\hat{\sigma}_a^2$= 0.1035			BIC=139.13

Where

a. stands for ARMA (1, 0)
b. stands for ARMA (3, 2)
c. stands for ARMA (3, 3)

Table-8 above displays the list of the parameters for each temporally entertained model. For each model parameter, the table presents the estimated value, standard error, *t*- value, p-value, AIC, AICc, BIC and variance ($\hat{\sigma}_a^2$) for the estimate. As indicated by McDowall et al., (1980), parameters must differ significantly from zero and all significant parameters must be included in the model. The t-ratios (t_{cal}) related to the parameter estimate were compared with the critical value of ± 1.96 obtained from the t-distribution at the 0.05 significance level. The test checks whether some terms in the model may be omitted or not. In our case, the estimated value and associated *t*- values for moving average parameters as well as autoregressive parameters are all significantly different from zero and they should retain in the model.

4.5.4. Model Selection

The AIC, AICc, BIC and variance of the estimate ($\hat{\sigma}_a^2$) in Table-8 are used to selected the best-fitted model to the monthly log return cereal price series. The model with the smallest information criteria is said to fit the data best. Model (b) has AIC, AICc, BIC and variance of estimate equal to 114.65, 115.14, 133.78 and 0.1033 respectively, we choose Model (b) because it has relatively smallest information criteria and fulfils the model adequacy. Hence, ARMA (3, 2) fits the monthly log return cereal price series best, and can be further analyzed. Therefore, the fitted model is

$$r_t = 1.2324_{(0.1942)} r_{t-1} - 0.0923_{(0.1292)} r_{t-2} - 0.3317_{(0.0870)} r_{t-3} - 1.8113_{(0.1828)} \hat{a}_{t-1} + 0.8886_{(0.1761)} \hat{a}_{t-2} + \hat{a}_t, \qquad \hat{\sigma}_a^2 = 0.1033$$

where the values in parenthesis are the corresponding standard errors.

4.5.5. Test of Model Adequacy

We will focus on ARMA(3,2) model which has the smallest information criteria. Figure 16 displays a plot of the standardized residuals $\{\tilde{a}_t\}$, the ACF of the residuals, a boxplot of the standardized residuals, and the p-values associated with the Ljung-Box-statistic, (eq. 25), at lags H = 3 through H = 20 (with corresponding degrees of freedom H - 2).

Inspection of the time plot of the standardized residuals, $\{\tilde{a}_t\}$, in Figure 16 shows no obvious patterns. The ACF of the standardized residuals shows no apparent departure from the model assumptions, and the Ljung-Box-statistic is never significant at the lags shown. Hence, the series does not have significant serial correlations so that it can be directly used to test for the ARCH effect. Indeed, the Ljung-Box test statistic of $\{\tilde{a}_t\}$ gives χ^2_{10} = 5.4043 with p-value 0.86, confirming no serial correlations in the data. On the other hand, the Ljung-Box test statistic $\{\tilde{a}_t^2\}$ shows strong ARCH effects with test statistic $\chi_{10}^2 \approx 43.76$, the p-value of which is close to zero. To address the issue of presence of ARCH/GARCH effect we use GARCH model.

4.6. Parameter Estimation of GARH Models for the Return Series

We have identified the ARMA model to be ARMA (3, 2) the next step is to choose the ARCH and GARCH model so that all the tests for the residuals and residual square will satisfy the assumptions stated.

Now we estimate the whole ARMA(P, Q)- GARCH (p, q) model. This cannot be done by simple OLS due to nonlinearity. Instead, maximum likelihood estimator must be used.

For each model parameter, the table presents the estimated value, standard error, *t*- value, p-value, AIC, AICc, BIC and HQIC for the estimate. Both ARMA (2, 1) +GARCH(1,1) and ARMA(2, 2)+GARCH(1,1) have very close information criteria with the first being the smallest. However, in the later it was found that one estimated parameter is (θ_2 = -0.17914, std. Error = 0.850 and P-value= 0.3955 >0.05) which is insignificant. Therefore, it should not be included in the model. Hence the model that has to be considered becomes ARMA (2, 1) +GARCH(1,1). A joint estimation of the ARMA (2,1)-GARCH(1, 1) gives (i.e. the fitted model is):

$$r_t = 0.335_{(0.001)} r_{t-2} + 0.203_{(0.021)} r_{t-2} - 0.834_{(<0.001)} a_{t-1} + a_t$$

$$a_t = \sigma_{t|t-1}\epsilon_t, \qquad \sigma_{t|t-1}^2 = 0.0001246_{(0.88)} + 0.066_{(0.0561)} a_{t-1}^2 + 0.926_{(<0.001)} \sigma_{t-1|t-2}^2$$

Where the values in parentheses are p-values.

Table 9. summary of Parameter Estimates and selection criteria						
Error analysis	Estimate	Std. Error	T-value	Pr(>|t|)	Criteria	Mean and variance equation
μ	0.0051	0.0039	1.303	0.1927	AIC =	ARMA (2, 2) + GARCH(1, 1)
ϕ_1	0.1557	0.2742	0.568	0.5702	0.51852	
ϕ_2	0.2900	0.1177	2.463	0.0138*	AICc	
θ_1	-0.6512	0.2783	-2.463	0.0193*	0.6610	
θ_2	-0.1791	0.2108	-0.850	0.3955	BIC	
ω	0.0001	0.0008047	0.146	0.8841	0.5148	
α_1	0.0630	0.0327	1.924	0.0544	HQIC	
β_1	0.9297	0.0332	28.021	<0.0001***	0.5766	
μ	0.0040	0.0028	1.42	0.15545	AIC	ARMA (2, 1) + GARCH(1, 1)
ϕ_1	0.3636	0.1069	3.402	0.0007**	0.5112	
ϕ_2	0.2151	0.0918	2.343	0.0191*	AICc	
θ_1	-0.8707	0.0748	-11.64	<2.e-16***	0.6358	
Ω	0.0001	0.0008146	0.141	0.887659	BIC	
α_1	0.0631	0.0330	1.910	0.0561	0.5083	
β_1	0.9263	0.0337	27.56	<2e-16***	HQIC	
					0.5617	
ϕ_1	0.2659	0.071	3.747	0.0002	AIC=	ARMA(3,2)+GARCH(1,1)
ϕ_2	-0.6167	0.061	-10.13	<0.001	0.529	
ϕ_3	-0.3629	0.070	-5.2	<0.001	AICc =	
θ_2	-0.7035	0.0364	-19.32	<0.001	0.672	
θ_1	0.9197	0.0615	14.95	<0.001	SIC =	
Ω	1.442e-07	7.3e-04	0.00	0.99	0.525	
α_1	0.00508	0.025	2.025	0.0423	HQIC =	
β_1	0.9432	0.0283	36.521	<0.001	0.587	

Table-9 above displays list of the parameters for each temporally entertained model.

From the variance equation the implied unconditional variance of a_t is

$$\frac{0.0001246}{1 - 0.066 - 0.926} = 0.01674664$$

The t ratios of the parameters in the mean equation suggest that all three AR coefficients are significant at the 5% level. Therefore, we should not drop any AR and MA parameters to refine the model.

4.7. Diagnostic Checking for the Selected GARCH Model:

In this section, we will assess how well the selected model, $\text{ARMA}(2, 1) - GARCH(1, 1)$, fit the actual return cereal price data. If the model fits the data well, the residuals of the fitted model are random (Chatfield, 1996). To do this we have used diagnostic checking for independence of the residuals, then tested ARCH effect left in the residuals and finally conducted test for normality of the residuals.

4.7.1. Testing for Independence:

In this section, we check if residuals of the selected model, $\text{ARMA}(2, 1) + GARCH(1, 1)$, are independent. Figure 17 displays a plot of the standardized residuals, the ACF of the residuals, and Table 20 displays the p-value associated with the Ljung-Box Q-statistics, for H=10, 15, and 20.

Inspection of the time plot of the standardized residuals in Figure 17 shows no obvious patterns. The ACF of the standardized residuals shows no apparent departure from the model assumptions, and the Ljung-Box- Q-statistic in Table 20 is never significant at the lags shown. Hence, the residuals are random. The other thing we need to test for independence of the residuals is to check if there are ARCH effects left.

4.7.2. Testing for ARCH Effect:

For the diagnostic checking of the presence of ARCH effect in residuals, LM ARCH test and the Ljung-Box Q-test for the squared residuals are used. The results are found in Table 20. The p-values of the test statistic for both tests are greater than the 5% level of significance. This result implies that we do not have enough evidence to reject the null hypothesis that there is no ARCH left in residuals. Therefore, we conclude that no Auto regressive conditional heteroscedasticity left in the residuals – a conclusion consistent with the independence tests given in 4.7.1. Finally, we check the normality assumption of the residuals.

4.7.3. Testing for Normality:

If the residuals are normally distributed, the histogram should be bell-shaped. In addition to this, Both the Jarque-Bera statistic and Shapiro-test statistic should not be significant.

