

Strathmore
UNIVERSITY

**MACROECONOMIC FACTORS AND THE PREDICTABILITY OF STOCK RETURNS
IN KENYA**

Mwangi Ebenezer Gachugu

072056

**Submitted in partial fulfillment of the requirements for the Degree of
Bachelor of Business Science in Actuarial Science at Strathmore University**

Strathmore Institute of Mathematical Sciences

Strathmore University

Nairobi, Kenya

December, 2016

Table of Contents

Table of Contents	3
List of tables	4
Table of figures	4
1. Introduction	5
1.1 Background of the study	5
1.2 Problem statement	6
1.4 Research objectives	7
1.3 Research questions	7
1.5 Scope of the study	7
1.5 Justification of the study	7
2. Literature review	8
2.1 Theoretical framework	8
2.2.1 The Efficient Market Hypothesis	8
2.2.2 The Arbitrage Pricing Theorem	8
2.3 Empirical review	12
2.3.1 How to identify predictability of stock returns	13
2.4 Recursive predictions of stock returns	16
3. Methodology	17
3.1 Forecasting	17
3.2 Recursive modelling	19
3.3.1 Model 1	19
3.3.2 Model 2	20
3.3.3 Model selection	20
3.3 Selection of Explanatory Variables	21
3.4 Data sources	22
3.5 Assumptions	22
4.1 Diagnostics	23
4.1.1 Unit root tests	23
4.1.2 Johansen tests of Cointegration	23
4.1 Model performance	23
4.2 Regression and forecasts	25
5. Conclusion	27
5.1 Recommendations for future study	27

References	28
------------------	----

List of tables

<i>Table 1: Test results summary</i>	25
--	----

Table of figures

<i>Figure 1: Forecast under model 1</i>	25
<i>Figure 2: Forecast under Model 2</i>	26

DECLARATION

I declare that this work has not been previously submitted and approved for the award of a degree by this or any other University. To the best of my knowledge and belief, the Research Project contains no material previously published or written by another person except where due reference is made in the Research Project itself.

© No part of this Research Project may be reproduced without the permission of the author and Strathmore University

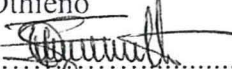
Mwangi Ebenezer Gachugu

Signature 

Date 8/12/2016

This Research Project has been submitted for examination with my approval as the Supervisor.

Ferdinand Othieno

Signature 

Date 8/12/2016

Strathmore Institute of Mathematical Sciences
Strathmore University

1. Introduction

1.1 Background of the study

The stock market is an anchor to the financial sector, providing a platform for users and suppliers of financial instruments for investment purposes in the stocks of companies. In a well-functioning market investors prefer the successful companies with high performing stocks more than those with lower margins. The shares of a high performing company have been seen to be related to the other companies listed in the same stock market where the expected dividend growth and the prices of the shares are well reflected in the stock market. (Ross, 1976)

In his study Fama (1965) realizes that one of the characteristics of investors in an efficient market is that they are rational and seek to maximize their profits. As a result of this nature, they actively compete to predict future stock prices. However, this prediction is only based on publicly available information

Stock markets exhibit volatility over a period of time which makes predictability a challenge to investors given the availability of useful information is limited. This means that stock markets experience movements up and down in response to shocks that are a result of numerous factors. This happens over a very short period of time. In the past, the excess volatility in stock markets has driven investors to shift their resources to risk-free assets and other securities with minimal risks. It is therefore necessary for investors and their advisors to understand the dynamic movements within the stock markets. Various studies in the past have explained that different variables differ in terms of the potential to explain the variations in the stock returns of different markets hence the importance of multifactor models.

For healthy growth of any economy to take place, efficient stock markets are critical determinants. It should be characteristic of the markets to reflect the information behind the share prices prevailing in the respective economies. Fama (1970) indicates that an efficient market is one that reflects relevant information in the share price. Investors in this case have are obligated to determine the intrinsic value of the stocks by applying fundamental and technical analysis (Ross, 1976). This helps them to know the sensitivity of the stock returns with respect to the macroeconomic variables.

Macroeconomic indicators should be used to measure the strength or weakness of the economy in a country, the most important being: inflation, money supply, interest rates, gross domestic product (GDP), exchange rates, industrial production and foreign direct investment. It is a field of great interest for investors, researchers, academicians, regulators and other relevant government bodies to identify the impact of these macroeconomic factors on stock market returns.

1.2 Problem statement

Efforts to forecast the performance of the stock market has attracted a big audience over the years. The interest in financial markets has led to numerous studies on the subject and different findings have been concluded for different markets. The Kenyan economy has witnessed relative macroeconomic stability over the past decade in terms of GDP growth, inflation and crude oil prices. However other factors like interest rates and stability of the Shilling/dollar exchange rate. This growth can be attributed to the growth of major sectors in the Kenyan economy including the capital and money markets.

Stock prices should reflect the expectations of the future performance of corporate profit considering the relationship between external business conditions and expected returns (Kuwornu, 2012). Taking these external business conditions under the umbrella of macro-economic uncertainty, investor expectations about future cash-flows vary resulting in a domino effect on stock prices and eventually stock returns. The predictive power of the macroeconomic factors has been observed by various studies to change through time and the variance is directly correlated with stock returns (Fama, 1977). It however becomes apparent that predictive models and paths taken by excess returns are either not popular, and have not been applied in emerging markets such as the NSE. This study therefore aims to find out the effectiveness of the multifactor models in predicting excess returns in an emerging economy (Kenya) with a strong emphasis on the movements of the NSE All Share Index on its sole bourse.

1.4 Research objectives

The objectives of the study are as follows:

- i. To determine which macroeconomic factors are important for predicting stock returns.

1.3 Research questions

The study sought to answer the following questions:

- i. What macroeconomic factors are important for predicting stock returns?

