SHORT TIME SERIES MODELLING: AN APPLICATION TO KENYAN POLITICAL OPINION POLLS AND THE NAIROBI STOCK EXCHANGE MARKET DATA

OTIENO MICHAEL ODUOR

A Thesis Submitted to the Graduate School in Partial Fulfillment for the Requirements of the Award of the Degree of Master of Science in Statistics of Egerton University

EGERTON UNIVERSITY

SEPTEMBER, 2014

DECLARATION AND RECOMMENDATION

DECLARATION

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

Michael Oduor Otieno

SM12/1504/05

Signature: ………………………… Date: …………………………

RECOMMENDATION

This thesis has been submitted for the degree of Master of Science in Statistics with the approval and recommendations of:

Dr. Jesse W. Mwangi

Senior Lecturer of Statistics

Mathematics Department

Egerton University

Signature: ………………………… Date: ………………………

Prof. Ali Salim Islam

Associate Professor of Statistics

Mathematics Department

Egerton University

Signature: ………………………… Date: ………………………

COPYRIGHT

DEDICATION

I dedicate this thesis to my loving parents, Councilor Eric Otieno Ogwambo and Clementine Otieno, for their unconditional support and encouragement throughout my study programme.

ACKNOWLEDGEMENT

I am thankful to the Almighty God for the gift of life and knowledge. Without Him, I would not have had the strength and will to journey this far.

I am greatly indebted to my supervisors, the late Dr. Mwangi and Prof. Ali for their support and guidance throughout my study. I am also thankful to my lecturer Professor Nassiuma.

I am also highly indebted to my parents, and siblings Boaz, Judy, Rose, Edward, Joyce and Christabel for their support and constant prayers. You all gave me the motivation to solder on even when the going was tough. May God bless you all abundantly.

To my friends and course mates, Wagalla, Thiga, Esha, Mwangasha and Anastacia, i convey my gratitude. I also acknowledge relatives and friends that accorded me support in one way or another but are not mentioned here. May God bless you all.

Finally, i wish to thank my loving wife Pamela and daughter Natasha for being there throughout the turbulent moments and for their support and understanding.

ABSTRACT

Modeling of time series with many observations has been a focus of considerable research both in theoretical and empirical applications over the last three decades. However, the problem of short time series modeling has not so far been adequately studied both in theory and practical applications, despite the fact that many real life situations involve fewer observations leading to short time series. This calls for making use of appropriate estimation techniques in order to come up with models that can capture the short time series properties and thus be adequately used for forecasting without losing the principle of parsimony. This study intended to determine efficient short time series models that would be able to capture the underlying characteristics of short time series (opinion polls and stock market data) so as to come up with good forecasts. Appropriate Autoregressive Moving Average (ARMA) and Autoregressive Fractionally Integrated Moving Average (ARFIMA) class of models were fitted to the short time series data. ARIMA-GARCH models were also fitted to the stock market data to model volatility. A model-selection strategy based on the corrected Akaike Information Criterion (AICC) was adopted to determine the correct model specification. Exact maximum likelihood estimation method was used to estimate the model parameters and the Root Mean Square Error (RMSE) used to evaluate the forecast performance of the models. The political opinion polls data used were obtained from the Infotrak Harris Research, Consumer Insight Research and Strategic Research for the period between September and December 2007. The stock market data were obtained from the Nairobi Stock Exchange. The weekly average company share prices for Access Kenya Group and Safaricom Limited were used. ARFIMA models are found to outperform ARMA models in forecasting the short time series polls data. ARIMA-GARCH model fitted better the Access Kenya data while for the Safaricom data, ARIMA model had the least RMSE values.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

ACF	Autocorrelation Function
AIC	Akaike Information Criterion
AICC	Corrected Akaike Information Criterion
AR	Autoregressive
ARCH	Autoregressive Conditional Heteroscedasticity
ARFIMA	Autoregressive Fractionally Integrated Moving Average
ARIMA	Autoregressive Integrated Moving Average
ARMA	Autoregressive Moving Average
BIC	Bayesian Information Criterion
CDSC	Central Depository and Settlement Corporation
CMA	Capital Markets Authority
EAR	Exponential Autoregressive
FCVECM	Fractionally Cointegrated Vector Error Correction Model
FPE	Final Prediction Error
I.I.D	Independently and Identically Distributed
LRS	Likelihood Ratio Statistic
LRT	Likelihood Ratio Test
MA	Moving Average
MLE	Maximum Likelihood Estimation
NASI	Nairobi All Share Index
NSE	Nairobi Stock Exchange
NYSE	New York Stock Exchange
PACF	Partial Autocorrelation Function
RCA	Random Coefficient Autoregressive
SACF	Sample Autocorrelation Function
SBC	Schwarz Bayesian Criterion
SIC	Schwarz Information Criterion
SPACF	Sample Partial Autocorrelation Function
TSMod	Time Series Modeling
WN	White Noise

CHAPTER ONE

INTRODUCTION

1.1 Background Information

Over the last few years there has been growing interest among political scientists in applying time series techniques to analyze the statistical properties of aggregate political popularity data in various formats such as approval levels and partisanship measures. For example, Box-Steffensmeir and Smith (1996, 1998), Byers, Davidson and Peel (1997, 2000), Eisinga, Franses and Ooms (1999), Dolado, Gonzalo and Mayoral (2003), Lebo, Walker and Clarke (2000), Clark and Lebo (2003), Asikainen (2003) and various articles included in a special issue of Electoral Studies (2000) report evidence for the United States, United Kingdom, Spain, Sweden, Finland and several other OECD countries. These studies indicate that the time series of poll ratings in those countries are well modeled by fractionally-integrated processes which present high persistence but that eventually revert to their mean. Byers *et al.,* (1997) have argued that voters can be grouped into the committed and the uncommitted. The committed individuals are those with strong party allegiances. The uncommitted individuals on the other hand, who are usually called "floating voters", tend to award their votes on the basis of performance. The voting intentions of the floating voters are on the whole a poorer predictor of future voting intentions than those of the committed voters (Byers *et al.*, 1997). The average voter forgets eventually, but not rapidly implying that voters have a long memory of events.

The stock market, on the other hand, is a place where shares representing ownership of corporate enterprises or documents in respect of corporation or loans made to the government can be traded (Bodicha, 2004). The performance in the stock market is influenced by the existing economic, political and socio-cultural environment. Changes in the macro economic variables such as interest rates, exchange rates and inflation rates would lead to wide share price fluctuations (Bodicha, 2004). Other factors that could influence the share prices include changes in bank rates, variations in hire purchase regulations, the publication of foreign trade Figures, expectation of dividends, quality of management and fiscal policies such as taxation. Forecasting or predicting the movement of stock prices is therefore important for planning and control of all business operations.

1.2 The opinion poll research companies

The Nation Media Group contracted three research companies to conduct research on the voting intentions of the voters prior to the 2007 general elections. The research companies were mandated with obtaining information from a sample of eligible voters on their voting intentions in relation to their preferred presidential candidate and political party support. The companies contracted were Strategic Research, Consumer Insight Research and Infotrak Harris Research. These research groups released their results on the popularity of the presidential candidates and the popularity of the parties on a weekly basis. They conducted their observations independent of one another using different samples and sample sizes. Infotrak Harris Research is a professional research company that was founded in October 2004; Consumer Insight Research was incorporated in Nairobi in 1988. The research companies carried their observations between the months of September and December 2007.

1.3 The Nairobi Stock Exchange

The vision of the Nairobi Stock Exchange is “to be a leading securities exchange in the world.” The mission is “to provide a world class trading facility for wealth creation.” The Nairobi Stock Exchange was established in 1954. There are about 58 listed companies in the NSE (52 equities and 6 corporates). There are over 70 Government of Kenya treasury bonds listed on the fixed income segment of the securities exchange. The NSE market is regulated by the Capital Markets Authority of Kenya (CMA (K)). The Authority is a government body mooted in 1989, under the Ministry of Finance and through the Capital Markets Authority Act Cap 485A. The instruments traded at the NSE are Equities, Preference Shares, Treasury Bonds and Corporate Bonds. The main indices are the NSE 20-share index and the Nairobi All Share Index. The NSE 20-share index is the geometric mean of 20 companies share prices while the Nairobi All Share Index is the geometric mean of all the companies share prices. Deliveries and settlements are done through the Central Depository and Settlement Corporation (CDSC).

The major role that the stock market plays is that it promotes a culture of thrift, or saving. People are encouraged to consume less and save more by the fact that there are institutions where they can safely invest their money and in addition earn a return. Stock markets promote higher standards of accounting, resource management and transparency in the management of businesses. This is made possible by the fact that there is separation of owners of capital, on the one hand, from managers of capital, on the other. The stock exchange improves the access to finance of different types of users by providing the flexibility for customization. The financial sector allows the different users of capital to raise capital in ways that are suited to meeting their specific needs. The stock exchange provides investors with an efficient mechanism to liquidate their investments in securities. The investors are certain of the possibility of selling out what they hold, as and when they want, and this guarantees mobility of capital in the purchase of assets.

1.4 Statement of the problem

The bulk of literatures available on time series are based on the principle that the series have many observations, yet there are many situations in life where very short time series occur, particularly in environmental and ecological studies. Forecasts from models fitted to such data may be poor due to the biases in the parameter estimates. The opinion polls data having been collected between September and December 2007 had few observations and this was the case with the NSE data as the companies considered were just recently listed in the stock market. This study therefore sought to find out if these two sets of data could adequately be used to forecast the outcome of the general election and share price movement respectively. A comparative study of the fitted models was also performed to come up with the most efficient model for forecasting purposes by fitting appropriate ARMA and ARFIMA class of models to the short time series data.

1.5 Objectives

1.5.1 Main objective

The main objective of the study was to apply and compare models fitted for short time series and come up with the most efficient model.

1.5.2 Specific objectives

i. To fit appropriate short time series ARMA models to the 2007 Kenyan political opinion polls data and the Nairobi Stock Exchange stock data.

ii. To fit appropriate short time series ARFIMA models to the 2007 Kenyan political opinion polls data and the Nairobi Stock Exchange stock data.

iii. To compare the efficiency of the fitted models for short time series.

1.6 Hypotheses

H_0 The 2007 Kenyan political opinion polls data and the Nairobi Stock Exchange stock data do not significantly fit the short times series ARMA models.

H_0 The 2007 Kenyan political opinion polls data and the Nairobi Stock Exchange stock data do not significantly fit the short times series ARFIMA models.

H_0 There is no significant difference in the efficiency of the fitted short time series models.

1.7 Justification

The dissemination of right information is the starting point for coming up with good policies and making wise decisions. Such information can be used by political scientists and political advisors to help in the steering of the campaigns and to help them focus on the major issues that need to be addressed. The stock market on the other hand plays an important role in the development of an economy, and as such, any information obtained on the stock market should be accurate enough to help in its development. Investors in the stock market mostly rely on the information they have about the possible share price movements to make important decision on the shares to hold, buy or sell. However, for emerging markets with relatively short history, conventional time series analysis would yield models that give poor forecasts. This study therefore intended to come up with an appropriate short time series model that would give accurate information for forecasting purposes that could aid the political advisors and investors in the stock market in policy formulation and decision-making processes.

1.8 Definition of terms

Equities	These are a company's ordinary shares that carry no fixed interest.
Preference shares	Business shares that give the owners the right to be paid interest before any money is paid to the owners of the ordinary shares.
Shares	These are financial instruments where one acquires ownership stakes of a company.
20-share index	This is an equi-weighted geometric mean of twenty large ordinary stocks traded at the NSE.

CHAPTER TWO

LITERATURE REVIEW

2.1 Opinion polls time series analysis

Following the seminal works of Goodhart and Bhansali (1970), numerous studies have examined the evolution of voting intentions, as measured by opinion polls, and in particular relationship between political popularity and economic variables such as inflation and unemployment. Byers *et al.*, (1997) tried to study the degree of persistence in political popularity. They examined the monthly Gallup data on party support in the UK and found that the series were virtually pure "fractional noise" processes. The works of Hall (1978), Holden and Peel (1985) and Chrystal and Peel (1987) found that the effect of news about the economy on voting intentions would be permanent. The implication of their model is that the time series of opinion data should behave like a random walk, with the autoregressive moving average representation of the time series containing an autoregressive root of unity. Byers (1991) rejected the unit root hypothesis in favour of stationary autoregressive moving average models although with autoregressive coefficients close to unity. Similar results are reported by Scott, Smith and Jones (1977). Such models would imply that the effect of news on voting intentions, although it could quite be persistent in practice, is in principle transitory. The works of Byers *et al.*, (1997, 2000, and 2002) have concluded that poll data series appear fractionally integrated; they are covariance nonstationary, but mean reverting and therefore not random walks. Box-Steffennsmeier and Smith (1996) obtained evidence in favour of long memory in the time-series behaviour of aggregate partisanship for Republicans and Democrats in the US. Dolado *et al.*, (2003) found evidence of long memory in Spanish poll data. Asikainen (2003) found long memory and structural breaks in Finnish and Swedish party popularity series. Clarke and Lebo (2003) employed Fractional Error Correction models on the governing party support in Britain. Davidson, Byers and Peel (2005) have presented two versions of a fractionally cointegrated vector error correction model (FCVECM) in the analysis of poll evidence from the UK.

2.1.1 Fractional integration in opinion polls series

Granger (1980) has shown that fractionally integrated data can be produced by two types of aggregation that are of interest to political scientists. First when data are aggregated across heterogeneous autoregressive processes the resulting series will be fractionally integrated. For example if the presidential approval time series shows a different pattern for highly politically sophisticated respondents than for lesser sophisticated respondents we would expect that the aggregated data would be fractionally integrated. Zaller (1992) has shown in detail the impact of political awareness and sophistication and one might reasonably hypothesize that the highly politically aware might use information from farther back in time to drive their evaluations of the presidency while those individuals low in political awareness might use their impressions of the current presidency or of what the presidency has done in the last one month to inform their overall evaluation. Modeling such data as stationary or integrated without testing for fractional integration would lead one to draw incorrect conclusions about the nature of the political process. Second, if the data involve heterogeneous dynamic relationships at the individual level, which are then aggregated to form a time series, that series will be fractionally integrated (Granger, 1980; Lebo *et al.*,, 2000). So if different sets of individuals evaluate Presidency or Parliament in different ways, aggregating those individuals will produce fractional integration. Zaller (1992) has shown that when elites are polarized on an issue, the public becomes polarized as well, usually along partisan lines. Zaller (1992) attributes this effect to cueing information from elites.