Figure 17 displays the histogram of the residuals. The histogram describes the distribution that is cloth to normality. Hence, we conclude, it is possible that the residuals follow a normal distribution. Next, conduct a normality test of hypothesis. We have use two different methods, Shapiro-Wilk and Jarque-Berra tests. The results are displayed in Table 20.

The Jarque-Bera statistic tests, χ_2^2=2.770888(0.250) the residuals of the fit for normality based on the observed skewness and kurtosis, and it appears that the residuals have normal skewness and kurtosis. The Shapiro-Wilk statistic, W=0.9892886(0.1970764), tests the residuals of the fit for normality based on the empirical order statistics. The values in parentheses are p-values. The p-value of the test statistics for both tests are greater than the 5% level of significance. These tests imply that we do not have enough evidence to reject the null hypothesis that the residuals are normal. Therefore, we conclude that the residuals of the fitted models are normally distributed.

4.8. Interpretations of Empirical Results from GARCH-Model

For the sake of further comparison of the coefficients, we present the coefficients in Table 9 even if AR, MA, GARCH or ARCH term were not significant and thus is probably not present in the real data generating process in some of the cases.

The fitted model is

$$r_t = 0.335_{(0.001)} r_{t-1} + 0.203_{(0.021)} r_{t-2} - 0.834_{(<0.001)} \hat{a}_{t-1} + \hat{a}_t$$

$$a_t = \sigma_{t|t-1}\epsilon_t, \qquad \sigma^2_{t|t-1} = 0.0001246_{(0.88)} + 0.066_{(0.0561)} \hat{a}^2_{t-1} + 0.926_{(<0.001)} \sigma^2_{t-1|t-2}$$

Where the values in parentheses are p-values.

Here in the return of cereal price ϕ_1, ϕ_2 , θ_1,and GARCH term are significantly different from zero at 5% and 10% level of significance. Whereas ϕ_0, α_0 and ARCH term, α_1, where all zero at 1% and 5% level of significance. Hence, ARCH term did not meet the positivity condition and had to be removed from the equations. Nevertheless, the ARMA and GARCH terms alone were sufficient to specify the conditional volatility equation and the standardized residuals fulfilled the iid condition. Furthermore, in our case the GARCH term had the strongest effect on volatility from all the analyzed time series. This is not surprising since the ARCH term was removed from the equation and all the explanatory power remained for the GARCH term. Value of the GARCH term, $\beta_1 = 0.926$, for the return of cereal price is close to one. This means slow convergence of volatility to a steady state and high persistence in volatility.

The constant term, in the mean equation was significant and thus it follows an ARMA (2, 1) Model for the return of cereal price.

4.9. Forecasting

There are two reasons why forecasting volatility could be important. Firstly, good forecast capability of volatility models provides a practical tool for price data analysis. Secondly, as proxy for risk, volatility is related to expected returns, hence good forecast models enable to device more appropriate securities pricing strategies (Akgiray, 1989). In the previous sections, we have performed the Seasonal ARIMA and GARCH models and some diagnosis checking of standardized residuals to compare the specifications. However, the in-sample and out-of-sample forecast evaluation potentially provides comparison that is more useful. So far, the models are estimated by using actual data generating process, but literature suggests that the good performance in parameter estimates and goodness of fit statistics do not guarantee to give accurate forecast models (Lopez, 2001). The satisfactory estimation of seasonal ARIMA and GARCH-models is further evaluated through multi-step forecasting.

4.9.1. In-Sample Forecast of Domestic Price Volatility by GARCH Models

Incorporating the most adequate choice of the volatility models for price of cereals, the volatility of domestic prices using variance as volatility measures was forecasted using the in-sample observations under static forecasting hence the results are presented in Figure 1 and Figure 2 below.

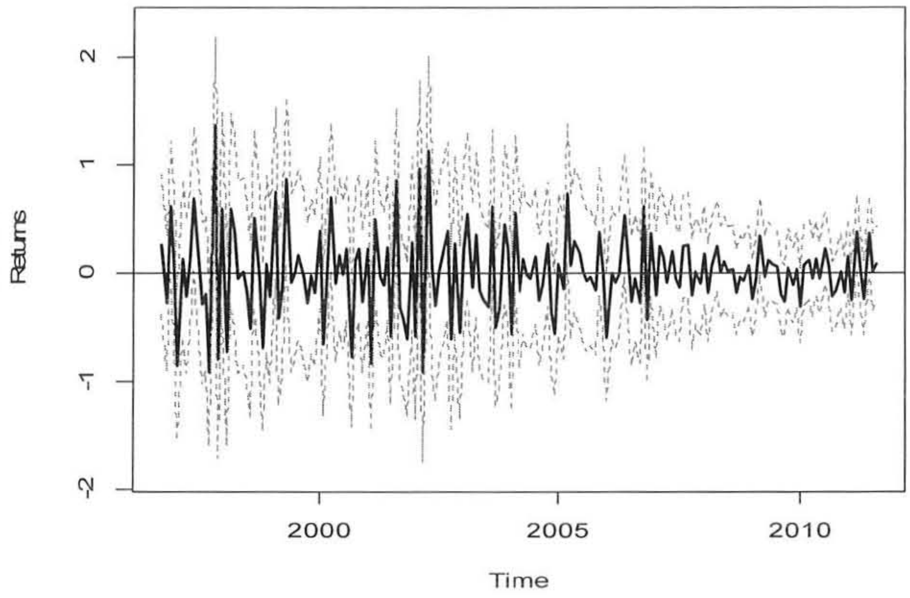

Fig.1. GARCH predictions of the cereal volatility, $\pm 2\hat{\sigma}_t^2$, displayed as dashed lines

To explore the GARCH predictions of volatility, we calculated and plotted all the observations from the data along with the one-step-ahead predictions of the corresponding volatility, $\pm 2\hat{\sigma}_t^2$. The results are displayed as the data $\pm 2\hat{\sigma}_t^2$as a dashed line surrounding the data in Figure 1.

The important point to note from Figure 1 and Figure 13 appendix is that, during 1997 to 2002, there was high volatility, when many large positive and large negative returns were observed during a short space of time. In addition, volatility decreases to the end of the 2011.

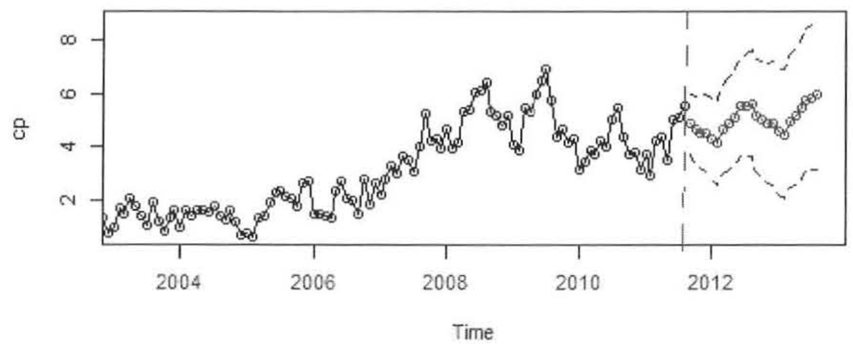

Fig.2. Twenty-four month forecasts for the cereal price series.

The vertical dashed line separates the data from the predictions. The actual data shown are from about September 1996 to July 2011, and then the forecasts plus or minus 2 standard error are displayed.

4.9.2. Out-Sample Forecast of Domestic Price Volatility by GARCH Models

The out-of-sample forecast is performed for the first year from August 2011, to July 2012 generating 12 observations. The 12 months period is chosen based on previous studies.

Table 10. Out sample forecast using GARCH Model

Month	Out-sample Data Return Cereal Price	Mean Forecast	Mean Error	St. Deviation
Aug-11	**0.082908**	-0.06535	0.173871	0.173871
Sep-11	**0.023934**	-0.00101	0.194661	0.173546
Oct-11	**-0.07102**	-0.01041	0.194371	0.173222
Nov-11	**-0.0255**	7.72E-06	0.19402	0.172901
Dec-11	**-0.02599**	0.001775	0.193662	0.172581
Jan-12	**-0.04625**	0.004659	0.193306	0.172263
Feb-12	**-0.00956**	0.006088	0.192951	0.171947
Mar-12	**-0.01182**	0.007228	0.192598	0.171633
Apr-12	**0.07254**	0.00795	0.192247	0.171321
May-12	**0.016241**	0.008458	0.191898	0.17101
Jun-12	**0.086265**	0.008798	0.191551	0.170701
Jul-12	**0.030293**	0.009031	0.191207	0.170394

Table. 11. Out sample forecast using the selected SARIMA model

Month	Out-sample Cereal Price	Mean Forecast	Standard Error
Aug-11	6.049121	5.653146	0.997934
Sep-11	6.195649	4.870892	0.571438
Oct-11	5.770895	4.658712	0.632531
Nov-11	5.625584	4.558733	0.683928
Dec-11	5.481245	4.504183	0.726568
Jan-12	5.233531	4.29897	0.7661
Feb-12	5.183739	4.138841	0.803401
Mar-12	5.12285	4.711113	0.838982
Apr-12	5.508272	4.921205	0.873095
May-12	5.598462	5.528471	0.90592
Jun-12	6.102859	5.100549	0.937594
Jul-12	6.290564	5.56189	0.968234

Incorporating the most adequate choice of the volatility models for producer price of cereals, the volatility of producer prices using variance as volatility measures was forecasted using the out-sample observations under static forecasting and the results are presented in Table 10 above. In addition, the Twelve-month forecasts for the cereal price series using the selected seasonal ARIMA are given in table 11 above.

