EFFICIENCY EVALUATION IN MODELLING STOCK DATA USING ARCH AND BILINEAR MODELS

ADOLPHUS WAGALA

A Thesis Submitted To The Graduate School In Partial Fulfillment For The Requirements Of The Master Of Science Degree In Statistics Of Egerton University

EGERTON UNIVERSITY

OCTOBER 2008

DECLARATION AND RECOMMENDATION

Declaration

This thesis is my original work and has not been presented to any other institution for award of any degree.

Signature---------------------- Date--------------------------

Adolphus Wagala

SM12/1479/05

Recommendations

This thesis has been submitted with our approval as the supervisors.

Signature ------------------------------ Date-------------------------------

Prof. Dankit K. Nassiuma,

Kabarak University-Nakuru.

Signature ------------------------------ Date-------------------------------

Dr. Ali S. Islam,

Egerton University-Njoro.

COPYRIGHT

DEDICATION

To

My dad William and late mum Abigail.

ACKNOWLEDGEMENT

I wish to thank the Lord God Almighty who has bestowed upon me good health, sane mind, strength, patience and sufficient grace to see me through this study. All the honor and glory be unto you oh Lord!

I am grateful to my lecturers and supervisors: Prof. Nassiuma and Dr. Ali for their commitment through out my study and their patience to go through the earlier manuscripts of this work plus the many resourceful suggestions thereafter that opened my eyes to see deeper statistically. Much thanks also to Dr. Mwangi of the Mathematics department who introduced me to time series and has been of much support and encouragement by providing literature materials.

To my family and in particular my sisters Nancy, Dorita, Lillian and my cousin Isaya Jakawaya for moral and financial support, I say thank you. Special appreciation to my dad William and step-mum Isabella, I thank God for you and I am grateful for your encouragements, inspirations, prayers and financial assistance you offered to me. You are simply the best parents one could ever have. Finally, I also wish to acknowledge my friend Juliana Chidumu, my classmates: Mike, Anastacia, Mwangasha, Esha and Thiga for the moral support and encouragements while in the turbulent waters of academics.

ABSTRACT

Modelling of stock market data has witnessed a significant increase in literature over the past two decades. Focus has been mainly on the use of the ARCH model with its various extensions due to its ability to capture heteroscedasticity prevalent in the financial and monetary variables. However, other suitable models like the bilinear models have not been exploited to model stock market data so as to determine the most efficient model between the ARCH and bilinear models. The underlying problem is that of identifying the most efficient model that can be applied to stock exchange data for forecasting and prediction. The purpose of this study was to determine the most efficient model between the two models namely, ARCH and bilinear models when applied to stock market data. The data was obtained from the Nairobi Stock Exchange (NSE) for the period between 3rd June 1996 to 31st December 2007 for the company share prices while for the NSE 20-share index data was for period between 2nd March 1998 to 31st December 2007.The share prices for three companies; Bamburi Cement, National Bank of Kenya and Kenya Airways which were selected at random from each of the three main sectors as categorized in the Nairobi Stock Exchange were used. Specifically, the different extensions of ARCH-type models were utilized with ARMA and bilinear models for modelling the weekly mean of the chosen data set. The model efficiency was determined based on the minimal mean squared error (MSE). The results show that the Bilinear-GARCH model with the normal distribution assumption and the AR-Integrated GARCH (IGARCH) model with student's t-distribution are the best models for modelling volatility in the Nairobi Stock Market data. The results also indicate that the volatility in Nairobi Stock Exchange is statistically significant and persistent with the positive return innovations having a greater impact than the negative ones. This implies that the leverage effect experienced in most developed countries is not applicable to Nairobi Stock Market. The results obtained are significant for planning, prediction and management of investments on shares in the Nairobi Stock Exchange. The chosen models are also helpful for decision making especially by the investors, stockbrokers and financial advisors regarding the trading in shares at the Nairobi Stock Exchange.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

AR	Autoregressive
ARMA	Autoregressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARCH	Autoregressive Conditional Heteroscedasticity
AIC	Akaike Information Criteria
ACF	Autocorrelation Function
BL	Bilinear
BIC	Bayesian Information Criteria
CLSE	Conditional Least Squares Estimate
EAA	East African Airways Corporation
FPE	Final Prediction Error
FTA	Financial Times Actuaries
EXPAR	Exponential autoregressive
FAR	Fractional Autoregressive
GARCH	Generalized Autoregressive Conditional Heteroscedasticity
GED	General Error Distribution
IID	Independent Identically Distributed
JB	Jarque-Bera statistics for normality
LSE	Least Squares Estimate
LR	Log likelihood Ratio test.
MA	Moving Average
MLE	Maximum Likelihood Estimate
MSE	Mean squared error
NEAR	Newer exponential autoregressive
NBK	National Bank of Kenya Limited
NSE	Nairobi Stock Exchange
NSSF	National Social Security Fund
NYSE	New York Stock Exchange
PACF	Partial Autocorrelation Function
S & P	Standard and Poor

RCA	Random coefficient autoregressive
SE	Standard Error
WN	White Noise

CHAPTER ONE

INTRODUCTION

1.1 Background

Stock market volatility is one of the most important aspects of financial market developments, providing an important input for portfolio management, option pricing and market regulation (Poon and Granger, 2003). An investor's choice of portfolio is intended to maximize his expected return subject to a risk constraint, or to minimize his risk subject to a return constraint. An efficient model for forecasting of an asset's price volatility provides a starting point for the assessment of investment risk. To price an option, one needs to know the volatility of the underlying asset. This can only be achieved through modelling the volatility. Volatility also has a great effect on the macro-economy. High volatility beyond a certain threshold will increase the risk of investor losses and raise concerns about the stability of the market and the wider economy (Hongyu and Zhichao, 2006).

Financial time series modelling has been a subject of considerable research both in theoretical and empirical statistics and econometrics. In recent literature, numerous parametric specifications of ARCH models have been considered for the description of the characteristics of financial markets. Engle (1982) introduced the Autoregressive Conditional Heteroscedasticity (ARCH) for modelling financial time series while Bollerslev (1986) came up with the Generalized ARCH (GARCH) to parsimoniously represent the higher order ARCH model. Owing to the empirical success of the ARCH and GARCH models, researchers have concentrated on the two models due to their ability to model heteroscedasticity. There is a significant amount of research on volatility of stock markets of developed countries. Gary and Mingyuon (2004) applied the GARCH model to the Shanghai Stock Exchange while Bertram (2004) modelled Australian Stock Exchange using ARCH models. Other studies include, Baudouhat (2004) who utilized the GARCH model in analyzing the Nordic financial market integration, Walter (2005) applied the structural GARCH model to portfolio risk management for the South African equity market. Hongyu and Zhichao (2006) forecasted the volatility of the Chinese stock market using the GARCH-type models.

The Sub-Saharan Africa has been under-researched as far as volatility modelling is concerned. Studies carried out in the African stock markets include, Frimpong and Oteng-Abayie (2006) who applied GARCH models to the Ghana Stock Exchange, Brooks *et al*., (1997) examined the effect of political change in the South African Stock market, Appiah-Kusi and Pascetto (1998) investigated the volatility and volatility spillovers in the emerging markets in Africa. More recently, Ogum *et al.,* (2006) applied the EGARCH model to the Kenyan and Nigerian Stock Market returns. From the available literature, the NSE just like other Sub Saharan Africa Equity Markets has been clearly under-researched as far as market volatility is concerned and therefore this study contributes to the limited literature available on the Nairobi stock market.

Mandelbrot (1963) utilized the infinite variance distributions when considering the models for stock market price changes. Fama (1965) when modelling stock market prices attributed their discrepancies to the possibility of the process having stable innovations and thus fitted an adequate model on this basis. These developments in financial econometrics suggest the use of nonlinear time series models in analyzing the stock market prices and the expected returns.

The focus of financial time series modelling has mainly been on the ARCH model and its various extensions thereby ignoring the other suitable nonlinear models like the bilinear class of models. As a matter of fact, the subject of the efficiency of the models for financial modelling has received little attention as far as econometric modelling is concerned. This study therefore aims at finding the most efficient model from amongst the nonlinear models namely, bilinear models and the autoregressive conditional heteroscedasticity models.

1.2 The Nairobi Stock Exchange

The Nairobi Stock Exchange (N.S.E) was formed in 1954 as a voluntary organization of Stock brokers. It is now one of the most active capital markets and a model for the emerging markets in Africa in view of its high returns on investments and a well developed market structure (Ogum *et al.*, 2005). The Nairobi Stock Exchange is a market place where shares (also known as equities) and bonds (also known as debt instruments) are traded. The ordinary shares are also known as variable income securities since they have no fixed rate of dividend payable, as the dividend is dependent upon both the profitability of the company and what the board of directors decides. The bonds are also known as the fixed income securities and include Treasury and Corporate

Bonds, preference shares, debenture stocks; these have a fixed rate of interest/dividend, which is not dependent on profitability.

As a capital market institution, the Stock Exchange plays an important role in the process of economic development. The major role that the Nairobi stock exchange has played, and continues to play is that it promotes a culture of thrift, or saving. The NSE also assists in the transfer of savings to investment in productive enterprises thereby utilizing the money that would otherwise lie idle in savings. This helps in avoiding economic stagnation. It also assists in the rational and efficient allocation of capital. An efficient stock market sector will have the expertise, the institutions and the means to prioritize access to capital by competing users so that an economy manages to realize maximum output at the least cost. The NSE promotes higher standards of accounting, resource management and transparency in the management of business. In addition, the stock exchange improves the access of finance to different types of users by providing the flexibility for customization. Finally, the stock exchange provides investors with an efficient mechanism to liquidate their investments in securities. The investors are able to sell out what they hold, as and when they want. This is a major incentive for investment as it guarantees mobility of capital in the purchase of assets.

1.3 Statement of the problem

The last two decades have witnessed an increase in the modelling of stock market data using the ARCH models and its various extensions. However, efficiencies of competing models such as the bilinear models has so far not been determined for modelling equity market data. This study therefore seeks to determine the most efficient model for application to the Nairobi stock market data from the two classes of models.

1.4 Objectives of the study

1.4.1 Main objective

The overall objective of this study was to determine the most efficient model from the two classes of models namely; ARCH, bilinear and bilinear-ARCH.

1.4.2 Specific objectives

The weekly mean for the NSE 20-share index and the average weekly share prices for the following companies; National Bank of Kenya, Kenya Airways and Bamburi Cement Ltd were used to achieve the following specific objectives;

i) To model the Nairobi Stock Exchange stock data using ARCH-type models.

ii) To apply Bilinear models in fitting the Nairobi Stock Exchange data.

iii) To compare the efficiency of the two classes of models and make recommendations regarding the best model for modelling volatility.

1.5 Justification

The establishment of an efficient stock market is indispensable for an economy that is keen on utilizing scarce capital resources to achieve its economic growth. It is therefore prudent to determine the most efficient model that will help in predicting volatility which in turn is important in pricing financial derivatives, selecting portfolios, measuring and managing risks more accurately. The efficient model will not only be useful in long term forecasting and short term prediction but also in helping the investors on decisions regarding which shares to sell, hold or buy.

1.6 Definition of terms used

Bonds are financial instruments that serve as an "i owe you"; an investor loans an issuer, and returns are fixed and guaranteed, no voting rights and no benefits from exceptional performance by a company.

Shares are financial instruments where one acquires ownership stakes of a company. Returns are neither fixed nor guaranteed. One acquires voting rights and shares the company's profits and losses.

Volatility is the variance or variation of a given time series data.

Returns are transformations given by $X_t = \ln(P_t) - \ln(P_{t-1})$, where X_t and P_t represents the return and weekly average value for each series respectively.

CHAPTER TWO

LITERATURE REVIEW

2.1 Financial Data

Financial time series data often exhibit some common characteristics. Fan and Yao (2003) summarizes the most important features of financial time series as; the series tend to have leptokurtic distribution, i.e they have heavy tailed distribution with high probability of extreme values. In addition, changes in stock prices tend to be negatively correlated with changes in volatility, that is; volatility is higher after negative shocks than after positive shocks of the same magnitude. This is referred to as the leverage effect. The sample autocorrelations of the data are small whereas the sample autocorrelations of the absolute and squared values are significantly different from zero even for large lags. This behaviour suggests some kind of long range dependence in the data. The distribution of log returns over large periods of time (such as a month, a half a year, a year) is closer to a normal distribution than for hourly or daily log-returns. Finally, the variances change over time and large (small) changes of either sign tend to be followed by large (small) changes of either sign (Mandelbrot, 1963). This characteristic is known as volatility clustering. These are facts characterizing many economic and financial variables.

2.2 Models for Stock Market Data

Researchers have applied different models to the stocks data from time to time. Mandelbrot (1963) utilized the infinite variance distributions when considering the models for stock market price changes. Fama (1965) similarly pointed out initially, their application in cases of economics particularly in modelling stock market prices. Fama *et al.,* (1969) used a random walk to model the price changes. Andrew and Whitney (1986) tested the random walk hypothesis for weekly stock market returns by comparing the variance estimators. Here the random walk model was strongly rejected. Omosa (1989) applied the ARIMA model to the NSE data and used the models for forecasting. Muhanji (2000) studied the efficiency of the Nairobi Stock Exchange and concluded that the NSE had a weak form of efficiency implying that the market is efficient at a particular period and becomes inefficient at another time.

In recent studies, various specifications of ARCH models have been considered for the description of the characteristics of financial markets. Some studies in which ARCH-type models were utilized include; Gary and Mingyuon (2004) who applied the GARCH model to Shanghai Stock Exchange, Bertram (2004) modelled Australian Stock Exchange using ARCH models and Baudouhat (2004) used the GARCH model in analyzing the Nordic financial market integration. In addition, Curto (2002) employed the GARCH model to explain the volatility of the Portuguese equity market, Walter (2005) applied the structural GARCH model to portfolio risk management while Frimpong and Oteng-Abayie (2006) modelled the Ghana Stock Exchange volatility using the GARCH models. More recently, Ogum *et al*., (2006) applied EGARCH model to the Kenyan and Nigeria daily stock market data.

Simple regression models have also been utilized in modelling stock market data. Bodicha (2003) applied regression models to the NSE data and found out that the regression models are only appropriate for short term prediction and not for long term forecasting. The analysis of the general linear regression model forms the basis of every standard econometric model. Mills (1999) applied the simple linear relationship in modelling the expected risk and return in holding a portfolio while Gujarati (2003) applied econometric modelling to the NYSE data.

2.3 Linear Time Series Models

Let ς be a subset of the real numbers. For every $t \in \varsigma$, let $X_t(\omega)$ be a random variable defined on a probability space $\{\Omega : \omega \in \Omega\}$; then the stochastic process $\{X_t(\omega) : t \in \varsigma\}$ is called a time series. Here, $\{X_t, t = 0, \pm 1, ...\}$ is a realization at time t for any given ω. A time series model for the observed data $\{X_t\}$ is specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables $\{X_t\}$. A time series model accounts for patterns in the past movements of a variable data and that information is used to control and predict its future movements.

The autoregressive moving average (ARMA) processes are the most widely known and applied set of linear time series models. For the ARMA(p,q) process, the observation X_t is linearly related to the p most recent observations ($X_{t-1}, ..., X_{t-p}$), q most recent forecast errors ($\varepsilon_{t-1}, ..., \varepsilon_{t-q}$) and the current disturbance ε_t by the relation:

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \,,\ \varepsilon_t \sim WN(0, \sigma^2) \qquad 2.3.1$$

where ϕ_i and θ_i are model parameters. The equation in 2.3.1 represents an ARMA (p,q) process. Alternatively, using the backshift operator, an ARMA (p,q) process is represented as $\Phi(B) w_t = \Theta(B)\varepsilon_t$,where B is a backshift operator such that $BX_t=X_{t-1}$, $\{\varepsilon_t\}$ is a sequence of uncorrelated random variables with zero mean and variance σ^2 .The polynomials:

$$\Phi(B) = 1 + \phi_1 B + \phi_2 B^2 + ... + \phi_p B^p \qquad 2.3.2$$

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + ... + \theta_q B^q \qquad 2.3.3$$

represent the autoregressive and the moving average operators of order p and q respectively. The coefficients in $\Phi(B)$ and $\Theta(B)$ represent some of the model parameters.

A special case of the ARMA(p,q) process is known as the autoregressive process which date back to the work of Yule (1927) where he developed the first order autoregressive process (AR(1)) which is given by the relation $X_t = \phi X_{t-1} + \varepsilon_t$ where ϕ is a model parameter and $\varepsilon_t \sim WN(0, \sigma^2)$. In general, the AR(p) process is represented as

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \varepsilon_t \qquad 2.3.4$$

where ϕ_i are constants and $\varepsilon_t \sim WN(0, \sigma^2)$

Another type of linear time series model is known as the moving average (MA) process which was developed by Slutzky (1937). The functional form for the first order moving average process (MA (1)) process is given by the equation $X_t = \theta\varepsilon_{t-1} + \varepsilon_t$, where θ is model parameter and $\varepsilon_t \sim WN(0, \sigma^2)$. The general representation of an MA(q) process is given as

$$X_t = \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} + \varepsilon_t \qquad 2.3.5$$

where θ_i are model parameters and $\varepsilon_t \sim WN(0, \sigma^2)$. Here ε_t is not observable.

When the stationarity condition is assumed, i.e. when the mean, variance and autocovariances of a process are invariant under time translations, then the process is modelled using the ARMA models. The ARMA models have been applied to modeling the UK interest rates, returns on the

FTA All share index, S&P 500 stock index and the dollar/sterling exchange rate (Mills, 1999). Fan and Yao (2003) also modelled the German Egg prices using the Autoregressive Integrated Moving Average (ARIMA) models. In addition Gujarati (2003) applied the Box Jenkins approach to model the money supply in the United States.

The stationarity condition restricts the mean and the variance to be constant and requires the autocovariances to depend only on the time lag. However, this is not true in many financial time series, they are certainly non stationary and have a tendency to exhibit time changing means and variances. Box and Jenkins (1976), suggested differencing as a means of transforming a non-stationary ARMA (p,q) process into a stationary ARMA(p,q) process known as the autoregressive integrated moving average (ARIMA) process. This is applicable to the financial time series model building.

