UNIVERSITY FOR DEVELOPMENT STUDIES

MULTI – SECTORAL TEMPORAL MODELS FOR GHANA STOCK RETURNS

AKOTO YAW OMARI – SASU

Thesis submitted to the Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies in partial fulfillment of the requirements for the award of Doctor of Phylosophy Degree in Applied Statistics

2011

UNIVERSITY FOR DEVELOPMENT STUDIES

MULTI – SECTORAL TEMPORAL MODELS FOR GHANA STOCK RETURNS

BY

AKOTO YAW OMARI – SASU

(BSc. MATHS, MSc. FIN. MATHS & STATS)

(DAS/0004/08)

Thesis submitted to the Department of Statistics, Faculty of Mathematical Sciences, University for Development Studies in partial fulfillment of the requirements for the award of Doctor of Phylosophy Degree in Applied Statistics

NOVEMBER, 2011

DECLARATION

I hereby declare that this dissertation/thesis is the result of my own original work and that no part of it has been presented for another degree in this university or elsewhere:

Candidate's Signature:.. Date: 02-11-2011

Name: AKOTO YAW OMARI-SASU

Supervisors' Declaration

I hereby declare that the preparation and presentation of the dissertation/thesis were supervised in accordance with the guidelines on supervision of dissertation/thesis laid down by the University for Development Studies.

Principal Supervisor's Signature:................................ Date: 02-11-2011

Name: PROF KAKU SAGARY NOKOE, FGA

Co-Supervisor's Signature:.. Date: 02-11-2011

Name: DR. KAZEEM ALAGBE GBOLAGADE

ABSTRACT

Stochastic and econometric models have earlier been used in studying or modeling movements of stock prices and returns depending on the interest of the researcher. But econometric models such as the capital asset pricing model have been found useful in explaining the determinants of returns on investments in stocks. Cross–sectional analyses were performed to identify the economic variables that contribute significantly to the expected returns of individual equities listed on the Ghana Stock Exchange. The least–squared second differencing approach of disaggregation gave the best results in transforming the data into quarterly form. Although the central limit theorem in probability theory justifies the approximation of large sample statistics to the normal distribution in controlled experiments, using the specified distribution of the residuals eliminate biased regression coefficients and Cauchy distribution gave the best fit for the residuals of stock returns. Finally, a longitudinal model with Banded Toeplitz variance–covariance structure which incorporated Cauchy as the residual distribution was proposed for the Ghana stock returns.

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisors Prof. K. S. Nokoe and Dr. K. A. Gbolagade for their valuable comments and useful suggestions in completing this Thesis.

I am particularly thankful to Prof. K. S. Nokoe, my principal supervisor, for his meticulous effort and guidance throughout this study.

I would also like to acknowledge Dr. A. O. Adebanji, Dr. J. N. Mojekwu and Dr. C Iwunor for their insight that added value to this study.

I am equally indebted to my internal and external assessors for their valuable comments and constructive suggestions.

I am also thankful to my family and friends for their encouragement and support. I would like to thank my wife, Dr. Akosua Omari-Sasu, for her unconditional support and understanding during these years.

Above all, I thank God for His unfailing love, protection and care.

DEDICATION

To my father, His Lordship Justice Rtd. Kwadwo Omari-Sasu, my late mother, Mad. Margaret Obese Apori, my wife, Akos and my sons, Kwaku Ayeboafo and Kofi Mireku.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE / SYMBOLS LIST

ABL	Accra Brewery Limited
ALW	Aluworks Limited
APT	Arbitrage Pricing Theory
ARCH	Autoregressive Continuous Heteroskedastic
BM	Book-to-Market ratio
BRIC	Brazil, Russia, India, and China
CAPM	Capital Asset Pricing Model
CDF	Cumulative Density Function
CFAO	CFAO Ghana Limited
CMLT	Camelot Ghana Limited
CRSP	Center for Research in Security Prices
DE	Debt to equity ratio
EGARCH	Exponential Generalized Autoregressive Continuous Heteroskedastic
EIC	Enterprice Insurance Company Limited
FML	Fan Milk Limited
G	Growth in Earnings
GARCH	Generalized Autoregressive Continuous Heteroskedastic
GCB	Ghana Commercial Bank Limited