The estimation and diagnostic analysis results of the seasonal ARIMA and the GARCH models revealed that the models are adequate. In particular, the residual analysis, which is important for diagnostic checking confirmed that there is no violation of assumptions in relation to model adequacy.

To illustrate the forecasting performance of the prior seasonal model and GARCH model, we estimate the models using the first 180 observations and reserve the last twelve data points for forecasting evaluation. We compute 1-step to 12-step ahead forecasts and their standard errors of the fitted model at the forecast origin $h = 180$. Table 10 and 11 show the forecasting performance of the models, where the observed data are shown in second column, mean forecast on second column and Standard Error on third column. The forecasts are close to the observed data. Finally, we have used accuracy measures to compare the selected models.

4.10. Forecasting Accuracy Evaluation for the selected models ARIMA and GARCH

In order to identify the superior in-sample forecast model, six accuracy evaluations of error statistics are employed. These are Mean percentage Error (MPE), Mean Absolute Percentage Error (MAPE), both Theil's U1 and U2 statistics, bias proportion and variance proportion. The results of the measures for SARIM(0,1,1)*(0,1,1) and ARMA(2,1)+GARCH(1,1) are given Table 12 below.

Table 12. In-sample forecast evaluation data from 1997:09 to 2011:07

Forecast evaluation statistics	SARIMA	GARCH
Mean percentage error	**-6.18024**	65.89529
Mean absolute percentage error	**24.64436**	72.11019
Theil's U1	**0.052917**	0.591611
Theil's U2	0.856539	**0.768701**
Bias Proportion	0.000591	**0.000121**
Variance Proportion	**0.004424**	0.40262
Covariance Proportion	**0.994985**	0.597259

Bold shows the butter by the criteria

Almost all the forecast evaluation of error statistic generate the same results which indicate the outperformance of SARIM(0,1,1)*(0,1,1) model. The bias and variance proportion for the ARIMA model are close to zero. While, the variance proportion for GARCH model is relatively large. Both tests have very close Theil's U2 test, which is less than one, indicating both models are good for In-sample forecasting. In general, regarding the forecasting ability of the models in the case of Ethiopian cereal price, the results indicate that Seasonal ARIMA model is better than the GARCH model. The results obtained are considerably decisive and consistent across six measures.

Table13. Out-sample Forecast Evaluation (Data from 2011:08 to 2012:07)

Forecast evaluation statistics	SARIMA	GARCH
Mean percentage error	**14.33640723**	86.5327
Mean absolute percentage error	14.33640723	**11.57745**
Theil's U1	0.071325257	**0.030353**
Theil's U2	3.208759133	**0.852948**
Bias Proportion	0.873856854	**0.061574**
Variance Proportion	0.016957267	**0.36333**
Covariance Proportion	0.009758	**0.2449**

Bold shows the butter by the criteria

To find the superior out-sample forecast model, the above forecast evaluation are used. Results are given in Table 13 above. Most of the Measurement error evaluation goes in favor of ARMA(2,1)+GARH(1,1) model. Hence, we conclude that the selected GARCH model is better in forecasting out-sample as compared to the seasonal ARIMA model. Therefore we conclude that GARCH model is better than seasonal ARIMA model for forecasting out sample while the seasonal ARIMA model is better than the GARCH(1,1) in forecasting in-sample cereal price in Ethiopia.

4.11. Discussion of Results

The ARIMA model we got is common among seasonal time series models and can generally be expressed as:

$$(1-B)(1-B^s)x_t = [1+\theta B][1+\Theta B^s]w_t$$

Where s is the periodicity of the series, w_t is a white noise series, $|\theta| < 1$, and $|\Theta| < 1$. The model is referred to as the airline model in the literature (Box, Jenkins, and Reinsel ,(1994), and Tsay, (2010)). The MA part of the model involves two parameters, whereas the AR part consists of the regular and seasonal differences. For comparison purpose the model is given below.

$$(1-B)(1-B^{12})x_t = [1+0.5385B][1+0.8634B]\hat{w}_t, \qquad \hat{\sigma}_w^2 = 0.3263$$

where standard errors of the two MA parameters are 0.0709 and 0.08934, respectively.

Consistent with our findings, Akgiray (1989), Bollerslev (1986) and others agree on that GARCH (1, 1) model fit financial data sufficiently enough for forecasting variance without the need of more sophisticated models.

CHAPTER FIVE: Conclusion and Recommendation

5.1. Conclusion

In this paper, we have examined the Seasonal ARIMA and GARCH models for the time series of cereal price. In the GARCH case, the conditional mean was estimated using the ARMA model and conditional variance using the GARCH models. We have estimated the standard Gaussian model ARMA(2,1)+GARCH(1,1) and then analyzed the asymmetric extensions of GARCH model with different error distribution. After examining the goodness of fit of competitive models it was concluded that the simple GARCH, GARCH (1,1), model with normal distribution provides the best fit and the prediction was made based on this model.

In the Seasonal ARIMA case, after we have examined the goodness of fit of competitive models we came to conclusion, that the Seasonal ARIMA(0,1,1)*(0,1,1) provides the best fit and the prediction was made using this model.

To detect the best fit estimation and the most accurate forecast model for cereal price the two selected models were further compared. To gain the choice of superior in-sample and out-sample forecast model, six measures were employed: Mean Percentage Error (MPE), Mean Absolute Percent Error (MAPE), and Theil's U statistics, bias proportion and variance proportion.

Almost all the in-sample forecast evaluation of error statistic generate the same results which indicate the outperformance of SARIM(0,1,1)*(0,1,1) model. And the out-sample forecast was in favor of the GARCH(1,1) with conditional mean ARMA(2,1). In general, regarding the forecasting ability of the models in the case of Ethiopian cereal price, the results indicate the outperformance of the seasonal ARIMA over GARCH (p, q) models for in-sample forecast. However, GARCH (1,1) model is found to be better than seasonal ARIMA model when it comes to making out-sample forecast. The results obtained are considerably decisive and consistent across six measures.

5.2. Limitations and Recommendations

Although the paper provides the results are relatively understandable and consistent with the previous studies, the results should be treated with caution due to the following reasons. Firstly, the monthly data period of fifteen years is quite limited for applying GARCH models comparable to the empirical studies with longer time horizon of estimation. Secondly, though the paper attempts to comprehensively investigate the producer cereal price market and its volatility by the means of various univariate time series models, it covers the most widely used models without consideration of the huge number of other GARCH extensions such as EGARCH , APPARCH and other approaches such as stochastic volatility models.

These limitations suggest further studies on the same subject to generate results that are more fruitful. Alternative Seasonal ARIMA specifications as well as other simpler or more sophisticate GARCH approaches may be conducted. Another suggestion is that further study may employ multivariate models such as seasonal or Dynamic Conditional Correlation Multivariate model to analyze the time-varying correlation of cereal price market with the others price markets in particular and international Price markets in general.

References

Ahmed, H.A. (2007). Structural Analysis of Price Drivers in Ethiopia. Mimeo, Ethiopian Development Research Institute, Addis Ababa.

Ahmed, S. (2008). Global Food Price Inflation: Implications for South Asia: Policy Reactions and Future Challenges, *WB Policy Research Working Paper* No. **4796**. Washington D.C.

Akaike, H. (1969). Fitting autoregressive models for prediction. *Ann. Inst. Stat. Math.*, **21**, 243-247.

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.

Akaike, H. (1974). A new look at statistical model identification. *IEEE Trans. Automat. Contr., AC*-**19**, 716-723.

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: evidence and forecasts. *Journal of Business*, **62**, 55-80.

Apergis , N. and Rezitis, A. (2003). Food Price Volatility and Macroeconomic Factors: Evidence from GARCH and GARCH-X Estimates. *Journal of Agricultural and Applied Economics,* **43**, 95–110.

Bekele, G. (2002). The Role of the Ethiopian Grain Trade Enterprise in Price Policy. In Agricultural Technology Diffusion and Price policy in Addis Ababa, ed. T.Bonger, E. Gabre-Madhin, and S. Babu. Proceedings of a Policy Forum in Addis Ababa: 2020 Vision Network for East Africa Report I. Addis Ababa, Ethiopia, and Washington, DC: Ethiopian Development Research Institute and International Food Policy Research Institute.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31**, 307–327.

Box, G. E. P, Jenkins, G. M. and G.Reinsel, 1985. *Time Series Analysis: Forecasting and Control.* Prentice-Hall, Englewood Cliffs.

Box, G. E. P. and Jenkins, G.M., 1976. *Time Series Analysis, Forecasting and Control.* Holden day.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994). *Time Series Analysis, Forecasting and Control*, Second Edition, New York: Prentice-Hall.