2.3.1 The Box and Jenkins Approach to ARIMA Modelling

Box and Jenkins (1976) proposed three major stages in ARIMA modelling, namely; identification, estimation, diagnostic checking and forecasting. The approach is as follows;

a) Identification Stage

In this stage the model selection is done. The simplest and most basic tool for identification is the time series plot, which is simply a graph in which data values are arranged sequentially in time. A plot is an effective way of quickly perceiving the evolution of a single or a group of time series. The plots are useful in detecting outliers, the seasonal, cyclic and the trend components of a time series data.

The next criterion for identification is the Autocorrelation Function (A.C.F). The autocorrelation is given by,

$$\rho = \frac{Cov(x_t, x_{t-k})}{[v(x_t).v(x_{t-k})]^{1/2}} = \frac{\gamma_k}{\gamma_0} \qquad 2.3.6$$

The autocorrelations considered as a function of k is referred to as the autocorrelation function ACF or sometimes the correlogram. The ACF plays a major role in modelling the dependencies among observations. It indicates, by measuring the extent to which one value of the process is correlated with the previous, the length and strength of the ‘memory’ of the process. In general, the correlation between two random variables is often due to both variables being correlated with a third one. In the context of time series, a large portion of the correlation between X_t and X_{t-k} can be due to the correlation between $X_{t-1}, X_{t-2}, \ldots, X_{t-k+1}$. To adjust for this correlation, the partial autocorrelation function (PACF) may be calculated. The PACF measures the additional autocorrelations between X_t and X_{t-k} after adjustments have been made for intervening lags. The ACF and PACF are useful in the identification of orders in the ARMA processes. An AR(p) process has a declining ACF, exponentially decaying to zero and the PACF is zero for lags greater than p. An MA(q) process on the other hand has an ACF that is zero for lags greater than q and PACF that declines exponentially. However, if the decay in the ACF starts after a few lags lags then the process could be an ARMA(p,q). If the series is non-stationary, then it is transformed by differencing to attain stationarity. The decision about differencing is based on the visual examination of the correlogram.

The adequacy of the fitted model or an indication of potential improvements is determined using the following diagnostic checks. These checks include the Final Prediction Error (FPE) criterion which was developed by Akaike (1969) for selecting the appropriate order of an AR process. The idea here is to select the model for $\{X_t\}$ in such a way as to minimize the one-step mean squared error when the model fitted to $\{X_t\}$ is used to predict an independent realization $\{Y_t\}$ of the same process that generated $\{X_t\}$. The FPE for order p is given as,

$$\text{FPEp} = \hat{\sigma}^2 \frac{n+p}{n-p} \qquad 2.3.7$$

To apply the FPE criterion, p is chosen such that it minimizes the value of FPEp.

Another criterion is the Akaike information criterion (AIC) which is more generally applicable for model selection. The AIC was developed by Akaike (1974).The AIC is defined by:

$$\text{AIC (p,q)} = \log \hat{\sigma}^2 + 2(p+q)T^{-1} \qquad 2.3.8$$

where $\hat{\sigma}^2$ is the estimate of the error variance of an ARMA (p,q) and T is the total number of observations. For fitting autoregressive models, the AIC has a tendency to overestimate p, i.e. for AIC an over parameterized model is more likely to be obtained.

Schwarz (1978) suggested the Bayesian Information Criteria (BIC) defined as

$$\text{BIC (p, q)} = \log \hat{\sigma}^2 + (p+q)T^{-1} \log T \qquad 2.3.9$$

The BIC attempts to solve the over parameterization of AIC and is thus strongly consistent, in that it determines the true model asymptotically.

These procedures entail the comparison of the sample values with the corresponding theoretical values.

b) Estimation of Parameters

Once a tentative formulation of the time series models has been accomplished, estimation of model parameters follows. The estimation techniques used include the Yule-Walker estimation criterion. The Yule-Walker estimates are obtained by matching patterns in the sample autocorrelations with theoretical patterns. For instance, consider AR(P) process,

$$X_t = \phi(B)X_t + \varepsilon_t \qquad 2.3.10$$

Here $E\left(X_{t-i}(X_t - \sum_{i=0}^{p} \phi_i X_{t-1}) \right) = E(X_{t-i}\varepsilon_t)$ for i=0,…,p $\Leftrightarrow$ $\gamma(0) - \phi'\gamma_p = \sigma^2$ and $\gamma_p - \Gamma_p \phi = 0$

where $\phi = (\phi_1, ..., \phi_p)'$. Now, choose ϕ so that $\gamma = \hat{\gamma}$. The Yule Walker equations can be written in a matrix form as $\rho = R_p \phi$ where R_p is a covariance matrix. The Yule Walker equations can also be presented as $\begin{cases} \hat{\Gamma}_p \hat{\phi} = \hat{\gamma}_p \\ \hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}'\hat{\gamma}_p \end{cases}$ where $\mathbf{\Gamma}_p$ is the covariance matrix $[\gamma(1-i)]_{j=1}^{p}$ and $\underline{\gamma}_p = [\gamma(1), \gamma(2), ..., \gamma(p)]$. This further leads to $\hat{\phi} = \hat{\Gamma}_p \underline{\gamma}_p$ where $\hat{\Gamma}_p = R_p$ and $\hat{\phi}$ is the Yule-Walker estimate.

The next estimation method is the maximum likelihood procedure. This method is more appropriate for small samples and especially when the parameter values approach the invertible boundaries. In the maximum likelihood estimation (MLE) of time series models, two types of MLEs are computed. The first type is based on maximizing the conditional log-likelihood function. These estimates are the conditional MLEs defined by;

$\hat{\phi}_{c.m.l.e} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln f(X_t / I_{t-1}, \theta)$. The second type is based on maximizing the exact log-likelihood function. These exact estimates are called exact MLEs, and defined by;

$$\hat{\phi}_{m.l.e} = \arg\max_{\theta} \sum_{t=p+1}^{T} \ln f(Y_t / I_{t-1}, \theta) + \ln f(y_1,..., y_p; \theta) \qquad 2.3.11$$

For stationary models, $\hat{\phi}_{c.m.l.e}$ and $\hat{\phi}_{m.l.e}$ are consistent and have the same limiting normal distribution. In finite samples however, $\hat{\phi}_{c.m.l.e}$ and $\hat{\phi}_{m.l.e}$ are generally not equal and may differ by a substantial amount if the data are close to being non-stationary or non-invertible.

Consider an AR (1) process $X_t = c + \phi X_{t-1} + \varepsilon_t$ where $\varepsilon_t \sim WN\ (0, \sigma^2)$ t=1,2,3,…,T and c a constant. The exact log-likelihood function is then;

$$\ln L(\phi / y) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\ln(\frac{\sigma^2}{1-\phi^2}) - \frac{1-\phi^2}{2\sigma^2}(x_1 - \frac{c}{1-\phi})^2 - \frac{(T-1)}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}\sum_{t=2}^{T}(x_t - c - \phi x_{t-1})^2$$

The exact log-likelihood function is a non-linear function of parameters ϕ and so there is no closed form solution for the exact MLEs.

The next estimation criterion is the conditional least squares estimation (CLSE) method. These estimates are easier to compute compared to MLE. This procedure entails minimizing the sum of squares $Q = \sum_{t=1}^{n} \varepsilon_t^2$. Thus for an AR (P) process, $Q = \sum_{t=1}^{n}(X_t - \sum_{i=1}^{p}\phi_i X_{t-i})^2$. Now, for *p*=1,

$$\hat{\phi}_1 = \frac{\sum_{t=1}^{n} X_t X_{t-1}}{\sum_{t=1}^{n} X_{t-1}^2} \qquad 2.3.12$$

The optimal estimation technique is also useful in estimation and was initially studied by Godambe (1960) in comparison to the MLE. The optimal estimates for stochastic process are obtained as follows: Consider an AR (1) process:

$$X_t = \phi X_{t-1} + \varepsilon_t \text{ where } \varepsilon_t \sim WN\ (0,\sigma^2) \qquad 2.3.13$$

Let $h_t = X_t - \phi X_{t-1}$ be a linear function such that $E(h_t / \Im_{t-1}) = 0$ where $\Im_{t-1}$ is the sigma algebra up to time *t-1* i.e. σ-algebra on which ($x_1, x_2, \ldots, x_n$) is defined.

The optimal estimating function (Godambe, 1985) is obtained as;

$$g_t^* = \sum_{t=1}^{n} a_{t-1}^* h_t \quad 2.3.14$$

where $$a_{t-1}^* = \frac{E(\frac{\partial h_t}{\partial \phi} / \Im_{t-1})}{E(h_t^2 / \Im_{t-1})} \quad \Rightarrow g_t^* = \frac{-\sum X_{t-1}}{\sigma^2}(X_t - \phi X_{t-1}) \quad 2.3.15$$

and by setting $g_t^* = 0$ leads to the optimal estimate of ϕ as

$$\hat{\phi} = \frac{\sum_{t=1}^{n} X_t X_{t-1}}{\sum_{t=1}^{n} X_{t-1}^2} \quad 2.3.16$$

which is the same as the CLSE.

c) Forecasting

When a time series model has been identified and parameter estimates obtained, it can be used for forecasting. The most common forecasting criterion is based on minimizing the mean square error, that is, for the process X_t, the aim is to obtain the forecast X_t, such that $E[X_{t+1} - \hat{X}_{t+1}]^2$ is minimized. ARIMA modelling is popular because of its success in forecasting. In many cases, the forecasts obtained by this method are more reliable than those obtained by econometric modelling (Gujarati, 2003).

Mathematically, linear time series models are the simplest type of difference equations and a complete theory of Gaussian sequences are readily understood. The theory of statistical inference is also most developed for linear Gaussian models. The computation time required for obtaining a parsimonious ARMA model for the data is well within the reach of most practitioners. These models have been reasonably successful in analysis, forecasting and control of various time series data. Linear time series models are easy to compute and manipulate. The assumption of stationarity makes modelling in them simple. The estimation procedures for the linear models are also not complicated.

Some of the limitations of linear time series models include the fact that linear difference equations do not permit stable periodic solutions independent of initial value. Having symmetric joint distributions, stationary linear Gaussian models are not ideally suited for data exhibiting strong asymmetry, i.e. they are dependent on the symmetric systems. The ARMA models are not ideally suited for data exhibiting sudden bursts of very large amplitude at irregular time epochs. Linear time series models do not capture series that exhibit cyclicity. In practice linear time series models are assumed to be Gaussian and thus have short tails. However, this is not true in most time series data especially the financial data which usually follow non-normal distributions. Linear models are also unable to utilize higher moments and assume only the first two moments while in some cases, there exist a third or even a fourth moment. In series that are time irreversible, linear models are of no use (Tong ,1990; Kantz and Schreiber, 2005; Zivot and Wang, 2005).

2.4 Non Linear Time Series Models

In trying to address the limitations of linear time series models, many non-linear time series models have been developed (Kantz and Schreiber, 2005; Zivot and Wang, 2005). Non-linear time series models can be used to model series that show cyclicity. They are also useful in series having outliers. Non-linear models are useful for higher moment utilization since they are able to capture higher moments. They are also useful in series that are time irreversible. Extreme non-stationarity especially in the variance makes non-linear models more appropriate. They are also useful in modelling data that are asymmetric (Tong, 1990).

In the recent years, a few non-linear time series models have been proposed. In many cases, the results are still incomplete and much research is going on at present. An example of a non-linear time series model is that of non-linear autoregression models. This class of models is motivated directly by dynamical system (Tong, 1990). The next class of non-linear models is the amplitude-dependent exponential autoregressive (EXPAR) models. These models were independently introduced by Jones (1976) and Ozaki and Oda (1978). The EXPAR models are useful in modelling ecological/population data (Ozaki, 1982), wolf's sunspot numbers (Haggan and Ozaki, 1981) and to a small extent, the economics data (Tong, 1990).

Another important class of non-linear time series models is the Fractional Autoregressive (FAR) models. This class of models has not so far been exploited conclusively to find their best area of application. Random coefficient autoregressive (RCA) models have been applied to areas such as, ecology/population (Nicholls and Quinn 1982) and Medical data (Robinson, 1978).There exist a subclass of RCA models with the marginal distribution that is exponential. They are known as the Newer exponential autoregressive (NEAR) models. The NEAR models were applied to Geophysics by Lawrence and Lewis (1985). The other class of non-linear time series models is the Threshold models which were introduced by Tong (1978).These models have a wide range of applications for instance in finance (Petruccelli and Davies, 1986; Wecker, 1981; Tyssedal and Tjøstheim ,1988), population dynamics (Stenseth *et al*., 1999), economics (Tiao and Tsay, 1994). They are also applicable in ecology/population data (Li and Liu, 1985). In addition, they have been used in geodynamics (Zheng and Chen, 1982) and also in neural science (Brillinger and Segundo, 1979).

2.5 Autoregressive Conditional Heteroscedasticity (ARCH) models

An ARCH process is a mechanism that includes past variances in the explanation of future variances (Engle, 2004). The Autoregressive property describes a feedback mechanism that incorporates past observations into the present while Conditionality implies a dependence on the observations of the immediate past and Heteroscedasticity means time-varying variance (volatility). These models were first introduced by Engle (1982) when modelling the United Kingdom inflation. In contrast to the ARMA models which focuses on modelling the first moment. ARCH models specifically take the dependence of the conditional second moments in modelling consideration. This accommodates the increasingly important demand to explain and to model risk and uncertainty in financial time series (Degiannakis and Xekalaki, 2004; Engle, 2004; Fan and Yao, 2003).

An ARCH process can be defined in terms of the distribution of the errors of a dynamic linear regression model. The dependent variable y_t is assumed to be generated by

$$y_t = x_t'\xi + \varepsilon_t \quad t = 1,\dots,\mathrm{T} \qquad 2.5.1$$

where x_t' is a kx1 vector of exogenous variables, which may include lagged values of the dependent variable and ξ is a kx1 vector of regression parameters. The ARCH model characterizes the distribution of the stochastic error ε_t conditional on the realized values of the set of variables $\psi_{t-1} = \{y_{t-1}, x_{t-1}, y_{t-2}, x_{t-2}, ...\}$. Specifically, Engle's (1982) model assumes

$$\varepsilon_t / \psi_{t-1} \sim N(0, h_t) \qquad 2.5.2$$

where

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_q \varepsilon_{t-q}^2 \qquad 2.5.3$$

with α_0>0 and $\alpha_i \geq 0, i = 1,\ldots,q$ to ensure that the conditional variance is positive.

An explicit generating equation for an ARCH process is

$$\varepsilon_t = \eta_t \sqrt{h_t} \qquad 2.5.4$$

where η_t ~*i.i.d* N (0,1) and h_t is given by equation (2.5.3). Since h_t is a function of ψ_{t-1} and is therefore fixed when conditioning on ψ_{t-1}, it is clear that ε_t as given in (2.5.4) will be conditionally normal with $E(\varepsilon_t / \psi_{t-1}) = \sqrt{h_t} E(\eta_t / \psi_{t-1}) = 0$ and $Var(\varepsilon_t / \psi_{t-1}) = h_t$, $Var(\eta_t / \psi_{t-1}) = h_t$. Hence the process (2.5.4) is identical to the ARCH process (2.5.2).

Engle (1982, 1983) found that a large lag q was required in the conditional variance function when applying the ARCH model to the relationship between the level and volatility of inflation. This would necessitate estimating a large number of parameters, subject to inequality restriction. To reduce the computational burden, Engle (1982, 1983) parameterized the conditional variance as;

$$h_t = \alpha_0 + \alpha_1 \sum_{i=1}^{q} w_i \varepsilon_{t-i}^2 \qquad 2.5.5$$

where the weights $w_i = \dfrac{(q+1) - i}{0.5q(q+1)}$ decline linearly and are constructed so that $\sum_{i=1}^{q} w_i = 1$. With this parameterization, a large lag can be specified and yet only two parameters are required in the conditional variance function.

Despite the importance of the ARCH model for many financial time series, a relatively long lag length in the variance equation with the problem of estimation of parameters subject to inequality restrictions is often called for to capture the long memory typical of financial data.

The ARCH model has been extended to various generalizations. Some of the generalizations are given in the following sections.

2.5.1 Generalized Autoregressive Conditional Heteroscedasticity (GARCH)

The GARCH model was developed by Bollerslev (1986) and is today one of the most widely used ARCH-type model (Engle, 2004). He proposed an extension of the conditional variance function (2.5.3) which he termed as the generalized ARCH (GARCH) and suggested that conditional variance be specified as,

$$h_t = \alpha_0 + \alpha_1\varepsilon_{t-1}^2 + ... + \alpha_q\varepsilon_{t-q}^2 + \beta_1 h_{t-1} + ... + \beta_p h_{t-p} \qquad 2.5.6$$

with the inequality conditions $\alpha_0 > 0$, $\alpha_i \geq 0$ for *i=1,…,q*, $\beta_i \geq 0$ for *i= 1,…,p* to ensure that the conditional variance is strictly positive. A GARCH process with orders p and q is denoted as GARCH (p,q) and this essentially generalizes the purely autoregressive ARCH to an autoregressive moving average model. The motivation for the GARCH process can be seen by expressing (2.5.5) as

$$h_t = \alpha_0 + \alpha(B)\varepsilon_t^2 + \beta(B)h_t \qquad 2.5.7$$

where $\alpha(B) = \alpha_1 B + ... + \alpha_q B^q$ and $\beta(B) = \beta_1 B + ... + \beta_q B^q$ are polynomials in the backshift operator B. Now, if the roots of $1 - \beta(Z)$ lie outside the unit circle, equation 2.5.7 be written as

$$h_t = \frac{\alpha_0}{1-\beta(1)} + \frac{\alpha(B)}{1-\beta(B)}\varepsilon_t^2 = \alpha_0^* + \sum_{i=1}^{\infty}\delta_i\varepsilon_{t-i}^2 \qquad 2.5.8$$

where $\alpha_0^* = \dfrac{\alpha_0}{[1-\beta(1)]}$ and the co-efficient δ_i is the co-efficient of B^i in the expression of $\alpha(B)[1-\beta(B)]^{-1}$.