1.5 Scope of the study

The proposed research covers the period between January 2009 and January 2016 to give a wide scope of investigation of monthly macro-economic changes over time. The research takes into consideration the effects of economic shocks during the period which may have led to surges in the economic variables such as oil price shocks. The research uses the market return indicators of the NSE as captured by the All Share Index and the NSE 20 Share Index. For the risk free asset returns, the study shall use returns of 91-day Treasury Bills.

1.5 Justification of the study

The limited literature on Kenya's stock market poses a significant challenge to investors wishing to model their returns in real time. Major studies on recursive modelling have been carried out on developed markets. This study therefore offers valuable contribution to the limited literature on the Kenyan market by examining how applicable a recursive approach using macroeconomic indicators is to investors. It aims to test the predictability of excess returns on emerging frontier markets such as the NSE. The study does not look to determine whether predictability by use of recursive models is doable. It aims to determine whether it can be a viable alternative to other common forecasting techniques, for instance the random walk.

2. Literature review

2.1 Theoretical framework

Various theoretical frameworks have been used by researchers to try linking microeconomic variables with excess stock market returns. Fama (1970) developed the Efficient Market Hypothesis and Ross (1976) came up with the Arbitrage Pricing Theorem. In light of the various macroeconomic variables identified from previous studies, Dropsy (1998) grouped all the macroeconomic variables that affect stock returns into three broad categories. The first are performance variables which include changes in interest rates, inflation rates which are measured by the Consumer Price Indices and foreign exchange fluctuations appreciation. Second are the external variables which are essentially beyond the control of the government such as oil prices. Finally there are the policy variables which encompass changes in the fiscal and monetary policy. They can be captured by rates of GDP and growth of money supply for monetary policy

2.2.1 The Efficient Market Hypothesis

It is commonly referred to as the random walk theory, where it assumes that the market prices should incorporate all publicly available historical information at any given point in time. According to Eugene Fama (1970), in an efficient market, on the average, competition will cause the full effects of new information on intrinsic values to be reflected instantaneously in actual prices. He described an efficient market as one where prices reflect all available information. I agree with this statement since making profits from price movements that have been predicted prior to the investment, is highly unlikely since according to the EMH theory, arrival of new information acts as a main factor influencing changes in the stock prices.

2.2.2 The Arbitrage Pricing Theorem

This is another way that can be used to link excess stock market returns to macroeconomic variables. It is based on the mean variances framework that develops the Capital Asset Pricing Model (CAPM) which is further extended to APT after taking into consideration other systematic macroeconomic variables. In other words, CAPM is a single factor model whereas APT uses a multifactor approach to explain asset pricing. The macroeconomic factors are denoted with factor specific coefficients that are used to measure the sensitivity of the assets to each factor. It is important to note that two similar items in an efficient market cannot be traded at different prices, otherwise arbitrage opportunities would exist. Ross (1976) further suggests that the returns on any stock should be linearly related to a set of indexes which include anticipated and unanticipated

factors. Most of the return realized by investors however, is the result of the unanticipated macroeconomic events.

Modern financial theories rely on three main assumptions; investors in the market exploit any potential arbitrage opportunity, they are rational and that markets are highly efficient. As a valuation process it also has a set of underlying principles that it is founded on: no arbitrage principle, linear positive state prices, the law of one price and market completeness. With these principles, there are two approaches to asset pricing. There is the absolute pricing which involves pricing an asset by reference to the assets exposure to fundamental sources of macroeconomic risk, for example the consumption-based asset pricing. The second approach is the relative pricing which involves deriving an assets value given the prices of other assets, an example is option pricing using the Black Scholes framework.

Grossman and Shiller (1981) try to show how historical movements can be justified by new information. They prove that stock price movements could be attributed to various macroeconomic factors including interest rate movements. Chen, Roll and Ross (1986) also try to investigate the impact macroeconomic variables could have on stock prices in the US stock market. They use various macroeconomic variables as regressors in their multifactor model. They concluded that industrial production, twists in the yield curve and change in risk premium could be linked to stock returns. They however found out that consumption market indices and oil prices are not related to financial markets

Gjerde and Sættem (1999) base a similar study in Norway. Their results contradict that of Chen, Roll and Ross (1986) showing that oil prices and real economic activity have a positive relationship with stock market returns. They however failed to show a significant relationship between the stock returns and inflation.

Another study by Fama and Schwert (1977) investigating the relationship between stock returns and interest rates. They found that periods of high short-term interest rates tend to be followed by lower subsequent stock-market returns. This can be explained by an increase in the SDF as required rates of return increase. Bhattacharya et. al. (2001) based their study in India and attempted to use the Granger non-causality. The aim of their study was to analyze the causal relationship and macroeconomic variables. Their model was based on three variables: exchange rate, trade balance

and foreign exchange reserves. His results concluded that there is no causal linkage between stock market returns and the three variables.

Doong et al (2012) also tried a similar study but based their study on six Asian countries. They attempted to investigate the relationship between stocks and exchange rate using the Granger causality test. They concluded that a significantly negative relationship existed between the stock returns and changes experienced in the exchange rate market for all (but one) the countries included in the study.

Uddin and Allan (2007) on the other hand attempt to examine the linear relationship between share price and interest rates. They further examine the association between changes of share prices and changes in interest rates. Their study also included exploring how changes in share prices affect interest rates and vice-versa. Their study was based in the Bangladesh market. Their findings show a negative relationship between interest rates and share price. Changes in the interest rate also proved to have significant negative relationship with changes in the share price.

Geetha, Mohidin, Chandran and Chong (2011) carry out a study which models the relationship between the stock market returns, inflation rate(expected and unexpected), interest rates, GDP and exchange rates. They based their study in the Malaysia, US and China stock markets. They apply the cointegration test to investigate the number of vectors which cointegrate. This however only proved the relationship in the long-run. The short-term relationship was determined by applying the vector error correction (VEC) model. They concluded for all the stock markets that there is a long-run cointegration relationship between stock markets and the macroeconomic variables applied. The VEC model however proved no short run relationship exists between the stock market and the named variables for the Malaysia and US market. China's market however proved that there is a short run relationship between expected inflation rates and the stock market.