DeBoef and Granato (1997) warn that fractional dynamics also become more likely when a variable is bounded at its upper and lower levels. Given that so many variables of interest to political scientists, such as approval levels and partisanship measures, are constructed by aggregating individual-level behaviour and are bounded within a narrow range (often 0-1 or 0-100), we should expect fractional dynamics to be prevalent in political time series.

2.1.2 Micro foundations of the popularity model

The Byers *et al.*, (1997) model is based on the idea that voters fall into two categories; the ‘committed’ and the ‘floating’ voters. Support of the ‘committed’ voters is determined mainly by conviction or group solidarity, and so is relatively insensitive to the current performance of the party. The ‘floating’ voters are more pragmatic, and their support is driven mainly by performance. It follows that the future voting behavior of the second group is typically less predicTable from current behavior than the first group. The degree of persistence of aggregate support depends on the distribution of these attributes in the voter population.

The Byers *et al.,* (1997) model assumes that the log-odds in favor of voter i supporting a given party is described, apart from a deterministic component, by an autoregressive process driven by news. In other words, if p_t^i represents the probability of voter i supporting the party at time t then

$$\log \frac{p_t^i}{1 - p_t^i} = C^i + y_t^i \qquad 2.1.1$$

where

$$y_t^i = \alpha^i y_{t-1}^i + \varepsilon_t^i \qquad 2.1.2$$

The term C^i is time-varying and it captures the effect of the election cycle. Equation (2.1.2) measures the degree of persistence of party support in the face of 'news', whose effect on the individual is measured by ε_t^i

Assuming ε_t^i to be a serially uncorrelated process, the case $\alpha^i = 1$ in (2.1.2) corresponds to a random walk process, which evolves with high probability towards $+\infty$ and $-\infty$, so that the probability of support p_t^i tends to unity or zero under (2.1.1) and p_t^i is defined on the open interval (0, 1). Thereafter, it changes only rarely. This represents the behavior of committed voters. On the other hand $\alpha^i < 1$ implies a reversion to mean, and hence of p_t^i migrating (in the particular case $C^i = 0$) to 1/2, in the absence of news. Because of the nonlinearity of the logistic transformation, support is also a lot more volatile in this case, in the face of the same news, than it is in the unit root case. This case represents the shorter 'memory' of pragmatic voters. The α^i are assumed to be distributed in the voting population over the interval [0, 1] according to the beta(u, v) density, where u and v are constant parameters, and $0 < v < 1$. For a suiTable choice of v, this distribution can concentrate a significant part of the probability mass very close to 1. Since the beta is a very flexible functional form, the distribution can assume a range of shapes on the rest of the interval, depending on the parameters. It can be approximately uniform.

Let $\overline{X}_t$ represent the arithmetic average of N independent binary (0-1) opinion poll responses, sampled from the population at time t, such that $100\,\overline{X}_t$ is the usual percentage support measure.

Consider the time series properties of log[$\overline{X}_t/(1-\overline{X}_t)$] when t represents a succession of time periods (monthly or quarterly). Byers *et al.*, (1997) show that this variable converges in probability as $N \to \infty$ to the same limit as $\overline{C} + \overline{y}_t$, the mean of the right hand side of (2.1.1), where $\overline{C}$ is converging to a constant and

$$\overline{y}_t = N^{-1}\sum_{i=1}^{N} y_t^i$$

$\overline{y}_t$ is a random variable in the limit, being a function of news variable that all voters observe, although the individual effects are averaged out. The key result, due to Granger (1980), is that under a beta(u, v) distribution for the α^i, the time series representation of $\overline{y}_t$ approximates (large N) to a process of the form

$$\overline{y}_t = \sum_{k=0}^{\infty} \alpha_k \overline{\varepsilon}_{t-k}$$

where $\alpha_k = O(k^{-v})$, and $\overline{\varepsilon}_t$ is a shock process depending on news. This says that averaging a mixture of sTable autoregressions and near-unit root processes yields in the limit a moving average process whose coefficients decline hyperbolically. This process has high persistence, or 'long memory', but is nonetheless mean-reverting for $v > 0$. The hyperbolic-decline property is shared by the fractionally integrated or ARFIMA(p, d, q) class of process, which take the form

$$x_t = (1-L)^{-d} u_t$$

where u_t is a stationary ARMA(p, q) process, with $d = 1 - v$. The ARFIMA model, plus a possible deterministic component, is accordingly proposed as a plausible model to represent the time series of log[$\overline{X}_t/(1-\overline{X}_t)$]. When d is close to 1 the series is accordingly more persistent, as is expected since the parameter v is close to 0 when the distribution of α^i is concentrated near 1. The degree of persistence of the aggregate process therefore depends on the proportion of committed voters in the population.

2.2 Financial time series analysis

The analysis of financial time series has been the focus of intense research in the last years (Mantegna and Stanley, 1999). The aim is to characterize the statistical properties of the series with the hope that a better understanding of the underlying stochastic dynamics could provide useful information to create new models able to reproduce experimental facts. In a further step such knowledge might be crucial to tackle relevant problems in finance such as risk management or the design of optimal portfolios, just to cite an example (Turiel and Perez-Vicente 2002). Another important aspect concerns concepts such as scaling and the scale invariance of return fluctuations (Mantegna and Stanley, 1995; Galluccio, Caldarelli, Marsili and Zhang, 1997). The cumulative distribution of the stock price fluctuations has a long tail. The distribution has been shown to be robust, retaining the same functional form for time scales of up to several days.

Stock markets are complex and dynamic. The random walk theory claims that stock price changes are serially independent, but traders and certain academics have observed that they are reasonably predicTable (Klassen, 2005). The two major types of analysis for predicting stock prices include fundamental and technical. The fundamental analysis measures the intrinsic value of a particular stock by studying everything from the overall economy and industry conditions, to the financial conditions and management of companies. It uses revenue earnings, future growth, return on equity, profit margins and other data to determine a company's underlying value and potential for future growth. Technical analysis is a method of evaluating stocks by analyzing statistics generated by market activity, past prices and volume. It looks for peaks, bottoms, trends, patterns and other factors affecting a stock's price movement. Future values of stock prices often depend on their past values and the past values of other correlated variables. Technical analysis looks for patterns and indicators on the stock charts that will determine a stock's future performance. The stock market time series analysis accounts for the fact that data points taken over time may have an internal structure (such as autocorrelation, trend or seasonal variation) that should be accounted for. It derives the stock's future movement from its historical movement, basing on the assumption that there exists strong enough correlation for predication (Klassen, 2005).

That economic time series can exhibit long range dependence has been a hypothesis of many early theories of the trade and business cycles (Lo, 1991). Such theories were often motivated by the distinct but nonperiodic cyclic patterns that typified plots of economic aggregates over time, cycles of many periods, some that seem nearly as long as the entire span of the sample. The presence of long memory components in asset returns has important implications for many of the paradigms used in modern financial economics. For example, optimal consumption or savings and portfolio decisions may become extremely sensitive to the investment horizon if stock returns were long range dependent. Tolvi (2003) defines long memory in time series as autocorrelation at long lags, of up to hundreds of time periods.

There are several studies where evidence of long memory has been detected in monthly, weekly and daily stock market returns, for example, Crato (1994), Cheung and Lai (1995), Barkoulas and Baum (1996), Barkoulas, Baum and Travlos (2000), Sadique and Silvapulle (2001), Henry (2002) and Tolvi (2003). Panas (2001) examined the daily returns of 13 Greek stocks, and found statistically significant long memory in most of the series. Wright (2001) also examined a number of emerging markets and found that long memory is more often found in them than in developed markets. Hiemstra and Jones (1997) reported that US stocks with heavy-tailed return distributions and high (risk-adjusted) average returns are more likely to have long memory. Greene and Fielitz (1977) claimed to have found long range dependence in the daily returns of many securities listed on the New York Stock Exchange (NYSE).

2.3 Linear time series models

A time series is a set of observations x_t each one being recorded at a specific time t (Brockwell and Davis, 1996). It is natural that each observation x_t is a realized value of a certain random variable X_t. A time series model for the observed data $\{x_t\}$ is a specification of the joint distributions (or possibly only the means and covariances) of a sequence of random variables $\{X_t\}$ of which $\{x_t\}$ is postulated to be a realization. The most widely applied set of linear time series models are the autoregressive moving average (ARMA) models. $\{X_t\}$ is an autoregressive moving average process of order (p,q) denoted as ARMA(p,q) if $\{X_t\}$ is stationary and if for every t

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = e_t + \theta_1 e_{t-1} + \cdots + \theta_q e_{t-q} \qquad 2.3.1$$

where $e_t \sim WN\left(0,\sigma^2\right)$.

It is convenient to use the more concise form of (2.3.1)

$$\Phi(B)X_t = \Theta(B)e_t \qquad 2.3.2$$

where $\Phi(\cdot)$ and $\Theta(\cdot)$ are the p^{th} and q^{th} degree polynomials

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \qquad 2.3.3$$

and

$$\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q \qquad 2.3.4$$

and B is the backward shift operator such that $BX_t = X_{t-1}$.

Therefore for the ARMA(p,q) process the observation X_t is linearly related to the p most recent observations $(X_{t-1}, \cdots, X_{t-p})$, q most recent forecast errors $(e_{t-1}, \cdots, e_{t-q})$ and the current disturbance e_t. These types of processes date back to the work of Yule (1927) where he developed the first order autoregressive process denoted as AR (1) and given by the relation

$$X_t = \phi X_{t-1} + e_t, \text{ where } e_t \sim WN\left(0, \sigma^2\right). \qquad 2.3.5$$

The general autoregressive model of order p (AR(p)) can be written as

$$X_t - \phi_1 X_{t-1} - \phi_2 X_{t-2} - \cdots - \phi_p X_{t-p} = e_t. \qquad 2.3.6$$

Rewriting it using the backshift operator

$$\left(1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p\right) X_t = e_t, \text{ where } e_t \sim WN\left(0, \sigma^2\right).$$

Another type of process known as the moving average process was developed by Slutzky (1937). The functional form of the first order moving average process denoted as MA (1) is given by the relation

$$X_t = \theta e_{t-1} + e_t \text{ where } e_t \sim WN\left(0, \sigma^2\right). \qquad 2.3.7$$

The general moving average model of order q (MA(q)) can be written as

$$X_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_p e_{t-q} \qquad 2.3.8$$

Rewriting it using the backshift operator

$$X_t = \left(1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q\right) e_t = \Theta(B) e_t, \qquad 2.3.9$$

where $e_t \sim WN\left(0, \sigma^2\right)$

Wold (1954) combined the AR and MA processes into the autoregressive moving average (ARMA) processes. The most common model of this type is the first order autoregressive moving average ARMA (1,1) process given by the relation

$$X_t - \phi X_{t-1} = \theta e_{t-1} + e_t, \text{ where } e_t \sim WN\left(0, \sigma^2\right). \qquad 2.3.10$$

Let $\{X_t\}$ be a time series with $\mathrm{E}(X_t^2) < \infty$ and mean function $\mu_x(t) = \mathrm{E}(X_t)$. The covariance function of $\{X_t\}$ is $\gamma_x(r,s) = \mathrm{cov}(X_r, X_s) = \mathrm{E}[(X_r - \mu_x(r))(X_s - \mu_x(s))]$ for all integers r, s and t. $\{X_t\}$ is said to be weakly stationary if $\mu_x(t)$ is independent of t and $\gamma_x(t+h,t)$ is independent of t for each h. However, the stationarity condition does not hold for many financial time series since most of them exhibit time changing means and / or variances. Box and Jenkins (1976) suggested differencing as a means of transforming a non-stationary ARMA(p,q) process into a stationary ARMA process known as the Autoregressive Integrated Moving Average abbreviated as ARIMA process. For example a non-stationary ARMA(p,q) process which requires differencing d times before it becomes stationary is said to follow an Autoregressive Integrated Moving Average of order (p,d,q) abbreviated as ARIMA(p,d,q). The difference operator ∇ when applied to the entry X_t yields the difference $\nabla X_t = X_t - X_{t-1}$.

Box and Jenkins (1976) developed a methodology for fitting data to models from the ARIMA class. The approach involves the iterative three-stage cycle namely; model selection, parameter estimation and model checking. This is finally followed by forecasting. Model selection stage involves making decision on the amount of differencing necessary to produce a stationary time series based on visual examination of a plot of the sample autocorrelations (SACF) and sample partial autocorrelation function (SPACF) then the determination of the autoregressive and moving average orders. At the parameter estimation stage, a number of procedures have been proposed which typically yield quite similar estimates when the sample size is large. Some estimation techniques employed include the Yule-Walker estimation criterion, the maximum likelihood criteria, the conditional and unconditional least squares method and the optimal estimation criterion. The model checking stage is where the fitted model is checked if it adequately represents the given data. An approach to model checking is based on the fact that, if the model is correctly specified, the error terms e_t will be white noise.

2.4 Nonlinear time series models

The linear time series models have certain intrinsic limitations. The nonlinear time series models have therefore been introduced to model series that show outliers, series that show cyclicity, series that are time irreversible, series that are asymmetric and they are also used to capture higher moments such as skewness and kurtosis.

One class of nonlinear time series models is the bilinear models. The general discrete time bilinear model takes the form

$$X_t + \sum_{i=1}^{p} a_i X_{t-i} = \sum_{j=0}^{q} c_j e_{t-j} + \sum_{i=1}^{m} \sum_{j=1}^{k} b_{ij} X_{t-j} e_{t-k} \qquad 2.4.1$$

where $\{e_t\}$ is a sequence of i.i.d random variables, usually but not always with zero mean and variance σ_e^2 and $c_0 = 1$, a_i, b_{ij} and c_j are model parameters. 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).