The slope parameter β measures the combined marginal impacts of the lagged innovations while α , on the other hand captures the marginal impact of the most recent innovation in the conditional variance. When $\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1$, then the process is weakly stationary and the conditional variance (σ_t^2) approaches the unconditional variance (σ^2) as time goes to infinity i.e $E(\sigma_{t+s}^2) \to \sigma^2$ as $s \to \infty$. However, when $\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j > 1$ then the process is non stationary.

There exists some situations whereby parameter estimates in GARCH (p,q) models are close to the unit root but not less than unit, i.e $\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j = 1$, for the GARCH process. Here, the multi-step forecasts of the conditional variance do not approach the unconditional variance.

These processes exhibit the persistence in variance/volatility whereby the current information remains important in forecasting the conditional variance. Engle and Bollerslev (1986) refer to these processes as the Integrated GARCH or IGARCH. The IGARCH process does not possess a finite variance but are stationary in the strong sense (Nelson, 1990).

From 2.5.8, it is easy to see that a GARCH (p,q) process is an infinite order ARCH with a rational lag structure imposed on the co-efficient. The intention is that the GARCH process can parsimoniously represent a high-order ARCH-process (Bera and Higgins, 1993; Engle, 2004; Degiannakis and Xekalaki, 2004). The simplest GARCH(1,1) is often found to be the benchmark of financial time series modelling because such simplicity does not significantly affect the preciseness of the outcome.

A GARCH model can be applied with the assumption of normal, student t or general error distributions. Besides the empirical success, GARCH models have two major draw backs: First, they are unable to model asymmetry because in a GARCH model, positive and negative shocks of the same magnitude produce the same amount of volatility (i.e only the magnitude and not the sign of the lagged residuals determines the conditional variance). However, volatility tends to rise in response to "bad" news and fall in response to "good" news (Nelson, 1991). The second disadvantage of GARCH models is the non-negativity constraints imposed on the parameters which are often violated by estimated parameters (Curto, 2002).

2.5.2 GARCH-in-Mean (GARCH-M) model

The GARCH-M model was developed by Engle *et al.,* (1987) whose key postulate was that time varying premia on different term instruments can be modelled as risk premia where the risk is due to unanticipated interest rates and is measured by the conditional variance of the one period holding yield. The GARCH (1,1)-M model is presented by,

$$x_t = y_{t-1}\beta + h_t\gamma + \varepsilon_t \qquad 2.5.9$$

where x_t and h_t are defined as before while y_{t-1} is a vector of additional explanatory variables. The residual ε_t can be decomposed as in equation 2.5.4.

Just like the GARCH model, the GARCH-M is unable to capture asymmetric characteristics of financial data.

2.5.3 Exponential GARCH (EGARCH) model

EGARCH models were introduced by Nelson (1991) in an attempt to address the two major limitations of the GARCH models. Here the volatility depends not only on the magnitude of the shock but also on their corresponding signs. The non-negativity restrictions are not imposed as in the case of GARCH since the EGARCH model describes the logarithm of the conditional variance which will always be positive. The specification for the conditional variance (Nelson, 1991) is given as,

$$Log\sigma_t^2 = \varpi_0 + \sum_{i=1}^{q} \beta_i \log\sigma_{t-i}^2 + \sum_{i=1}^{p} \alpha_i \frac{|\varepsilon_{t-i}|}{|\sigma_{t-i}|} + \sum_{i=1}^{p} \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} \qquad 2.5.10$$

Note that $\varepsilon_t = \eta_t \sqrt{\sigma_t}$ where $\eta_t \sim$ i.i.d N(0,1).

The parameter (α_i) in equation (2.5.10) measures the impact of innovation on volatility at time t while parameter (β_i) is the auto-regressive term on lagged conditional volatility, reflecting the weight given to previous period's conditional volatility *t*. It measures the persistence of shocks to the conditional variance. The stationarity requirement is that the roots of the auto-regressive polynomial lie outside the unit circle. For EGARCH (1,1) this translates into β_1 *<1* (Ogum *et al*., 2006). Unlike the linear GARCH, in the EGARCH model a negative shock can have a different impact compared to a positive shock if the asymmetry parameter γ_i is non-zero.

2.5.4 Quadratic GARCH (QGARCH) model

The QGARCH model was introduced by Sentana (1995). The model can be interpreted as a second order Taylor approximation to the unknown conditional variance function and hence it is called a quadratic GARCH. The model of order p=q=1 is as follows,

$$h_t = \gamma + \xi\varepsilon_{t-1} + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} \qquad 2.5.11$$

where $\varepsilon_t = \eta_t \sqrt{h_t}, \eta_t \sim$ i.i.d N(0,1) and γ, ξ, α and β are parameters to be estimated. In this model, if $\varepsilon_{t-1} > 0$, its impact on h_t is greater than in the case if $\varepsilon_{t-1} < 0$ (assuming γ, ξ, α and β are positive). Thus it captures the asymmetric effects from another point of view.

The stationarity in QGARCH is covariance based whenever the sum of α and β is less than one. This sum also provides a measure of persistence of shocks to the variance process.

2.5.5. Threshold GARCH (TGARCH) model

Threshold GARCH models were introduced by Zakoian (1994). The generalized specification of the conditional variance equation is given by,

$$h_t = \alpha_0 + \sum_{j=1}^{q} \beta_j h_{t-j} + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{k=1}^{r} \gamma_k \varepsilon_{t-k}^2 I_{t-k}^- \qquad 2.5.12$$

where $\varepsilon_t = \eta_t \sqrt{h_t}$ and $I_t^- = 1$, if $\varepsilon_t < 0$ and zero otherwise. In this model, good news, $\varepsilon_{t-i} > 0$, and bad news $\varepsilon_{t-i} < 0$, have differential effects on the conditional variance. Good news has an impact of α_i, while bad news has an impact of $\alpha_i + \gamma_i$. If $\gamma_i > 0$, bad news increases volatility while if $\gamma_i \neq 0$, the news impact is asymmetric. When the threshold term I_{t-k}^- is set to zero, then equation 2.5.12 becomes a GARCH (p,q) model.

2.5.6 Glosten, Jagannathan and Runkle (GJR) model

This is a modified GARCH-M model developed by Glosten *et al.,* (1993). The GJR model allows positive and negative innovations to returns to have different impacts on the conditional variance. This is achieved by the introduction of a dummy variable into the conditional variance equation. The GJR-GARCH(1,1) model is given by,

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \varpi S_{t-1}^- \varepsilon_{t-1}^2 + \beta h_{t-1} \qquad 2.5.13$$

where $\varepsilon_t = \eta_t \sqrt{h_t}$,$\eta_t \sim$ i.i.d (0,1) and S^- denotes an indicator (dummy) function that takes the value one when $\varepsilon_{t-1} \geq 0$ and zero otherwise.

2.5.7 Parameter Estimation of ARCH models

In this study, the focus is on the Maximum likelihood approach which is the most commonly used estimation procedure for the ARCH models. This is based on the normality assumption of the conditional distribution. Following Bera and Higgins (1993), consider the standard ARCH-regression model $y_t / \psi_{t-1} \sim N(x_t' \xi, h_t)$ with its log likelihood function is given by, $l(\theta) = \frac{1}{T}\sum_{t=1}^{T} l_t(\theta)$,where $l_t(\theta) = C - \frac{1}{2}\log(h_t) - \frac{\varepsilon_t^2}{2h_t^2}$ and $\theta = (\xi', \gamma')$. Here ξ and γ denote the conditional mean and conditional variance parameters respectively. One attractive feature of this normal likelihood function is that the information matrix is block diagonal between the parameters ξ and γ. Now, the $(i,j)^{th}$ element of the off-diagonal block of the information can be written as

$$\frac{1}{T}\sum_{t=1}^{T}E\left[\frac{\partial^2 l_t}{\partial\xi_i\partial\gamma_j}\right]=\frac{1}{T}\sum_{t=1}^{T}E\left[\frac{1}{2h_T^2}\frac{\partial h_t}{\partial\xi_i}\frac{\partial h_t}{\partial\gamma_j}\right]$$

If h_t is an asymmetric function of the lagged errors in the sense of Engle (1982), then the last expression in the square brackets is anti-symmetric and therefore has expectation zero. When the block is diagonal, under the likelihood frame work, estimation and testing for the mean and variance parameters can be carried out separately (Engle, 1982; Bollerslev, 1986; Bera and Higgins, 1993; Fan and Yao, 2003; Davidson, 2008).

Most of the applied work on ARCH models uses the Berndt *et al,* (1974) algorithm to maximize the log likelihood function [$l(\theta)$]. Starting from estimates of the r^{th} iteration, the $(r+1)^{th}$ step of the algorithm can be written as

$$\xi^{(r+1)}=\xi^{(r)}+\left[\sum_{t=1}^{T}\left(\frac{\partial l_t}{\partial\xi}\right)\left(\frac{\partial l_t}{\partial\xi}\right)'\right]^{-1}\sum_{t=1}^{T}\frac{\partial l_t}{\partial\xi}$$

and

$$\gamma^{(r+1)}=\gamma^{(r)}+\left[\sum_{t=1}^{T}\left(\frac{\partial l_t}{\partial\gamma}\right)\left(\frac{\partial l_t}{\partial\gamma}\right)'\right]^{-1}\sum_{t=1}^{T}\frac{\partial l_t}{\partial\gamma}$$

where derivatives are evaluated at $\xi^{(r)}$ and $\gamma^{(r)}$.

When ε_t is Student's t-distributed with v>2 degrees of freedom, the criterion function maximized is given by;

$$l(\theta)=T\log\frac{\Gamma((v+1)/2)}{\Gamma(v/2)}-\frac{T}{2}\log\pi(v-2)-\frac{1}{2}\sum_{t=1}^{T}\left(\log h_t+(v+1)\log\left(1+\frac{\varepsilon_t^2}{(v-2)h_t}\right)\right) \quad 2.5.14$$

where $v>2$ controls the tail behaviour. The student's t distribution approaches normality as $v\to\infty$. To improve numerical stability, the parameter estimated is $v^{\frac{1}{2}}$ (Davidson, 2008). For the GED, 2.5.14 can be written as

$$l(\theta)=-\frac{1}{2}\log\left(\frac{\Gamma(1/v)^3}{\Gamma(3/v)(v/2)^2}\right)-\frac{1}{2}\log h_t\left(\frac{\Gamma(3/v)(y_t-X_t'\theta)^2}{h_t\Gamma(1/v)}\right)^{v/2} \quad 2.5.15$$

where $v > 0$ controls the tail behaviour . The GED corresponds to the Gaussian distribution if $v = 2$ and is leptokurtic (fat tailed) when $v < 2$.

2.5.8 Model identification for the ARCH-type models

The selection of the appropriate model is one of the most challenging areas in statistical modelling using ARCH models. This area has had very little development.

The portmanteau Q-test statistics based on the squared residuals is used to test for the independence of the series (McLeod and Li, 1983). This Q-statistics is used to test the ARCH effects present in the residuals. Since it is calculated from the squared residuals, it can be used to identify the order of the ARCH process. The Lagrange Multiplier test proposed by Engle (1982) is also used in a similar manner as the Q-statistic.

The Akaike information criterion (Akaike, 1974) and the Schwarz Bayesian criterion (Schwarz, 1978) model selection methods have widely been used in the ARCH literature, despite the fact that their statistical properties in the ARCH context are unknown. These selection methods are based on the maximized value of the log-likelihood function and evaluate the ability of the models to describe the data as discussed in section 2.3.

2.6 Bilinear Models (BL)

Following Granger and Anderson (1978) , Subba (1981), Subba and Gabr (1984), a time series $\{X_t\}$ is said to follow a bilinear time series denoted by BL(p,q,m,k) if it satisfies the equation

$$X_t + \sum_{i=1}^{p} a_i X_{t-i} = \sum_{j=0}^{q} c_j \varepsilon_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-j} \varepsilon_{t-k} + \varepsilon_t \qquad 2.6.1$$

where $\{\varepsilon_t\}$ is a sequence of *i.i.d* random variable, usually but not always with zero mean and variance σ_ε^2 and c_0=1, a_i ,b_{ij} and c_j are model parameters. It is easy to see that bilinear model is a special case of ARMA (p,q) model.

Using lag (Backshift) operator, equation 2.6.1 can be specified as

$$\phi(B) w_t = (\lambda(B) \varepsilon_{t-1}) \psi(B) w_{t-1} + \theta(B) \varepsilon_t \qquad 2.6.2$$

where

$$w_t = (1-B)^{d_1} X_t ,\ \varepsilon_t \sim WN(0,\sigma_\varepsilon{}^2) \qquad 2.6.3$$

In equation 2.6.2, $\psi(B) = \psi_1 + \psi_2 B + ... + \psi_p B^{p-1}$ and $\lambda(B) = 1 + \lambda_1 B + ... + \lambda_k B^{k-1}$ .This is equivalent to the Subba (1981) BL(p,q,m,k) class of models (Davidson, 2008).

Bilinear models have been applied in geophysics data (Subba, 1988), Spanish economic data (Maravall, 1983) and in solar physics data by (Subba and Gabr, 1984). These models are particularly attractive in modelling processes with sample paths of occasional sharp spikes (Subba and Gabr, 1984).These phenomena are found in financial time series data.

2.6.1 Estimation of parameters in bilinear models

Consider the BL(p,0,m,k) model of the form

$$X_t + a_1 X_{t-1} + ... + a_p X_{t-p} = \alpha + \sum_{i=1}^{m}\sum_{j=1}^{k} b_{ij} X_{t-i}\varepsilon_{t-j} + \varepsilon_t \qquad 2.6.4$$

where $\{\varepsilon_t\} \sim \mathrm{N}(0,\sigma_\varepsilon^2)$. (Here the, MA terms have been dropped and a constant α has been added to the R.H.S to facilitate the fitting of such models to non mean corrected data).

The likelihood function of the unknown parameters is constructed, given N observations $X_1, X_2,..., X_N$. Since the model involves lagged values of the $\{X_t\}$, one cannot evaluate the residuals for the initial stretch of data. The conditional likelihood based on $X_{\gamma+1}, X_{\gamma+2},..., X_N$, given $X_1, X_2,..., X_N$ where $\gamma = \max(p,m,k)$ is thus considered.

Let $\underline{\theta} = (\theta_1, \theta_2,...,\theta_n)$ denote the complete set of parameters {a_i}, {b_{ij}}, α i.e. set $\theta_i = a_i, i = 1,2,..., p$, $\theta_{p+1} = b_{11}, \theta_{p+2} = b_{12},...,\theta_{p+mk} = b_{mk}, \theta_{p+mk+1} = \alpha$ and write n=p+mk+1 to denote the total number of parameters.

The joint probability density function of $\varepsilon_{\gamma+1}, \varepsilon_{\gamma+2},...,\varepsilon_N$ is given by

$$(1\!\!\diagup_{2\pi\sigma_\varepsilon^2})^{(N-\gamma)/2} \exp\{\frac{1}{2\sigma_\varepsilon^2}\sum_{t=\gamma+1}^{N}\varepsilon_t^2\} \qquad 2.6.5$$

and since the Jacobian of the transformation from $\{X_t\}$ to $\{\varepsilon_t\}$ is unity, equation 2.6.5 also represents the likelihood function of $\boldsymbol{\theta}$, given $\{X_t; t = \gamma + 1, ..., N\}$. The (conditional) maximum likelihood estimates of $\theta_1, \theta_2, ..., \theta_n$ are thus given by maximizing (2.6.5) or equivalently by minimizing

$$Q(\boldsymbol{\theta}) = \sum_{t=\gamma+1}^{N} \varepsilon_t^2 \qquad 2.6.6$$

The minimization is performed numerically: for a given set of values $(\theta_1, \theta_2, ..., \theta_n)$ then $\{\varepsilon_t\}$ is evaluated recursively from equation 2.6.1 and then the Newton-Raphson method is used to minimize Q(**θ**) (Subba, 1981). The Newton-Raphson iterative equations for minimization of Q (**θ**) are given by,

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)} - \mathbf{H}^{-1}(\boldsymbol{\theta}^{(i)})\mathbf{G}(\boldsymbol{\theta}^{(i)}) \qquad 2.6.7$$

where $\boldsymbol{\theta}^{(i)}$ is the vector of parameter estimates obtained at the i^th^ iteration, and gradient vector **G** and Hessian matrix **H** are given respectively as,

$\left[\frac{\partial Q}{\partial \theta_1}, \frac{\partial Q}{\partial \theta_2}, ..., \frac{\partial Q}{\partial \theta_n}\right]'$ and $\mathbf{H}(\boldsymbol{\theta}) = \left\{\frac{\partial^2 \theta}{\partial \theta_i \partial \theta_j}\right\}$. The partial derivatives of Q with respect to $\{\boldsymbol{\theta_i}\}$ are shown to be $\frac{\partial Q}{\partial \theta_i} = 2\sum_{t=\gamma+1}^{N} \varepsilon_t \frac{\partial \varepsilon_t}{\partial \theta_i}$, i=1,2,…,n

$$\frac{\partial^2 Q}{\partial \theta_i \partial \theta_j} = 2\sum_{t=\gamma+1}^{N} \frac{\partial \varepsilon_t}{\partial \theta_i} \frac{\partial \varepsilon_t}{\partial \theta_j} + 2\sum_{t=\gamma+1}^{N} \frac{\partial^2 \varepsilon_t}{\partial \theta_i \partial \theta_j}, \; i,j = 1,2,\ldots,N \qquad 2.6.8$$

Subba and Gabr (1984) developed a neat set of recursive equations for these derivatives as follows. Differentiating equation (2.6.1) with respect to each of the parameters the following are obtained

$$\frac{\partial \varepsilon_t}{\partial a_i} + \phi(a_i) = X_{t-i}, \qquad i=1,2,\ldots,p$$

$$\frac{\partial \varepsilon_t}{\partial a_i} + \phi(b_{ij}) = -X_{t-i}\varepsilon_{t-i}, \qquad i=1,2,\ldots,m \quad j=1,2,\ldots,k$$

$$\frac{\partial \varepsilon_t}{\partial a_i} + \phi(\alpha) = -1 \text{ ,where } \phi(\theta_1) = \sum_{i=1}^{m}\sum_{j=1}^{k} b_{ij} X_{t-i} \frac{\partial \varepsilon_{t-i}}{\partial \theta_i}$$

Assuming the initial conditions

$$\varepsilon_t = \frac{\partial \varepsilon_t}{\partial \theta_i} = 0 \quad t=1,2,\dots,\gamma \qquad\qquad i=1,2,\dots,n$$

the second order derivatives satisfy

$$\frac{\partial^2 \varepsilon_t}{\partial a_i \partial a_{i'}} = 0 \,,\, \frac{\partial^2 \varepsilon_t}{\partial a_i \partial \alpha} = 0 \,,\; i,j'=1,\dots,p$$

$$\frac{\partial^2 \varepsilon_t}{\partial a_i \partial b_{j1}} + \psi(a_i, b_{j1}) = -X_{t-j} \frac{\partial \varepsilon_{t-1}}{\partial a_i},$$

$$\frac{\partial^2 \varepsilon_t}{\partial b_{ij} \partial b_{i'j'}} + \psi(b_{ij}, b_{i'j'}) = -X_{t-i} \frac{\partial \varepsilon_{t-j}}{\partial b_{i'j'}} - X_{t-i'} \frac{\partial \varepsilon_{t-j'}}{\partial b_{ij}}$$

$$\frac{\partial^2 \varepsilon_t}{\partial b_{ij} \partial \alpha} + \psi(b_{ij}, \alpha) = -X_{t-j} \frac{\partial \varepsilon_{t-j}}{\partial \alpha},\; \frac{\partial^2 \varepsilon_t}{\partial \alpha^2} = 0$$

where $\psi(\theta_r, \theta_l) = \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-i} \frac{\partial \varepsilon_{t-i}}{\partial \theta_r \partial \theta_l}$.