SCB	Standard Chartered Bank Ghana Limited
SPPC	Super Paper Products Company Limited
SSB	SG-SSB Ghana Limited
T	Time
TGARCH	Threshold Generalized Autoregressive Continuous Heteroskedastic
UNIL	Unilever Ghana Limited
+	Addition Sign
α	Alpha
β	Beta
=	Equal Sign
γ	gamma
>	Greater than
$\geq$	Greater than or equal to
∞	Infinity
λ	Lambda
<	Less than
$\leq$	Less than or equal to
μ	mu

π	pi
σ	sigma

CHAPTER 0NE

INTRODUCTION

1.0 General

The need for an informed investment decision cannot be over emphasized. Most investors and financial analysts have shown considerable interest during the last few decades in the new security markets that have emerged around the world. This interest has undoubtedly been spurred on by the large and in some cases extraordinary returns offered by these markets. All over the world practitioners use a lot of models in their portfolio selection process and in their attempt to assess the risk exposure to different investment options. This calls for the collection and analysis of available information on the behaviour of the security market.

Various data sets have been collected, collated and analyzed for the purpose of making sound investment decisions. Several models have also been used in studying or modeling movements of stock prices and returns on stocks. These include stochastic and econometric models depending on the interest of the researcher.

Econometric models have been found useful in explaining the determinants of returns on investments in stocks. One of such models is the Capital Asset Pricing Model (CAPM) proposed by Sharpe (1964), Lintner (1965) and Mossin (1966).

Several authors have attempted to develop different versions of the CAPM. Most of these attempts have failed to develop a suitable model (in terms of the explanatory powers of the independent variables included in the model and providing an appropriate model for describing changes in returns on investment in stocks as well as for predictive purposes). The search for an adequate asset pricing model is the motivation for this study.

1.1 Justification of the Study

One of the many things people always want to know about the stock market is, "How do I make money investing?" A number of factors should be considered by an investor before selecting any security on the stock market. Some of the factors include the past performance of companies, prevailing conditions in the sector in which the companies operate as well as the future outlook in terms of profit growth and dividend payment.

These research information can be obtain either by contacting a stockbroker for advice and/or reading research reports since sound investment is typically based on research.

Existing financial literature on most emerging markets including the Ghana Stock Exchange are inadequate and it is the goal of this study to widen the theoretical analysis of this market by using modern finance theory and to provide useful insights for future analyses of this market.

1.2 Aims and Objectives

The aim of this study is to develop a multi-sectoral temporal model for the Ghana Stock Market. Now considering various debates on investment returns, there is the need to develop a multi factor model that inculcates all the variables which may capture significant variation in stock returns.

Therefore the objectives are to:

- Identify the economic variables that contribute significantly to the expected returns of individual equities listed on the Ghana Stock Exchange,
- Develop a general linear model for the Ghanaian market,
- Investigate the behavior of residual returns and
- Model the changes in returns over time.

1.3 Organization of the Thesis

This thesis is structured into five chapters. The first, being the introductory chapter followed by the review of the literature in chapter two. Here, the various models including the capital asset pricing model, the ARCH and GARCH models, models from emerging markets and some few in Ghana are discussed. Also the various economic factors which have been found to contribute significantly to returns on stocks are discussed.

Chapter three gives the description of the various mathematical and statistical techniques that were employed in the research. These included: the data source and collection techniques; disaggregation techniques, general linear models and general linear mixed models.

The analysis of results are presented in chapter four and the last chapter gives the conclusions, summary of findings and recommendation for further study.

CHAPTER TWO

LITERATURE REVIEW

2.0 Introduction

Researchers and financial analysts all over the world have used a lot of models for both matured markets in developed economies and emerging markets in developing economies. Notable among them are the capital asset pricing model and the ARCH/GARCH models for stock returns volatilities. Most of the existing models of stock returns will be discussed in this chapter.