Another class of nonlinear models is the exponential autoregressive models introduced by Ozaki and Oda (1978) and Haggan and Ozaki (1981) in an attempt to construct models which reproduce certain features on nonlinear random vibrations theory. The general EAR models take the form

$$X_t = \left(\phi_1 + \pi_1 e^{\gamma X_{t-1}^2}\right) X_{t-1} + \ldots + \left(\phi_k + \pi_k e^{\gamma X_{t-1}^2}\right) X_{t-k} + e_t \ , \ \gamma > 0 \qquad 2.4.2$$

A sufficient condition for the existence of a limit cycle is then $\dfrac{1 - \sum_i \phi_i}{\sum_i \pi_i} > 1$ or < 0 .

Other examples of nonlinear models include the amplitude-dependent exponential autoregressive (EAR) models. These models were independently introduced by Jones (1976) and Ozaki and Oda (1978). The EAR models are useful in modelling ecological and population data (Ozaki, 1982), wolf's sunspot numbers (solar physics), (Haggan and Ozaki, 1981) and to some extent the economics data (Tong, 1990). Random coefficient autoregressive (RCA) models have been applied to areas such as, ecology and population (Nicholls and Quinn 1982) and Medical data (Robinson, 1978).

The most widely applied nonlinear time series models are the Autoregressive Conditional Heteroscedasticity (ARCH) models introduced by Engle (1982). The model has been applied to model the risk and uncertainty in the financial time series (Fan and Yao, 2003).

2.5 Short-time series

There are many situations in practice where one observes several very short time series (Cox and Solomon, 1988; and Rai, Abraham and Peiris, 1995). Though the theory for time series is well developed to deal with series containing many observations, in this case one cannot rely on the usual estimation or asymptotic theory (Peiris, Allen and Thavaneswaran, 2004). Given a limited amount of data, it is easy to find a model with a large number of parameters to achieve a good fit to the sample data, but post-sample forecasts developed from such models are likely to be disappointingly poor (Brockwell and Davis, 1996). This is because an excellent strategy in building models for forecasting is to seek the simplest possible model, that is, the model with the fewest parameters that appear to provide an adequate description of the major features of the data. This is the principle of parsimony.

The exact likelihood procedure appears more suiTable than approximate procedures when working with small data sets and particularly when estimating models whose characteristic equations have roots close to the boundary of the unit circle (Nicholls and Hall, 1979). Interest in the closed form of the likelihood function of the model stems from the need to make inferences in the following situations; when the sample is small, when the parameters are close to the invertibility boundary, and as an inference function for robust and missing observation problems (Haddad, 1995). Sowell (1992) has also applied exact MLE technique to estimate the parameters of a univariate fractionally integrated time series. He also looked at the small sample properties of the estimators.

It is well known that when data are collected sequentially in time the usual assumption of independence of errors is not guaranteed (Bence, 1995). The uncritical treatment of such data, as though they were a random sample, has been termed “pseudoreplication in time” (Hurlbert, 1984). Often, autocorrelation is positive, so that errors close in time are similar. The effect of ignoring positive autocorrelation is (1) to produce nominal confidence intervals about parameter estimates that are smaller in size than they should be, or in a hypothesis testing context to make too many type I errors (Cochrane and Orcutt, 1949; Hurlbert, 1984) and (2) potentially to produce less efficient estimates of parameters than could be obtained if the autocorrelation were taken into account (Cochrane and Orcutt, 1949). A reasonable approach, especially when dealing with the relatively short time series is to estimate the extent of first order autocorrelation and to adjust estimates and hypothesis tests for the estimated autocorrelation (Stewart-Oaten, 1987; Carpenter, Frost, Heisey and Kratz, 1989).

In a situation where one suspects the serial correlation among observations in a short realization, it is reasonable to begin with a standard autoregressive moving average (ARMA) type model given by the relation

$$\Phi(B)X_t = \Theta(B)e_t$$

where $\Phi(B) = I - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\Theta(B) = I + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$ are stationary autoregressive and invertible moving average polynomials in B, I is the identity operator; $\{e_t\}$ is a sequence of uncorrelated random variables (not necessarily independent) with zero mean and variance σ^2. When the sample size, n, is small one can handle the situation by taking m independent repeated measurements on (2.3.2). In this case write (2.3.2) as autoregressive moving average (ARMA) type model for $t = 1,2,\cdots,n$

$$\Phi(B)X_{it} = \Theta(B)e_{it}, \text{ where } i = 1,2,\cdots,m. \qquad 2.5.1$$

2.6 Autoregressive (AR) model for short time series

The first order autoregressive (AR) model with a nonzero mean is given by the relation

$$(I - \phi B)(X_{it} - \mu) = e_{it}, \text{ where } i = 1,2,\cdots,m \qquad 2.6.1$$

where $|\phi| < 1$ and μ are constants and e_{it} are independent and identically distributed (i.i.d.) $N(0,\sigma^2)$. It is assumed that ϕ and μ remain unchanged for each series, thus $\mu = \mathrm{E}(X_{it})$.

2.6.1 Model selection for AR process in short time series

One of the leading model selection methods is the Akaike Information Criterion (Akaike, 1973). This was designed to be an approximately unbiased estimator of the expected Kullback-Leibler information of a fitted model. As m, the dimension of the candidate model increases in comparison to the n, the sample size, AIC becomes a strongly negative biased estimate of the information. This bias can lead to overfitting even if a maximum cut-off is imposed (Hurvich and Tsai, 1989). A bias correction to the AIC should therefore be considered. The correction is of particular use when the sample size is small, or when the number of fitted model parameters is a moderate to large fraction of the sample size. The bias reduction of AICC compared to AIC is quite dramatic, as is the improvement in the selected model orders. Furthermore, a maximum model order cut-off is not needed for AICC

The AIC criterion for selecting an AR model is given by

$$\text{AIC} = n\left(\log \hat{\sigma}_n^2 + 1\right) + 2(m+1). \qquad 2.6.2$$

If the approximating family includes the operating model, an approximate unbiased estimator of the Kullback-Leibler discrepancy is given by

$$\text{AICC} = n \log \hat{\sigma}_n^2 + n \frac{1+\frac{m}{n}}{1-\frac{(m+2)}{n}}. \qquad 2.6.3$$

The value of the maximum cut-off has no effect on the model chosen by AICC. For many of the other criteria such as FPE, BIC, AIC, SIC, increasing the value of the maximum cut-off tends to lead to increased overfitting of the model.

2.6.2 Estimation of AR parameters in short time series

The estimation of the autoregressive parameters can be done by the conditional maximum likelihood estimation procedures or the exact maximum likelihood estimation procedures. Dahlhaus (1988) has observed that in small samples the exact maximum likelihood estimates are optimal in the model than the conditional likelihood estimates. It is well known that Bartlett adjustment reduces level-error of the likelihood ratio statistic from order n^{-1} to order $n^{-\frac{3}{2}}$. Barndorff-Nielsen and Hall (1988) have shown that the level-error of the adjusted statistic is actually n^{-2}.

The exact maximum likelihood estimation of the autoregressive parameters is computed in the following way: The covariance stationary first order Gaussian Autoregressive AR(1) process is

$$x_t = \rho x_{t-1} + \varepsilon_t \qquad 2.6.4$$

$$\varepsilon_t \sim N\left(0, \sigma^2\right)$$

where $|\rho| < 1$, $t = 1,2,\ldots, T$. The likelihood may be factored into the product of $T-1$ conditional likelihoods and an initial marginal likelihood. Specifically,

$$L(\theta) = l_T\left(x_T / \Omega_{T-1}; \theta\right) l_{T-1}\left(x_{T-1} / \Omega_{T-2}; \theta\right) \ldots l_2\left(x_2 / \Omega_1; \theta\right) l_1\left(x_1; \theta\right) \qquad 2.6.5$$

where $\theta = \left(\rho;\sigma^2\right)'$ and $\Omega_t = \{x_t,..., x_1\}$. The initial likelihood $l_1(x_1;\theta)$ is known in closed form and is given as

$$l_1(x_1;\theta) = (2\pi)^{-\frac{1}{2}}\sqrt{\frac{1-\rho^2}{\sigma^2}}\exp\left[-\frac{1-\rho^2}{2\sigma^2}x_1^2\right] \qquad 2.6.6$$

The remaining likelihood terms are;

$$l_T\left(x_t/\Omega_{t-1};\theta\right) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}}\exp\left[-\frac{1}{2\sigma^2}\left(x_t-\rho x_{t-1}\right)^2\right] \qquad 2.6.7$$

$t = 2,3,..., T$

Beach and MacKinnon (1978) show that small-sample bias reduction and efficiency gains are achieved by maximizing the exact likelihood, which includes the initial likelihood term, as opposed to the approximate likelihood, in which the initial likelihood is either dropped or treated in an ad hoc manner.

2.6.3 Estimation of MA parameters in short time series

Ansley and Newbold (1980) used simulations to investigate the small sample properties of various estimators in autoregressive moving average models. For values of θ close to the invertibility boundaries their simulations show a considerable concentration of $\hat{\theta}$ values at or very near the boundaries ± 1. For large sample sizes the pile up effect becomes less pronounced. Similar observations were made by Cooper and Thompson (1977). However, Cryer and Ledolter (1981) have discovered, by looking at the complete likelihood function, that the global maximum frequently occurs precisely at plus or minus one when θ is near ± 1, even when the sample size is moderate. They show that the exact distribution for $\hat{\theta}$ for a small sample size is of mixed type, a continuous unimodal part over (-1, +1) and nonzero probabilities for $\{\hat{\theta} = -1\}$ and $\{\hat{\theta} = +1\}$.

Following Cryer and Ledolter (1981), the exact likelihood function of the parameters $\left(\theta,\sigma^2\right)$ in the first order moving average process is given as

$$L\left(\theta,\sigma^2/X\right) = \left(2\pi\sigma^2\right)^{-\frac{1}{2}n}\left|\Omega\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}X'\Omega X/\sigma^2\right) \qquad 2.6.8$$

where,

$X' = (X_1,..., X_n)$ and Ω is an $n \times n$ band matrix with elements $\omega_{ii} = 1 + \theta^2$, $\omega_{i,i+1} = \omega_{i+1,i} = -\theta$ and $\omega_{ij} = 0$ otherwise. Maximizing over σ^2 we obtain the concentrated likelihood function

$$L(\theta/X)\alpha \{g(\theta)\}^{-\frac{1}{2}n}, \qquad 2.6.9$$

where $g(\theta) = |\Omega|^{\frac{1}{n}} X'\Omega^{-1}$. Evaluating the determinant and the inverse of Ω we can show that $g(\theta) = X'A(\theta)X$ where the elements of the symmetric matrix $A(\theta)$ are given by

$$a_{ij}(\theta) = \theta^{j-i}\left(\sum_{k=0}^{n}\theta^{2k}\right)^{\frac{1}{n-1}}\left(\sum_{k=0}^{i-1}\theta^{2k}\right)\left(\sum_{k=0}^{n-1}\theta^{2k}\right), \ (i \leq j) \qquad 2.6.10$$

Maximum likelihood estimates can then be found by minimizing $g(\theta)$.

Since $a_{ij}(\theta) = a_{ij}(\theta^{-1})$, it follows immediately that $L(\theta/x) = L(\theta^{-1}/x)$ and that $L(\theta/x)$ always has zero derivative at the invertibility boundary +1 and -1.

2.7 Autoregressive Fractionally Integrated Moving Average (ARFIMA) model

The fractionally integrated ARMA model denoted ARFIMA(p, d, q) has become increasingly popular to describe time series that exhibit long memory (Lieberman, Rousseau and Zucker, 2000). In many cases it provides a more parsimonious description of economic time series data than ARMA models (Doornik and Ooms, 2003). The fractional integration model has also been quite widely employed by political scientists as a way of capturing the characteristics of series such as opinion polls and other indices of political interactions. Works of Box-Steffensmeier and Smith (1996), Clarke and Lebo (2003) which analyses UK approval series and Lebo and Moore (2003) which analyses indices of foreign policy interactions have all utilized the fractional integration model.

Granger and Joyeux (1980) and Hosking (1981) introduced fractional differencing and the general class of autoregressive fractionally integrated moving average (ARFIMA) models. Mandelbrot and Van Ness (1968) presented the fractional Gaussian noise model and Mandelbrot (1969, 1972) promoted the relevance of nonsummable autocorrelation models to economics.

Let B denote the lag operator $BX_t = X_{t-1}$. Then the stationary and invertible ARFIMA(p, d, q) model is written as

$$\Phi(L)(1-L)^d(y_t - \mu) = \Theta(L)\varepsilon_t, \ \varepsilon_t \sim NID(0, \sigma_t^2) \qquad 2.7.1$$

where d is the fractional integration parameter, $\Phi(L) = 1 - \phi_1 L - ... - \phi_p L^p$ specifies the AR lag polynomial, and $\Theta(L) = 1 + \theta_1 L + ... + \theta_q L^q$ the MA polynomial. For stationarity and invertibility, the roots of $\Phi(z)$ and $\Theta(z)$ must lie outside the unit circle.

In ARFIMA(p, d, q) process, p and q are known having been chosen according to Schwarz Information Criterion (SIC).