For a given set of parameter values α ,{a_i}, {b_{ij}} the first and the second derivatives of Q can be evaluated from the above equations and hence the vector **G** and matrix **H** evaluated. The iteration equation 2.6.7 is then implemented.

When the final parameter estimate $\underline{\hat{\theta}}$ have been obtained, σ_e^2 is obtained as,

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{(N-\gamma)} Q(\hat{\theta}) = \frac{1}{(N-\gamma)} \sum_{t=\gamma+1}^{N} \hat{\varepsilon}_t^2 \qquad 2.6.9$$

2.6.2 Least squares estimation of model parameters for bilinear models

Following the approach of Tong (1990), consider the bilinear model (equation 2.6.1). Rewriting it in a Markovian representation with monor changes in notation. Set p=max(p,q,m,k).

$$\xi_t = (\mathbf{A} + \mathbf{B}\varepsilon_t)\xi_{t-1} + \mathbf{c}\varepsilon_t + \mathbf{d}(\varepsilon_t^2 - \sigma^2)$$

$$\mathbf{X}_t = \mathbf{H}\xi_{t-1} + \varepsilon_t \qquad 2.6.10$$

where ξ_t is a p-vector and

$$\mathbf{A}=\begin{bmatrix} a_1 & 1 & 0 & \cdots 0 \\ a_2 & 0 & 1 & \cdots 0 \\ \vdots & \vdots & \vdots & \vdots \\ a_{p-1} & 0 & 0 & \cdots 1 \\ a_p & 0 & 0 & \cdots 0 \end{bmatrix} \qquad \mathbf{B}=\begin{bmatrix} b_1 & 0 & \cdots & 0 \\ \vdots & & & \vdots \\ b_p & 0 & \cdots & 0 \end{bmatrix}$$

$$\mathbf{c}=\begin{bmatrix} a_1+c_1 \\ \vdots \\ a_p+c_p \end{bmatrix} \qquad \mathbf{d}=\begin{bmatrix} b_1 \\ \vdots \\ b_p \end{bmatrix} \qquad \mathbf{H}=\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$$

By convention, $c_j=0$ if $j>q$, $b_{jk}=0$ if $j>m$. Diagonality of the model implies the relation **B=dH.** The converse is not true.

Following Guegan and Pham (1989), take $\boldsymbol{\theta}=(a_1,\ldots,a_p,\ c_1,\ldots,c_p,\ b_1,\ldots,b_p)$ as the fundamental parameter vector and assume that the representation in equation (2.6.6) is quasiminimal in the sense that there is no other Markovian representation with the same noise structure but with a state vector which is a linear transformation of the original state vector and has a smaller dimension. Further**,** assume that the model is invertible and stationary.

Let $(X_1,\ldots,X_N)$ denote the observations. It is plausible to estimate $\boldsymbol{\theta}$ by minimizing $\sum_{t=1}^{N}\varepsilon_t^2(\tilde{\theta}/\xi_0)$ w.r.t $\tilde{\theta}\in\Theta_N$ where $\varepsilon_t(\tilde{\theta}/\xi_0)$ is the $\varepsilon_t(\theta/\xi_0)$ given by $\varepsilon_t(\tilde{\theta}/\xi_0)=\mathbf{X}_t-\mathbf{H}\xi_{t-1}(\xi_0)$ with $\tilde{\theta}$ as the parameter vector. Intuitively, for this to make sense the effect of ξ_0 on $\varepsilon_t(\tilde{\theta}/\xi_0)$ should diminish as $t\to\infty$ (see Tong,1990). Let there exist a stationary time series $\{\varepsilon_t(\tilde{\theta})\}$ such that $\varepsilon_t(\tilde{\theta}/\xi_0)-\varepsilon_t(\tilde{\theta})\to 0$ as $t\to\infty$. Note that $\varepsilon_t(\tilde{\theta})$ is then measurable w.r.t σ-algebra generated by X_s, $s\le t$. Here, the model (2.6.6) is said to be invertible at $\tilde{\theta}$ relative to the observation $\{X_t\}$. A sufficient condition for this is

$[E_\theta(\ln\|\tilde{A}-\tilde{c}\tilde{H}-\tilde{d}\tilde{H}X_t\|_\phi)]<0$, where $\tilde{A}$, $\tilde{c}$ **,** $\tilde{d}$ and $\tilde{H}$ are given by $\boldsymbol{\theta}$.

A reasonable choice of Θ_N is suggested by the above sufficient condition and given as

$$\Theta_{N,\delta}=\{\tilde{\theta}\in\Theta_0:\prod_{t=1}\|\tilde{A}-\tilde{c}\tilde{H}-\tilde{d}\tilde{H}X_t\|_\phi<(1-\delta)^N\} \qquad 2.6.11$$

where Θ_0 is a given compact set and δ is a small positive number. The set Θ_0 is chosen large enough to include **θ**, the true parameter and the set of parameters satisfying the stationarity condition.

Let $$Q_N(\widetilde{\theta}) = \sum_{t=1}^{N} \varepsilon_t^2(\widetilde{\theta} / \xi_0) \qquad 2.6.12$$

Let $\hat{\theta}_N$ be the minimizer of $Q_N(\widetilde{\theta})$ over the set $\Theta_{N,\delta}$. This is the LSE of **θ** (Tong,1990).

2.6.3 Order selection for Bilinear models

Suitable values of p,m,k are determined by fitting a range of models covering various values of p,m,k and then selecting the model with the minimum value of the AIC defined as;

$$AIC = (N - \gamma)\log \hat{\sigma}_e^2 + 2 \text{ (number of fitted parameters) (Akaike, 1977).}$$

Note that (N-γ) is the effective number of observations to which each model has been fitted. In using the AIC criterion, the goal is to strike a balance between reducing the magnitude of the residual variance and increasing the number of model parameters. This method requires that the upper bounds be set to p,m,k and then search for the various combinations within this bound.
This is clearly a nested search procedure. The Subba (1981) algorithm is used to accomplish this as follows;

1. For a given value of p, fit a linear AR(p) model.
2. Using the AR coefficients as initial values for the $\{a_i\}$, $\{b_{ij}\}$ and α, and setting initially b_{11}=0, fit a BL(p,0,1,1) model using the Newton-Raphson technique.
3. Fit BL(p,0,1,2) and BL(p,0,2,1) models using the parameters of the BL(p,0,1,1) model as the initial values and setting initially the remaining bilinear coefficients to zero.
4. Of the two models fitted in step 3, choose the one which has the smaller residual variance and use its parameters as starting values of fitting the BL (p, 0,2,2) model.
5. The procedure is continued until *m,k* have reached a common upper bound Γ. At each stage, a bilinear term of order (*m,k*) is fitted by considering bilinear terms of orders (*m-1,k*) , (*m,k-1*) , and choosing whichever model has smaller residual variance to provide the starting values, with initial values for the remaining coefficients set to zero.
6. All previous steps are repeated for p=1,2,…,Γ. and the procedure terminates when the residual variance $\hat{\sigma}_e^2$ starts to increase as *m,k* increases. As a working rule, Γ should be at

least as large as the order of the best AR model selected by the AIC criterion. The final choice of model is then made by selecting the model for which the AIC value is smallest.

CHAPTER THREE

METHODOLOGY

3.1 The scope of the study

This study was focused on modelling the weekly NSE 20-share index and share prices for the three chosen companies namely, National Bank of Kenya Limited (NBK), Bamburi Cement and Kenya Airways from Nairobi Stock Exchange using ARCH, Bilinear (BL) and BL-ARCH models. The companies selected have been consistent in the NSE and are representative of the three sectors namely, Finance & Investment, Industrial & Allied and Commercial & Services categorized in the NSE. The two models chosen are able to capture the properties of financial data discussed in section 2.1.

3.2 Data collection

Secondary data was collected from the NSE. The NSE 20-share index was used in addition to the individual company share prices because the behaviour of the volatility of individual stocks has received far less attention in the literature when compared with studies on market indices. Furthermore, individual investors are more interested in the specific risk of the securities they hold rather than the market index; this justifies the need to study stock level data. Moreover, it has also been identified in the literature that basing an analysis on index data can lead to false perceptions of price change dependence, even when price changes of individual shares represented by the index are independent, because stocks which are not traded frequently affect the market index (Baudouhat, 2004). The three companies were randomly selected from the three sectors. These are the major sectors which are consistent and contribute a lot to the Nairobi stock market. The average weekly share prices for the following companies were used: National Bank of Kenya Limited (NBK), Bamburi Cement and Kenya Airways (KQ) for the period between 3rd June 1996 to 31st December 2007. The NSE 20-share index was for period between 2nd March 1998 to 31st December 2007 was also modelled.

3.3 Data analysis

Time plots for the data were obtained in order to check the empirical characteristics of the data. MLE procedure assuming Gaussian, t-distributions and GED were tested for each series when fitting ARCH models while for the Bilinear models, the maximum likelihood estimation assuming a normal distribution was utilized. The models were diagnosed using the Log likelihood ratio test, AIC and the BIC. Model adequacy was carried out for all cases by examining the standardized residuals and squared residual correlations through Ljung-Pierce Q-statistics. The MSE was used to check on the efficiency of various models in addition to the residual plots. The analysis was facilitated by use of computer softwares namely E-Views 5.0 (Quantitative Micro Software, 2004) and TSMod 4.25 (Davidson, 2008).

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Preliminary Analysis

In this study, four sets of data were used for modelling. They include the weekly average share prices for Bamburi Cement Ltd, National Bank of Kenya Limited (NBK), Kenya Airways (KQ) Ltd as well as the weekly average NSE 20 share index.

The NSE 20-share index is a weighted mean with 1966 as the base year at 100. It was originally based on 17 companies and was calculated on a weekly basis. In 1992, the number of companies was increased to 20 to represent nearly 90% of the NSE market capitalization and computation changed from weekly to a daily basis. The index is useful in determining the performance of the NSE by measuring the general price movement in the listed shares of the stock exchange.

Bamburi Cement, Ltd. was founded in 1951 and manufactures cement in sub-Saharan Africa. The company reported earnings results for the six months ended June 30, 2006. For the period, the company reported that the profit increased by 32.7% from KES 1.03 billion that was posted in a similar period the previous year. The firm improved in profitability with the turnover rising by 12% to KSH. 7.9 billion from KSH. 7 billion.

The Kenya Airways' principal activities include passengers and cargo carriage. It was incorporated in 1977 as the East African Airways Corporation (EAA). The company was listed in the NSE in 1996 and has been a major player in the Nairobi stock market.

The National Bank of Kenya Limited (NBK) was incorporated on 19th, June 1968 and officially opened on Thursday 14th, November 1968. Its main objective was to help Kenyans to get access to credit and control their economy after independence. NBK's current shareholding is distributed as: National Social Security Fund (NSSF) 48.06%, General Public 29.44%, and Kenya Government 22.5%. During the 34th, AGM held on 25th, April 2003, the bank increased its Share Capital by Kshs. 6 Billion. NBK is a major player in Kenya's banking industry and is one of the largest bank in the country giving financial services to all sectors of the economy.

NBK is also involved in the stock market playing multiple roles as an arranger, underwriter and placing agent.

The preliminary analysis was done by use of time plots for the various series. Figures 4.1 to 4.4 represents the time plots for the four series.