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, Newbold, P., 1971. Some Comments on a Paper by Chatfield and Kendal. *Journal Royal Statistics .Soc,* **9**, 229.

Brockwell, P.J. and Davis, R. A.,1996. *Introduction to Time Series and Forecasting,* Second Edition, Springer, New York.

Brooks, C. (2002). *Introductory Econometrics for Finance*. Cambridge University Press.

Chatfield, C. (1996). *The Analysis of Time Series*, Fifth Edition, Chapman & Hall, London.

Cromwell, J.B., Labys, W.C. and Terraza, M. (1994). *Univariate Tests for Time Series Models*. A Sage Publications, London.

Cryer. J. D and Chan .K.S. (2008). *Time Series Analysis: (With Applications in R),* Second Edition, Iowa City: Springer.

Diao X., Belay F., Steven H., Alemayehu S., Kassu W., Bingxin Y. (2007). Agricultural Growth Linkages in Ethiopia: Estimating using Fixed and Flexible Price Models. IFPRI discussion paper no 00695, Addis Ababa

Diebold F.X. , Kilian, L., and Nerlove M.(2006). Time Series Analysis, *Working Paper No.* **06-01** University of Maryland, College Park.

Dorosh, P.A. and Subran, L. (2007). Food Markets and Food Price Inflation in Ethiopia: Background paper for the Food Price Inflation Policy Note. World Bank, Washington DC.

Durbin, J. (1960). Estimation of parameters in time series regression models. *J. R. Stat. Soc. B*, **22**, 139-153.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**, 987-1007

Ferguson, T.S., Genest, C. and Hallin, M.,(2000). Kendal's Tau for Serial Dependence. *The Canadian Journal of Statistics,* **28**, 587-604.

Gebremedhin, B., Hoekstra, D., and Tegegne, A. (2006). Commercial of Ethiopian Agriculture: Extension Service from Input Supplier to Knowledge Broker and Facilitator. Improving Productivity and Market Success (IPMS) of Ethiopian

Farmers Project Working Paper 1. Nairobi, Kenya; International Livestock Research Institute.

Gibbons, R.D., (1994). *Statistical Methods for Groundwater Monitoring.* John Wiley & Sons, New York.

Gilbert L. C. (2009). How to understand high food prices. presented on the 2009 ICABR conference. Ravello, Italy, 17-19.

Gourieroux, C. (1997). *ARCH Models and Financial Applications.* Springer Series in Statistics, Springer- Verlag New York.

Granger and Newbold, (1986). Forecasting Economic Time Series. Academic Press.

Greene, W.H. (2000). *Econometric Analysis. Prentice Hall International.* Inc., 1004 p., New Jersey-USA.

Hansen, P.R and Lunde, A. (2005). A forecast comparison of volatility models: does anything beat a GARCH (1,1), *Journal of applied econometrics*, **20**, 873-889.

Harvey, A.C. (1993). *Time Series Models.* Cambridge, MA: MIT Press.

Human Development Report (2010) . The Real Wealth of Nations: Pathways to Human Development. 20th Anniversary Edition Published for the United Nations Development Program.

Human Development Report (2011). Sustainability and Equity: A Better Future for All. the United Nations Development Program 1 UN Plaza, New York, NY 10017, USA.

Hurvich, C.M and Tsai, C.L. (1989). Regression and time series model selection in small samples. *Biometrika,* **76**, 297-307.

IMF (2008b). Federal Democratic Republic of Ethiopia: Selected Issues. IMF Country Report 08/259. Washington, DC: International Monetary Fund.

IMF. (2008a). Federal Democratic Republic of Ethiopia: Article IV Consultation Staff Report. IMF Country Report 08/264. Washington, DC: International Monetary Fund.

Ivanic, M. & Martin, W. (2008). Implications of Higher Global Food Prices for Poverty in Low- Income Countries. *Policy Research Working Paper WPS4594.* The World Bank. Washington D.C.

Janacek, G. and Swift, L. (1993). *Time Series Forecasting, Simulation, Application.* Ellis Horwood, New York,USA.

Jarque, C. M.; Bera, A. K. (1981). Efficient tests for normality, homoscedasticity and serial independence of regression residuals: Monte Carlo evidence. *Economics Letters* 7 (4): 313–318. doi:10.1016/0165-1765(81)90035-5.

Johnson, R.A. and Wichern D.W. (1992). *Applied Multivariate Statistical Analysis,* Third Edition, Englewood Cliffs, NJ: Prentice-Hall.

Klugman, J. (2007). Explaining Food Price Inflation Policy Note. *Discussion paper*, Ethiopia.

Lehmann, A. and Rode, M. (2001).Long-Term Behaviour and Cross-Correlation Water Quality Analysis of the River Elbe, Germany. *Journal of Water Resources*, **35**, 2153-2160.

Levinson, N. (1947). The Wiener (root mean square) error criterion in filter design and prediction. *J. Math. Phys.*, **25**, 262-278.

Ling, S. and McAleer, M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, **106**, 109–117.

Lirenso, A. (1994). Liberalizing Ethiopian Grain Markets. *Papers of the 12th International Conference of Ethiopian Studies.* East Lansing: Michigan State University, p. 2-24.

Liu, L.M., and William, J. (2001). Data mining on time series: An illustrative using fast food restaurant franchise data. Scientific Computing Associate Corp.

Ljung, G.M. and Box, G.E.P., 1978. On a measure of lack of fit in time-series models. *Biometrika* **65**, 297-303.

Lopez J.A. (2001), Evaluating the Predictive Accuracy of Volatility Models, *Journal of Forecasting*, **20**, 87-109.

Makridakis, S., Wheelwright, S.C. and Hyndman, R. J. (1998). *Forecasting methods and Application*. New York: John Wiley & Sons.

McDowall, D., McCleary, R.., Meidinger. E., Hay, R.A. (1980). *Interrupted Time Series Analysis*. Sage Publications, Inc.

McLeod, A.I and Li, W.K. (1983). Diagnosis checking ARMA time series models using squared-residual autocorrelation. *Journal of Time Series Analysis*, **4**, 269-273.

McLeod, A.I. (1978). On the distribution of residual autocorrelations in Box-Jenkins models. *J. R. Stat. Soc.* B, **40**, 296-302.

McQuarrie, A.D.R. and Tsay, C.L. (1998). *Regression and Time Series Model Selection*, Singapore: World Scientific.

Meron, A. and Rashid, S. (2007). Cereal Price Instability in Ethiopia: an Analysis of Sources and Policy Options. Paper prepared for the Agricultural Economics Association for Africa, Ghana.

Mulat, D., Atlaw, A., Bilisuma, B., Saba, Y. and Tadele, F. (2007). Exploring Demand and Supply Factors behind the New Developments in Grain prices in Ethiopia: Key Issues and Hypotheses. A report submitted to DFID_ ETHIOPIA.

Pindyck, R. S. (2004). Volatility and Commodity Price Dynamic. *The Journal of Futures Markets*, **24**, 1029-1047.

Rashid, S. (2010). Variation in staple food prices: Causes, consequence, and policy options. Paper presented at the COMESA policy seminar held on Maputo, Mozambique, January 2010.

Schwarz, F. (1978). Estimating the dimension of a model. *Ann. Stat.*, **6**, 461-464.

Shapiro, S.S. and Wilk, M.B. (1965). An Analysis of Variance Test for Normality (complete samples). *Biometrika,* **52**, 591-611pp.

Shiferaw, y. A. (2012). Modeling Volatility of Price of Some Selected Agricultural Products in Ethiopia: ARIMA-GARCH Application(June 7, 2012) Available at SSRN: http:// ssrn.com/abstract = 2125712 or http:// dc.doi.org/10.2139/ssrn.2125712.

Shumway, R. H. and Stoffer, D. S. (2010). *Time Series Analysis and Its Applications (with R Examples),* Third Edition, New York, Springer.

Shumway, R. H., and Stoffer, D. S. (2006). *Time Series Analysis and Its Applications (with R Examples),* Second Edition. New York, Springer.

Song, H., Liu, X., and Romilly, P. (1998). Stock returns and volatility: An empirical study of Chinese Stock Market. *International Review of Applied Economics*, **12** (1), 129-140.

Stephan, Pf. Paul, Newbold, and Anthony, R. (2009). The Development of Cereal Prices during the 20th Century. February 24, 2009, University of Liverpool, Liverpool L69 7ZH.

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.

Taffese A. S., (2008). Understanding the constraints to continued rapid growth in Ethiopia: the role of agriculture. paper prepared for the DFID funded study, Ethiopian Economics Association, 2008.

Taylor, S. J. (1986). *Modeling Financial Time Series*. Chichester: John Wiley & Sons.

Tsay, R.S. (2002). *Analysis of Financial Time Series*. New York: Wiley. [11.1.4, 13.7.9]

Tsay, S.R. (2005). *Analysis of Financial Time Series,* Second Edition, John Wiley & Sons, Inc., Hoboken, New Jersey.