2.7.1 Maximum Likelihood Estimation of the ARFIMA model

The autocovariance function of a stationary ARMA process with mean μ,

$$\gamma_i = \mathrm{E}[(y_t - \mu)(y_{t-i} - \mu)]$$

defines the variance matrix of the joint distribution of $y = (y_1,..., y_T)'$:

$$V[y] = \begin{bmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{T-1} \\ \gamma_1 & \gamma_0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \gamma_1 \\ \gamma_{T-1} & \cdots & \cdots & \gamma_0 \end{bmatrix} = \Sigma \,, \qquad 2.7.2$$

which is a symmetric Toeplitz matrix, denoted by $\tau[\gamma_0,..., \gamma_{T-1}]$. Under normality:

$$y \sim N_T(\mu, \Sigma)$$

and combined with a procedure to compute the autocovariances in (2.7.2), the log likelihood (writing $z = y - \mu$) is given by

$$\log\ L(d, \phi, \theta, \beta, \sigma_\varepsilon^2) = -\frac{T}{2}\log\,(2\pi) - \frac{1}{2}\log|\Sigma| - \frac{1}{2} z' \Sigma^{-1} z\,. \qquad 2.7.3$$

Additional regression parameters in μ are denoted by β, but can be ignored initially. The autocovariances of a stationary ARMA process scaled by the error variance, $r_i = \gamma_i / \sigma_\varepsilon^2$ is $R = \tau[\gamma_0/\sigma_\varepsilon^2,..., \gamma_{T-1}/\sigma_\varepsilon^2]$. In order to allow maximum likelihood estimation, the autocovariance function must be evaluated in order to construct R. An algorithm for the computation of autocovariances of the ARFIMA process (2.7.1) is derived in Sowell (1992):

$$\gamma_i = \sigma_\varepsilon^2 \sum_{k=-q}^{q} \sum_{j=1}^{p} \Psi_k \zeta_j C(d, p + k - i, \rho_j) \qquad 2.7.4$$

where $\rho_1,..., \rho_p$ are the (possibly complex) roots of the AR polynomial and

$$\Psi_k = \sum_{s=|k|}^{q} \theta_s \theta_{s-|k|} \,, \qquad 2.7.5$$

$$\zeta_j^{-1} = \rho_j \left[\prod_{i=1}^{p} (1 - \rho_i \rho_j) \prod_{\substack{m=1 \\ m \neq j}}^{p} (\rho_j - \rho_m) \right] \qquad 2.7.6$$

where $\theta_0 = 1$. C is defined as

$$C(d, h, \rho) = \frac{\Gamma(1-2d)}{[\Gamma(1-d)]^2} \frac{(d)_h}{(1-d)_h} \times \left[\rho^{2p} F(d+h, 1; 1-d+h; \rho) + F(d-h; 1; 1-d-h; \rho) - 1 \right] \qquad 2.7.7$$

and Γ is the gamma function, ρ_j are the roots of the AR polynomial and $F(a,1;c;\rho)$ is the hypergeometric function

$$F(a, b; c; \rho) = \sum_{i=0}^{\infty} \frac{(a)_i (b)_i}{(c)_i} \frac{\rho^i}{i!} \qquad 2.7.8$$

and we use Pochhammer's symbol

$$(a)_i = a(a+1)(a+2)\ldots(a+i-1), \; (a)_0 = 1 \,. \qquad 2.7.9$$

Computation of $F(a,1;c;\rho)$ can be done recursively,

$$F(a,1;c;\rho) = \frac{c-1}{\rho(a-1)} \left[F(a-1,1;c-1;\rho) - 1 \right] \qquad 2.7.10$$

In the absence of AR parameters (2.7.4) reduces to

$$\gamma_i = \sigma_\varepsilon^2 \sum_{k=-q}^{q} \Psi_k \frac{\Gamma(1-2d)}{[\Gamma(1-d)]^2} \frac{(d)_{k-i}}{(1-d)_{k-i}} \qquad 2.7.11$$

the ratio $\frac{(d)_h}{(1-d)_h}$ for $h = p - q - T + 1, \ldots, 0, \ldots, p + q$ can be computed using a forward recursion for $h > 0$

$$(d)_h = (d + h - 1)(d)_{h-1}, \; h > 0 \,, \qquad 2.7.12$$

and a backward recursion otherwise:

$$(d)_h = \frac{(d)_{h-1}}{(d-h)}, \; h < 0 \,. \qquad 2.7.13$$

2.8 The Autoregressive Conditional Heteroscedasticity (ARCH) model

The ARCH model is the first model of conditional heteroscedasticity. The 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 property implies a dependence on the observations of the immediate past and Heteroscedasticity means time-varying variance also known as volatility. Engle (2004) said that the original idea was to find a model that could assess the validity of the conjecture of Friedman (1977) that the unpredictability of inflation was a primary cause of business cycles. Engle(1982) applied arch model to parameterizing conditional heteroscedasticity in a wage-price equation for the United Kingdom. These models take the dependence of the conditional second moments in modeling considerations unlike ARMA models that focus on modeling the first moment. ARCH models are therefore used for modeling the risk and uncertainty in financial time series (Degiannakis and Xekalaki, 2004; Engle, 2004; Fan and Yao, 2003). Let ε_t be a random variable that has a mean and a variance conditional on the information set ζ_{t-1} (the $\sigma-$ field generated by $\varepsilon_{t-j}, j \geq 1$). The ARCH model of ε_t has the following properties. First, $\mathrm{E}\{\varepsilon_t / \zeta_{t-1}\} = 0$, and, second, the conditional variance $h_t = \mathrm{E}\{\varepsilon_t^2 / \zeta_{t-1}\}$ is a nontrivial positive-valued parametric function of ζ_{t-1}. The sequence $\{\varepsilon_t\}$ may be observed directly, or it may be an error or innovation sequence of an econometric model. In the latter case,

$$\varepsilon_t = y_t - \mu_t(y_t) \qquad 2.8.1$$

Where y_t is an observable random variable and $\mu_t(y_t) = \mathrm{E}\{y_t / \zeta_{t-1}\}$, the conditional mean of y_t given ζ_{t-1}.

Engle (2004) assumed that ε_t can be decomposed as follows:

$$\varepsilon_t = z_t h_t^{1/2} \qquad 2.8.2$$

where $\{z_t\}$ is a sequence of independent, identically distributed (iid) random variables with mean zero and unit variance. This implies $\varepsilon_t / \zeta_{t-1} \sim D(0, h_t)$ where D stands for the distribution (typically assumed to be normal or a leptokurtic one). The following conditional variance defines an ARCH model of order q:

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 \qquad 2.8.3$$

Where $\alpha_0 > 0$, $\alpha_j \geq 0, \quad j = 1, ..., q-1$, and $\alpha_q > 0$.

The parameter restrictions in (2.8.3) form a necessary and sufficient condition for positivity of the conditional variance. Suppose the unconditional variance $\mathrm{E}(\varepsilon_t^2) = \sigma^2 < \infty$. The definition of ε_t through the decomposition of (2.8.2) involving z_t then guarantees the white noise property of the sequence $\{\varepsilon_t\}$, since $\{z_t\}$ is a sequence of iid variables.

2.8.1 The Generalized ARCH model

In application, the ARCH model has been replaced by the GARCH model that Bollerslev (1986) and Taylor (1986) proposed independently of each other. In GARCH model, the conditional variance is also a linear function of its own lags and has the form

$$h_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \qquad 2.8.4$$

The conditional variance defined by (2.8.4) has the property that the unconditional autocorrelation function of ε_t^2, if it exists, can decay slowly, albeit still exponentially.

A sufficient condition for the conditional variance to be positive with probability one is $\alpha_0 > 0$, $\alpha_j \geq 0, \quad j = 1, ..., q$; $\beta_j \geq 0, \quad j = 1, ..., p$.

The GARCH(p, q) process is weakly stationary if and only if $\sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j < 1$, that is, the conditional variance approaches the unconditional variance as time goes to infinity. However, when $\sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j > 1$ then the process is nonstationary.

There exists some circumstances when the parameter estimates in GARCH(p, q) models are close to the unit root but not less than unit, that is, $\sum_{j=1}^{q} \alpha_j + \sum_{j=1}^{p} \beta_j = 1$. In such situations the multi-step forecasts of the conditional variance do not approach the unconditional variance. Such processes exhibit the persistence in variance or volatility in which the current information remains important in forecasting the conditional variance. Engle and Bollerslev (1986) refer to these processes as the Intergrated GARCH or IGARCH. The IGARCH processes do not possess a finite variance but are stationary in the strong sense (Nelson, 1990). The simplest

GARCH(1, 1) is often found to be the benchmark of financial time series modeling because such simplicity does not significantly affect the preciseness of the outcome.

2.8.2 Parameter estimation of the ARCH – type models

The commonly used estimation technique for the ARCH – type models is the maximum likelihood estimation method. Following Bera and Higgins (1993), consider the standard ARCH regression model $y_t / \psi_{t-1} \sim N(x_t'\xi, h_t)$. The log likelihood function 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 $\theta = (\varepsilon', \gamma')$. ε and γ represent the conditional mean and conditional variance parameters respectively. The $(i, j)^{th}$ element of off-diagonal block of the information can therefore be written as

$$\frac{1}{T}\sum_{t=1}^{T} E\left[\frac{\partial^2 l_t}{\partial \xi_i \partial \gamma_j}\right] = \frac{1}{T}\sum_{t=1}^{T} E\left[\frac{1}{2h_T^2}\frac{\partial h_t}{\partial \xi_i}\frac{\partial h_t}{\partial \gamma_j}\right] \qquad 2.8.5$$

When the block is diagonal, estimation and testing of the mean and variance parameters can be carried out separately (Bollerslev, 1986; Higgins and Bera, 1993; Fan and Yao, 2003; Engle, 1982; Davidson, 2008). The log likelihood function is maximized by an algorithm suggested by Berndt, Hall, Hall and Hausman (1974). Starting from the estimates of the r^{th} iteration, the $(r+1)^{th}$ step of algorithm can be written as

$$\xi^{(r+1)} = \xi^{(r)} + \left[\sum_{t=1}^{T}\left(\frac{\partial l_t}{\partial \xi}\right)\left(\frac{\partial l_t}{\partial \xi}\right)'\right]^{-1} \sum_{t=1}^{T}\frac{\partial l_t}{\partial \xi} \qquad 2.8.6$$

and

$$\gamma^{(r+1)} = \gamma^{(r)} + \left[\sum_{t=1}^{T}\left(\frac{\partial l_t}{\partial \gamma}\right)\left(\frac{\partial l_t}{\partial \gamma}\right)'\right]^{-1} \sum_{t=1}^{T}\frac{\partial l_t}{\partial \gamma} \qquad 2.8.7$$

where the 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

$$l(\theta) = T \log \frac{\Gamma((v+1)/2)}{\Gamma(v/2)} - \frac{T}{2}\log \pi(v-2) - \frac{1}{2}\sum_{t=1}^{T}\left(\log h_t + (v+1)\log\left(1 + \frac{\varepsilon_t^2}{(v-2)h_t}\right)\right) \quad 2.8.8$$

where $v > 2$ controls the tail behavior. 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, equation () can be written as

$$l(\theta) = -\frac{1}{2}\log \frac{\Gamma(1/v)^3}{\Gamma(3/v)(v/2)^2} - \frac{1}{2}\log h_t \left(\frac{\Gamma(3/v)(y_t - X_t{}'\theta)^2}{h_t \Gamma(1/v)}\right)^{v/2} \quad 2.8.9$$

where $v > 0$ controls the tail behavior. The GED corresponds to the Gaussian distribution if $v = 0$ and is leptokurtic when $v < 2$.

CHAPTER THREE

MATERIALS AND METHODS

3.1 The scope of the study

The study was confined within the limits of modelling political opinion polls data obtained from Infotrak Harris Research, Consumer Insight Research and Strategic Research and the weekly average company share prices for Access Kenya LTD and Safaricom LTD. Appropriate ARMA and ARFIMA class of models were fitted to the short time series and the model that best captured the underlying characteristics of the short time series was chosen to be used for forecasting purposes.

3.2 Data collection

The study used secondary data obtained from the political opinion pollsters and the Nairobi Stock Exchange. The political opinion polls data used was collected for the period between September and December 2007 and was obtained from Infotrak Harris Research, Consumer Insight Research and Strategic Research. The polls data consisted of 12 observations collected on a weekly basis. The stock exchange data for Access Kenya LTD and Safaricom LTD used was from June 2007 up to 31st December 2008 and was obtained from the NSE. For the selected companies, the weekly average share prices were used.

3.3 Data analysis

The models fitted to the short time series data were ARMA and ARFIMA. The model selection was done by use of corrected AIC while the estimation of parameters was done by use of Exact Maximum Likelihood Estimation. The appropriateness of the models specification was checked by use of a modified Portmanteau statistic. The presence of residuals correlation was tested by use of the Ljung-Box (1978) test while the Jarque-Bera (1987) test was used to test the normality of the residuals. The selection of the most efficient model was done using RMSE. The analysis of the data was done using Excel, SPSS (2003), SAS, TSMod 4.2.5 (Davidson, 2007) and R.

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.1 Preliminary analysis

The preliminary analysis of the data was done by use of time plots which provided basic characteristics of the series. Nine data sets were used for the modeling of the polls data obtained from three pollsters. Each pollster tracked the approval ratings for three main presidential contenders during the 2007 Kenyan general elections. The pollsters used were Consumer Insight Research, Infotrak Harris Research and Strategic Research. For the stock market data, two data sets were used for modeling, that is, Safaricom Kenya Limited and Access Kenya Limited.

Safaricom Ltd is a leading mobile network operator in Kenya. It was formed in 1997 as a fully owned subsidiary of Telkom Kenya. In May 2000, Vodafone group Plc of the United Kingdom acquired a 40% stake and management responsibility for the company. Safaricom's initial public offering of stock, on the NSE, closed in mid April 2008.

Access Kenya is an Internet Service Provider located in Kenya. It was founded in 1995 by brothers David and Jonathan Somen to provide Information and Communications Technology for corporate clients within Kenya, under the name "Communications Solutions Limited". The company changed its name in 2000 to "Access Kenya". In 2007, Access Kenya performed an initial public offering (IPO) of stock on the Nairobi Stock Exchange, which ended in April 2007. Access Kenya then became Kenya’s first publicly listed ICT Company in the NSE.

Figures 1 - 3 below represent the time plots for the opinion polls series while Figure 4 and Figure 5 represent the time plots for Safaricom Ltd and Access Kenya respectively.