2.1 Capital Asset Pricing Model (CAPM)

Since the pioneering work of Markowitz (1952) in normative portfolio selection, one of the most important developments in modern capital theory is the capital asset pricing model (CAPM) developed by Sharpe (1964), Lintner (1965) and Mossin (1966). CAPM suggests that high expected returns are associated with high levels of risk. Simply stated, CAPM postulates that "the expected return on an asset above the risk-free rate is linearly related to the non-diversifiable risk as measured by the asset's beta". The assumptions under which the CAPM was developed are as follows: (i) Investors are risk-averse individuals who maximize the expected utility of their end-of period wealth; (ii) Investors are price takers and have homogenous expectations about asset returns that have a joint normal distribution; (iii) Investors behave in a normative sense and desire to hold a portfolio that lies along the efficient frontier; (iv) There exists a risk-free asset such that investors may borrow or lend unlimited amounts at the risk-free rate; (v) The quantities of assets are fixed and all assets are also marketable and perfectly divisible; (vi) Asset markets are frictionless and information is costless and simultaneously available to all investors; (vii) There are no market imperfections such as taxes, regulations, or restrictions on short selling.

2.2 Robustness of CAPM

A vast amount of empirical work has been done to verify or refute the CAPM model. Initial empirical tests such as those by Black, Jensen and Scholes (1972) and Fama and Macbeth (1973, 1974) have focused on the linear relationship between a security's expected return and its beta. Although some support has been found for the model, in that no significant nonlinearities exist, it does not validate the theoretical relationship because the slope of beta estimates tend to be flatter and the intercept higher (i.e., Black's zero beta concept) than the model predicts.

Later empirical work has shifted to the anomalies in the CAPM framework. The anomalies in the stock returns literature have stirred voluminous empirical studies, since the CAPM has been put into question. The evidence supporting the fact that stock returns are predictable by variables besides beta has become overwhelming. What differs among researchers is the interpretation to their findings.

One school of thought criticizes the CAPM based on several simplifying assumptions and, because most of these assumptions appear to be unrealistic and do not hold true in the real world, it has been argued that they are the cause of flaws in the CAPM (Watson and Head 1998, Harrington 1987). Several of the CAPM assumptions have been criticized. For instance, the assumptions that there are no taxes and no transaction costs do not conform to reality. In addition, the assumption of homogeneous expectations is also open to doubt, because investors usually: have divergent expectations, apply various investment holding periods, differ in respect of their decision-making processes *et cetera* (Levy et. al., 2000). Furthermore, the assumption that either share returns are normally (symmetrically) distributed or investors are only interested in the mean and the variance of returns (and therefore do not care about upside potential or downside risk) is also deemed unsatisfactory, because portfolio returns are generally distributed asymmetrically and investors do view risk as more than merely the mean and variance of returns. Therefore beta is viewed as an incomplete risk measure (Ward 2000; Leland 1999).

However, Moyer *et al* (2001) and Reilly and Brown (1997) note that the CAPM has stood up well to the realization of many of the assumptions and that, in general, the apparent unrealistic assumptions do not have a significantly negative effect on its implications. In addition, it is important to note that the CAPM is an expectation model and that it should not be judged on the realism of its assumptions, but rather on how well it explains the relationship between variables and predicts expected behaviour (Pike and Neale 1996). Radcliffe (1997) concludes that the validity of the CAPM can only be assessed by investigating how well it predicts real-world phenomena, and such an assessment requires empirical testing.

Furthermore, there are those who take the route of CAPM misspecification while assuming that the market is efficient. The argument here is that if stocks are priced rationally in the long run (assuming investors are rational), then systematic differences in expected returns must be due to differences in risk. Hence, such anomalous effects must proxy other dimensions of risks that are not captured by beta risk. The goal in this approach is to find the underlying fundamental factors that make up these proxies. As a result these researchers have suggested that a multi-factor model, such as the Arbitrage Pricing Theory (APT) that was developed by Ross (1976), as cited in Laubscher (2001), may represent a better description of returns. The APT suggests that returns are a function of various macroeconomic risk factors and not of only one risk factor, beta, as suggested by the CAPM.