Tsay, S.R. (2010). *Analysis of Financial Time Series, Third Edition,* Wiley & Sons, Inc., Hoboken, New Jersey.

Wei, W.W.S. (1990). *Time Series Analysis*. Addison-Wesley Publishing Company, Inc., 478 P., New York-USA.

Weiss, A.A. (1986). Asymptotic theory for ARCH models: estimation and testing, Economet. Theory **2**, 107-131.

World Bank, (2007). Explaining Sources of Food Price Inflation in Ethiopia: A Just in Time Policy Note. World Bank (draft).

Zhang, M. Y., Russell, J. R., and Tsay, R. S. (2008). Determinants of bid and ask quotes and implications for the cost of trading. *Journal of Empirical Finance* **15**: 656–678.

Ruitenberg, r., and Dreibus(2011). France's Sarkozy Urges Action Against the 'Plague' of Food Price Surges to Group of 20 farm ministers in France on June 22, 2011. rruitenberg@bloomberg.net; tdreibus@bloomberg.net ccarpenter2@bloomberg.net

Appendix

List of Tables

Table 15. Serial correlation check for residuals of SARIMA (0, 1, 1)*(0, 1, 1)$_{12}$ model using Ljung-Box Q-statistic test

To lag	Chi-Square	Degree of freedom	Pr> Chi-Square
7	0.8669	5	0.9726
12	6.4389	10	0.7771
17	18.1816	15	0.2532
22	25.3879	20	0.187

Table 16. Serial correlation check for squared residuals of SARIMA (0, 1, 1)*(0, 1, 1)$_{12}$ model using Ljung-Box Q-statistic test

To lag	Chi-Square	Degree of Freedom	Pr> Chi-Square
7	1.4971	5	0.9134
12	8.056	10	0.6263
17	15.0414	15	0.4484
22	19.5615	20	0.4856

Table 17. Ljung-Box Q-statistic test for squared residuals of ARMA (3, 2)

To lag	Chi-Square	Degree of Freedom	Pr> Chi-Square
10	35.07	7	<0.0001
15	47.19	12	< 0.0001
20	59.47	17	< 0.0001

Table 18: Summary of ADF unit-root test (the log Return Cereal price Series)

Series	ADF test statistic	1%crit.Value	5%crit. Value	10% crit. Value	Decision
Return Cereal Price Series	-4.8706	-4.04	-3.45	-3.15	Reject H_0 at 1, 5 and 10% level of sig.

Table 19 Standardized Residuals Test for ARMA (3, 2)

			Statistics	p-Value
Jarque-Bera Test	R	Chi^2	2.6785	0.262
Shapiro-Wilk Test	R	W	0.9909	0.3145
Ljung-Box-test	R	Q(10)	5.4043	0.8626
Ljung-Box-test	R	Q(15)	11.9618	0.6819
Ljung-Box-test	R	Q(20)	20.136	0.4495
Ljung-Box-test	R^2	Q(10)	43.7614	3.633e-06
Ljung-Box-test	R^2	Q(15)	55.5888	1.422e-06
Ljung-Box-test	R^2	Q(20)	66.8134	5.982e-07
LM ARCH Test	R	TR^2	0.2695306	0.873921

Table 20. Standardized Residuals Tests: For ARIM(2, 1)-GARH(1, 1)

```
Standardised Residuals Tests:
                                Statistic p-Value
 Jarque-Bera Test   R    Chi^2  2.770888  0.2502127
 Shapiro-Wilk Test  R    W      0.9892886 0.1970764
 Ljung-Box Test     R    Q(10)  5.473972  0.8573551
 Ljung-Box Test     R    Q(15)  19.94413  0.1740863
 Ljung-Box Test     R    Q(20)  30.11338  0.06803722
 Ljung-Box Test     R^2  Q(10)  14.09703  0.1686114
 Ljung-Box Test     R^2  Q(15)  19.2135   0.2042216
 Ljung-Box Test     R^2  Q(20)  20.53145  0.4251564
 LM Arch Test       R    TR^2   14.54867  0.2670503

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
0.5112023 0.6358486 0.5082944 0.5617454
```

Table 21. Standardized Residuals Tests: For ARIM(2, 2)-GARH(1, 1)

```
Standardised Residuals Tests:
                                    Statistic p-Value
 Jarque-Bera Test   R    Chi^2  3.16114    0.2058577
 Shapiro-Wilk Test  R    W      0.989061   0.1840032
 Ljung-Box Test     R    Q(10)  6.16647    0.8010886
 Ljung-Box Test     R    Q(15)  19.75929   0.1813624
 Ljung-Box Test     R    Q(20)  29.59376   0.076709
 Ljung-Box Test     R^2  Q(10)  12.0962    0.2786696
 Ljung-Box Test     R^2  Q(15)  17.48648   0.2906236
 Ljung-Box Test     R^2  Q(20)  19.26484   0.5046759
 LM Arch Test       R    TR^2   12.03809   0.4426258

Information Criterion Statistics:
      AIC       BIC       SIC      HQIC
0.5185237 0.6609767 0.5147520 0.5762872
```

Table 22. Parameter Estimates for Suggested SARIMA Models continuation of table 6

Model	Parameter	Estimate	Std. error	t-value	p-value	Criteria
f.	θ_1	-0.5400	-0.0089	-0.7387	<0.001	
	Θ_2	0.0731	0.1011	0.08452	0.9326*	
	SARIMA(0, 1, 1)*(0, 1, 2)			Deliberately Added		
g.	ϕ_3	-0.1347	0.0789	-1.7072	0.08778*	
	Θ_2	-0.8692	0.917	-9.47835	<0.001	
	SARIMA(3, 1, 1)*(0, 1, 2)			Deliberately Added		

Lists of Figures

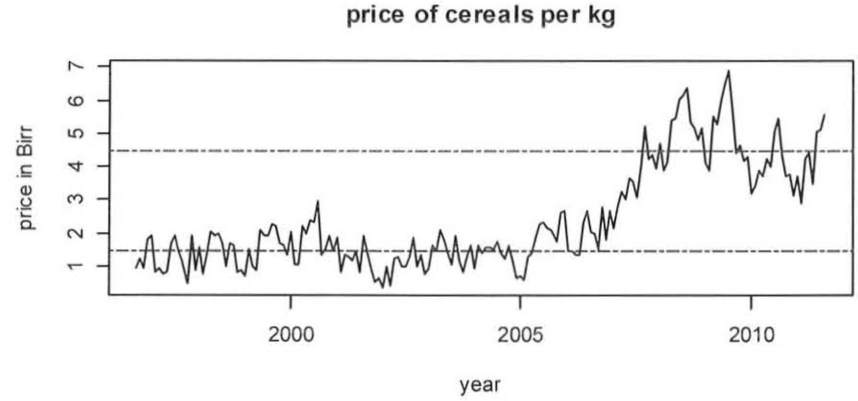

Fig. 3. Time plot of Cereal price, the lower dashed line is the mean of the months from September 1996 to December 2006. The upper dashed line is 3 times the mean.

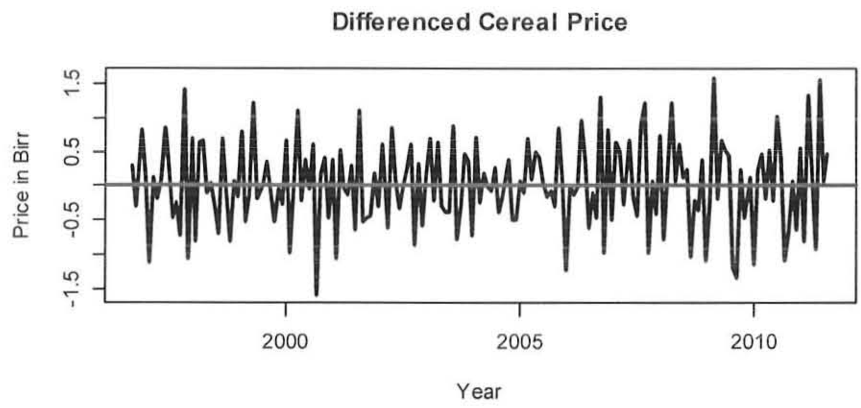

Fig. 4. First difference of cereal price

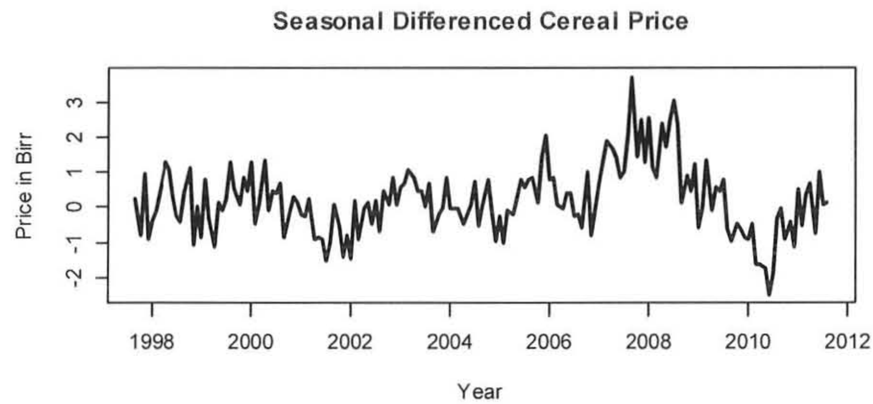

Fig. 5. Seasonal Differenced of Cereal Price series.