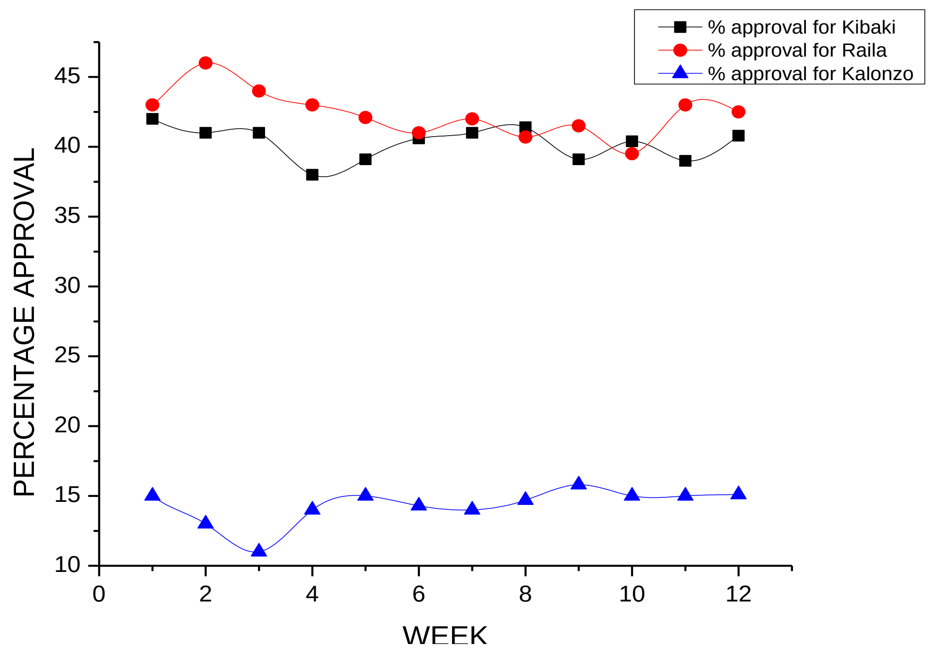

Figure 1: Time series plot for the Consumer Insight series

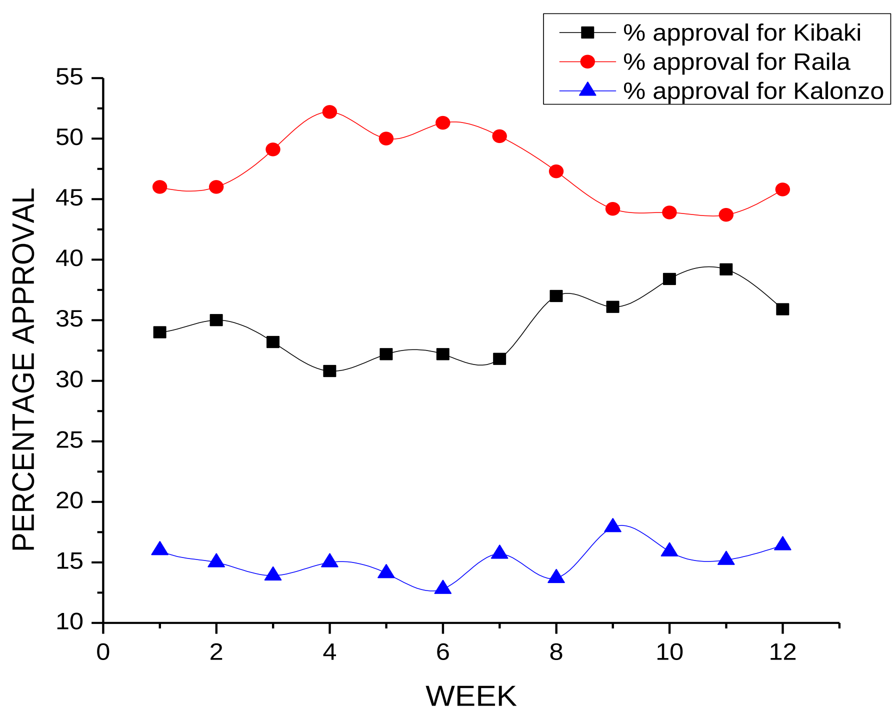

Figure 2: Time series plot for Infotrak Harris series

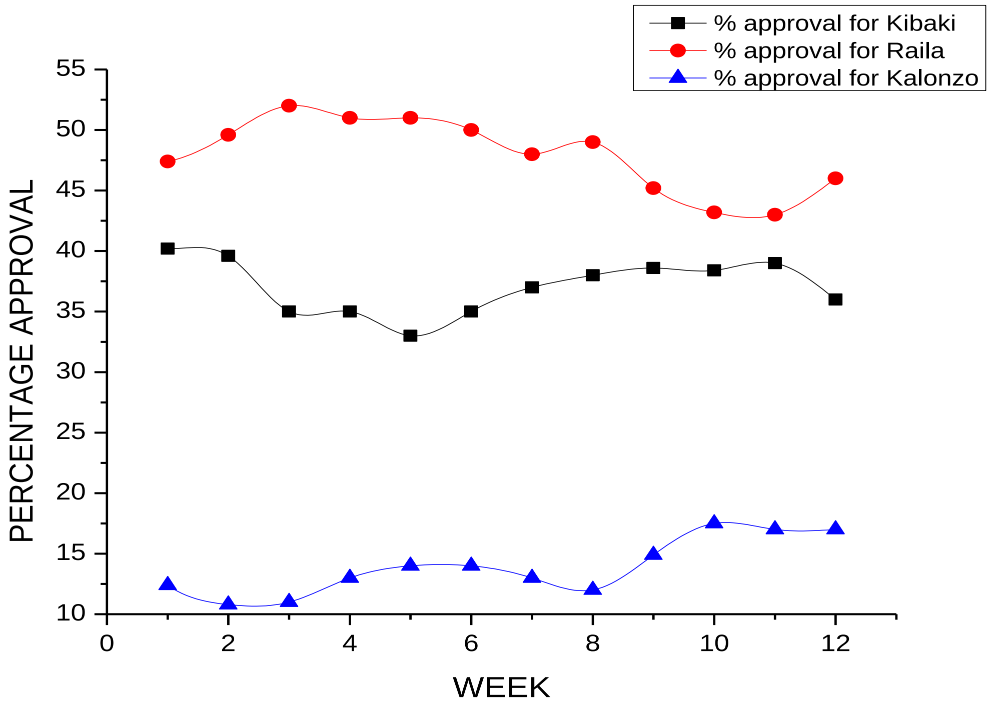

Figure 3: Time series plot for Strategic Research series

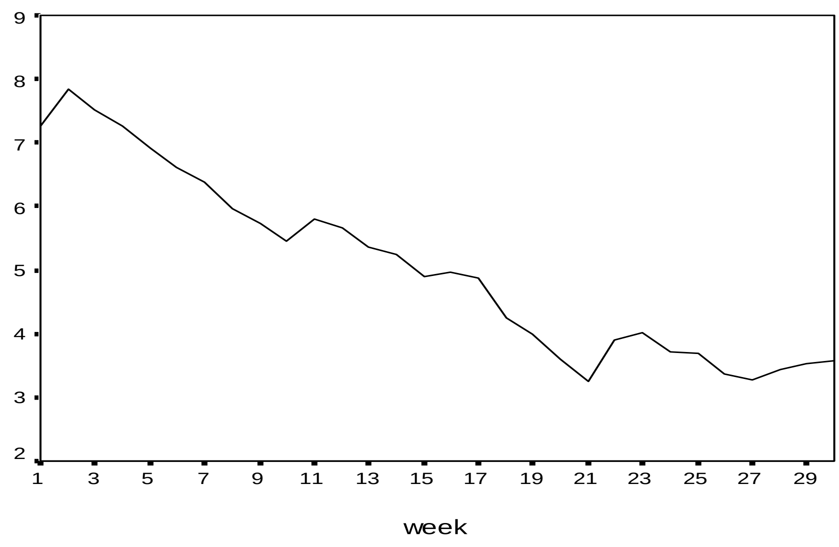

Figure 4: Time series plot for Safaricom weekly average share prices

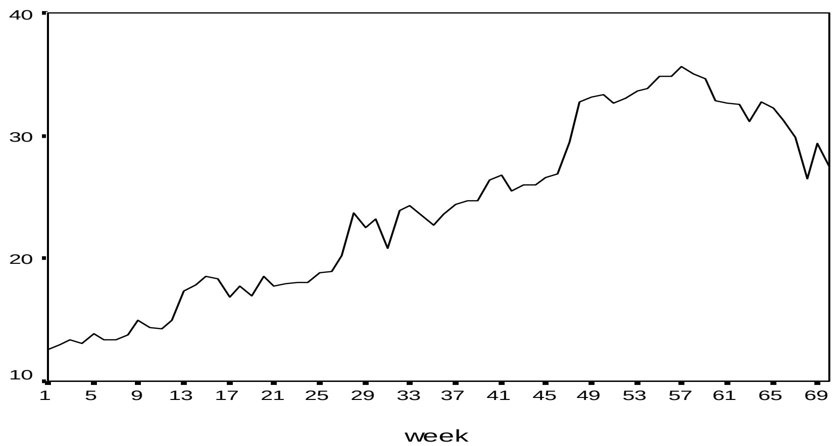

Figure 5: Time series plot for Access Kenya weekly average share prices

A visual inspection of the time plots for the opinion polls series reveals that it is not easy to establish the stationarity or non-stationarity due to short length of the series. However, even though most of the opinion poll series appear to be stationary, it is necessary to estimate the exact decay rate to establish whether the series are stationary or not. The stock market data on the other hand show a very clear trend. The average weekly share prices for Safaricom Kenya show a downward trend while the average weekly share prices for Access Kenya show an upward trend. This is indicative of the fact that the stock market data are not stationary and therefore need some transformation to make them stationary.

The stock market data were transformed by taking the first differences. This transformation was done in order to attain stationarity in the first moment. The time plots for the differenced series are presented in Figures 6 and 7 below.

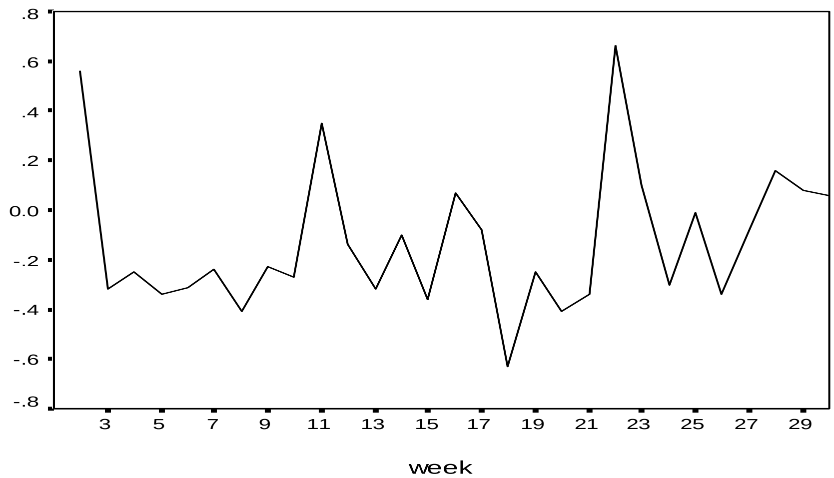

Figure 6: Time series plot for first-differenced Safaricom series

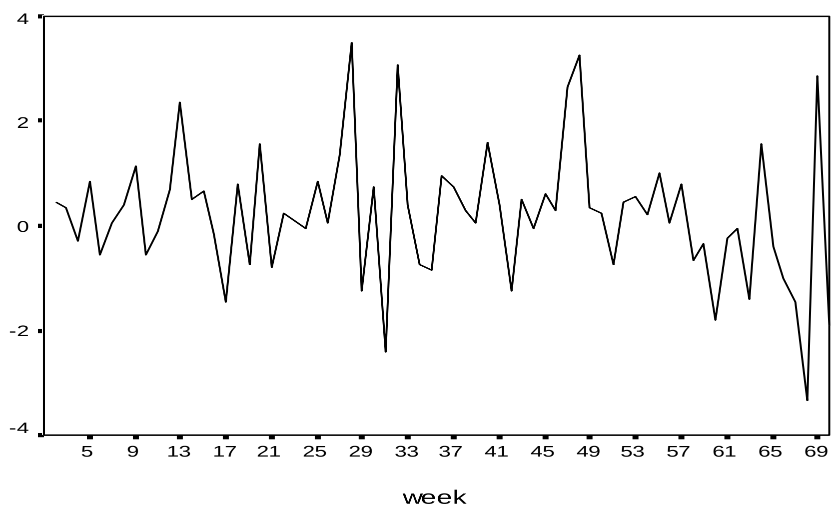

Figure 7: Time series plot for first-differenced Access Kenya series

The first-differenced Access Kenya series shows a characteristic that is prevalent in the financial data, that is, "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(1969). Such series exhibit ARCH effects prevalent in many financial time series data.

4.2 ARFIMA modeling of the opinion polls series

To apply the ARFIMA and ARIMA models the series were first examined for unit root and stationarity. Among the widely applied stationarity tests which include an option for fractional unit roots are the variance ratio test, Rescaled range test, Schmidt-Phillips test and the KPSS test (Asikainen, 2003, Shittu and Yaya, 2009). However these tests share the severely limiting weakness that a long time series ($n \geq 1000$) is needed to distinguish long memory from short memory reliably. An adapTable stationarity test for a series with less than 200 observations is the ADF test. It does not directly indicate whether the series has a fractional unit root but this weakness can be covered if it can be concluded that a series possibly has a fractional unit root when both alternatives are excluded (Asikainen, 2003).

The most exact information on the memory decay process, however, is obtained by estimating the decay rate, d (Asikainen, 2003). Fractional integration estimates also simplify the analysis of time series data by ending debates over the best way to test for unit roots, where one needs to choose among many different tests, such as Dickey Fuller, ADF, variance ratio, or KPSS, and by assumption choose the null hypotheses of d = 1 or d = 0, that is, instead of running multiple tests and looking for patterns suggesting stationarity, one can instead rely upon the point estimates of d (Box-Steffensmeier and Smith 1998). There are three methods of doing this: semiparametric estimation (Geweke and Porter-Hudak, 1983), the approximate maximum likelihood in the frequency domain (Li and Mcleod, 1986, Fox and Taqqu, 1986) and the exact maximum likelihood in the time domain (Sowell, 1992). The first two do not perform well in small samples (Sowell, 1992).

In this study the maximum likelihood method was used to estimate the decay rate d. The estimated long memory parameter estimates were used to transform the series by fractionally differencing them. The equation representing the transformation is given by e. The transformed series were then modeled as ARFIMA processes.