Gay (2008) focuses his study on Brazil, Russia, India and China. The variables applied are exchange rates and oil prices for the stated economies. He uses the Box-Jenkins ARIMA model. His findings however show no significant relationship between the respective oil prices and exchange rate on the stock market returns. He develops a conclusion that the markets under speculation a weak form of market efficiency.

In his study, Mohammad (2011) examines the impact of both microeconomic and macroeconomic variables on stock market returns in Bangladesh. He uses Multivariate Regression model which he computes on standard OLS formula and Granger causality test. He focuses on monthly data for all the variables under speculation. The results of the study show that stock returns and inflation exhibit a negative relationship as well as foreign remittance. The variables that show a positive relationship are growth in market capitalization and Price/Earnings ratio. No unidirectional Granger causality is however found between the stock returns and any of the variables which are assumed to be independent. This proves that the market is informally inefficient.

Also in India, Ray Sarbapriya (2012) studies the relationship between foreign exchange reserves and stock market capitalization. Results of the simple linear regression model applied and Granger causality test show that causality unidirectional and it runs from the foreign exchange reserve to the stock market capitalization. Results also show that the regressors used have a positive impact on stock market capitalization.

Many studies have shown that excess returns can be predicted by variables such as dividend-earnings ratios, earnings-price ratios and an array of other financial indicators. It can also be concluded that stock returns can be predicted by means of publicly available information, such as time series data on financial and macroeconomic variables with important business cycle components (Pesaran & Timmermann, 1995). This includes the use of macroeconomic models using variables such as inflation, changes in short term interest rates, changes in industrial production, monetary growth rates and important business cycle components. This conclusion has proved to hold across stock markets globally and also across different time horizons.

In his paper, Fama (1965) mentions that there has been increasing interest in testing for stock price behavior. He finds that stock prices follow the random walk theory and concludes that stock prices are unpredictable. Two approaches that have come up for predicting stock price behavior are the fundamental and technical analysis. Technical analysis involves designing trading rules to exploit profitable opportunities implied by the patterns in asset prices. Fundamental analysis implies predicting asset prices by identifying the underlying determinants of rates of return. When shocks are applied to these factors, a direct effect is felt on the stock returns within that market or even on the global market.

Most literature finance report excess return regressions estimated on the basis of entire sample of available observations or on substantial subsamples of the data, which for the purpose of trading, is inappropriate. This means that no investor in reality could have obtained parameter estimates based on the entire sample. When it comes to the choice of the forecasting models to be used, the same assumption applies. Any analysis of stock market predictability that focuses on a particular forecasting model, taken as known with certainty over a whole the whole sample period, can be criticized for ignoring the problem of uncertainty which should be incorporated in all forecasting models and could impact negatively on investors' portfolio strategies in "real time". When the same forecasting model is used over the entire sample period, it inevitably raises the possibility that the choice of the model could have been made with the benefit of hindsight.

Pesaran and Timmermann (1995) try to assess the economic significance of the predictability of stock returns, explicitly accounting for the forecasting uncertainty faced by investors who only have access to historical information. It challenges the assumption that investors historically may have insight on how specific forecasting models would perform. A much weaker assumption is made about investors' beliefs over the sort of business cycle and economic variables thought as being potentially important in forecasting stock returns. Based on these beliefs, we assume that agents establish a base set of potential forecasting variables and, at each point in time, search for a reasonable model specification, capable of predicting stock returns across this set.

2.3 Empirical review

According to Pesaran and Timmermann (1995), this procedure assumes that, at each point in time investors only use historically available information to select a model according to model selection criteria that should be predefined. The model chosen should then be used to make one period ahead predictions of excess returns. The recursive forecasts are then applied in a portfolio switching strategy according to which shares or other assets are held depending on the excess returns recorded. In this case no target value is set. The constraints however can be set to be positive or negative so as to affect the decision of the portfolio manager. Transaction costs are also factored in. scenarios are created in the models to analyze whether the predictable components in stock returns can be exploited economically net of transaction costs.

2.3.1 How to identify predictability of stock returns

Considering the price of an asset at two time periods t and $t+h$ such that h is very small, the stochastic discount factor, m , would be equal to 1 (since the interest rate is approximately equal to 0). If the asset's price at time t is P_t , the absence of arbitrage implies that $E(P_{t+1}) = P_t$. This implies that the price of the asset may go up or down but these movements (returns) will be randomly distributed: they will follow a random walk. Hence, the price of the asset will be a martingale.

Most studies test Fama's three forms of informational efficiency: the weak, semi-strong and strong forms. Findings have generally been that there is very little predictability: stock prices generally follow a random walk. In his PhD dissertation from 1963 proposed three types of tests of the random walk hypothesis: tests for serial correlation, run tests (tests whether successive price movements occur so frequently that they cannot be attributed to chance) and filter tests (tests whether investors can consistently make abnormal profits by specifying their orders in a certain manner).

A different approach that could be taken is the use of event studies. This is based on the theoretical belief that if asset prices incorporate all public information then relevant news should have an immediate price impact as soon as it is announced but price movements after the announcement should be random. Event studies are affected by the fact that many factors affect asset prices at the same time making it difficult to attribute price changes to particular information or events. The joint hypothesis problem also arises as a model to conduct the test needs to be specified. The original study by (Fama, Fisher, Jensen, & Roll, 1969) studied a stock split and found that, consistent with theory, the stock price did not exhibit abnormal returns after the announcement. Other studies in this field include those by Asquith and Mullins (1968) and Ball and Brown (1968).

These studies test whether it is possible to predict asset prices over a longer term horizon. They have to directly confront the joint-hypothesis problem: the results of tests of long term predictability can indicate whether the asset pricing model is correct. They have to directly confront the joint-hypothesis problem. The results of tests of long term predictability can indicate whether the asset pricing model is correct.