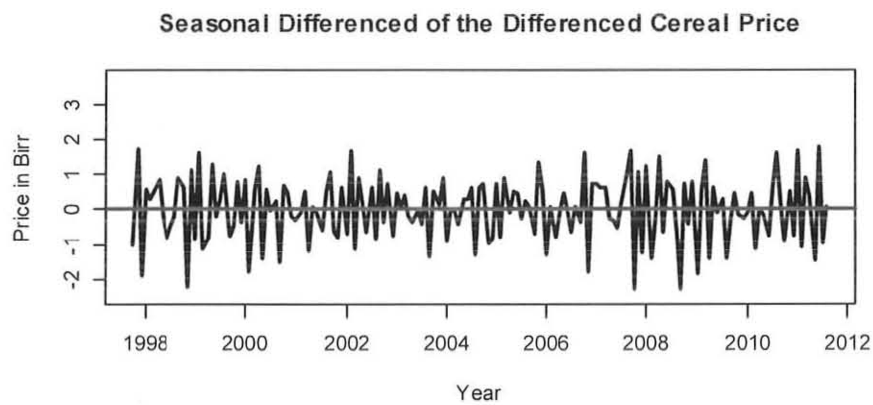

Fig. 6. Seasonal Differenced of the differenced Cereal Price Series

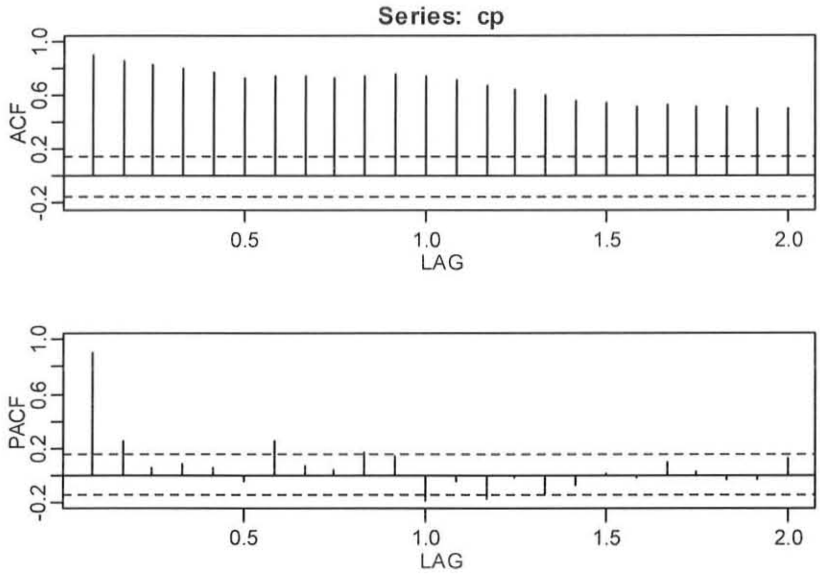

Fig. 7. Correlogram of cereal Price (Autocorrelation above and Partial ACF Below)

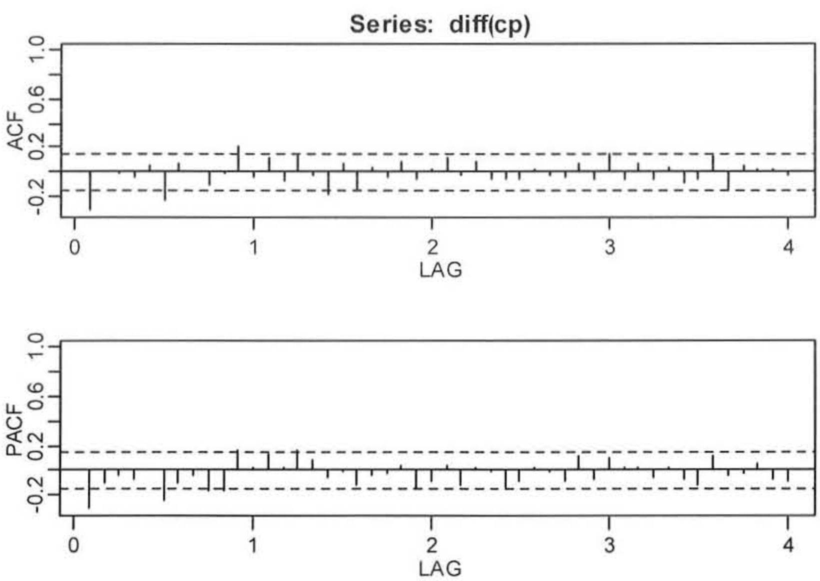

Fig. 8. Correlogram of Differenced cereal Price (Autocorrelation function above and Partial ACF Below)

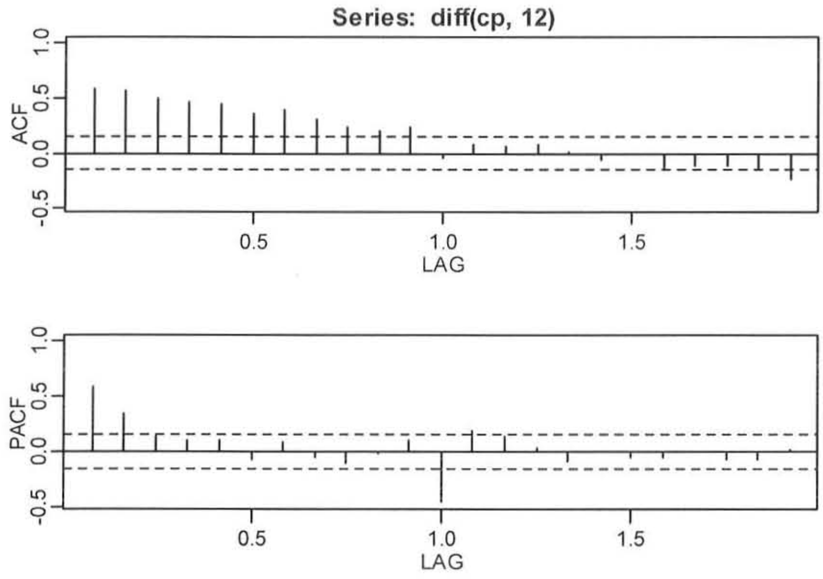

Fig. 9. Correlogram of Seasonal Differenced cereal Price (ACF above and Partial ACF Below)

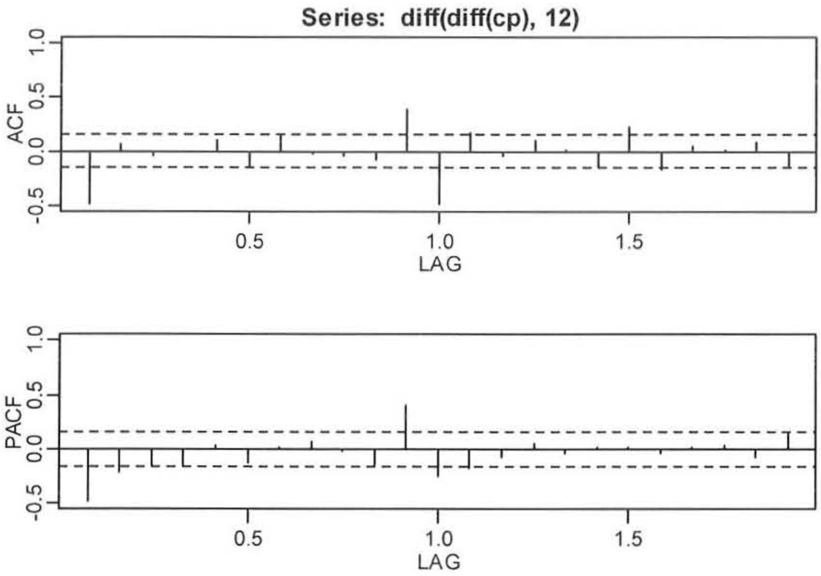

Fig. 10. correlogram of Seasonal Differenced the Differenced cereal Price (ACF above and Partial ACF Below)