The relative fit of the models was evaluated by use of the corrected Akaike Information Criterion (AICc). AICc was chosen due to its suitability in small sample time series. The optimal models for the fractionally differenced series were the ones that had the least values of the AICc. Optimal models for the fractionally differenced polls data together with the AICc values and 95% confidence intervals for the parameter estimates are given in Tables 1 – 3. Out of the nine series seven of them exhibited long memory characteristics with the value of the fractional differencing parameter ranging between 0.1 and 0.4. Six of the models were pure fractional processes, ARFIMA(0, d, 0), while one was ARFIMA(0, d, 1) process. The short memory moving average parameter part of the ARFIMA(0, d, 1) model was estimated by the exact MLE. The estimates of the fractional differencing parameter d all fell within the 95% confidence interval. Of interest was the fact that for all ARFIMA models, d = 0 and d = 1 did not fall within the confidence intervals. Taking the first differences would therefore lead to overdifferencing. Therefore, in spite of the small sample sizes, the series were found to be stationary and exhibited long-range dependence due to the fact that d was less than 0.5.

Table 1: Optimal ARFIMA(p, d, q) models for the Consumer Insight data

Series	Fitted models	95% CI for d	AICc
Raila series	ARFIMA(0, 0.3395859, 0)	0.3395855 - 0.3395863	42.66
Kalonzo series	ARFIMA(0, 0.09802818, 0)	0.09802783 - 0.09802854	38.96

Table 2: Optimal ARFIMA(p, d, q) models for the Infotrak Harris data

Series	Fitted models	95% CI for d	AICc
Kibaki series	ARFIMA(0, 0.2990022, 0)	0.2990017 - 0.2990027	49.71
Raila series	ARFIMA(0, 0.3528283, 0)	0.3528278 0.3528288	52.51

Table 3: Optimal ARFIMA(p, d, q) models for the Strategic Research data

Series	Fitted models	95% CI for d	q	AICc
Kibaki series	ARFIMA(0, 0.3062491, 0)	0.3062487 0.3062496	-	49.21
Raila series	ARFIMA(0, 0.368705, 0)	0.3687045 0.3687055	-	50.4
Kalonzo series	ARFIMA(0, 0.3342121, 1)		-0.8346534	42.77

4.2.1 Diagnostic tests for the ARFIMA(p, d, q) models

Standard diagnostic tests were done using the residual ACF and PACF of the models. The models passed these tests since their residual values lied within the 95% confidence interval band. Residual plots of the ARFIMA(p, d, q) models were also plotted to examine whether the models were white noise or not. The presence of residuals correlation was tested by use of the Ljung-Box test. The Jarque-Bera test was used to test the normality of the residuals. These are tests for model adequacy and the results are reported in Tables 4 - 6. The Ljung-Box test rejected the presence of serial correlation since all the p-values were greater than 0.05. The models were therefore considered adequate at the 5% level since their residuals were not significantly correlated. The Jarque-Bera tests also showed that the residuals were normally distributed since the p-values were all greater than the 0.05 level of significance.

Table 4: Diagnostic tests for ARFIMA(p, d, q) models for the Consumer Insight data

Series	**Fitted models**	**Diagnostic test**	**Test statistic**	**p-value**
Raila series	ARFIMA(0, 0.3395859, 0)	Ljung-Box (lag 3)	1.7221	0.632
		Jarque-Bera	2.2411	0.3261
Kalonzo series	ARFIMA(0, 0.09802818, 0)	Ljung-Box (lag 3)	2.9545	0.3987
		Jarque-Bera	3.5051	0.1733

Table 5: Diagnostic tests for ARFIMA(p, d, q) models for the Infotrak Harris data

Series	**Fitted models**	**Diagnostic test**	**Test statistic**	**p-value**
Kibaki series	ARFIMA(0, 0.2990022, 0)	Ljung-Box (lag 3)	2.4105	0.4917
		Jarque-Bera	0.6202	0.7334
Raila series	ARFIMA(0, 0.3528283, 0)	Ljung-Box (lag 3)	4.6635	0.1982
		Jarque-Bera	0.5212	0.7706

Table 6: Diagnostic tests for ARFIMA(p, d, q) models for the Strategic Research data

Series	**Fitted models**	**Diagnostic test**	**Test statistic**	**p-value**
Kibaki series	ARFIMA(0, 0.3062491, 0)	Ljung-Box (lag 3)	4.7183	0.1936
		Jarque-Bera	0.847	0.6548
Raila series	ARFIMA(0, 0.368705, 0)	Ljung-Box (lag 3)	3.4343	0.3294
		Jarque-Bera	0.5154	0.7728
Kalonzo series	ARFIMA(0, 0.3342121, 1)	Ljung-Box (lag 3)	3.807	0.2831
		Jarque-Bera	0.6192	0.7338

4.2.2 Forecasting evaluation of the fitted ARFIMA models

After the diagnostic tests, the forecast values were studied and the in-sample forecast values were computed. The forecast performance of the models was evaluated by use of the Root Mean Square Error (RMSE). The model with the lowest value of these forecast evaluation tools was considered to have the best prediction power. The in-sample forecast values together with RMSE values are shown in Tables 7 – 13.

Table 7: Out-of-sample forecasts for the ARFIMA(0, 0.3, 0) model fitted to Raila series from Consumer insight

Week	Optimal forecast	Actual	Error	RMSE
11	41.19523	43	1.80477	0.647994
12	43.47045	42.5	-0.97045	

Table 8: Out-of-sample forecasts for the ARFIMA(0, 0.1, 0) model fitted to Kalonzo series from Consumer insight

Week	Optimal forecast	Actual	Error	RMSE
11	14.31977	15	0.68023	0.341219
12	14.26231	15.1	0.83761	

Table 9: Out-of-sample forecasts for the ARFIMA(0, 0.3, 0) model fitted to Kibaki series from Infotrak Harris

Week	Optimal forecast	Actual	Error	RMSE
11	35.49741	39.2	3.70259	1.205786
12	34.98894	35.9	0.91106	

Table 10: Out-of-sample forecasts for the ARFIMA(0, 0.4, 0) model fitted to Raila series from Infotrak Harris

Week	Optimal forecast	Actual	Error	RMSE
11	46.36690	43.7	-2.6669	0.913922
12	46.91363	45.8	-1.11363	

Table 11: Out-of-sample forecasts for the ARFIMA(0, 0.3, 0) model fitted to Kibaki series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	37.50152	39	1.49848	0.630206
12	37.31383	36	-1.31383	

Table 12: Out-of-sample forecasts for the ARFIMA(0, 0.4, 0) model fitted to Raila series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	46.42137	43	-3.42137	1.138672
12	47.12248	46	-1.12248	

Table 13: Out-of-sample forecasts for the ARFIMA(0, 0.3, 1) model fitted to Kalonzo series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	16.16380	17	0.8362	0.772989
12	14.70307	17	2.29693	

4.3 ARMA modeling of the opinion polls data

Appropriate ARMA models were also fitted to the Kenyan presidential approval polls data. Model selection was done by use of the corrected AIC. After the best models were chosen, the parameters of the models were examined next. The parameters were estimated by the exact maximum likelihood estimation method which is known to perform better in small samples. The results of the parameter estimates of the optimal models together with the values of the corrected AIC are shown in Tables 14 – 16.

Table 14: Optimal ARIMA(p, d, q) models for the Consumer Insight data

Series	Fitted models	p	q	AICc
Kibaki series	ARMA(0, 1)	-	0.2154	41.23
Raila series	ARMA(1, 0)	0.5775598	-	46.12
Kalonzo series	ARMA(0, 1)	-	0.9819	39.46

Table 15: Optimal ARIMA(p, d, q) models for the Infotrak Harris data

Series	Fitted models	p	q	AICc
Kibaki series	ARMA(1, 0)	0.5552499	-	52.7
Raila series	ARMA(0, 1)	-	0.7345	53.24
Kalonzo series	ARMA(1, 0)	-0.084131	-	44.86

Table 16: Optimal ARIMA(p, d, q) models for the Strategic Research data

Series	Fitted models	p	q	AICc
Kibaki series	ARMA(1,0)	0.6202688	-	50.25
Raila series	ARMA(0, 1)	-	0.9930	51.19
Kalonzo series	ARMA(0, 1)	-	1.0000	44.06

4.3.1 Diagnostic tests for the ARMA models

After the optimal models had been fitted, the next step was checking for the adequacy of the models. The model adequacy was tested using two diagnostic tests, Ljung-Box test and Jarque-Bera test. Ljung-Box test was used to check for serial correlation in the residuals while the Jarque-Bera test was used to check for normality. As shown in Tables 17 – 19, the residuals appeared to be white noise and the series were normal since the p-values for the two diagnostic tests were all greater than 0.05 level of significance.

Table 17: Model checking of ARMA(p, q) models for the Consumer Insight data

Series	Fitted models	Diagnostic test	Test statistic	p-value
Kibaki series	ARMA(0, 1)	Ljung-Box (lag 3)	4.2491	0.2358
		Jarque-Bera	0.9109	0.6342
Raila series	ARMA(1, 0)	Ljung-Box (lag 3)	0.7100	0.7010
		Jarque-Bera	0.5109	0.7746
Kalonzo series	ARMA(0, 1)	Ljung-Box (lag 3)	0.6298	0.8896
		Jarque-Bera	3.3988	0.1828

Table 18: Model checking of ARMA(p, q) models for the Infotrak Harris data

Series	Fitted models	Diagnostic test	Test statistic	p-value
Kibaki series	ARMA(1, 0)	Ljung-Box (lag 3)	0.9000	0.6373
		Jarque-Bera	0.7514	0.6868
Raila series	ARMA(0, 1)	Ljung-Box (lag 3)	2.198	0.5323
		Jarque-Bera	0.7986	0.6708
Kalonzo series	ARMA(1, 0)	Ljung-Box (lag 3)	0.5100	0.7747
		Jarque-Bera	0.3434	0.8422

Table 19: Model checking of ARMA(p, q) models for the Strategic Research data

Series	Fitted models	Diagnostic test	Test statistic	p-value
Kibaki series	ARMA(1, 0)	Ljung-Box (lag 3)	1.6800	0.4324
		Jarque-Bera	0.6758	0.7133
Raila series	ARMA(0, 1)	Ljung-Box (lag 3)	1.1723	0.7597
		Jarque-Bera	1.0311	0.5972
Kalonzo series	ARMA(0, 1)	Ljung-Box (lag 3)	3.3816	0.3364
		Jarque-Bera	1.0428	0.5937

4.3.2 Forecasting evaluation of the fitted ARMA models

The out-of-sample forecast values of the fitted models are shown in Tables 20 – 28 below. The Root Mean Square Error values are also displayed in the Tables. RMSE values were used to evaluate the forecast performance of the fitted models.

Table 20: Out-of-sample forecasts for the ARMA(0, 1) model fitted to Kibaki series from Consumer Insight

Week	**Optimal forecast**	**Actual**	**Error**	**RMSE**
11	40.46827	39	-1.46827	0.481318
12	40.39893	40.8	0.40107	

Table 21: Out-of-sample forecasts for the ARMA(1, 0) model fitted to Raila series from Consumer Insight

Week	**Optimal forecast**	**Actual**	**Error**	**RMSE**
11	40.67438	43	2.32562	0.820055
12	41.35266	42.5	1.14734	

Table 22: Out-of-sample forecasts for the ARMA(0, 1) model fitted to Kalonzo series from Consumer Insight

Week	Optimal forecast	Actual	Error	RMSE
11	14.29901	15	0.70099	0.364903
12	14.18340	15.1	0.9166	

Table 23: Out-of-sample forecasts for the ARMA(1, 0) model fitted to Kibaki series from Infotrak Harris

Week	**Optimal forecast**	**Actual**	**Error**	**RMSE**
11	36.47423	39.2	2.72577	0.876065
12	35.40495	35.9	0.49505	

Table 24: Out-of-sample forecasts for the ARMA(0, 1) model fitted to Raila series from Infotrak Harris

Week	Optimal forecast	Actual	Error	RMSE
11	46.72746	43.7	-3.02746	1.149908
12	47.81429	45.8	-2.01429	

Table 25: Out-of-sample forecasts for the ARMA(1, 0) model fitted to Kalonzo series from Infotrak Harris

Week	Optimal forecast	Actual	Error	RMSE
11	14.92428	15.2	0.27572	0.449247
12	15.00637	16.4	1.39363	

Table 26: Out-of-sample forecasts for the ARMA(1, 0) model fitted to Kibaki series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	37.86078	39	1.13922	0.602285
12	37.52632	36	-1.52632	

Table 27: Out-of-sample forecasts for the ARMA(0, 1) model fitted to Raila series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	46.95964	43	-3.95964	1.468564
12	48.42653	46	-2.42653	

Table 28: Out-of-sample forecasts for the ARMA(0, 1) model fitted to Kalonzo series from Strategic Research

Week	Optimal forecast	Actual	Error	RMSE
11	15.34	17	1.66	1.242143
12	13.44	17	3.56	

The observed and forecasted values were closer for the ARFIMA models than for the ARMA models. This showed that ARFIMA models forecasted better the given polls data than the ARMA models. Besides, the RMSE values for the ARFIMA models were found to be lower than those of the ARMA models in five of the seven compared models. ARMA(1, 0) models seemed to be forecasting better than ARFIMA(0, d, 0) models for the Kibaki series obtained from Infotrak Harris and Strategic Research. Therefore, ARFIMA models are the best models for the Kenyan presidential approval data than the ARMA models.

4.4 Modeling of the stock market data

Weekly average prices data were used for modeling. For Access Kenya, the data was made up of 83 observations. However 70 observations were used for fitting the model while the remaining 13 observations were preserved to be used for checking the forecasting ability of the fitted models. For Safaricom series, a total of 30 observations were used for model fitting and 6 observations preserved for use in checking the forecasting ability of the fitted models. The first-differenced series were used for model fitting.

4.5 ARIMA modeling of the stock market data

The corrected AIC was used to select the order of the models. The models that minimized the corrected AIC were considered to be the best. Table 29 below shows the models that minimized the AICC.

Table 29: Optimal ARIMA(p, d, q) models for the stock market data

	Model	**AICC**
ACCESS KENYA	ARIMA(2, 1, 2)	219.2201
SAFARICOM	ARIMA(0, 1, 2)	11.54175

The model parameters were estimated using the exact maximum likelihood estimation method. Exact maximum likelihood estimation method was preferred due to the fact that it works best with short time series. Table 30 below shows the results of the parameter estimates.