One of the main tests of long term predictability is the use of variance ratio tests studied by Shiller (1979). The problem tested is the nature of expected market returns in terms of being constant over time or varying in a predictable way.

Pesaran and Timmermann (1995) first make an assumption that the investor believes that stock returns can be predicted using a set of macroeconomic indicators. However the investor does not know the true form of the underlying specification. With time, the historical observations available to the investor increase. This causes the investor to modify the forecasting equation or change the models completely. In this case the investor should be open minded with no bias to any model in particular.

On the selection of linear regression models, Pesaran and Timmermann (1995) found that recursive predictions based on the \bar{R}^2 and $TC_{t,i}$ criteria are equivalent, in the sense that they select the same model. They both performed better than recursive predictions from the model that included all regressors. This directly translates to the ability of the investors to engage in active search processes in search of adequate forecasting equations rather than being hell-bent on fixed model specifications including entire sets of regressors.

The models used in this paper to investigate the economic significance of stock returns predictability assumes that investors use historically available information according to a predefined model selection. The investors then predict one period into the future and switch between bonds and equities depending on the predictions of excess returns. An investor would hold more stock if the excess returns are positive.

The model selection criteria include:

Statistical model selection. It is based on the likelihood and usually assigns weights to the parsimony and fit of the model which is measured by maximizing the log-likelihood function. Examples of the statistical model selection criteria are: Akaike's Information Criteria (AIC), Bayesian Information Criteria (BIC) AND R^2 .

Recursive wealth criterion which maximizes the cumulated wealth obtained using forecasts from the model in a switching portfolio.

Recursive Sharpe criterion which maximizes the ratio of the mean excess returns on portfolio i to its standard deviation.

The measure of directional accuracy. This predicts the switching of the sign of excess returns. An example is the Sign Criterion (SC) which maximizes the proportion of correctly predicted signs.

The model with the highest value for the criterion function is chosen to forecast excess returns at period $t+1$. There are k set of regressors that are used to make one period ahead forecasts of excess returns. The benchmark set of regressors that are incorporated in the models include: dividend yield, earnings-price ratio, 1-month T-bill rate, 12-month T-bond rate, inflation rate, rate of change in industrial output and growth rate in the narrow money stock.

Equation 1: Simple regression model

$$X_T = \{YSP_{t-1}, EP_{t-1}, I1_{t-1}, I1_{t-2}, I12_{t-1}, I12_{t-2}, \pi_{t-2}, \Delta IP_{t-2}, \Delta M_{t-2}\}$$

where: YSP_{t-1} is the dividend yield

EP_{t-1} is the earnings – price ratio

$I1_{t-1}$ is the 1 – month T – bill rate

$I12_{t-1}$ is the 12 – month T – bond rate

π_{t-2} is the inflation rate

ΔIP_{t-2} is the rate of change in industrial output

ΔM_{t-2} is the growth rate in the narrow money stock

Excess returns on the stock are computed using the formula:

Equation 2: Excess returns

$$P_t = \frac{P_t + D_t - P_{t-1}}{P_{t-1} - I1_{t-1}}$$

Where: P_t is the stock price

D_t is the dividends

$I1_{t-1}$ is the 1 – month T – bill rate

2.4 Recursive predictions of stock returns

We first link movements in stock returns to business cycle indicators. According to (Angas, 1936), the major determinant of price movements on the stock exchange is the business cycle. Various studies have suggested reliable variables to be used in the forecasting models that can be systematically linked with stock returns include, short-term interest rates, long-term interest rates, dividend yields, industrial productivity, and rates of inflation within the specified economy, liquidity measures and company earnings of the specified stock. The variables determine various sets of benchmarks to be used as regressors over which the search for a satisfactory prediction model could be conducted by a potential investor.

Pesaran and Timmermann (1995) argue that the efficient market hypothesis is often interpreted as the impossibility of constructing a trading rule, based on publicly available information, which is capable of yielding excess profits (discounted at an appropriate risk-adjusted rate). According to Jensen (1978), a market is efficient with respect to information set Ω_t if it is impossible to make economic profits by trading on the basis of information Ω_t . However the theory of predictability critiques this statement by claiming that it does not guarantee that an investor can earn profits from a trading strategy based on forecasts made using historical information that is publicly available. The monthly excess returns do not follow a standard distribution but are considered more leptokurtic than normal in terms of distribution. Pesaran and Timmermann (1995) find that predictability of stock returns of a magnitude that can be exploited economically depends on both the evolution of the business cycle and the shocks experienced in the market. The robustness of the models however seem to be heavily linked to macroeconomic events.

As a disclaimer for investors, the standard measures of predictive performance may not be reliable when it comes to the indication of opportunities for profit making. On the other hand investment strategies based on recursive forecasts of excess returns tend to display far much higher transaction costs and may not be as profitable as the passive investment strategies when transaction costs are appropriately taken into account. This means that transaction costs may erode the profits made from the financial markets.

3. Methodology

This chapter is set out to explain the methodology used to achieve the objectives set out earlier on in the first chapter. The succeeding section shows how to identify explanatory variables and the other section identifying the predictive power of the variables.

3.1 Forecasting

The study first tries to understand the ideology behind conditional expectation. This would be expressed in the form of

Equation 3: Conditional expectation

$$E(y_{t+1}|\Omega_t)$$

This means that the expectation of y is taken for time $t+1$, on condition that all information is available up to and inclusive of time $t(\Omega_t)$. Excess returns follow a random walk, hence the zero mean white noise process for the optimal mean is equal to zero and is given as,

Equation 4: White noise process

$$E(y_{t+1}|\Omega_t) = 0 \forall s > 0$$

Since the forecasting is based on real time analysis with minimal usage of hindsight, it will need a model with a finite memory but one that is capable of giving us a historic mean value. One model that would be viable is the autoregressive [1] moving average [1] (ARMA [1, 1]) process, with lags extending up to one previous (financial) period only. The reason behind this is, we are implying that today's values of parameters determine the values of tomorrow since we are forecasting in real time. Let $f_{t,1}$ indicate the forecasting model made using ARMA [1, 1] at time t for 1 step into the future for a series y . a forecast function is used to generate the forecasted values. Typically it can be written in the form