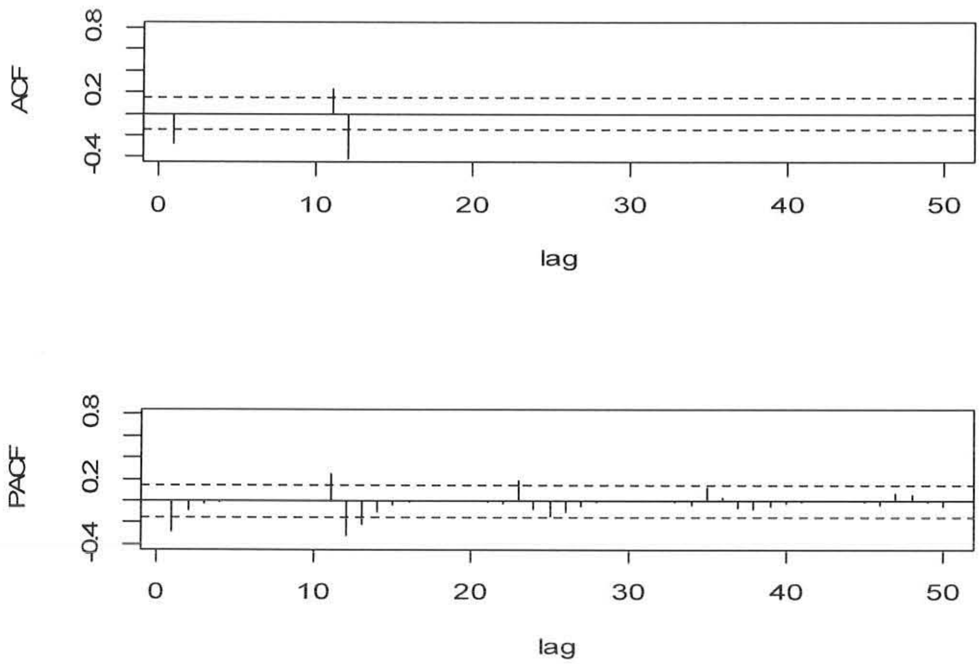

Fig. 11. Theoretical graph of with same simple alpha 1 and seasonal alpha 1 as the above.

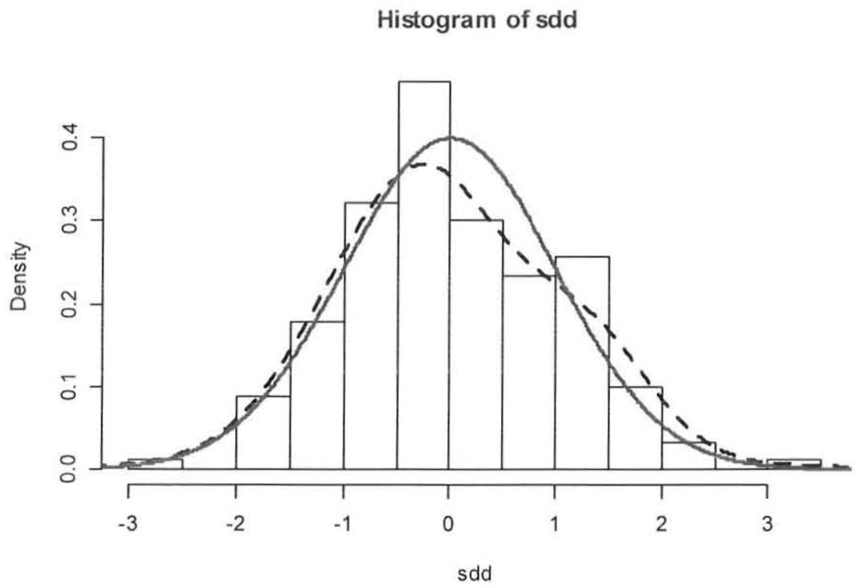

Fig. 12. Histogram of the residuals from SARIMA(0, 1, 1)*(0, 1, 1) fit on Cereal Price.
---------Kernel smoothing with bandwidth 0.5

______Normal ~N(0,1)

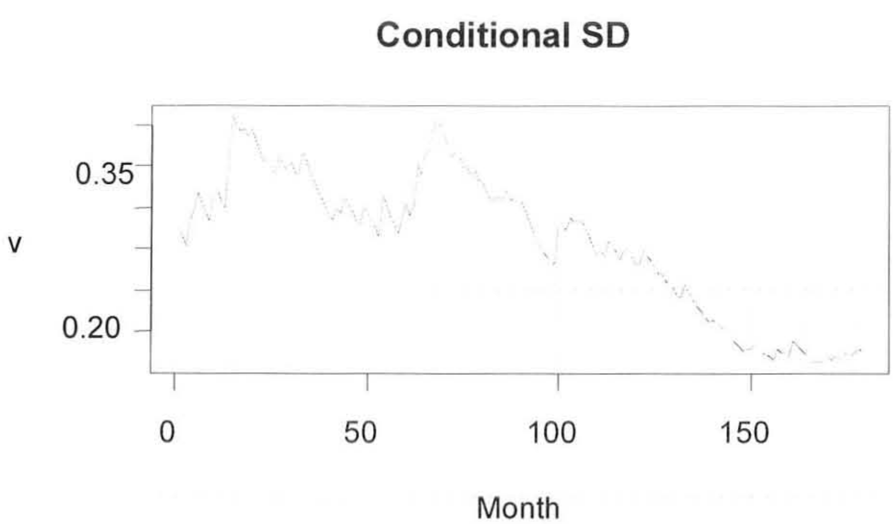

Fig. 13. Conditional Standard Deviation

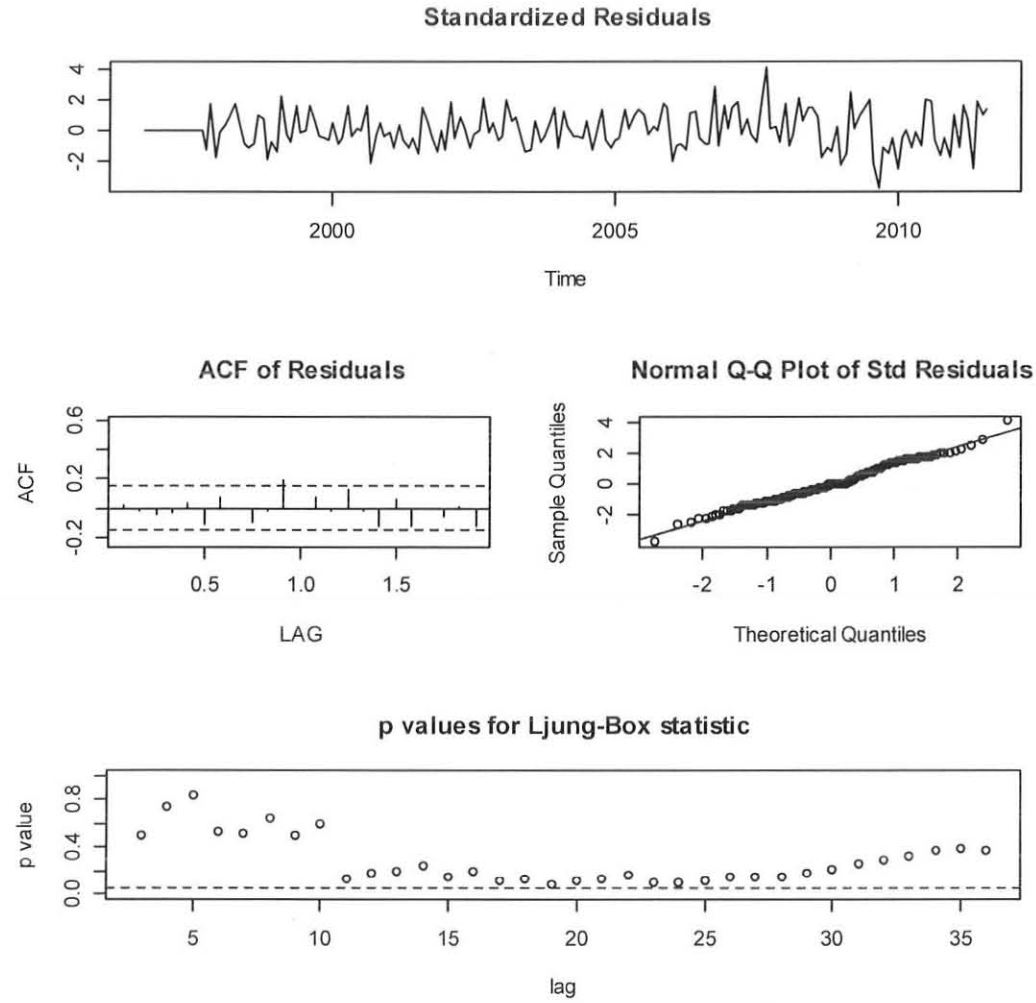

Fig. 14. Diagnostics of the residuals from SARIMA (0, 1, 1)*(0, 1, 1) fit on Cereal Price.