Table 30: ARIMA(p, d, q) parameter estimates for stock market data

Parameters	ACCESSKENYA ARIMA(2, 1, 2)	SAFARICOM ARIMA(0, 1, 2)
α	0.2228 (0.1582)	-0.1388 (0.0357)
ϕ_1	-0.6522 (0.1253)	
ϕ_2	-0.1772 (0.1400)	
θ_1	0.5982 (6.1557)	0.3569 (0.2091)
θ_2	0.3999 (0.1212)	-0.6432 (0.1874)

NB: Standard errors of the parameter estimates are given in parentheses

4.5.1 Diagnostic tests

Final models were selected only after checking their adequacy. The models that passed the diagnostic tests were considered as the best. Standard diagnostic tests were done using the residual ACF and PACF of the models. The models passed these tests since their residual ACF values lied within the 95% confidence interval band. The model adequacy was tested using two diagnostic tests, Ljung-Box test and Jarque-Bera test. Ljung-Box test was used to check for serial correlation in the residuals while the Jarque-Bera test was used to check for normality. The results of the diagnostic tests are shown in Table 31 below. The models were adequate since their standardized residuals were not significantly correlated based on the Ljung-Box Q statistics. The squared residuals were also not significantly correlated. The Jarque-Bera test statistic rejected the normality assumption for Safaricom series but failed to reject the same assumption for Access Kenya series.

Table 31: Diagnostics tests for ARIMA(p, d, q) models for the stock market data

	ACCESS KENYA ARIMA(2, 1, 2)	**SAFARICOM ARIMA(0, 1, 2)**
$Q(10)$	13.2812(0.2084)	5.3597 (0.8659)
$Q^2(10)$	9.0296 (0.5293)	12.0415 (0.2823)
JB	3.2116 (0.2007)	6.2717 (0.04346)

JB – represents the Jarque-Bera test statistic for normality

$Q(10)$ - represents the Ljung-Box Q statistics for the standardized residuals

$Q^2(10)$ - represents the Ljung-Box Q statistics for the squared standardized residuals

P – values are given in the parentheses

4.6 ARFIMA modeling of the stock market data

The maximum likelihood method was used to estimate the decay rate d. The results are shown in Table 32 below. The two series both generated a decay rate value of 0.5. This indicated the non-stationarity of the two series since the decay rates fell within the non-stationarity region. The decay rate is used to capture the long range-dependence prevalent in financial data. The estimated long memory parameters were used to transform the series by fractionally differencing them. The transformed series were then modeled as ARFIMA processes.

The relative fit of the models was evaluated by use of the corrected Akaike Information Criterion (AICC). AICC was chosen due to its suitability in small sample time series. The optimal models for the fractionally differenced series were the ones that had the least values of the AICC. Optimal models for the fractionally differenced stock data together with the AICC values are shown in Table 32 below.

Table 32: Optimal ARFIMA(p, d, q) models for the stock market data

	Model	AICC
ACCESSKENYA	ARFIMA(2, 0.5, 4)	261.0835
SAFARICOM	ARFIMA(1, 0.5, 3)	25.66235

The short memory parameters were estimated using the exact maximum likelihood estimation method. The parameter estimates are shown in Table 33 below.

Table 33: ARFIMA(p, d, q) parameter estimates for stock market data

Parameters	ACCESS KENYA ARIMA(2, 0.5, 4)	SAFARICOM ARIMA(1, 0.5, 3)
α	-0.0423(2.8649)	0.0353(0.5786)
ϕ_1	0.6149(0.1354)	0.8383(0.1297)
ϕ_2	0.3297(0.1711)	
θ_1	0.1940(0.1061)	0.5366(0.0347)
θ_2	0.2967(0.0616)	0.0461(0.0238)
θ_3	-0.6257(0.1880)	0.3631(0.0275)
θ_4	0.1401(0.0628)	

NB: Standard errors of the parameter estimates are given in parentheses

4.6.1 Diagnostic tests

The standard diagnostic tests done on the fitted models indicated that the models were all adequate their standardized residuals were not significantly correlated based on the Ljung-Box Q statistics. The squared residuals were also not significantly correlated. The Jarque-Bera test statistics rejected the assumption of normality for the two series. The residual ACF values also lied within the 95% confidence interval band. Since the models were adequate they could then be used for forecasting. The diagnostic results are shown in Table 34 below.

Table 34: Diagnostics tests for ARFIMA(p, d, q) models for the stock market data

	ACCESS KENYA ARIMA(2, 1, 2)	**SAFARICOM ARIMA(0, 1, 2)**
$Q(10)$	14.7172(0.1427)	13.2106(0.2121)
$Q^2(10)$	16.2653(0.09229)	10.4237(0.4041)
JB	135.1218(0.0)	38.9319(0.0)

JB – represents the Jarque-Bera test statistic for normality

$Q(10)$ - represents the Ljung-Box Q statistics for the standardized residuals

$Q^2(10)$ - represents the Ljung-Box Q statistics for the squared standardized residuals

P – values are given in the parentheses

4.7 GARCH modeling of stock market data

GARCH(p, q) models are also fitted to the data to take care of the volatility prevalent in financial data. Before the GARCH models were fitted, ARMA models were applied in order to capture the autocorrelation present in the series. The relative fit of the models were assessed using the Log Likelihood, the Akaike Information Criterion, the Schwarz Information Criterion, the Bayesian Information Criterion and the Hannan-Quinn Information Criterion. The models that maximized the Log Likelihood and on the other hand minimized the Information Criterion were considered to be the best. The model specifications were based on the Student's t-distribution, the Gaussian distribution and the General Error Distribution. The results are displayed in Table 35 below. The General Error distribution worked well for the Access Kenya series while the Student's t-distribution worked well for the Safaricom series. Financial data is highly leptokurtic and heavy tailed which can be best captured by the Student's t-distribution and the General Error Distribution. GARCH(1, 1) was found to be the best model for both series and this was consistent with most studies involving GARCH modeling.

Table 35: Goodness of fit statistics for GARCH models

		GARCH(1, 1)		
		Gaussian	**GED**	**t**
SAFARICOM	**L**	0.512028	0.3414362	5.950181
	AIC	0.3784808	0.4592113	0.07240131
	BIC	0.6613696	0.7892482	0.40243823
	SIC	0.3109640	0.3703570	-0.01645303
	HQIC	0.4670781	0.5625747	0.17576476
ACCESS KENYA	**L**	-100.7327	-97.6907	-99.43724
	AIC	3.151673	3.092484	3.143108
	BIC	3.410700	3.383889	3.434513
	SIC	3.128334	3.063416	3.114040
	HQIC	3.254437	3.208094	3.258718

The parameters for the models were estimated using the Maximum Likelihood Method. The results of the parameter estimates are displayed in Tables 36 and 37 below.

Table 36: Maximum Likelihood parameter estimates for ARIMA(p, q) models

	SAFARICOM **ARIMA(0, 1, 2)-GARCH(1, 1)**	**ACCESS KENYA** **ARIMA(2, 1, 2)-GARCH(1, 1)**
C	-0.24621763(0.0)	0.21300(0.0)
ϕ_1		-0.89869(0.0)
ϕ_1		-0.44102(0.0)
θ_1	0.27240075(0.0495)	0.99609(0.0)
θ_2	-0.10586109(0.4525)	0.90408(0.0)

NB: p – values of the parameter estimates are given in the parenthesis

Table 37: Maximum Likelihood parameter estimates for GARCH(1, 1) model

	SAFARICOM **ARIMA(0, 1, 2)-GARCH(1, 1)**	**ACCESS KENYA** **ARIMA(2, 1, 2)-GARCH(1, 1)**
C	8.58564039(0.4217)	0.72152(0.3061)
α_1	0.00000001(1.0)	0.33352(0.3626)
β_1	0.99999999(0.0)	0.11779(0.8072)

NB: p – values of the parameter estimates are given in the parenthesis

When fitting GARCH models the parameters α and β must satisfy the condition $(\alpha + \beta) < 1$ for stationarity. This condition was satisfied by the GARCH parameters for the Access Kenya series, implying that the model was weakly stationary and that the conditional variance did approach the unconditional variance. The series therefore did have a finite unconditional variance.

However, for the Safaricom series, $\alpha + \beta = 1$. This implies that the multi-step forecasts of the conditional variance do not approach the unconditional variance. In other words, the unconditional variance is infinite. The GARCH model for the Safaricom data was therefore strongly stationary. This was indicative of the fact that the data exhibited persistence in variability or volatility, a case in which the current information remains important in forecasting the conditional variance.

The models also passed the standard diagnostic tests as indicated in Table 38 below. The models were adequate since their standardized residuals were not significantly correlated as indicated by the Ljung-Box Q statistics. The squared residuals were also not significantly correlated.

Table 38: Diagnostics tests for ARIMA-GARCH models for the stock market data

	ACCESS KENYA **ARIMA(2, 1, 2)-GARCH(1, 1)**	**SAFARICOM** **ARIMA(0, 1, 2)-GARCH(1, 1)**
$Q(10)$	8.662302 (0.5644239)	13.11514 (0.2173054)
$Q^2(10)$	4.539743 (0.919732)	2.783844 (0.9860632)
JB	4.485289 (0.1061773)	13.56813 (0.001131666)

JB – represents the Jarque-Bera test statistic for normality

$Q(10)$ - represents the Ljung-Box Q statistics for the standardized residuals

$Q^2(10)$ - represents the Ljung-Box Q statistics for the squared standardized residuals

p – values are given in the parentheses

4.8 Forecasting performance of the models for the stock market data.

In-sample forecast performance of all the models fitted to the stock market data was evaluated. This evaluation was done using the Root Mean Square Error (RMSE). The model with the lowest RMSE value is considered as the best. As displayed in Table 39 below, ARIMA-GARCH models had the lowest RMSE values for the Access Kenya data while ARIMA model had the lowest value for Safaricom data. Therefore ARIMA-GARCH model and ARIMA model were chosen for Access Kenya and Safaricom respectively.

Table 39: RMSE values for models fitted to the stock market data

	Model fitted	**RMSE**
	ARFIMA(2, 0.5, 4)	1.412725
ACCESS KENYA	ARIMA(2, 1, 2)	1.2055
	ARIMA(2, 1, 2)-GARCH(1, 1)	1.1007
	ARFIMA(1, 0.5, 3)	0.307615
SAFARICOM	ARIMA(0, 1, 2)	0.248799
	ARIMA(0, 1 2)-GARCH(1, 1)	0.262097

CHAPTER FIVE

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

In this study, two different data sets were used for analysis, the Kenyan opinion polls data and the Nairobi Stock Exchange data. Both ARMA and ARFIMA class of models were fitted to the opinion polls data while ARIMA, ARFIMA and ARCH-type models were fitted to the stock market data. An attempt was made in trying to identify the models that could best capture the characteristics of the two data sets and thereby be adequately used for forecasting purposes.

5.1.1 Opinion polls data modeling

Long memory ARFIMA(p, d, q) models were used to fit the Kenyan presidential approval data. Seven out of the nine series were found to exhibit stationary long-range dependence with the estimates of the memory decay parameter d ranging between 0.1 and 0.4. Six of these models were pure fractional models while one had the short-memory Moving Average component. AR(1) and MA(1) models were also fitted to the same data. Even though all the models fitted the data well, the forecasts obtained using the long memory ARFIMA models resembled the actual values better than the forecasts using the short memory ARMA models in five of the seven models compared. The forecast values for the fitted ARFIMA(p, d, q) models had smaller RMSE values than the forecast values for the fitted ARMA models.

5.1.2 Stock market data modeling

ARIMA, ARFIMA and ARCH – type of models were fitted to the stock market data. All the models fitted to the two data sets were adequate since they passed the diagnostic tests. The best model among the three was considered to be the one with the best forecasting performance. This was done by looking at the Root Mean Square Error (RMSE) values. The out-of-sample forecast values were generated and the model with the least RMSE value was considered as the best. For the Access Kenya data, ARIMA-GARCH models had the best forecasting performance, while for the Safaricom data ARIMA model had the best forecasting performance. ARFIMA models were the poorest in both data sets.

5.2 Conclusions

In popularity series the assumption of stationarity means very stable popularity shares because of mean reversion. The campaigns during the election cycle matter because they inform potential voters, and as potential voters become more informed, their preferences begin to change. If the series is nonstationary it indicates that the public's presidential preferences during the general election campaign did not simply bounce around a constant mean but rather trended somewhere. It is not, as the election forecasting perspective might suggest, that voters knew the final answer right from the start, but instead voters underwent a process whereby they eventually reached the final answer in the end. If the series was stationary, it would suggest that voter's preferences really did not move much during the general election. ARFIMA models outperformed ARMA models for the polls data. ARFIMA models for Kenya's polls data fell within the stationary regime which suggested that Kenyan voters seemed to have a prior choice of their preferred presidential candidate and even the campaigns do not change very much their opinions. These fractional differencing parameter values were less than those found in mature democracies but close to those found in smaller regional parties in Spain. The lower values of long memory parameter implied the existence of some significant differences in the behaviour of Kenyan voters and those voters in more mature democracies. The implication was that the presidential candidates had a small share of 'non-militant' supporters with the 'militant' or die-hard supporters taking the most prominence.

As for the stock market data, ARIMA-GARCH models outperformed all other models for the Access Kenya series. This model was able to capture the volatility present in financial time series data. For the Safaricom data, ARIMA models seemed to outperform all other models even though the model could not be used to capture any volatility present in the data. The most efficient models were chosen based on the Root Mean Squared Error (RMSE). The models with least values of the RMSE were selected.

5.3 Recommendations

To our knowledge, ARFIMA and ARMA models have not been used to model opinion poll data in Kenya. This is therefore one of the contributions of this research. The second contribution of the research is to be able to predict the outcome of the elections. The study recommends ARFIMA models for the Kenyan opinion polls data. However, further research should be carried out on the Kenya's polls data for the influence of such factors as election cycle, structural break and the regional support influence. Kenyan opinion polls have not been tested before for fractional integration. Therefore, there is no comparable evidence for these results. This leaves the possibility of confusion between long/perfect memory and structural break still open.