Equation 5: Auto Regressive Moving Average

$$f_{t,1} = \sum_{i=1}^1 a_i f_{t,1-i} + \sum_{j=1}^1 b_j u_{t+1-j}$$

$$\begin{aligned} \text{Where } f_{t,1} &= y_{t+s}, s \leq 0; u_{t+1} = 0, s > 0 \\ &= u_{t+s}, s \leq 0 \end{aligned}$$

And a_i and b_j are the autoregressive and moving average coefficients respectively. The moving average component has a memory of length 1 and can be represented as

Equation 6: Moving average component

$$y_{t+1} = \mu + \theta_t + \mu_t$$

The constancy of the parameters over time is assumed meaning if the relationship aforementioned holds at time t , it will also hold at time $t+1, t+2, t+3, t+4, \dots, t+n$ for n periods. The study still holds the assumption that one period is in reference to a financial period, hence the model can be used through multiple periods repeatedly or repeatedly. At time $t+1$ the forecast will be

Equation 7: Moving average model

$$f_{t,1} = E(y_{t+1} | \Omega_t) = \mu + \theta_t + \mu_t$$

This linear combination of disturbance terms is used to forecast for y , one step ahead and made at time t . The values of the error terms at time $t+1$ are unknown given we have information only up to time t . $E(y_{t+1} | \Omega_t) = 0$ since μ_{t+1} is unknown at time t . All forecasts more than one period ahead collapse to the intercept except for the constant term.

The study will however only adopt an autoregressive process for our modelling strategy so as to rely on the benefit of hindsight as much as possible. The purpose of the AR [1] process can be derived from the fact that the stochastic element that is characterized by stock returns can be hardly ignored. The autoregressive process can be written as an estimate in the form

Equation 8: Autoregressive component

$$y_{t+1} = \mu + \phi_1 y_t + \mu_t$$

It also assumes here the stability of the parameters. This makes the production of one step ahead forecasts rather easy since at time t , all information that may be required would be available. It

then develops $f_{t+1,1}$ given expectations to the previous equation are applied. It will therefore come up with

Equation 9: Autoregressive model

$$f_{t+1,1} = E(y_{t+1}|\Omega_t) = E(\mu + \phi_1 y_t + \mu_t|\Omega_t)$$

It will then try to break down the y_t function.

3.2 Recursive modelling

This section focuses on understanding y_t function and breaking down regressors. Rates of return are regressed against lagged values of a stochastic explanatory variable (Harrison et al., 2005). Three such significant and predetermined models were highlighted by Keim and Stambaugh (1986). Simulation of the forecasting model is done by first establishing a set of regressors over which a search is to be conducted to determine the functional form of the estimated models as well as the criteria used to select particular regression models. The same set of base regressors may fail to be re-used to construct intertemporal forecasting equations due to the restrictiveness they cause in the inclusion of new variables in the future, this is because future relevant variables cannot be known a priori. Expanding the number of base regressors on the other hand does not solve the problem as it reduces the computational feasibility of the model through having a multiplier effect on the number of models that would need to be recursively estimated according to the frequency of the data points (in our case monthly). The main advantage of using a recursive modelling technique for data that exhibits large periods of non-stationary behavior is that it allow the forecasting model to change through time.

3.3.1 Model 1

The study uses the difference between the long-term low grade corporate bonds and short-term T-Bills as the first variable. The yield is stayed on a monthly basis hence it has to be divided by 12. It is normally used since it is believed to be a representation for changes in expected risk premiums.

The other variable to be regressed in this model is

Equation 10: variation in the NSE with changes in expected future discount rates

$$-\log\left\{\frac{NSE_{t-1}}{NSE_{t-1}}\right\}$$

Where NSE_{t-1} is the benchmark index and is equal to the deflated level of the index at time $t-1$, using consumer price index. The denominator on the other hand represents the long run level which is the mean of the end of year real index over the period of data collection prior to $t-1$. This variable is significant in that it reflects the variation in the NSE with changes in expected future discount rates or excess returns.

The other variable tries to capture the small firms which comprise of the most volatile element in the stock market. These firms exhibit the greatest ex-post sensitivity to overall change in expected risk premiums in that they are characterized by a risk premium that has the highest volatility. This variable however has no trend that can be detected, but it captures the variation in small stock prices.

3.3.2 Model 2

This is our main model. The recursive model will use an econometric forecasting model. The regressors that seem to be the most recursive in forecasting models in that they are significant and relevant in most financial and economic literature are employed. This can be illustrated in the second chapter by equation 1.

3.3.3 Model selection

The R-Squared, Adjusted R-Squared Akaike Information Criterion (AIC) Akaike (1973) and Schwarz's Bayesian Information Criterion (BIC) Schwarz (1978) are model selection criteria. In addition, the t statistics of particular variables will be considered for economic interpretation. These model selection criteria will be used to test for the best "fit". The study will take a recursive approach hence the regression is conducted through the period and rolled over.

The study will then use AIC and the R^2 values that result from the recursive models to comment on the predictability of excess returns on the NSE stock based on regression statistics covering the entire study period. The values will also be used to select the model specification that best predicts the excess returns. They assign different weights to the "parsimony" and "fit" of the models. The fit is measured by the maximum value of the log-likelihood function and the parsimony by the number of freely estimated coefficients. Pesaran and Timmermann (1995) also implies on the importance of model selection based on a measure of directional accuracy, this is grounded on the assumption that investors in practice are interested in predicting the switches in the sign of the

excess return function and not necessarily the magnitude of the changes in the excess returns. It is for this reason that we derive the sign criterion.