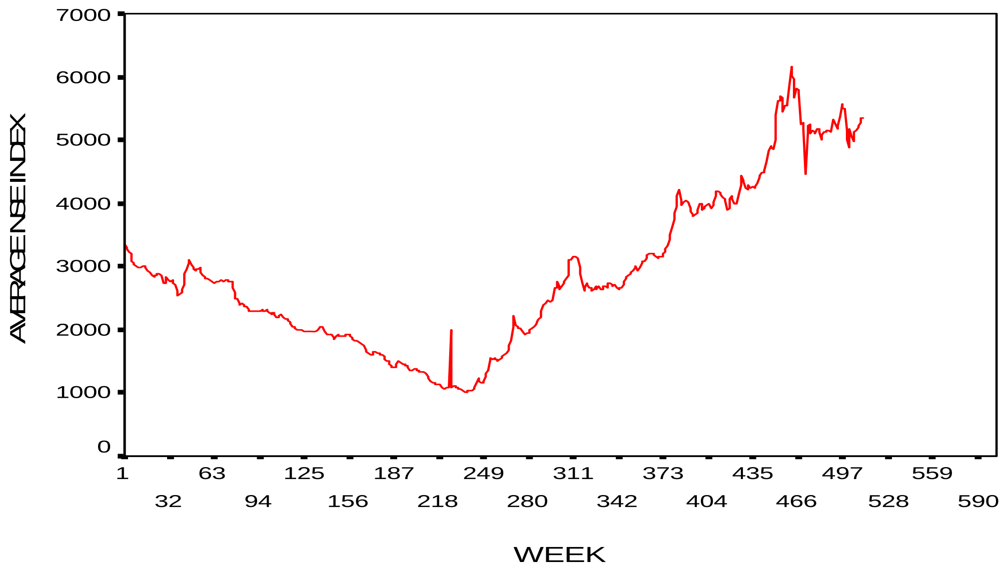

Figure 4.1: Time plot for weekly NSE- Index

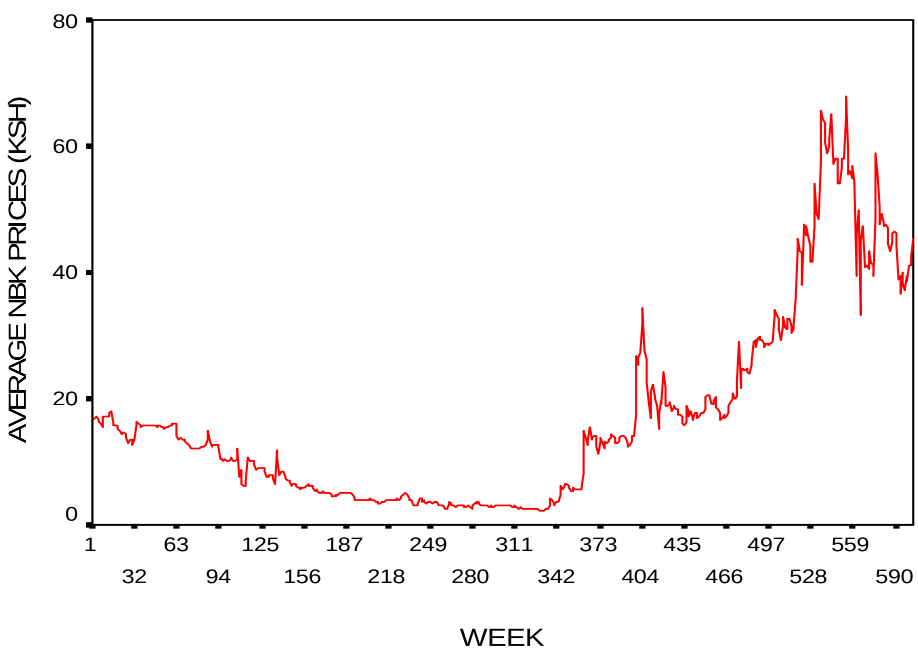

Figure 4.2: Time plot for weekly average NBK prices

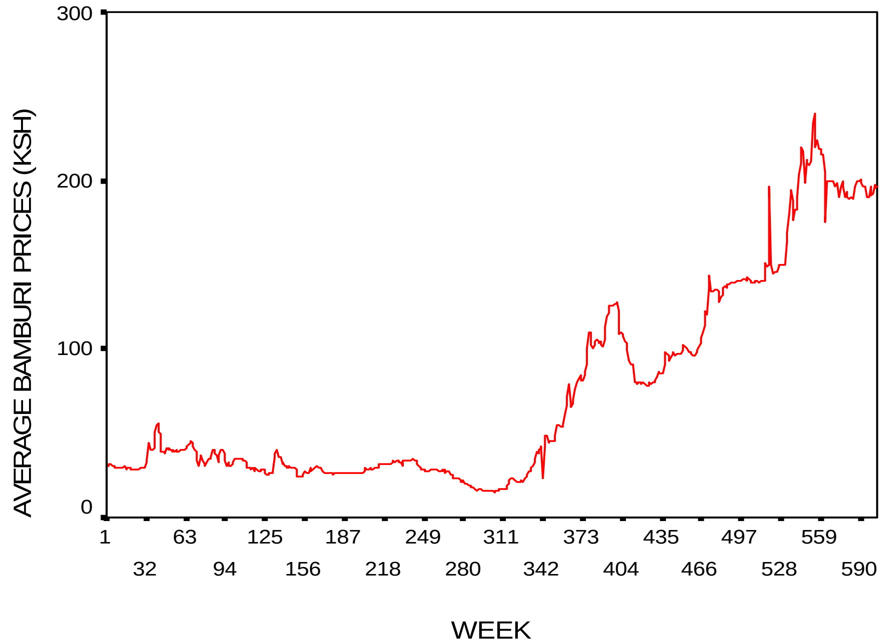

Figure 4.3: Time plot for weekly Bamburi prices

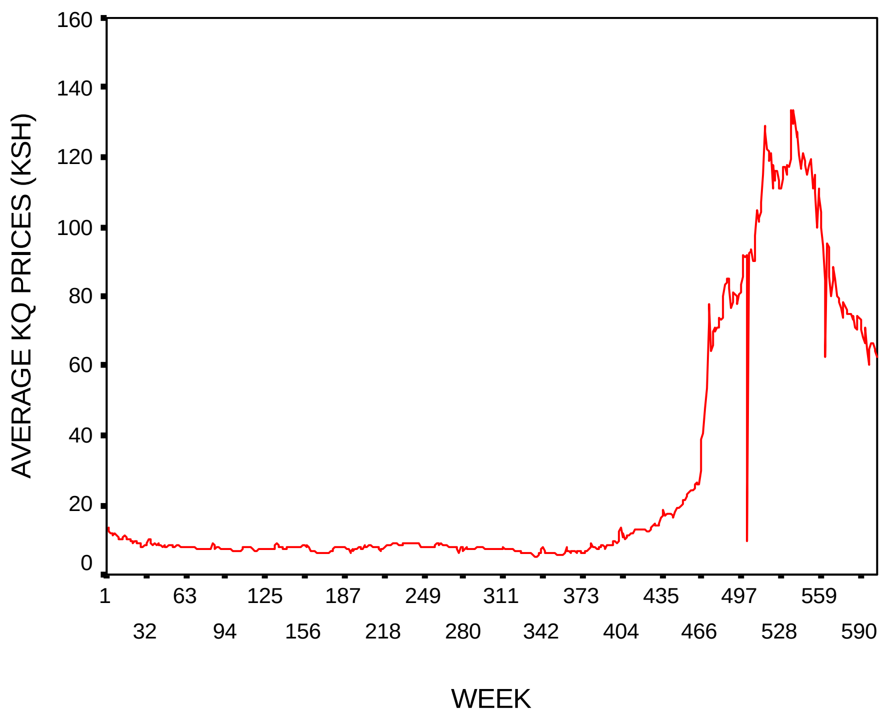

Figure 4.4: Time plot for weekly Kenya Airways (KQ) prices

A visual inspection of the time plots clearly shows that the mean and variance are not constant, implying non-stationarity of the data. The non-constant mean and variance suggests the utilization of a nonlinear model and preferably a non-normal distribution for modelling the data.

The series were transformed by taking the first differences of the natural logarithms of the values in each the series. The transformation was aimed at attaining stationarity in the first moment. The equation representing the transformation is given by $X_t = \ln(P_t) - \ln(P_{t-1})$, where P_t represents the weekly average value for each series. The sequence plots for the returns are presented in Figures 4.5 to 4.8.

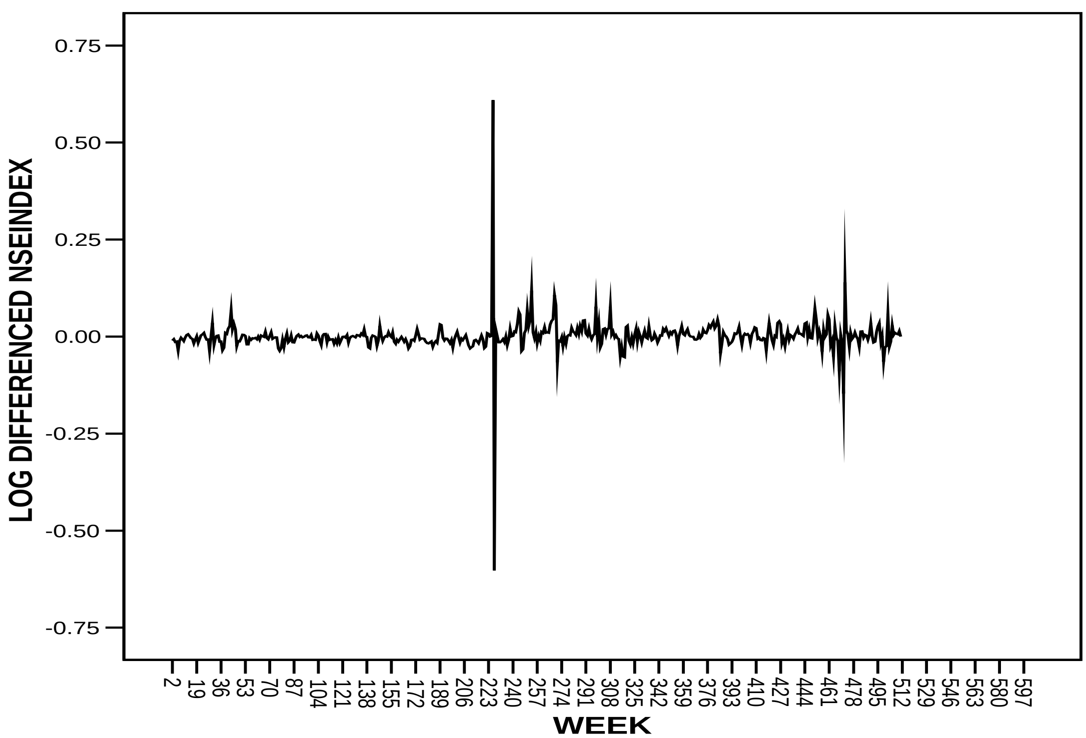

Figure 4.5: Time plot for the Log differenced NSE-Index

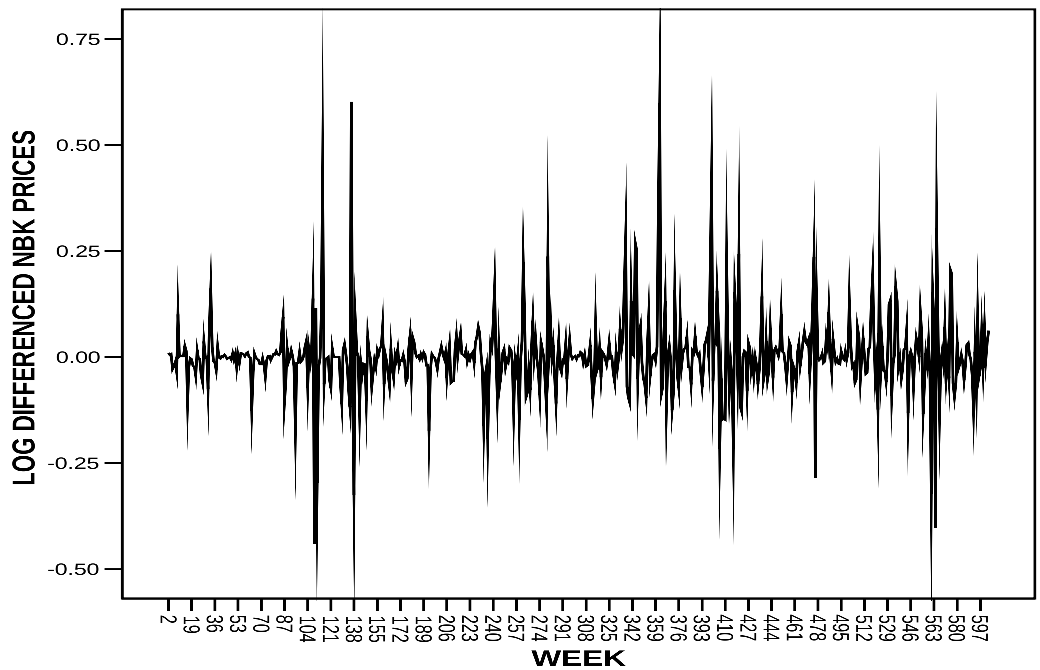

Figure 4.6: Time plot for the Log differenced NBK prices

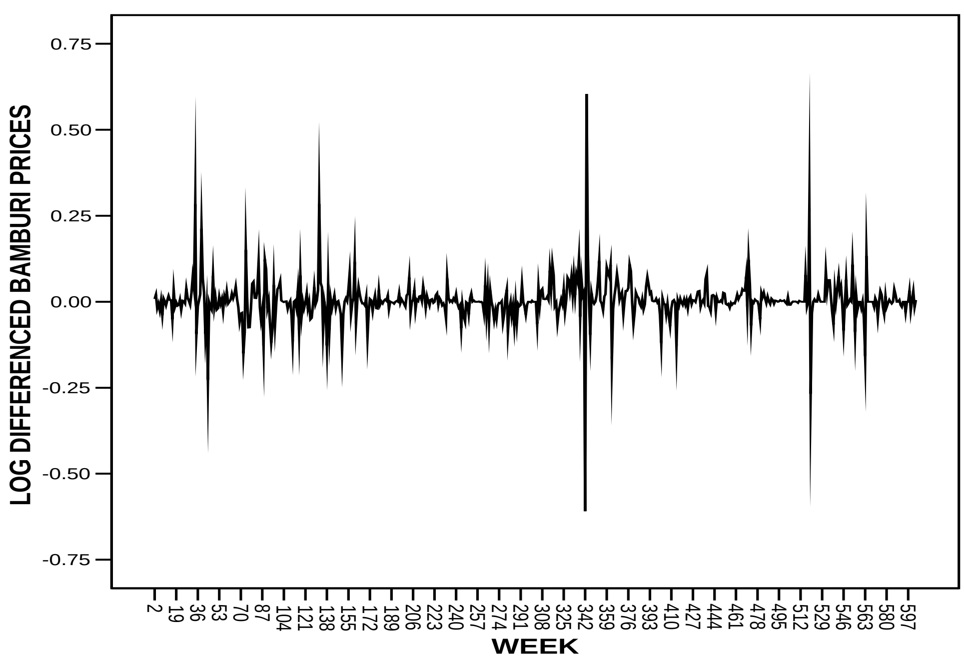

Figure 4.7: Time plot for the Log differenced Bamburi prices

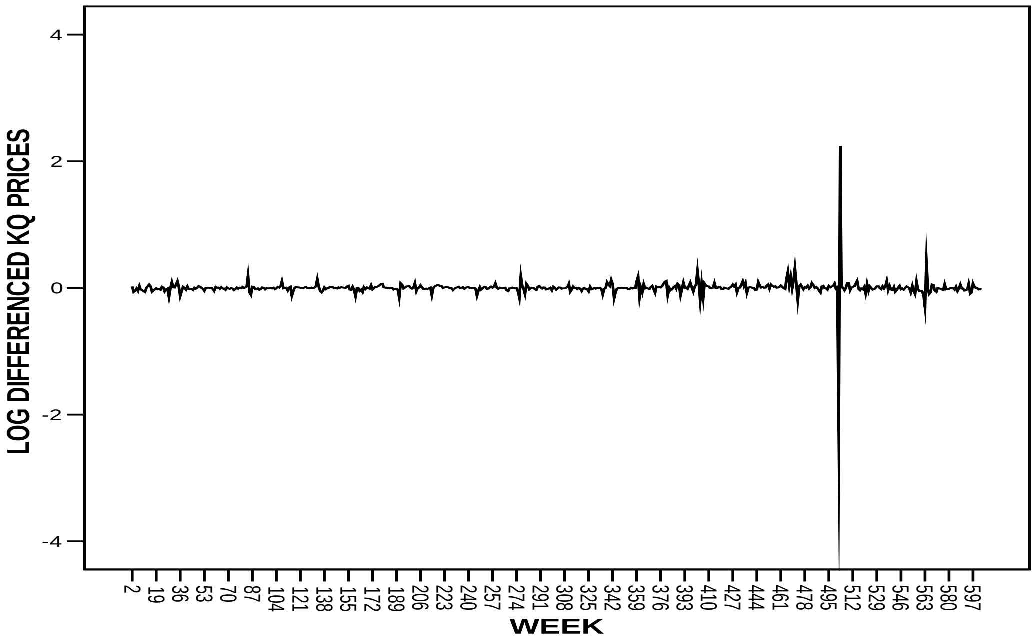

Figure 4.8: Time plot for the Log differenced Kenya Airways (KQ) prices

A closer examination of the return series plots reveals well known characteristics of high frequency data. It is easy to see that "large changes tend to be followed by large changes of either sign, and small changes tend to be followed by small changes of either sign, a characteristic first noted by Mandelbrot (1963). There is also an alternation between periods of high and low volatility. Large (small) changes are followed by large (small) changes, but of unpredictable sign as noted by Fama (1965). In general, all the series under study exhibits ARCH effects (also referred to as the heteroscedasticity) prevalent in many financial time series data.

The basic statistical properties of the data are presented in Table 4.1. The mean returns are all positive and close to zero a characteristic common in the financial return series. All the four series have very heavy tails showing a strong departure from the Gaussian assumption. The Jarque-Bera test also clearly rejects the null hypothesis of normality. Notable is the fact that all the four series exhibit positive Skewness estimate. This means that there are more observations on the right hand side.

Table 4.1: Basic statistical properties of returns

	NSE INDEX	NBK	BAMBURI	KENYA AIRWAYS
Mean	0.000919	0.001650	0.003070	0.002587
Median	-0.000116	0.000000	0.000395	0.000536
Maximum	0.609314	0.601761	0.604465	2.244758
Minimum	-0.602485	-0.440545	-0.608999	-2.240928
Std. Dev.	0.044259	0.086715	0.058789	0.139903
Skewness	0.208853	1.287552	0.165014	0.046429
Kurtosis	139.1245	15.20180	42.97123	220.4956
Jarque-Bera	393763.8	3900.838	40078.36	1186550
Probability	0.000000	0.000000	0.000000	0.000000
Observations	510	602	602	602

The series having exhibited heteroscedasticity as shown by the time plots were tested for the ARCH disturbances using Engle's (1982) Lagrange Multiplier (LM) while the Portmanteau Q test (McLeod and Li, 1983) based on the squared residuals was used to test for the independence of the series. Since both the Q statistic and the LM are calculated from the squared residuals, they were used to identify the order of the ARCH process.

For all the return series, the Q statistics and the Lagrange Multiplier (LM) tests indicated strong heteroscedasticity for all the lags from 1 to 12.This suggested an ARCH model of order q=8.

4.2 Empirical Results and Discussions

a) ARCH models

Before an ARCH model was fitted, the ACF was used to detect autocorrelation while the PACF was used to determine the orders of the ARMA models that could capture any autocorrelation present in the data. The four series studied exhibited some autocorrelation up to different lags and it was therefore important to model the series using both the AR models in addition to the variance equation (ARCH model). The ARCH orders were determined by the PACF of the squared returns in addition to the LM test.

The first set of models implemented in this study was the original Engle's (1982) ARCH models. Tables 4.2 and 4.3 presents the MLEs for the parameters of the Autoregressive and ARCH models respectively for the best selected models with the student's t distribution in all cases.

Table 4.2: Maximum likelihood estimates for the AR(p)

	NSE INDEX	NBK	BAMBURI	KQ
C	0.001280(0.4606)	-0.000556 (0.6463)	0.001653(0.1456)	0.000351(0.8502)
θ_1	0.287924(0.0001)	0.063694 (0.1424)	0.13071 (0.0005)	0.138328(0.0084)
θ_2	0.009555(0.8777)	-0.096911 (0.0005)	0.045581(0.1739)	-0.09848(0.0269)
θ_3	-0.01552(0.8442)	------------	-0.01648(0.5277)	-0.00712 (0.822)
θ_4	-0.0303(0.6621)	------------	-------------	-----------
θ_5	-0.02445(0.7031)	------------	-------------	--------------

P-values in the brackets

Table 4.3: Maximum likelihood estimates for the Variance Equation (ARCH)

	NSE INDEX	NBK	BAMBURI	KQ
C	0.000468(0.0000)	0.002493 (0.275)	0.001281 (0.0386)	0.001084 (0.000)
α_1	0.223091(0.0026)	4.282614 (0.285)	0.830275 (0.0574)	0.40954 (0.0000)
α_2	-0.032588(0.354)	0.530357 (0.352)	0.245715 ((0.1192)	0.121081 (0.006)
α_3	0.135125(0.0183)	-0.10449 (0.430)	0.074612 (0.2819)	-0.00096 (0.923)
α_4	0.009933(0.7744)	0.615525 (0.322)	0.027413 (0.5522)	-0.00608 (0.564)
α_5	0.000338(0.9878)	0.06252 (0.6357)	0.045551 (0.2838)	0.00726 (0.5863)
α_6	-0.002823(0.922)	0.18422 (0.4053)	0.007415 (0.7293)	-0.00052 (0.915)
α_7	0.023609(0.4975)	0.12481 (0.3839)	-0.003954 (0.8243)	-0.00026 (0.889)
α_8	0.001971(0.8204)	-0.04717 (0.065)	-0.000210 (0.9878)	-0.00012 (0.894)
α_9	--------------	0.132368 (0.331)	--------------	--------------

P-values in the brackets

As shown in Table 4.3, the ARCH model required a relatively bigger lag to model the volatility. The relative fit of the models were assessed using the Log likelihood, Schwarz Bayesian Information Criterion (BIC) , Akaike Information Criterion (AIC) and the Likelihood Ratio (LR) test, i.e the models that minimized the BIC and AIC but maximized the log likelihood were considered to be the best. The final model specifications however, were decided on by looking at the standardized residuals and the squared standardized residuals through Ljung-Box Q statistics in addition to the residual time plots and the residual correlogram. The goodness of fit statistics are presented in Table 4.4. The key for Table 4.4 and subsequent Tables is given as;

LR- Represents Log likelihood Ratio test
JB- Represents Jarque-Bera statistics for normality
Q(12) - Represents Ljung-Box Q statistics for the standardized residuals
$Q^2(12)$ - Represents Ljung-Box Q statistics for squared standardized residuals
P-Values are given in the brackets

Table 4.4: The goodness of fit statistics for ARCH models

		ARCH (q)	
		t	**GED**
NSE INDEX	LR	1234.183	1186.69
	AIC	-4.8249	-4.63639
	BIC	-4.69064	-4.50254
NBK	LR	888.5304	894.4222
	AIC	-2.9151	-2.93474
	BIC	-2.8125	-2.83215
BAMBURI	LR	1173.448	1131.428
	AIC	-3.87128	-3.73098
	BIC	-3.76855	-3.62825
KQ	LR	1056.460	877.2675
	AIC	-3.48067	-2.88236
	BIC	-3.37794	-2.77964

Diagnostic checks are presented in Table 4.5. The student's t-distribution and the General Error Distribution (GED) were tested for all the series. In all the cases, the student's t distribution assumption provided a better model than the GED. This could be due to the fact the financial data is highly heavy tailed and is better captured by the student's t-distribution since the GED distribution has a higher peak than the student's t-distribution. Although the GED distribution may be better able to capture peaks, it is far worse for capturing fat tails. The Jarque-Bera (1980) statistic also strongly rejected the normality assumption in the standardized residuals for all the series. The fitted models were adequate since their standardized residuals were not significantly correlated in all the four series basing on the Ljung-Box Q statistics. The squared residuals were also not significantly correlated for lags up to 12 for all the four series.