As for the stock market data, the study recommends ARIMA-GARCH models over ARFIMA models. Further study needs to be done in modeling short-time stock market data that exhibit persistence in volatility as well as long term dependence. Appropriate models need to be generated especially when such data are strongly stationary.

REFERENCES

Akaike, H. (1973). Maximum likelihood identification of Gaussian autoregressive moving average models. *Biometrika*, **60**, 255-265.

Ansley, C.F. and Newbold, P. (1980). Finite sample properties for autoregressive moving average models. *Journal of econometrics*, **13**, 159-83.

Asikainen, A.L. (2003). Long memory and structural breaks in Finnish and Swedish party popularity series. Available at *http://ethesis.helsinki.fi/julkaisut/val/kansa/disc/586/. pdf.*

Barkoulas, J.T. and Baum, C.F. (1996). Long term dependence in stock returns. *Economic letters*, **53**, 253-259.

Barkoulas, J.T., Baum, C.F. and Travlos, N (2000). Long memory in the Greek stock market. *Applied financial economics*, **10**, 177-184.

Barndorff-Nielsen, O.E. and Hall, P. (1988). On the level-error after Bartlett adjustment of the likelihood ratio statistic. *Biometrika*, **75**, 374-378.

Bartlett, M.S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal statistical society*. Series A, **160**, 268-282.

Beach, C.M. and MacKinnon, J.G. (1978). A maximum likelihood procedure for regression with autocorrelated errors. *Econometrica*, **46,** 51-58.

Bence, J.R. (1995). Analysis of short time series: correcting for autocorrelation. *Ecology*, **76**, 628-639.

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

Bodicha, H.H. (2004). *Forecasting of market share prices for selected companies at the Nairobi Stock Exchange based on econometric models*. MSc. Thesis, Egerton Univ. Njoro, Kenya.

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

Box, G.E.P. and Jenkins, G.M. (1976). *Time Series Analysis Forecasting and Control*. San Francisco: Holden. Day.

Box-Steffensmeier, J.M. and Smith, R.M. (1998). Investigating political dynamics using fractional integration methods. *American journal of political science*, **42(2)**, 661-689.

Box-Steffensmeier, J.M. and Smith, R.M. (1996). The dynamics of aggregate partisanship. *American political science review*, **90**, 567-580.

Brockwell, J.P. and Davis, R.A. (1996). *Introduction to Time Series and Forecasting*. Springer-Verlag, New York.

Byers, J.D. (1991). Is political popularity a random walk? *Applied economics*, **23**, 967-974.

Byers, J.D., Davidson, J. and Peel, D.A. (1997). Modelling political popularity: An analysis of long range dependence in opinion poll series. *Journal of the Royal statistical society*. Series A, **160**, 471-490.

Byers, J.D. Davidson, J. and. Peel, D.A. (2000). The dynamics of aggregate political popularity: evidence from eight countries. *Electoral Studies*, **19**, 49-62.

Byers, J.D. Davidson, J. and. Peel, D.A, (2002). Modelling political popularity: a correction. *Journal of the Royal Statistical Society* Series A **165**, 187-189.

Carpenter, S.R., Frost, T.M., Heisey, D.M. and Kratz, T.K. (1989). Randomized intervention analysis and the interpretation of whole-ecosystem experiments. *Ecology*, **70**, 1142-1152.

Cheung, Y.W. and Lai, K.S. (1995). A search for long memory in international stock market returns. *Journal of international money and finance*, **14**, 597-615.

Chong, T.T.L. (2006). The polynomial aggregated AR(1) model. *Econometrics journal*, **9,** 98-122.

Chrystal, K.A. and Peel, D.A. (1987). What can economics learn from political science and vice versa? *American economic review* (papers and proceedings), **76**, 62-65.

Clarke, H.D. and Lebo, M. (2003). Fractional (Co)integration and governing party support in Britain. *British journal of political science*, **33**, 283-301

Cochrane, D. and Orcutt, G.G. (1949). Application of least squares regression to relationships containing autocorrelated error terms. *Journal of the American statistical association*, **44**, 32-61.

Cooper, D.M. and Thompson, R. (1977). A note on the estimation of the parameters of the autoregressive moving average process. *Biometrika*, **64**, 625-8.

Cox, D.R. and Solomon, P.J. (1988). On testing for serial correlation in large numbers of small samples. *Biometrika*, **75**, 145-148.

Crato, N. (1994). Some international evidence regarding the stochastic behaviour of stock returns. *Applied financial economics*, **4**, 33-39.

Cryer, J.D. and Ledolter, J. (1981). Small sample properties of the maximum likelihood estimator in the first order moving average model. *Biometrika*, **68**, 691-4.

Dahlhaus, R. (1988). Small sample effects in time series analysis:A new asymptotic theory and a new estimate. *The annals of statistics*, **16**, 808-841.

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

Davidson, E.H.J. (2007). http://www.timeseriesmodelling.com. (*Accessed on 20/07/08*)

Davidson, E.H.J., Byers, D. and Peel, D. (2005). *Support for Governments and Leaders: Fractional cointegration analysis of poll evidence from the UK*, 1960-2004.

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

Diebold, F.X. and Schuermann, T. (1996). Exact maximum likelihood estimation of observation-driven econometric models, NBER Technical Working Papers 0194, National Bureau of Economic Research, Inc., USA.

Dolado, J.J., Gonzalo, J. and Mayoral, L. (2003). Long range dependence in Spanish political opinion poll series. *Journal of applied econometrics.* **18(2),** 137-155.

Doornik, J.A. and Ooms, M. (2003). Computational aspects of maximum likelihood estimation of Autoregressive Fractionally Integrated Moving Average models. *Computational statistics and data analysis,* **42,** 333-348.

Durbin, J. (1960). Estimation of parameters in time series regression models. *J. Roy. Stat. Soc*. Ser, B, **22**, 139-53.

Eisinga, R., Franses, P.H. and Ooms, M. (1999). Forecasting long memory left-right political orientations. *International journal of forecasting*, **15,** 185-199.

Electoral Studies (2000). Advances in the analysis of political time series. *Electoral Studies,* **19 (1),** 447-658

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

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

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

Fan, J. and Yao, Q. (2003). *Nonlinear Time Series: Nonparametric methods and Parametric methods*, 1st Ed. New York: Springer Series in Statistics.

Fox, R. and Taqqu, M.S. (1986). Large sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Annals of statistics,* **14,** 517-532.

Franklin, C.H. and Jackson, J.E. (1983). The dynamics of party identification. *American political science review*, **77(4),** 957-973.

Galluccio, S., Caldarelli, G., Marsili, M. and Zhang, Y.C. (1997). Scaling in currency exchange. *Physica* A, **245**, 423-436.

Gelman, A. and King, G. (1993). Why are American presidential election campaign polls so variable when votes are so PredicTable?, *British Journal of Political Science*, 23:409–451.

Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long-memory time series models. *Journal of time series analysis*, **4**, 221-38.

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

Godambe, V.P. (1976). Conditional likelihood and unconditional optimum estimating equations. *Biometrika*, **63**, 277-84.

Goodhart, C.A.E. and Bhansali, R.J. (1970). Political economy. *Political studies*, **18**, 265-279.

Granger, C.W.J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of time series analysis*, **1**, 15-29.

Granger, C.W.J. (1966). The typical spectral shape of an economic variable. *Econometrica*, **34**, 150-61.

Greene, M. and Fielitz, B. (1977). Long-term dependence in common stock returns. *Journal of Financial Economics*, **4**, 339-49.

Haddad, J.N. (1995). On the closed form of the likelihood function of the first order moving average model. *Biometrika*, **82,** 232-234.

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

Hall, P. (1978). Stochastic implications of the life-cycle-permanent income hypothesis: theory and evidence. *Journal of political economy*, **86**, 971-987.

Henry, O.T. (2002). Long memory in stock returns: some international evidence. *Applied financial economics*, **12**, 725-729.

Hiemstra, C. and Jones, J.D. (1997). Another look at long memory in common stock returns. *Journal of empirical finance*, **4**, 373-401.

Holden, K. and Peel, D.A. (1985). An alternative approach to explaining political popularity. *Electoral studies*, **4**, 231-239.

Hosking, J.R.M. (1981). Fractional differencing. *Biometrika*, **68**, 165-76.

Hosking, J.R.M. (1996). Asymptotic distributions of the sample mean, autocovariances, and autocorrelations of long-memory time series, *journal of econometrics*, **73,** 261-284.

Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, **54**, 187-211.

Hurvich, C.M. and Ray, B.K. (1995). Estimation of the memory parameter by non-stationary and non-invertible fractionally integrated processes. *Journal of time series analysis*, **16,** 17-41.

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

Jarque, C.M. and Bera, A.K.(1987). A test of normality of observations and regression residuals. *Int. Stat. Rev.*, **55,** 163-172.

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

Kendall, M.G. (1949). The estimation of parameters in linear autoregressive time series. *Econometrica*, **17**, 44-57.

Klassen, M. (2005). *Investigation of some technical indexes in stock forecasting using neural networks*. Vol. **5** of Proceedings of World Academy of Science, Engineering and Technology.

Lawley, D.N. (1956). A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika*, **43**, 295-303.

Lebo, M.L. and Moore, W.H. (2003). Dynamic foreign policy behaviour. *The journal of conflict resolution*, **47**, 13-32.

Lebo, M.J., Walker, R.W. and Clarke, H.D. (2000). You must remember this: Dealing with long memory in political analysis. *Electoral Studies*, **19**, 31-48.

Lieberman, O., Rousseau, J. and Zucker, D.M. (2000). Small-sample likelihood-based inference in the ARFIMA model. *Economic theory*, **16**, 231-248.

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

Lo, A.W. (1991). Long-term Memory in Stock Market Prices. *Econometrica* **59**, 1279-313.

Mandelbrot, B.B. and Van Ness, J.W. (1968). Fractional Brownian motion, fractional Brownian noise and applications. *SIAM review*, **10**, 422-437.

Mandelbrot, B.B. (1969). Long-run linearity, locally Gaussian processes, H-spectra and infinite variances. *International economic review*, **10**, 80-111.

Mandelbrot, B.B. (1972). Statistical methodology for non-periodic cycles: From the covariance to R/S analysis. *Annals of economic and social measurement*, **1**, 259-290.

Mantegna, R.N. and Stanley, H.E. (1995). Scaling behaviour in the dynamics of an economics index. *Nature*, **376**, 46-49.

Mantegna, R.N. and Stanley, H.E. (1999). *Introduction to econophysics: correlations and complexity in finance*. Cambridge univ. press, Cambridge, UK.

Maravall, A. (1983). An application of nonlinear time series forecasting. *J.Bus.Econ.Stat*.,**1**,66-74.

Mayoral, L. (2006). Minimum distance estimation of stationary and non-stationary ARFIMA processes. *Journal of econometrics,* **10,** 124-148.

Mills, T.C. (2006). Time series modelling of two millennia of Northern Hemisphere temperatures: long memory or shifting trend? *Journal of Royal statistical society*, **170(a),** 83-94.

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

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

Nicholls, D.F. and Hall, A.D. (1979). The exact likelihood function of multivariate autoregressives moving average models. *Biometrika,* **66,** 259-264.

Orcutt, G.H. (1948). A study of the autoregressive nature of the time series used for Tinbergen's model of the economic system of the United States, 1919-1932. *Journal of the royal statistical society,* B-**10**, 1-53.

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 Anal*. **3**, 29-41.

Panas, E. (2001). Estimating fractal dimension using sTable distributions and exploring long memory through ARFIMA models in Athens stock exchange. *Applied financial economics*, **11**, 395-402.

Peiris, S., Allen, D. and Thavaneswaran, A. (2004). An introduction to generalized moving average models and applications. *Journal of applied statistical science,* **13**, 251-267.

Rai, S.N., Abraham, B. and Peiris, S. (1995). Analysis of short time series with an overdispersion model. *Commun. Statist. Theory. Meth*., **24**, 335-348.

Robinson, P.M. (1978). Statistical inference for a random co-efficient autoregressive model. *Scand.J.Stat*.,**5**,163-168.

Sadique, S. and Silvapulle, P. (2001). Long term memory in stock market returns: international evidence. *International journal of finance and economics,* **6**, 59-67.

Scott, A.J., Smith, T.M.F. and Jones, R.G. (1977). The application of time series methods to the analysis of repeated surveys. *International statistical review* **45**, 13-28.

Shittu, O.I. and Yaya, O.S. (2009). Measuring forecast performance of ARMA and ARFIMA models: an application to US Dollar/UK Pound foreign exchange rate. *European journal of scientific research,* **32,** 167-176.

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

Sowell, F.B. (1992). Maximum likelihood estimation of stationary univariate Fractionally Integrated time series models. *Journal of econometrics,* **53**, 165-188.

Stewart-Oaten, A. (1987). *Assessing effects on fluctuating populations: tests and diagnostics*. In ASA/EPA conferences on interpretation of environmental data.

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.

Taniguchi, M. (1986). Third order asymptotic properties of maximum likelihood estimators for Gaussian ARMA processes. *Journal of multivariate analysis,* **18**, 1- 31.

Taniguchi, M. (1988). Asymptotic expansions of the distributions of some test statistics for Gaussian ARMA processes. *Journal of multivariate analysis*, **27**, 494-511.

Tolvi, J. (2003). Long memory in a small stock market. *Economics bulletin,* **7**, 1-13.

Tong, H. (1990). *Non- Linear Time Series. A Dynamical System Approach*. Vol. 6 of Oxford Statistical Science Series. Oxford Science Publications.

Turiel, A. and Perez-Vicente, C.J. (2002). Multifractal geometry in stock market time series. *Physica* A, **322**, 629-649.

Taylor, S.: (1986). *Modelling Financial Time Series*, Wiley, Chichester.

Wold, H. (1954). *A Study in the Analysis of Stationary Time Series (2nd Ed.).* Almquist and Wicksell, Uppsala.

Wright, J.H. (2001). Long memory in emerging market stock returns. *Emerging markets quarterly*, **5**, 50-55.

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