3.3 Selection of Explanatory Variables

The significant variables identified by Pesaran and Timmermann (1995) and are applicable to the NSE include: inflation, growth in the gross domestic product (GDP) , exchange rate (USD-KES) ,dividend yield, PE ratio, 3-month treasury bills, 10 year Treasury bond and growth of money stock. The study intends to test the robustness of these variables in predicting the return on the equity market index. The study will mainly focus on significant variables in the Kenyan economy. As witnessed however, by the many studies on predictability of stock returns, financial theory only provides little guidance for which state variables should have predictive power over excess return especially in emerging economies. Finance theories suggest that markets with risk averse agents, stock returns vary with the state of the business cycle (Balvers et al., 1990)

Angas (1936) identifies the major determinants of price movements on most emerging markets in Africa as long term interest rates, dividend yields, industrial production, liquidity measures and inflation rates. However, Kirui, Wawire and Onono, (2014) find that for the NSE, the most significant variables affecting stock returns are exchange rates and inflation.

The study will remove some variables based on their likelihood to have a significant contribution to the model and potential obedience to Ordinary Least Squares regression. The likelihood will be measured using the correlation of each variable with the explained variable. For the OLS regression, the study will reduce the potential for multi-collinearity as signaled by a variance-covariance matrix of the explanatory variables. Based on these criteria, some variables will be dropped.

Consider the conventional pricing relationship in which the price of an asset tomorrow is determined by determined by the investors' expectations of that price (given the information they hold today) in addition to a random error term:

Equation 11: Conditional Expectation

$$P_{t+1} = E[P_{t+1}|\Omega] + e_{t+1}$$

This implies that the only information that is incorporated into the return at time t is the information that is available at time t-1. Predictive regression of non-stationary variables may also prove

difficult to interpret. Due to this, the use of lags is consistent with economic theory. Therefore the regression model will incorporate lagged values of explanatory variables to add stationarity and in consideration of Lo and Mackinlay (1987) we make an assumption of a less stationary stance by predicting one period ahead returns. The study assumes a linear model.

Equation 12: Linear regression Model

$$M_i = \alpha_0 + \sum_{\forall i} b_i x_{i,t-1} + \varepsilon_t \quad , \text{ Where } t = 1, 2, \dots, t-1$$

Where: M is the return on the market index

α_0 is a constant term

b_i is the coefficient of the i th explanatory variable

x is the value of the i th explanatory variable

ε is a randomly distributed error term ($\varepsilon \sim N(0, \sigma^2)$)

3.4 Data sources

Published time series data were sourced from the Nairobi Securities Exchange Limited (NSE) databases, Kenya Statistical Abstracts, Kenya National Bureau of Statistics (KNBS) and the Central Bank of Kenya (CBK).

3.5 Assumptions

The main assumption in the study is that prices, significantly reflect all available information and that rational expectations by investors exist. The intuition that underlies the theoretical model arises from the maximization of wealth for a rational investor. This however goes beyond the precipice of maximization of utility whereby the study assumes that both the rational and sophisticated investor will look for god trades continuously and will not get satisfied after a number of successful profits.

Given that the study covers data from a Kenyan context prior to the presence of derivative markets, we assume the absence of short selling and that investors are not able to sue leverage when selecting their portfolio. The final assumption is that we take account of existence of low or medium transaction costs to find out whether the predictable elements in stock returns are economically exploitable net of transaction costs

4. Test results

4.1 Diagnostics

4.1.1 Unit root tests

Under the first recursive model conducted, the variable “deflated index” was found to be statistically significant upon differencing once; bond spread was statistically significant upon the inclusion of a one lag; and small firms was found to be statistically significant upon differencing once.

Under the second model, the variable 10 year treasury bond was statistically significant for a unit root test with a trend and intercept, upon first differencing; GDP growth was also found to be statistically significant, the rest of the variables recording statistical significance being growth of money stock and inflation.

4.1.2 Johansen tests of Cointegration

The main purpose of running these tests is to ascertain the credibility of the variables proposed by finance and economic theory, by testing whether they at all possess any long run relationship before we develop our forecasts.

Under Model 1, the study finds that there may be at most three cointegrating relationships at the 5% significance level, with evidence suggesting that there may be at least one cointegrating relationship between the variables at the same level of significance.

Under Model 2, the study finds that there may be at most 4 cointegrating relationships at the 5% significance level, with evidence suggesting that there may be at least one cointegrating relationship between the variables at the same level of significance.

4.1 Model performance

The above models were the best performing for each stock. They had the highest R-squared values and recorded significantly low Root Mean Squared Error. However with more variables added to the model, the forecasting power reduced for all models based on the adjusted R-Squared criteria which penalizes R-Squared for the additional regressors which do not contribute to the explanatory power of the model.

Upon testing for serial correlation, we reject the null hypothesis of no serial correlation between the residuals of the variables for five models. This leaves the index returns as the best model with significant variables for our study.

Based on the data utilized for the study, the variables had weak explanatory powers. The r-squared values for all the models were below 20%, implying that the models could only account for less than 20% of the variation in excess return. As a consequence of the low explanatory power of individual variables as well as the generated models, predicted values for excess returns were bound not to be a good fit to the actual. The Akaike Information Criterion and the Bayesian Information Criterion produce the same outcome. Tabular results of this model can be summarized in the table 1

Dependent Variable: INDEX_RETURNS
 Method: Stepwise Regression
 Date: 10/13/16 Time: 09:42
 Sample (adjusted): 2009M02 2016M01
 Included observations: 84 after adjustments
 Number of always included regressors: 1
 Number of search regressors: 6
 Selection method: Swapwise - Max R-squared
 Number of search regressors: 3

Variable	Coefficient	Std. Error	t-Statistic	Prob.*
C	19.87061	7.063416	2.813173	0.0062
GDP_GROWTH	-0.824508	0.254043	-3.245544	0.0017
ER_USDKES	-0.172447	0.078145	-2.206760	0.0302
MONEY_GROWTH	0.397630	0.563904	0.705138	0.4828
R-squared	0.191802	Mean dependent var		1.008565
Adjusted R-squared	0.161495	S.D. dependent var		5.598974
S.E. of regression	5.126976	Akaike info criterion		6.153357
Sum squared resid	2102.871	Schwarz criterion		6.269110
Log likelihood	-254.4410	Hannan-Quinn criter.		6.199889
F-statistic	6.328562	Durbin-Watson stat		2.360520
Prob(F-statistic)	0.000661			

Table 1: Test results summary

4.2 Regression and forecasts

The study employs AR (12) and MA (12) components into the regression estimates given that the data is monthly. This is because of existence of lower order correlations between the first 11 AR and MA terms resulting from the frequency of the data.