Table 4.5: Diagnostic Tests for Standardized Residuals for ARCH models

Series	Statistics	ARCH(q)
NSE INDEX	Skewness	13.10924
	Kurtosis	247.6270
	JB	1273648(0.0)
	Q(12)	9.1250 (0.244)
	$Q^2(12)$	0.2561(1.00)
NBK	Skewness	0.440471
	Kurtosis	12.88964
	JB	2464.525 (0.00)
	Q(12)	12.153 (0.275)
	$Q^2(12)$	7.7550 (0.653)
BAMBURI	Skewness	-2.206949
	Kurtosis	46.14442
	JB	46944.72(0.00)
	Q(12)	13.073 (0.159)
	$Q^2(12)$	0.5454 (1.000)
KQ	Skewness	-19.56340
	Kurtosis	445.6521
	JB	4928567(0.00)
	Q(12)	3.8591 (0.920)
	$Q^2(12)$	0.0250(1.00)

b) GARCH models

The next class of models implemented was the GARCH models. The autoregressive models were applied to capture the autocorrelation present in the series. The GARCH models for different values of p and q were fitted to the data, diagnosed and from the diagnosis and goodness of fit statistics, the GARCH (1,1) was found to be the best choice. This is consistent with most empirical studies involving the application of GARCH models in financial time series data. The Maximum Likelihood Estimation (MLE) method was employed in the parameter estimation. The model parameter estimates are presented in Tables 4.6 and 4.7.

Table 4.6: Maximum Likelihood Parameter Estimates for the AR(p) model

	NSE INDEX	NBK	BAMBURI	KQ
C	-0.00067 (0.511)	0.000451 (0.4341)	6.45E-06 (0.984)	0.000311 (0.489)
θ_1	0.397699(0.0000)	0.066623 (0.0025)	0.000203 (0.984)	0.13101 (0.0000)
θ_2	-0.041552(0.328)	-0.098898 (0.0000)	-2.21E-05 (0.997)	-0.07271 (0.000)
θ_3	0.068505(0.1041)	-------------	-2.80E-05 (0.997)	0.022187 (0.025)
θ_4	0.025138(0.4778)	-------------	-------------	---------
θ_5	-2.22E-05(0.999)	------------	-------------	---------

P-values are given in the brackets.

Table 4.7: Maximum likelihood estimates for the Variance Equation (GARCH)

	NSE INDEX	NBK	BAMBURI	KQ
α_0	0.000188(0.0301)	0.000839 (0.0002)	0.000476 (0.0005)	0.00068 (0.0000)
α_1	1.012761(0.0528)	0.820734 (0.0002)	1.746631 (0.0007)	1.28820 (0.0003)
β_1	0.337890(0.0000)	0.402620 (0.0000)	0.246922 (0.0006)	0.15924 (0.0298)

P-values are given in the brackets.

The GARCH parameter estimates for the variance equation was significant for all the series except for the NSE Index in which α_1 was not statistically significant. In the GARCH model, the parameters α and β must satisfy $\alpha_1 + \beta_1 < 1$ for stationarity. However, the GARCH (1,1) estimates violated the restriction imposed i.e. in all cases, $\alpha_1 + \beta_1 > 1$. This implies that the fitted GARCH model is not weakly stationary and the conditional variance (σ_t^2) does not approach the unconditional variance (σ^2) and thus the series might not have finite unconditional variance. This calls for the implementation of Integrated GARCH (1,1) model since it is capable of being stationary in the strong sense even though $\alpha_1 + \beta_1 = 1$ (Nelson,1990).

Two distributions were tested (i.e student's t and GED) for the specific GARCH (p,q) model and the best distribution choice was determined based on the BIC, AIC and the Log likelihood Ratio test in all the cases (see Table 4.8). For the NSE index, the distribution of choice was the student's t-distribution while for NBK, Bamburi and KQ, the Generalized Error Distribution was chosen. This shows that the NSE index data had fatter tails as compared to NBK, Bamburi and KQ. The model adequacy was checked using the Ljung-Box Q statistics for residuals and squared residuals in which the null hypothesis of no significant correlations was not rejected for all the series implying that the fitted models were adequate. The JB statistics rejected the null hypothesis of normality in the standardized residuals. This implies that the models with the respective distributions failed to normalize the residuals. The Goodness of fit statistics and Diagnostic tests are presented in Tables 4.8 and 4.9 respectively.

Table 4.8: Goodness of fit Statistics for GARCH(1,1)

		GARCH(1,1)	
		t	**GED**
NSE INDEX	LR	1330.525	1311.209
	AIC	-5.22980	-5.1533
	BIC	-5.14615	-5.06965
NBK	LR	874.3381	886.1993
	AIC	-2.89113	-2.93066
	BIC	-2.83983	-2.87937
BAMBURI	LR	1183.842	1224.82
	AIC	-3.92602	-4.06284
	BIC	-3.8672	-4.00414
KQ	LR	1121.227	1102.199
	AIC	-3.71695	-3.65342
	BIC	-3.65825	-3.59472

Table 4.9: The Diagnostic tests on the standardized residuals for the GARCH models

Series	Statistics	GARCH(1,1)
NSE INDEX	Skewness	16.55941
	Kurtosis	336.4305
	JB	2362405(0.00)
	Q(12)	1.7414 (0.973)
	$Q^2(12)$	0.0362 (1.00)
NBK	Skewness	0.826570
	Kurtosis	12.59127
	JB	2368.132 (0.00)
	Q(12)	13.956 (0.175)
	$Q^2(12)$	3.3475 (0.972)
BAMBURI	Skewness	-1.431525
	Kurtosis	36.27621
	JB	27841.11 (0.00)
	Q(12)	16.192 (0.063)
	$Q^2(12)$	1.0873 (0.999)
KQ	Skewness	-20.21615
	Kurtosis	465.1099
	JB	537054(0.00)
	Q(12)	1.8585 (0.994)
	$Q^2(12)$	0.0264 (1.00)

c) Integrated GARCH (IGARCH) Model

Since the parameter estimates in GARCH (1,1) models were close to the unit root but not less than unit, i.e $\sum_{i=1}^{p}\alpha_i + \sum_{j==1}^{q}\beta_j = 1$,the IGARCH model was fitted. The MLE method was utilized for parameter estimation. The parameter estimates for the mean and variance equations are presented in Tables 4.10 and 4.11 respectively.

Table 4.10: Maximum Likelihood Parameter Estimates for the AR(p) model

	NSE INDEX	NBK	BAMBURI	KQ
C	-0.0057 (0.0017)	-0.003828 (0.2012)	0.000229 (0.9138)	-0.002113 (0.3421)
θ_1	-0.3821 (0.0001)	-0.0789 (0.0409)	-0.0863 (0.0303)	-0.1530 (0.0003)
θ_2	0.0538 (0.2101)	0.1326 (0.0002)	-0.0262 (0.4602)	0.0907 (0.0081)
θ_3	-0.0413 (0.3398)	-------------	0.004685 (0.8632)	-0.0173 (0.5225)
θ_4	-0.0234 (0.5312)	-------------	-------------	---------
θ_5	0.0128 (0.7034)	------------	-------------	---------

P-values are given in the brackets.

Table 4.11: Maximum likelihood estimates for the Variance Equation (IGARCH)

	NSE INDEX	NBK	BAMBURI	KQ
α_0	0.00013 (0.0004)	0.000988 (0.0005)	0.000542 (0.0001)	0.000686 (0.0001)
α_1	0.6022 (0.0001)	0.5034 (0.0001)	0.6762 (0.0001)	0.7948 (0.0001)
β_1	0.3978 ((0.0001))	0.4966 (0.0001)	0.3238 (0.0001)	0.2052 (0.0014)

P-values are given in the brackets.

The mean equation was applied to capture the autocorrelation in the data. The parameter estimates for the variance equation were statistically significant at 0.05 significance level in all the series. In addition, $\alpha_1 + \beta_1 = 1$ for all the cases; implying that multi-step forecasts of the conditional variance do not approach the unconditional variance (i.e the unconditional variance is infinite). Despite the infinite unconditional variance, one attractive feature of the IGARCH model is that it is strongly stationary even though it is not weakly stationary. The results indicate that the data sets used exhibit the persistence in variance/volatility whereby the current information remains important in forecasting the conditional variance, i.e. the current information in the NSE remains important in forecasting the conditional variance.

Two distribution assumptions namely, the normal and t-distributions were tested. The Gaussian distribution was implemented since it has been used as a benchmark for IGARCH models. The student's t-distribution provided the best fit for the data and captured the heavy tail properties more adequately than the Gaussian distribution. The models were fitted and diagnosed using the AIC, BIC and the Log likelihood ratio test. However, the final model was considered adequate if its standardized residuals and squared residuals were not significantly correlated at 5% significance level. The residual correlation was tested using the Ljung-Box Q statistics. All the fitted IGARCH models were adequate since their residuals were not significantly correlated. Further, the standardized residuals were still non-normal as shown by the JB statistics for normality. The goodness of fit statistics for the IGARCH(1,1) model is presented in Table 4.12 while the diagnostic tests are presented in Table 4.13.

Table 4.12: Goodness of fit Statistics for IGARCH (1,1)

		IGARCH(1,1)	
		Gaussian	**t**
NSE INDEX	LR	1025.48985	1639.12862
	AIC	-2032.9797	-3258.257
	BIC	-1994.87	-3215.9131
NBK	LR	750.632822	1219.04801
	AIC	-1489.2656	-2424.096
	BIC	-1462.8541	-2393.2826
BAMBURI	LR	978.83866	1521.84383
	AIC	-1943.6773	-3027.6877
	BIC	-1912.8755	-2992.4856
KQ	LR	652.581675	1463.09986
	AIC	-1291.1634	-2910.1997
	BIC	-1260.3615	-2874.9977

Table 4.13: The Diagnostic tests on the standardized residuals for the IGARCH models

Series	Statistics	IGARCH(1,1)
NSE INDEX	Skewness	-4.5528489
	Kurtosis	164.221066
	JB	2372988.14 (0.0001)
	Q(12)	181.363 (0.500)
	$Q^2(12)$	127.096 (0.444)
NBK	Skewness	1.25458426
	Kurtosis	12.1746324
	JB	2209.2269 (0.0001)
	Q(12)	16.48 (0.170)
	$Q^2(12)$	40.21 (1.0000)
BAMBURI	Skewness	0.56088418
	Kurtosis	45.6729539
	JB	52802.3806 (0.0001)
	Q(12)	26.49 (0.670)
	$Q^2(12)$	136.98 (0.988)
KQ	Skewness	2.82266654
	Kurtosis	221.9130
	JB	5313541.27 (0.0001)
	Q(12)	197.66(0.664)
	$Q^2(12)$	87.453 (0.442)

In order to capture the leverage effects, two asymmetric ARCH-type models; the Exponential GARCH (EGARCH) and Threshold GARCH (TGARCH) were fitted.

d) EGARCH models

Despite the popularity and apparent success of GARCH models in practical applications, they cannot capture asymmetric response of volatility to news since the sign of the returns play no role in the model specification. Statistically, the asymmetric effect occurs when an unexpected decrease in price resulting from bad news increases volatility more than an unexpected increase in price of similar magnitude following good news.

Accordingly, Nelson's (1991) EGARCH model was fitted. Unlike the GARCH (p,q) model, a negative shock can have a different impact on future volatility when compared to the positive shock if asymmetry parameter γ_1 is not zero for the EGARCH model. It also does not need restrictions to be imposed on the parameters to ensure the non-negativity.

In the EGARCH model estimation, the MLE criterion was employed. Different orders for p and q in the variance equation were tested with the best results being achieved for p=q=1. The student's t-distribution also emerged to be the best for the NSE Index and Kenya Airways series while the Generalized Error Distribution provided the best results for NBK and Bamburi series. In this case, the distributional implication is that the NSE index and Kenya Airways series have long tails and are closer to symmetry while the Bamburi and NBK series have long tail but are asymmetric. Model parameter estimates for the mean and variance equations are presented in Tables 4.14 and 4.15 respectively.

Table 4.14: Maximum Likelihood Parameter Estimates for the AR(p) model

	NSE INDEX	NBK	BAMBURI	KQ
C	-0.00034 (0.752)	0.001272 (0.0628)	1.57E-06 (0.995)	0.000524 (0.6492)
θ_1	0.367256 (0.000)	0.072864 (0.0004)	0.000736 (0.949)	0.153554 (0.0000)
θ_2	-0.01234 (0.741)	-0.087534 (0.000)	8.36E-05 (0.981)	-0.081792 (0.0119)
θ_3	0.054319 (0.153)	------------	4.77E-05 (0.993)	6.46E-05 (0.9981)
θ_4	0.036511 (0.249)	-----------	-----------	-----------
θ_5	0.005672 (0.854)	----------	----------	----------

P-values are given in the brackets

Table 4.15: Maximum likelihood estimates for the Variance Equation (EGARCH (1, 1))

	NSE INDEX	NBK	BAMBURI	KQ
ω	-2.31720 (0.000)	-1.923399 (0.0000)	-1.93642 (0.000)	-1.698545 (0.0020)
β_1	0.71009 (0.0000)	0.736037 (0.0000)	0.764633 (0.000)	0.703361 (0.000)
α_1	0.43426 (0.0022)	0.837917 (0.000)	0.90602 (0.000)	1.429443 (0.2394)
γ_1	-0.0309 (0.7116)	0.085966 (0.3613)	0.18791 (0.1125)	0.809205 (0.2497)

P-values are given in the brackets

The EGARCH model parameter estimates also reveal the persistence in volatility of the Nairobi equity market. This is because the sum of α_1 and β_1 is approximately 1 in all the data sets. The asymmetric parameter γ_1 is positive and not significant for three series namely NBK, Bamburi and KQ. The positivity of γ_1 indicates that positive shocks increase volatility more than the negative shocks of an equal magnitude. This shows that the concept of "leverage effect" (i.e the negative shocks increasing volatility more than a positive shock of the same magnitude) is not applicable to the individual company stocks. However, for the NSE Index, the asymmetric parameter γ_1 is negative implying that negative shocks increase volatility more than a positive shock of the same magnitude. This contradicts the earlier studies on the Nairobi Stock Exchange for instance Ogum *et al.*, (2005, 2006) who found the asymmetry parameter γ_1 to be positive when modelling the daily NSE 20 Share Index using the EGARCH models. This could arise from the fact that the weekly return series were used in this study while Ogum *et al.*, (2005, 2006) modelled the daily returns. Some information could have been lost when using the weekly average for the NSE index and the share prices for the companies. In addition, the flow of information in NSE might not be as efficient as in the developed equity markets.

The model diagnostics and goodness of fit statistics are presented in Tables 4.16 and 4.17 respectively. The diagnostics included the autocorrelation of the standardized residuals and squared residuals respectively. The Ljung-Box Q statistics represented by Q(12) and $Q^2(12)$ for residuals and squared residuals respectively were used which were not significant in all cases confirming the adequacy of the fitted models. The models could thus explain the non-linear dependence in the residuals i.e the models captured the dependence in the variance shown by the original series of returns.