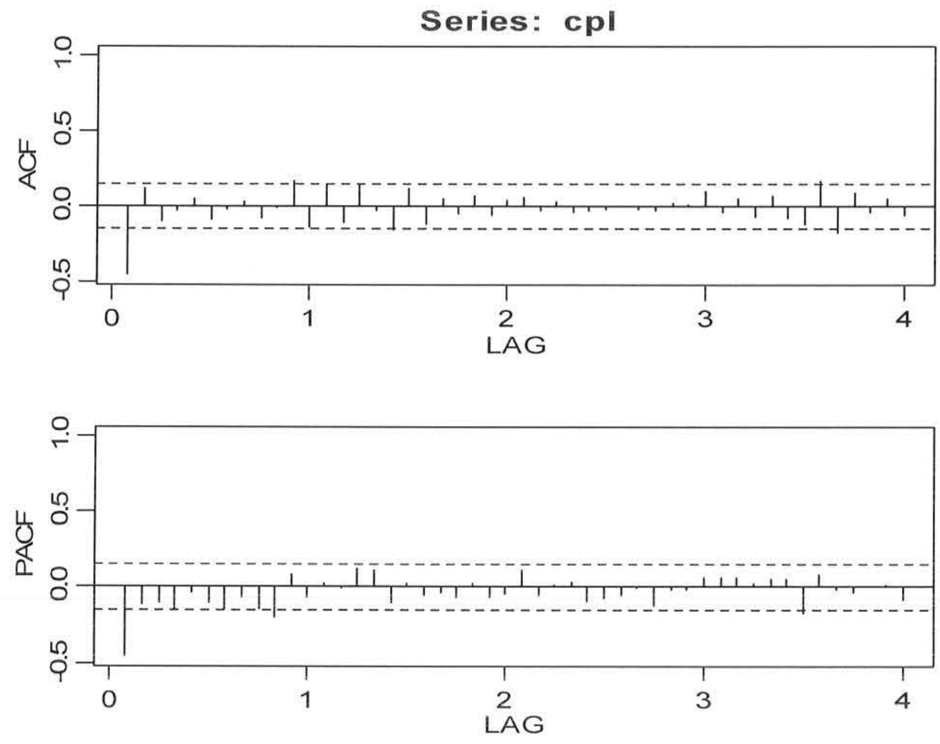

Fig. 15. Correlogram of Log return Cereal Price (ACF above and Partial ACF Below)

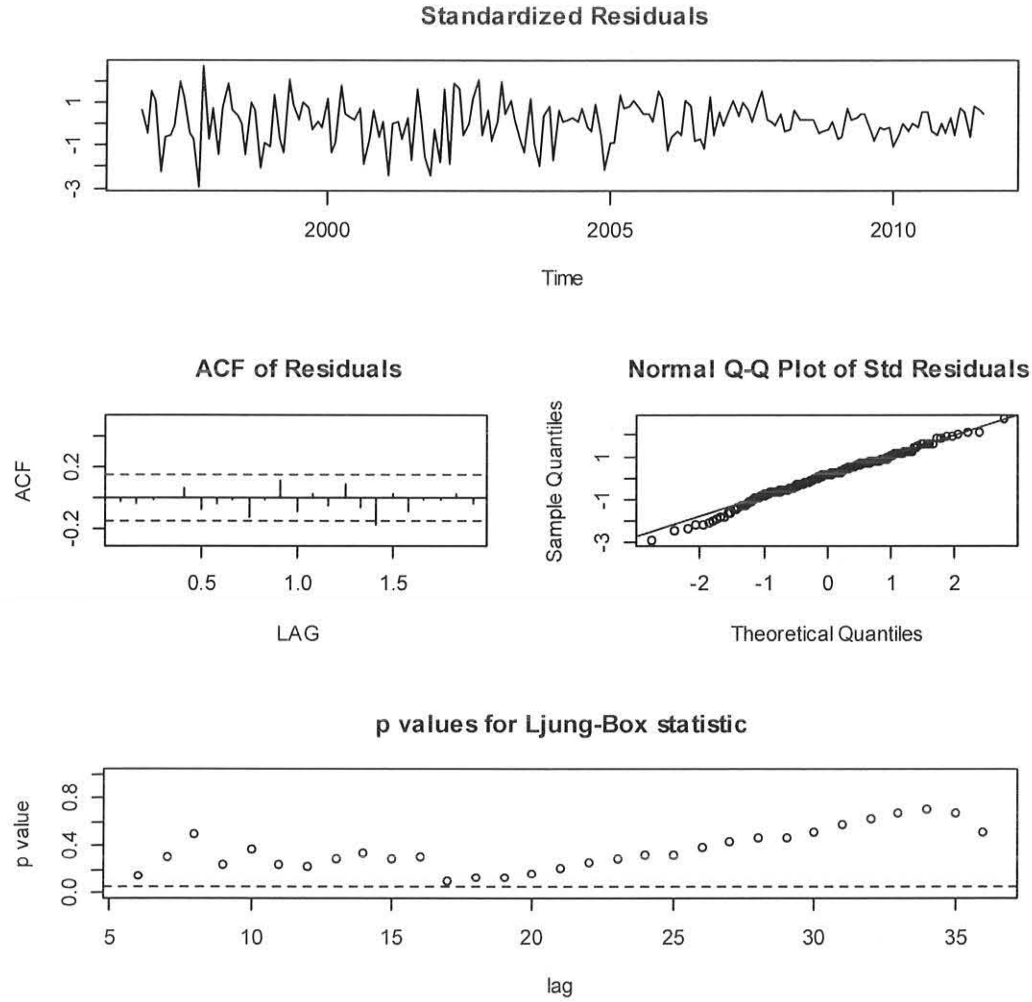

Fig.16. Diagnostics of the residuals from ARMA (3, 2) fit for the log return cereal price

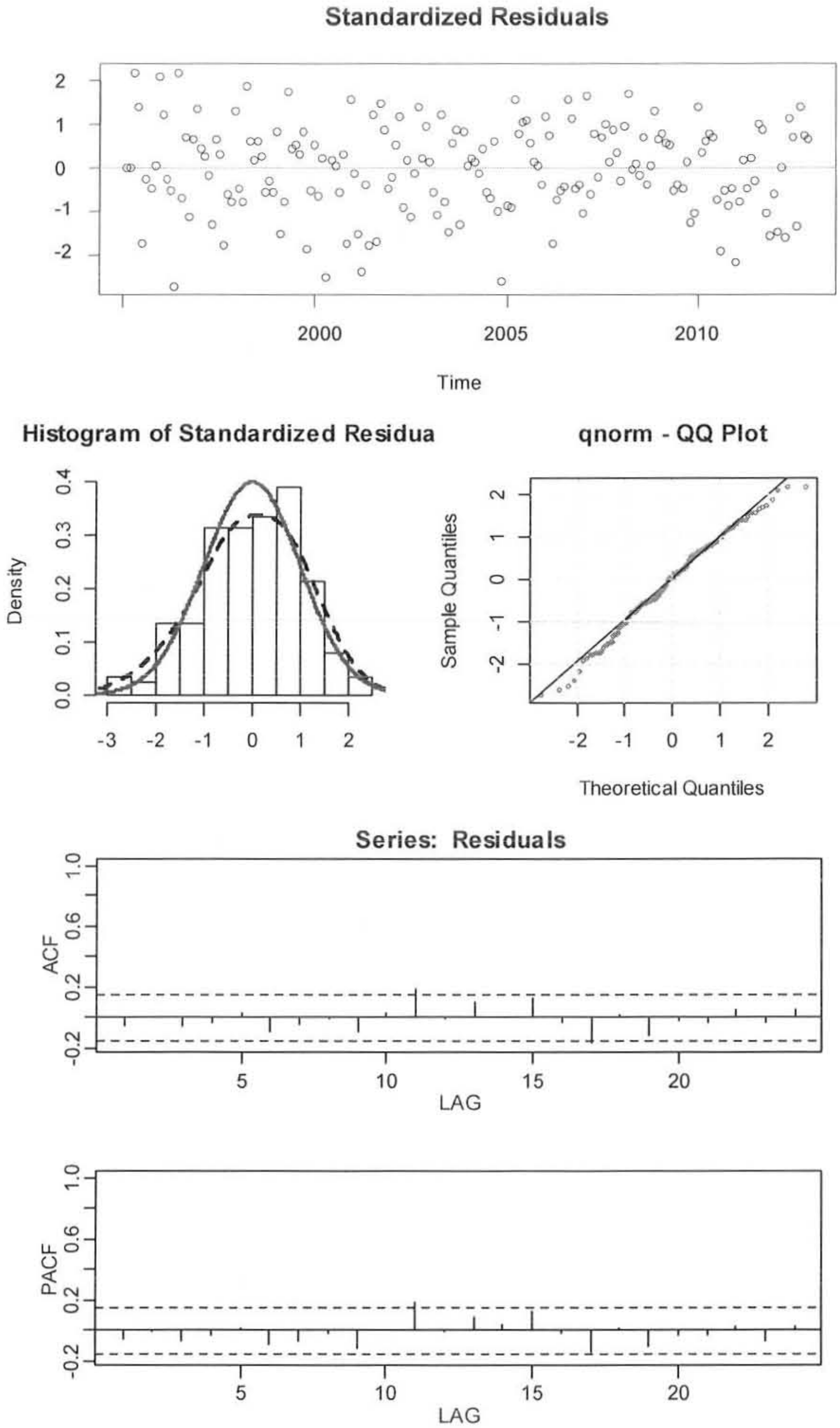

Fig.17. Diagnostics of the residuals from ARMA (2, 1)+ GARCH(1, 1) fit on Return of Cereal Price (Upper Standardized Residuals, next is Histogram and QQ-plot of the Residuals followed by ACF and PACF of the residuals).