Under Model 1, the study finds all variables as statistically significant with the model being specified as follows:

Equation 13: Model 1 Regression

$$dNSE = -8540.98 - 3663.10SF - 45116.89bs - 0.5846Lresid1 + 0.7845ar(12) - 0.9578ma(12)$$

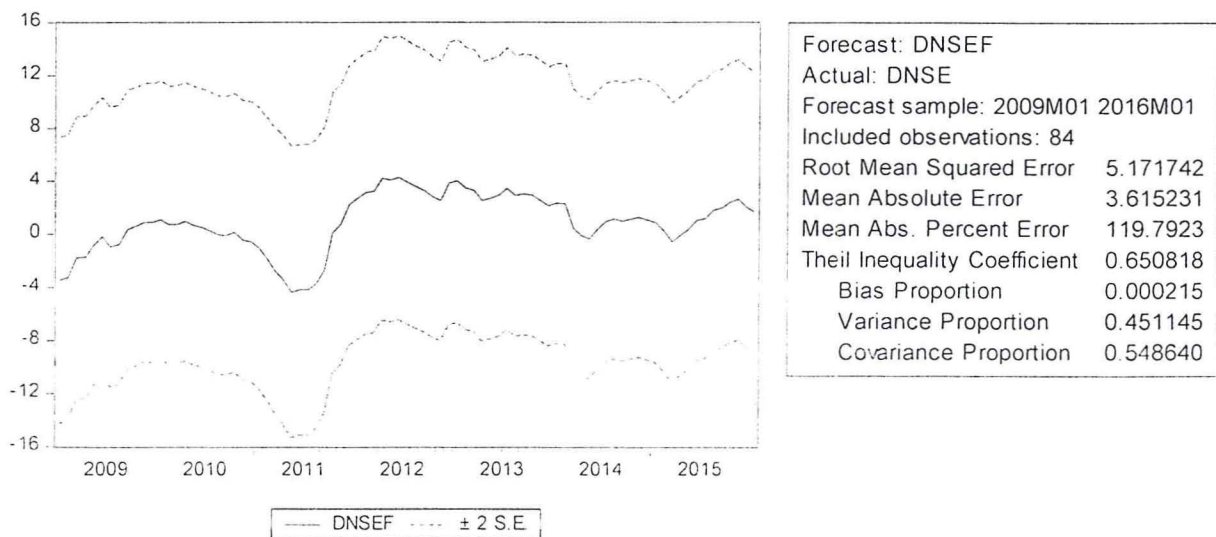
where: *SF* is small firms

bs is bond spread

Upon checking for serial correlation, the study fails to reject the null hypothesis of no serial correlation, as it finds a probability value of 0.1207 at 5% significance interval.

Forecast

Figure 1: Forecast under model 1



The forecast falls between the two standard deviations of the forecast as shown by *figure 1*, indicating a satisfactory predictive power of the models at the 95% confidence interval. The results indicate that the actual and the forecasted returns move closely given the Root Mean Squared Error of 5.171742.

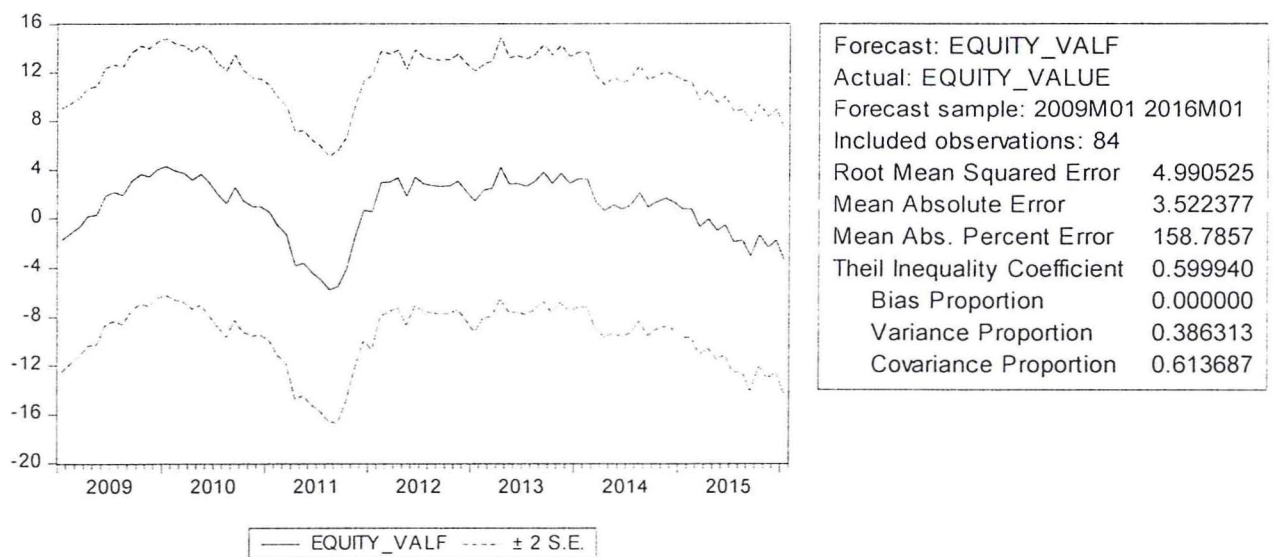
Under Model 2, the study finds the following variables in the regression, with the exception of the AR and MA terms to be significant. The model is specified as follows

$$\text{Equity value} = 7691.36 - 2623.32\text{money} + 5473.88 \text{inf} + 4114.24\text{GDP} + 0.2658\text{Lresid1} + 0.7112\text{ar}(12) - 0.9961\text{ma}(12)$$

Upon checking for serial correlation, the study fails to reject the null hypothesis of no serial correlation between the residuals of the variables, as it finds a probability value of 0.1207 at 5% significance interval.