Table 4.16: The Diagnostic tests on the standardized residuals for the EGARCH models

Series	Statistics	EGARCH(1,1)
NSE INDEX	Skewness	14.48218
	Kurtosis	281.5813
	JB	1650644 (0.000)
	Q(12)	1.9366 (0.963)
	$Q^2(12)$	0.0488 (1.00)
NBK	Skewness	0.829323
	Kurtosis	12.18049
	JB	2175.813 (0.00)
	Q(12)	12.569 (0.249)
	$Q^2(12)$	2.5460 (0.990)
BAMBURI	Skewness	-1.465635
	Kurtosis	25.31122
	JB	12638.47 (0.00)
	Q(12)	18.489 (0.300)
	$Q^2(12)$	2.1275 (0.989)
KQ	Skewness	-16.57644
	Kurtosis	357.2894
	JB	3160227 (0.00)
	Q(12)	2.5214 (0.98)
	$Q^2(12)$	0.391 (1.00)

Table 4.17: Goodness of fit Statistics for the EGARCH models

		EGARCH(1,1)	
		t	**GED**
NSE INDEX	LR	1329.191	1311.049
	AIC	-5.22056	-5.14871
	BIC	-5.12854	-5.05669
NBK	LR	880.8376	891.695
	AIC	-2.90946	-2.94565
	BIC	-2.85083	-2.88703
BAMBURI	LR	1188.800	1229.910
	AIC	-3.93923	-4.07650
	BIC	-3.87319	-4.01046
KQ	LR	1112.958	1100.172
	AIC	-3.68600	-3.64331
	BIC	-3.61996	-3.57727

The EGARCH model, in all cases showed a smaller Kurtosis compared to the ARCH and GARCH models. In addition, the student's t-distribution and Generalized Error Distributions also captured the tail properties of the data better than the Gaussian distribution in all the four cases. The JB statistics also strongly rejected the null hypothesis of normality in the standardized residuals in all the series under consideration.

e) Threshold GARCH (1,1)

The TGARCH (1,1) model which falls in the asymmetric class of ARCH-type models was also used. The model was fitted, estimated and diagnosed just like the previous models. From the two distributions tested, the student's t distribution emerged the best for the NSE index while GED was considered the best for the NBK, Bamburi and KQ. This is because the GED and the students's t-distributions were able to capture the tail properties of the data. It is worth noting that under the student's t distribution, the convergence during estimation was a major problem. The algorithm converged very slowly and sometimes weakly. This casts doubts on the stability of the parameter estimates.

The parameter estimates for the mean and variance equations are presented in Tables 4.18 and 4.19 respectively. In the variance equation, the asymmetry parameter γ_1 was less than zero for all the four series. This implies that good news increases volatility more than bad news. This is consistent with the findings of Ogum *et al*., (2005, 2006) who applied EGARCH models to the daily NSE 20 Share Index. Hence the leverage effect experienced in developed markets might not be a universal phenomenon.

Table 4.18 : Maximum Likelihood Estimates for the AR(p)

	NSE INDEX	NBK	BAMBURI	KQ
C	-0.00029 (0.781)	0.000230 (0.7662)	4.89E-07 (0.999)	0.000321 (0.4778)
θ_1	0.398891 (0.000)	0.074302 (0.0025)	0.091467 (0.000)	0.130972 (0.0000)
θ_2	-0.03898 (0.350)	-0.092697 (0.0000)	-0.00016 (0.987)	-0.072987 (0.0000)
θ_3	0.06793 (0.1052)	0.006025 (0.6790)	5.97E-05 (0.983)	0.022427 (0.0872)
θ_4	0.02378 (0.4985)	----------	-----------	-----------
θ_5	0.00251 (0.9347)	----------	----------	----------

P-values are given in the brackets

Table 4.19: Maximum likelihood estimates for the Variance Equation (TGARCH (1, 1))

	NSE INDEX	NBK	BAMBURI	KQ
α_0	0.00019 (0.0215)	0.000755 (0.0002)	0.00042 (0.0004)	0.000682 (0.0000)
α_1	1.40972 (0.0488)	0.916490 (0.0029)	1.92241 (0.0027)	1.800201 (0.0013)
γ_1	-0.82608 (0.148)	-0.127722 (0.7188)	-0.90175 (0.176)	-0.865160 (0.1709)
β_1	0.318458 (0.000)	0.406304 (0.0000)	0.28212 (0.0001)	0.157956 (0.0333)

P-values are given in the brackets

The diagnostic tests and goodness of fit statistics for the TGARCH models are presented in Tables 4.20 and 4.21. Just like the previous models, the best distributions were GED and the student's t-distribution. Also, based on the Ljung-Box Q statistics, both the residuals and the squared residuals were not significantly (5% level) correlated implying that the models were adequate. The JB statistic for normality also rejected the normality assumption in the standardized residuals.

Table 4.20 : The goodness of fit statistics for the TGARCH models

		TGARCH(1,1)	
		t	GED
NSE INDEX	LR	1332.741	1312.929
	AIC	-5.234619	-5.156153
	BIC	-5.142599	-5.064133
NBK	LR	871.7255	883.5855
	AIC	-2.880553	-2.920152
	BIC	-2.814514	-2.854113
BAMBURI	LR	1185.886	1217.043
	AIC	-3.929502	-4.033532
	BIC	-3.863463	-3.967493
KQ	LR	992.3657	1104.419
	AIC	-3.283358	-3.657492
	BIC	-3.217319	-3.591453

Table 4.21: The Diagnostic Tests in Standardized Residuals for the TGARCH models

Series	Statistics	TGARCH(1,1)
NSE INDEX	Skewness	16.46073
	Kurtosis	334.0355
	JB	2328646 (0.00)
	Q(12)	2.050 (0.957)
	Q^2(12)	0.357 (1.00)
NBK	Skewness	0.788546
	Kurtosis	12.80236
	JB	2460.232 (0.0)
	Q(12)	13.468 (0.143)
	Q^2(12)	3.3146 (0.951)
BAMBURI	Skewness	-1.488381
	Kurtosis	30.07995
	JB	18523.7 (0.00)
	Q(12)	14.448 (0.107)
	Q^2(12)	1.4251(0.998)
KQ	Skewness	-19.87371
	Kurtosis	454.4961
	JB	5127155
	Q(12)	1.8733 (0.993)
	Q^2(12)	0.0273 (1.00)

4.3 Efficiency Comparison between the ARCH-type Models

Model efficiencies for each of the ARCH-type models implemented were evaluated using the various MSE. The MSE for the chosen models are presented in Table 4.22.

Table 4.22: MSE for the fitted ARCH-type models

Series	ARCH	GARCH(1,1)	EGARCH(1,1)	TGARCH(1,1)	IGARCH(1,1)
NSE INDEX	0.002481	0.002804	0.002695	0.0028045	0.00274
NBK	0.007418	0.007379	0.0073909	0.007386	0.00736
BAMBURI	0.003593	0.003453	0.003476	0.0035650	0.00354
KQ	0.022491	0.022411	0.022929	0.022412	0.02290

Considering the MSE values in Table 4.22, it is clear that ARCH, GARCH, EGARCH and IGARCH are all equally efficient in modelling volatility based on the MSEs only, since the different ARCH-type models are almost equal for the respective data sets. The disadvantage with the ARCH model is that so many parameters are to be estimated. The GARCH, IGARCH, EGARCH and TGARCH models are able to parsimoniously model the series and hence are preferred to the original ARCH model. Considering the asymmetric properties of the data and the respective MSEs, the EGARCH (1,1) emerged as the best model for the NSE Index and Bamburi. For the NBK, both the EGARCH and the TGARCH are equally good but EGARCH is considered the best since the parameter estimates for the TGARCH are unstable due to weak convergence. The best model for Kenya Airways was the GARCH model.

The respective models chosen are justified by their relatively lower values of residual Kurtosis and MSE in addition to the other diagnostics considered as well as the asymmetric parameter that captures the leverage effect. However, in terms of stationarity, the IGARCH model emerged as the best ARCH-type model since it was strongly stationary thus being more stable. This makes the IGARCH model to be the preferred model from the ARCH-type models for modelling the Nairobi Stock Exchange data for the periods between 2nd March 1998 to 31st December 2007 for NSE 20-Share index while and 3rd June 1996 to 31st December 2007 for company share prices (i.e NBK, Bamburi and Kenya Airways).

e) Bilinear models.

The next class of models under consideration was the bilinear models. The MLE method assuming a Gaussian distribution was used for parameter estimation for the data studied. Order selection was done using the ACF and PACF.

The model adequacy was checked via Ljung-Box Q statistics as well as checking the residual and squared residuals, ACF and PACF which all showed that all the residuals for bilinear models were not significantly correlated to lag 12 except the squared residual for NBK which was significantly correlated at lag 12. This implies that the fitted bilinear models were adequate except the one for NBK.

The Jarque-Bera (1980) statistics rejected the null hypothesis of normality in all the residuals i.e the residuals are not normally distributed. Residual plots were further employed to check the model adequacy. In all the series, the residuals showed sharp spikes outside the standard error bands. This suggests that there could be problems ahead and casts doubts on the models' stability (Tong, 1990). The fitted bilinear models and model diagnostics are presented in Tables 4.23 and 4.24 respectively.

Table 4.23: Maximum Likelihood Estimates for the bilinear models

NSE 20-SHARE INDEX

$$X_t = \underset{(0.81)}{-0.05227} X_{t-1} + \underset{(0.535)}{0.02052} X_{t-2} - \underset{(0.033)}{0.3308}\ \varepsilon_{t-1} - \underset{(0.973)}{0.00279}\ \varepsilon_{t-2} - \underset{(0.00)}{2.03221} X_{t-1}\varepsilon_{t-1} - \underset{(0.991)}{0.01669} X_{t-2}\varepsilon_{t-1} + \varepsilon_t$$

$, \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$

NBK

$$X_t = \underset{(0.094)}{-0.25431} X_{t-1} - \underset{(0.04)}{0.10973} X_{t-2} + \underset{(0.000177)}{0.3713} X_{t-1}\varepsilon_{t-1} + \underset{(0.000495)}{0.2566} X_{t-1}\varepsilon_{t-2} - \underset{(0.0717)}{0.852222} X_{t-2}\varepsilon_{t-1} - \underset{(0.002)}{0.5889} X_{t-2}\,\varepsilon_{t-2} + \varepsilon_t$$

$\varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$

BAMBURI CEMENT LTD

$$X_t = \underset{(0.025)}{0.41221} X_{t-1} - \underset{(0.048)}{0.31609} X_{t-2} - \underset{(0.033)}{0.44521} \varepsilon_{t-1} + \underset{(0.019)}{0.39864} \varepsilon_{t-2} + \underset{(0.028)}{1.4483} X_{t-1}\varepsilon_{t-1} + \underset{(0.033)}{0.44521} X_{t-2}\varepsilon_{t-1} + \varepsilon_t$$

$, \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$

KENYA AIRWAYS (KQ)

$$X_t = \underset{(0.804)}{0.01995} X_{t-1} - \underset{(0.106)}{0.13001} X_{t-2} + \underset{(0.4838)}{0.4534} X_{t-1}\varepsilon_{t-1} - \underset{(0.226)}{0.03278} X_{t-2}\varepsilon_{t-1} + \underset{(0.644)}{0.06848} X_{t-1}\,\varepsilon_{t-2} - \underset{(0.2172)}{0.08098} X_{t-2}\,\varepsilon_{t-2} + \varepsilon_t$$

$\varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$

The p-values for the parameter estimates are given in the parenthesis.

From Table 4.23, basing on a significance level of 0.05, it is clear that for the NSE index, the dependence is significant for errors at lag 1 only and not the observations. The bilinear terms are also significant at lag (1,1) for the same series. The remaining bilinear term at lag (2,1) is not significant. For the NBK, the observations are significant at the second lag and also the interaction between the observations and errors at lags (1,1), (1,2) and (2,2) i.e. the bilinearity is significant at (1,1),(1,2) and (2,2).

In case of the Bamburi series, the observations and the errors as well as their interactions are significant at all the lags for the fitted model. However, for the Kenya Airways series, all the estimated parameters are not statistically significant i.e the dependence on the observations and errors are not significant. This implies that the model for the Kenya Airways is not useful and hence should be discarded but it was kept for comparison purposes with the other models.

It is worth noting that in the bilinear model, a lot of parameters have been estimated. This goes against the principle of parsimony where by models with fewer parameter estimates are preferred. The Goodness of fit statistics and the Diagnostics for the bilinear models are presented in Tables 4.24 and 4.25 respectively. All the fitted models are adequate except the one s for NBK and Kenya Airways which had a significant correlation in the squared standardized residuals

Table 4.24: Goodness of fit statistics for the bilinear models

		BILINEAR (GAUSSIAN)
NSE INDEX	LR	976.549
	AIC	-969.549
	BIC	-954.742
NBK	LR	639.686
	AIC	-631.686
	BIC	-614.098
BAMBURI	LR	893.821
	AIC	-886.821
	BIC	-871.432
KQ	LR	499.608
	AIC	-492.608
	BIC	-477.218

Table 4.25: Diagnostic Tests for the bilinear models

Series	Statistics	BILINEAR
NSE INDEX	Skewness	10.6232
	Kurtosis	178.462
	JB	661208 (0.00)
	Q(12)	4.9609 (0.959)
	$Q^2(12)$	0.0658 (1)
NBK	Skewness	1.3731
	Kurtosis	15.1073
	JB	3853.19 (0.00)
	Q(12)	12.3522 (0.418)
	$Q^2(12)$	42.3188 (0.00)
BAMBURI	Skewness	-2.4333
	Kurtosis	35.7359
	JB	27383.1 (0.00)
	Q(12)	11.9768 (0.448)
	$Q^2(12)$	8.3504 (0.757)
KQ	Skewness	-16.0448
	Kurtosis	345
	JB	2949840 (0.00)
	Q(12)	2.2569 (0.999)
	$Q^2(12)$	0.0348 (1)

Since there was a significant correlation in the squared standardized residuals and parameter instability for the fitted bilinear models, Bilinear-GARCH (BL-GARCH) models were fitted to the respective series employing the MLE method with the Gaussian distribution assumption. The variance equation was aimed at capturing the second order correlation thereby improving the model adequacy and stability. Table 4.26 presents the estimated BL-GARCH models.

Table 4.26: Estimated Bilinear-GARCH models

NSE 20-SHARE INDEX

$$X_t = \underset{(0.00)}{0.42042} X_{t-1} + \underset{(0.89)}{0.00705} X_{t-2} - \underset{(0.00)}{1.61722} X_{t-1}\varepsilon_{t-1} - \underset{(0.00)}{3.30624} X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \ \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$$

$$h_t = 0.01435 + \underset{(0.084)}{0.93004} \varepsilon_{t-1}^2 + \underset{(0.038)}{0.27537} h_{t-1}$$

NBK

$$X_t = \underset{(0.1)}{0.1402} X_{t-1} - \underset{(0.004)}{0.202297} X_{t-2} + \underset{(0.871)}{0.06996} X_{t-1}\varepsilon_{t-1} + \underset{(0.992)}{0.00501} X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$$

$$h_t = 0.03495 + \underset{(0.017)}{0.7798} \varepsilon_{t-1}^2 + \underset{(0.17)}{0.33243} h_{t-1}$$

BAMBURI

$$X_t = \underset{(0.072)}{0.18462} X_{t-1} - \underset{(0.645)}{0.0497} X_{t-2} + \underset{(0.1633)}{1.8581} X_{t-1}\varepsilon_{t-1} + \underset{(0.2024)}{2.0889} X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \ \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$$

$$h_t = 0.01289 + \underset{(0.028)}{0.68802} \varepsilon_{t-1}^2 + \underset{(0.00)}{0.57624} h_{t-1}$$

KENYA AIRWAYS (KQ)

$$X_t = \underset{(0.64)}{0.54092} X_{t-1} - \underset{(0.621)}{0.74813} X_{t-2} + \underset{(0.00)}{0.6948} X_{t-1}\varepsilon_{t-1} + \underset{(0.202)}{0.9503} X_{t-2}\varepsilon_{t-1} + \varepsilon_t, \varepsilon_t \sim WN(0, \sigma_\varepsilon{}^2)$$

$$h_t = 0.0309 + \underset{(0.423)}{2.9182} \varepsilon_{t-1}^2 + \underset{(0.739)}{0.06072} h_{t-1}$$

The probability values for the parameter estimates are given in the parenthesis

At 0.05 significance level, the parameter estimates for the fitted bilinear equations reveal that for the NSE index, observation dependence is significant for lag 1 while the interactions (bilinearity) of observations and errors are significant for the lags (1,1) and (2,1). The variance equation for the NSE index shows that the estimate for β is significant at 5% significance level while α is not significant. For NBK the dependence on the observations is significant for lag 2 but not significant for the interactions between observations and errors. For the variance equation, α is significant while β is not significant.

The Bamburi series exhibited non significant dependence on the observations, errors and the interaction between observations and errors in the mean equation. However, the parameter estimates for the conditional variance equation were all significant at 0.05 significance level. The KQ series showed that the dependence in the interaction at lag (1,1) was significant at 5% level.

The estimates in the variance equation were not significant and this once again shows that the model is not useful for assessing volatility but was kept for the sake of comparison with other models.

The variance equation in all cases gives the sum of α_1 and β_1 approximately equal to 1 or slightly greater than 1. The sum of the parameters α and β gives the rate at which the response function decays (Frimpong and Oteng-Abayie, 2006). This implies that the volatility in the Nairobi Stock Market is highly persistent and all the information is important in forecasting a given stock.

The fitted BL-GARCH models were diagnosed using AIC, BIC and the log likelihood ratio test. The Gaussian MLE criterion was used in parameter estimation for the BL-GARCH models. The BL-GARCH models fitted are adequate since the standardized residuals and squared residuals are not significantly correlated as shown by the Ljung-Box Q statistics. In addition, the J-B statistics strongly rejected the null hypothesis of normality in the residuals for all the series. The Goodness of fit statistics and the diagnostic checks for the bilinear models and the BL-GARCH models are presented in Tables 4.27 and 4.28 respectively.

Table 4.27: Goodness of fit statistics for the BL-GARCH models

		BL-GARCH GAUSSIAN
NSE INDEX	LR	1339.18
	AIC	-1331.18
	BIC	-1314.2
NBK	LR	744.066
	AIC	-737.066,
	BIC	-721.677
BAMBURI	LR	998.459
	AIC	-990.459
	BIC	-972.871
KQ	LR	722.82
	AIC	-714.82
	BIC	-697.232

Table 4.28: Diagnostic tests for the Bilinear-GARCH models

Series	Statistics	BL-GARCH(1,1)
NSE INDEX	Skewness	16.5653
	Kurtosis	336.162
	JB	2372670 (0.00)
	Q(12)	1.8724 (1.00)
	$Q^2(12)$	0.0376 (1.00)
NBK	Skewness	0.7773
	Kurtosis	11.0246
	JB	1670.27 (0.00)
	Q(12)	11.22 (0.51)
	$Q^2(12)$	2.3501 (0.999)
BAMBURI	Skewness	0.1698
	Kurtosis	15.2302
	JB	3742.31 (0.00)
	Q(12)	19.9837 (0.067)
	$Q^2(12)$	3.2934 (0.993)
KQ	Skewness	-5.0986
	Kurtosis	93.705
	JB	208284
	Q(12)	19.9507 (0.068)
	$Q^2(12)$	0.3643 (1.00)

4.4 Efficiency Evaluation in Bilinear and Bilinear-GARCH Models

The model efficiencies were once more evaluated using the Mean Squared Errors. The models that had the minimal MSE were considered the most efficient. However, other statistical properties especially the diagnostics and goodness of fit tests were considered in choosing the most efficient models. The MSE for the bilinear and bilinear-GARCH models are presented in Table 4.29.

Table 4.29: MSE for bilinear and bilinear-GARCH models

Series	BILINEAR	BL-GARCH(1,1)
NSE INDEX	0.0012526	0.0014120
NBK	0.006942	0.0075091
BAMBURI	0.0029757	0.003564
KQ	0.0110732	0.015044

Despite the bilinear models having a relatively smaller MSE, they are unstable as manifested by the residual time plots and hence could be unsuitable for modelling stocks data. However, this problem is easily solved by the inclusion of GARCH models**.** This is because the GARCH model captures the heteroscedastic properties of the series.

The Kurtosis for the BL-GARCH models are the lowest compared to the ones for ARMA-ARCH and bilinear models. This means that the BL-GARCH has successfully captured the heavy tail in the conditional variance of the stock market data. This could be due to interactions between past shocks and volatility in the data. The BL-GARCH is a better alternative to the bilinear models.
In conclusion, the BL-GARCH models are better than the bilinear models as far as the efficiency and statistical properties (diagnostics and goodness of fit) are concerned when applied to the Nairobi Stock data.

4.5 Comparison between ARMA-GARCH, Bilinear and Bilinear-GARCH models

A comparison between the three classes of models was done based on the diagnostic test, goodness of fit statistics in addition to the MSE which showed the efficiency for each model. The AR-IGARCH with student's t-distribution and the bilinear-GARCH models assuming Gaussian distribution emerged as the most efficient models for modelling stock market data while the pure bilinear models were the worst in terms of model adequacy and efficiency. The AR-IGARCH emerged as the most stable because it was strongly stationary while the bilinear-GARCH was the most efficient. This is an indication that the non-linearity in the data sets are best modelled by the bilinear models while the non-stationarity are best captured by the IGARCH.

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

In this study, two classes of models namely bilinear, ARCH and their combinations were applied to the analysis of stock market data. This was motivated by the increasing need to explain the volatility experienced in the Nairobi Stock Market as well as determine the most efficient model for prediction and forecasting the stock market volatility. A comprehensive summary is given below.

5.1.1 ARCH-type models

The original Engle's (1982) ARCH (p) model and its three extensions namely, standard GARCH (p,q), IGARCH(p,q), EGARCH (p,q) and TGARCH (p,q) were applied to the data. Different orders for ARCH(p) were tested in all cases where p=8 was found to be the most adequate for NSE index, Bamburi and KQ while for the NBK series, p=9 provided the best order for ARCH model. Four different p and q values were tested for GARCH(p,q), EGARCH (p,q) and TGARCH (p,q): (1,1), (1,2), (2,1) and (2,2). The order p, q equal to (1,1) is by far the most used values in GARCH research today and results obtained is also consistent with this. In all the four series, the order (1,1) was the best choice. Comparing the diagnostics and the goodness of fit statistics, the IGARCH (1,1) outperformed the ARCH, EGARCH and TGARCH models majorly due to its stationarity in the strong sense. However, the IGARCH model is unable to capture the asymmetry exhibited by the stock data. The EGARCH (1,1) and the TGARCH (1,1) are the preferred models to describe the dependence in variance for all the four series studied since they were able to model asymmetry and parsimoniously represent a higher order ARCH(p). However, the standardized residuals still displayed non-normality in all cases.

Judging from the asymmetry parameter ($\gamma_1 <0$) in the EGARCH model, the volatility increases more with the bad news (negative shocks) than the good news (positive shocks) of the same magnitude for the NSE Index. This is not consistent with the findings of Ogum *et al*., (2005, 2006). However, for the individual stocks the asymmetric parameter ($\gamma_1 >0$) meaning that volatility increases for good news more than bad news of the same magnitude. This implies that the leverage effect may not be a universal phenomenon after all. From the different distributions tested and estimated, the student's t distribution was the best choice for NSE index while GED was the best for NBK, Bamburi and Kenya Airways. The Gaussian assumption provided the poorest results and in some cases had convergence failures.

5.1.2 Bilinear models

Considering the bilinear model, the Gaussian assumption was more appropriate when employing the MLE criterion. In addition, the models seemed to have many cases of convergence problems when the MLE was implemented. The residual time plots for the bilinear models manifested sharp spikes outside the standard error band. This implies the instability of the bilinear models. Interestingly, the MSE for the bilinear emerged the lowest in all cases. This is quite challenging since the models seem very efficient but could not be considered due to their instability.

5.1.3 Bilinear-GARCH models

To address the instability manifested by the residuals for bilinear models, a GARCH (1,1) was fitted to the residuals of the bilinear models. The results indicated an improvement as far as the residuals are concerned especially in the reduction of residual Kurtosis. The BL-GARCH captured the asymmetry better than the bilinear and the ARMA-GARCH models.

5.2 Conclusions

From the results obtained both in terms of efficiency in analyzing volatility and statistical properties, the ARCH class of models outperformed the bilinear models. The bilinear model has been unable to efficiently capture the volatility present in the four series. The bilinear models in all cases had the lowest MSE compared to the ARCH-type models but their lack of stationarity and thus invertibility implies that there might be problems in using them especially for long range forecasting. In fact subsequent analysis revealed that they gave poor long-term forecasts. The forecasts diverged and were not able to predict the periodic behavior observed, because bilinear models are not designed to reflect such behavior (Priestly, 1988).

A comparison of ARMA-GARCH and Bilinear-GARCH shows that the BL-GARCH models and the AR-IGARCH had the lowest MSE and strong convergence during the estimation process hence ensuring efficiency in the models and stability in the estimated model parameters. In essence the BL-GARCH performs as good as the AR-IGARCH. This is because in this case, the bilinear model captured the conditional mean while the GARCH model captured the conditional variance.

5.3 Recommendations

When modelling the NSE Index, NBK and Bamburi series, the AR-IGARCH and BL-GARCH emerged as the best models while for the Kenya Airways it was the AR-IGARCH model that provided the best fit to the series. However, considering the models' stationarity, the AR-IGARCH assuming the student's t-distribution is recommended for modelling the volatility of the Nairobi Stock Market both for the NSE index and the company share prices.

5.4 Further Research

Modelling the NSE data using the Infinite Variance stable process is recommended for further research. In addition, the application of Bilinear-IGARCH models using various distribution assumptions could also be carried so as to determine their efficiency.

6.0 REFERENCES

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

Akaike, H. (1974). A new look at the statistical model identification. *IEE. Autom. control* **19**, 716-723.

Akaike, H. (1977). An entropy maximization principle. *Proc. Symp. on Applied Statistics* (Ed. P. Krishnaiah). North-Holland: Amsterdam.

Andrew, W and Whitney, K. (1986). A large sample chow for the single linear simultaneous equation. *Economic letters* **19**,351-353.

Appiah-Kusi, J. and Pescetto, G. (1998). Volatility and Volatility Spill-Overs in Emerging Markets,CERF Discussion Paper Series No. 98-04.

Baudouhat, A.V. (2004). *Nordic Financial Market Integration:An Analysis with GARCH.* Unpublished Msc Thesis, Göteborg University

Bera, A.K. and Higgins, M.L. (1993). On ARCH models: Properties, Estimation and Testing, *Journal of Economic Surveys*, **7**, 305-366.

Berndt, E.K, Hall, B.H., Hall, R.E and Hausman, J. (1974). Estimation and Inference in Nonlinear Structural Models. *Annals of Economic and Social Measurement*, **3**,653-665

Bertram, W. (2004). *An empirical investigation of Australian Stock Exchange Data*. School of Mathematics and Statistics, University of Sydney.

Bodicha, H.H. (2003). *Forecasting of market share prices for selected companies at the Nairobi Stock Exchange based on econometric models.* Unpublished Msc Thesis, Egerton University.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of econometrics,* **31**, 307-327.

Box,G.E.P and Jenkins, G.N. (1976). *Time series analysis forecasting and control*. San Francisco: Holden Day.

Brillinger, D.R.and Segundo, P. (1979). Empirical examination of the threshold model of neuron firing. *Biol. Cybern.,* **35**, 213-220.

Brooks, R. D., Davidson, S. and Faff, R. W. (1997). An Examination of the Effects of Major Political Change on Stock Market Volatility: The South African Experience, *Journal of International Financial Markets, Institutions & Money*,**7**, 255 - 275.

Curto, J.J.D.(2002). *Modelling the volatility in the Portuguese Stock Market: A Comparative Study with German and US markets*. Unpublished Phd Thesis, Lisbon University.

Davidson, E.H.J. (2008). Time series modelling in http://www.timeseriesmodelling.com/

Degiannakis, S. and Xekalaki, E. (2004). Autoregressive Conditional Heteroscedasticity Models: A Review. *Quality Technology and Quantitative Management*, **1(2),** 271-324.

Engle, R.F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K inflation, *Econometrica*, **50**, 987-1008.

Engle, R.F. (1983). Estimates of the Variance of US Inflation Based on the ARCH Model. *Journal of Money Credit and Banking*, **15**, 286–301.

Engle, R.F. (2004). Risk and Volatility: Econometric Models and Financial Practice. *The American Economic Review*, **94**, 405-420.

Engle, R.F.,Lilien, D.and Robins,R. (1987). Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M model, *Econometrica*, **55 (2),** 391-407.

Engle, R.F., and T. Bollerslev (1986). Modelling the persistence of Conditional Variances. *Econometric Reviews* **5**, 1-50.

Quantitative Micro Software (2004), LLC. http://www.eviews.com/

Fama, E. (1965). The behaviour of stock market prices. *J.Business,* **38**, 34-104.

Fama, E, Fisher, L., Jensen, M., Rolls, R. (1969). The adjustment of stock prices to new information. *International Economic Review,* **69**, 1-21.

Fan, J. and Yao, Q.(2003). *Nonlinear Time Series: Nonparametric methods and Parametric methods.* New York: Springer Series in Statistics.

Frimpong, J.M and Oteng-Abayie, E.F.(2006). Modelling and Forecasting Volatility of Returns on the Ghana Stock Exchange Using Garch Models. *American Journal of Applied Sciences,* **3 (10)**, 2042-2048.

Gary, T and Mingyuan, G. (2004). *Intraday Data and Volatility Models: Evidence from Chinese Stocks.* University of Western Sydney.

Glosten, L.,Jagannathan, R, and Runkle, D. (1993). On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance,* **48**, 1779-1801.

Godambe, V.P. (1960). An optimum property of regular maximum likelihood estimation. *Ann. Math. Stat.*, **31**,1208-12.

Godambe, V.P. (1985). The foundations of finite sample estimation in stochastic processes. *Biometrika,* **72**, 419-28.

Guegan, D and Pham, D. T. (1989). A note on estimation of the parameters of the diagonal bilinear model by the least squares method. *Scand. J. Stat.,***16**, 129-136.

Gujarati, D.N.(2003). *Basic Econometrics*. Tata McGraw-Hill Publishing Company, New Delhi

Granger, C.W.J and Andersen, A.P. (1978). *An introduction to bilinear time series models.* Vandenhoeck and Ruprecht. Gottingen.

Haggan, V. and Ozaki, T. (1981). Modelling non-linear random vibrations using an amplitude-dependent autoregressive time series model. *Biometrika,* **68**, 189-196.

Hongyu, P and Zhichao, Z.(2006). Forecasting Financial Volatility: Evidence from Chinese Stock Market. Working paper in economics and finance No. 06/02. University of Durham. United Kingdom.

Jarque, C.M and Bera, A.K. (1980). Efficient Tests for Normality, Homoskedasticity and Serial Independence of Regression Residuals *Economics Letters*, **6**, 255-259.

Jones, D.A. (1976). *Non-linear autoregressive processes*. Unpublished PhD Thesis, University of London

Kantz, H. and Schreiber, T. (2005). *Nonlinear Time Series Analysis*, 2nd Ed., Cambridge University Press, Cambridge.

Lawrence, A. J. and Lewis, P.A.W.(1985). Modelling and residual analysis of non-linear autoregressive time series in exponential variables. *J.R.Stat.Soc*. **B47**, 165-202.

Li, L.-Y. and Liu, D-F. (1985). Threshold modelling on the earthquake magnitude series. *Acta. Geophys.Sin*., **28**, 303-310.

Mandelbrot,B. (1963). The variation of speculative prices. *J. Bus*, **36**, 394-419.

Maravall, A. (1983). An application of nonlinear time series forecasting. *Journal of Business and Economics Statistics ,**1**,66-74.*

McLeod, A.I. and Li, W.K. (1983). Diagnostic checking ARMA Time Series Models Using Squared-Residual Autocorrelations, *Journal of Time Series Analysis*, **4**, 269-273.

Mills, T.C. (1999). *The econometric Modelling of Financial Time Series*, Cambridge Univ.Press.

Muhanji,S.I (2000). *An assessment of investment in securities in emerging capital markets: A case study of the Nairobi Stock Exchange.* Unpublished Msc Thesis, Kenyatta University

Nelson, D. B. (1990). Stationarity and persistence in the GARCH(1,1) models. *Econometric Theory* **6**, 318-334.

Nelson, D.B. (1991): Conditional Heteroskedasticity in Asset Returns: a New Approach. *Econometrica,* **59.2,** 347-370.

Nicholls, D.F. and Quinn, B.G (1982). *Random coefficient autoregressive Models: An introduction,* Lecture Notes in Statistics, Vol. No. **11**. Springer, New York.

Ogum, G., Beer, F., and Nouyrigat, G.(2005). Emerging Equity Market Volatility An Empirical Investigation of Markets on Kenya and Nigeria. *Journal of African Business*, **6**:1/2, 139 – 154.

Ogum, G., Beer, F. and Nouyrigat, G. (2006). An Empirical Analysis of the Kenyan and the Nigerian daily returns using EGARCH Models. In http://www.euromedia-marseille.com/media/Nouyrigat_-_ISINI.PDF

Omosa, F.Y.B. (1989). *Predictive ability of selected asset pricing models on the Nairobi Stock Exchange.* Unpublished MBA Thesis, University of Nairobi.

Ozaki, T. and Oda, H. (1978). Non-linear Time series model identification by Akaike's information criterion. *Proc.IFAC Workshop on Information and Systems, Compeign, France, October 1977.*

Ozaki, T. (1982). The statistical analysis of perturbed limit cycle processes using non-linear time series models. *Journal of Time series Analysis,*. **3**, 29-41.

Petruccelli, J,D. and Davies, N.(1986). Experience with testing for SETAR-type nonlinearity in time series. *Res.Rep.,* Department of Mathematics, Trent Polytechnic, Nottingham.

Poon, S.H. and C. Granger (2003). Forecasting volatility in financial markets: a review, *Journal of Economic Literature*, **XLI**, 478-539.

Priestly, M.B. (1988). *Non-Linear and Non-Stationary Time Series Analysis*, Academic Press, New York

Quantitative Micro Software (2004), LLC. http://www.eviews.com/

Robinson, P.M. (1978). Statistical inference for a random co-efficient autoregressive Model, *Scandinavian Journal of Statistic* ,5, 163-168.

Schwarz, G. (1978).Estimating the Dimension of a Model. *Ann*. *Inst.statistic. Math.***6**,461-464.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies*. **62 (4)**, 639-661.

Subba, R.T.(1981). On the theory of bilinear models. *J.Roy.Statistic. Soc. Ser*.B,43,244-255.

Subba , R.T (1988). Spectral and Bispectral methods for the analysis of non-linear (non-Gaussian) time series signals. *In Nonlinear time series and signal processing,*(ed Mohler, R.R), Lecture Notes in Control and Information Sciences. New York: Springer.

Subba, R. T. and Gabr, M.M (1984). *An introduction to Bispectral Analysis and Bilinear Time Series Models*. Lecture notes in Statistics, 24. New York: Springer.

Slustzky, H.(1937). The summation of random causes as the source of cyclic process. *Econometrica.* **5**, 105-146.

Stenseth, N.C., Falk, W., Chan, K.S, Bjørstad, O.N, O'Donoghue, M., Tong, H., Boonstra, R., Boutin, S., Krebs, C.J., and Yoccoz, N.G (1999). From ecological patterns to ecological processes: Phase and density-dependencies in Canada lynx cycle. *Proceedings of the National Academy of Sciences USA*, **95**, 15430-15435.

Tiao, G.C. and Tsay, R.S (1994). Some advances in nonlinear and adaptative modeling in time series. *Journal of forecasting*, **13**, 109-131.

Tong, H. (1978). On a threshold model. *In Pattern recognition and signal processing*,(ed, Chen, C.H). Amsterdam: Sijthoff and Noordhoff.

Tong, H. (1990). *Non- Linear Time Series. A Dynamical System Approach*. Oxford Science Publications.

Tyssedal, J. S. and Tjøstheim, D. (1988). An autoregressive model with suddenly changing parameters and application to stockmarket prices. *Appl.Stat.*, **37**,353-369.

Walter,A.D. (2005). A structural GARCH model: An application to portfolio risk management Unpublished PhD Thesis, University of Pretoria.

Wecker, W.E.(1981). Asymmetric time series. *Journal of American Statistics Assoc*.**76**,16-21.

Yule, G. U (1927). On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philos, Trans. Roy. Soe. London. Ser.A,***226**, 267-298.

Zakoian, J-M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control.* **18**, 931-955.

Zheng, D.W. and Chen, Z.G. (1982). Predictions on the earth's rotation parameters. *Ann. Shanghai Obs., Acad. Sin.,***4**, 116-120.

Zivot, E. and Wang, J. (2005). *Modelling Financial Time Series with S-Plus*, 2nd ed., Springer, New York.