Forecast

Figure 2: Forecast under Model 2



The forecast falls between the two standard deviations of the forecast as shown by *figure 2*, indicating a satisfactory predictive power of the models at the 95% confidence interval. The results indicate that the actual and the forecasted returns move closely given the Root Mean Squared Error of 4.990525.

5. Conclusion

The study attempted to establish the viability of recursive modelling and forecasting techniques in predicting stock returns using movements in macroeconomic factors. The study used different models with different combinations of variables to find the model with the highest explanatory power. All the variables were included in running the regressions and the best combination selected for each model given a certain number of variables are required in a given model. The major assumption to this effect was that the best combination was characterized by the highest R-squared.

The study used six models in the end to make comparisons of performance in predicting the dependent variable. The model that engaged exchange rates, growth in money stock and GDP growth was indeed the most powerful and reliable.

5.1 Recommendations for future study

Market agents may significantly benefit from future studies that aim to establish the variables that have the highest power in modelling and predicting stock returns. If the core factors are established, future model to model comparison of the performance of static and dynamic forecasts may yield more conclusive evidence of the superiority of one over the other. As at now their use would depend largely on the information that an investment manager seeks to obtain.

References

- Akaike, H. (1998). Information theory and an extension of the maximum likelihood principle. *Selected Papers of Hirotugu Akaike*, pp. 199-213.
- Andrew Lo, C. M. (1997). *The Econometrics of Financial Markets*. Princeton, N.J: Princeton University Press.
- Angas, L. (1936). *Investment for Appreciation. Forecasting movements in Security Prices. Techniques of trading in shares for profit*. London: Macmillan.
- Asquith, P., & Mullins, D. (1986). Signalling with dividends, stock repurchases, and equity issues. *Financial Management Autumn*, pp. 27 - 40.
- Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research* 6, pp. 159 - 178.
- Basabi, J. M. (2001). *The Nature of the causal relationship between stock market and macroeconomic aggregates in India: An empirical analysis*. Kolkata: Jadavpur University.
- Caroline, R. M. (2011). The relationship between inflation and stock market: Evidence from Malaysia, United States and China . *International Journal of Economics and Management Sciences*, Vol. 1, No. 2, pp. 01-16.
- Donald, R. S. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, vol. 17, issue 2, pages 357-39.
- Dropsy, V. (2002). The Asian Currency Crises: Vulnerability, Contagion, or Unsustainability. *Review of International Economics*, Vol. 10, pp. 79-91.
- Evans, K. H. (2014). Macroeconomic Variables, Volatility and Stock Market Returns: A Case of Nairobi Securities Exchange, Kenya. *International Journal of Economics and Finance*, Vol. 6, No. 8.
- Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, pp. 383-417.

- Fama, E. F. (1965). The Behavior of Stock-Market Prices. *The Journal of Business*, Vol. 38, pp. 34-105.
- Fama, E. F., & Schwert, G. W. (1977). Asset returns and inflation. *Journal of Financial Economics* 5, pp. 115-146.
- Fama, E., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review* 10, pp. 1 -21.
- Gay, R. (2008). Effect Of Macroeconomic Variables On Stock Market Returns For Four Emerging Economies: Brazil, Russia, India, And China. *International Business and Economics Research Journal*, Volume 7.
- Harrison, H. T. (2005). *Do Industries Lead Stock Markets?* Princeton: Princeton University Press.
- Jensen, M. (1978). Some anomalous evidence regarding market efficiency. *Journal of Financial Economics*, vol. 6, issue 2-3, pages 95-101.
- Kuwornu, J. K. (2012, November 2). Effect of Macroeconomic Variables on the Ghanaian Stock Market. *Agris on-line Papers in Economics and Informatics*, volume IV.
- Mahmudul, A. S. (2009). Relationship between Interest Rate and Stock Price: Empirical Evidence from Developed and Developing Countries . *International Journal of Business and Management* , pp43-51.
- Nai-Fu Chen, R. R. (1986). Economic Forces and the Stock Market. *The Journal of Business*, vol. 59, issue 3, pages 383-403.
- Pesaran, M. H., & Timmermann, A. (1995). Predictability of Stock Returns: Robustness and Economic Significance. *The Journal of Finance*, 1201-27.
- Ph.D. CFA Cristiana, T. P.-D. (2012, October 9). The Bucharest Academy of Economic Studies. *On the Causal Relationship between Stock Returns and Exchange Rates Changes for 13 Developed and Emerging Markets*, pp. pp275-282.
- Ronald. B. C. (1990). Predicting Stock Returns in an Efficient Market. *Journal of Finance*, vol. 45, issue 4, pages 1109-28.

- Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, pp 341-360.
- Sættem, Ø. G. (1999). Causal relations among stock returns and macroeconomic variables in a small, open economy. *Journal of International Financial Markets, Institutions and Money*, vol. 9, issue 1, pages 61-74.
- Sanford, J., Grossman, R. J. (1981). The Determinants of the Variability of Stock Market Prices. *The American Economic Review*, pp. 222-227.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, Volume 6, pp461-464.
- Schwert, E. F. (1977). Asset returns and inflation. *Journal of Financial Economics*, vol. 5, issue 2, pages 115-146.
- Shiller, R. (1979). The volatility of long term interest rates and expectations models of the term structure. *Journal of Political Economy* 87, pp. 1190 - 1219.
- Zaheer, A. R. (2014). Time Series Analysis of the Relationship between Macroeconomic Factors and the Stock Market Returns in Pakistan. *Journal of Yasar University*, 6261 - 6380.