According to Jones (1998) and Arnott (1993), there are, however, some difficulties associated with the APT which could be summarized as follows: it does not identify the risk factors; it is descriptive by nature (i.e. it explains what is and not what should be); and it is possible for the risk factors and their effect on share returns to change rapidly. Karnosky (1993) and Sharpe (1985) argue that, as a result of the fundamental difference between the CAPM and APT, it would be futile to compare the two models on the basis of their ability to replicate history and to forecast market prices. The two models should not be seen as alternatives, because the CAPM attempts to describe the underlying relationships of the market, while the APT attempts to provide an explanation of current market conditions.

Some other researchers attribute the anomalies to the errors of measuring beta or market portfolio. The idea here is that since beta and market portfolio are unobservable and the CAPM leaves no guidance as to how to measure them, improper measurements may cause errors-in-variables problems and statistical artifacts that are associated with the empirical regularities found in the studies. For instance, beta estimates are found to be heavily dependent on the return intervals that are used to compute them. In order to obtain better estimates of the value of beta coefficients, we need monthly or quarterly data since longer time period (eg. annually) might result in changes of beta over the examined period introducing biases in the beta estimate. On the other hand, high frequency data such as daily observations covering a relatively short and stable time span can result in the use of very noisy data and thus yield inefficient estimates (Michailidis et al, 2006). Failure to consider these issues may seriously invalidate the significant role of beta in rational market pricing. As reported by Laubscher (2001), the study of Roll (1977) was the first to question the empirical testing of the validity and usefulness of the CAPM on the grounds of its reliance on the existence of a market portfolio. This criticism is based on the fact that empirical tests have shown that, when incomplete measures of the market portfolio are used as proxies, beta is mismeasured and share returns are predicted inaccurately (Keogh, 1994 and Van Rhijn, 1994). The studies by Roll and Ross (1994) and Ross (1993), as reported in Reilly and Brown (1997), extended the 1977 critique by Roll and contended that, because a true market portfolio does not exist, the empirical testing of the validity of the CAPM and its use to evaluate investment performance is a meaningless exercise.

Despite the above-mentioned criticism, Reilly and Brown (1997) have noted that there are strong arguments in favour of the contention that the absence of a market portfolio does not invalidate the CAPM. The dilemma concerning the market portfolio and choice of market proxy only represents a measurement problem in respect of the testing of the CAPM, or in using it to evaluate investment performance. The challenge therefore lies in identifying and developing better proxies for the market portfolio and/or finding improved measures to adjust investment performance measurement to reflect this dilemma.

Finally, others have pursued the argument along the line of market inefficiency. In their interpretation, the market is inefficient because systematic excess returns can be achieved by forming portfolios that mimic size, book-to-market equity and other effects. Investors under this hypothesis are assumed to behave irrationally and always overreact to new information and hence, the strategy of buying losers and selling winners always yields abnormal returns. Market overreaction, information asymmetry or transaction costs, which are often viewed as consistent with market inefficiency, are used to explain the predictability in variation of stock returns. Motivated by the popularity among security analysts, many financial or accounting variables besides the beta risk have been found, contrary to the capital asset pricing model, to explain the constituents of average stock returns.

Among them, an empirical study by Banz (1981) identifies that firm size is also an important variable. Firm size as measured by market equity (price per share times the number of outstanding shares) captures the cross-sectional variation in average stock returns. The study finds that smaller firms have had higher risk adjusted returns, on average, than larger firms. Size effect is not linear and is concentrated in the very small sized firms since the risk adjusted returns between the average sized and the larger firms are found to be of little difference. Banz (1981) later concluded from such evidence that the CAPM is misspecified.

The earnings to price ratio (E/P) is one of the earliest variables identified. It has been consistently found that after earnings announcements, securities with high E/P ratio or low P/E securities seem to yield, on average, higher return than those with low E/P ratios. Furthermore, the abnormal returns of securities are a monotonic increasing function of the securities' E/P ratios. Hence, the systematic excess returns earned by high earnings yield securities have become anomalous (Basu, 1977).

Bhandari (1988) later found that in addition to beta and firm size, leverage (the ratio of debt to equity) is also positively related to the expected common stock returns. He reasoned that since beta may not be an adequate measure of risk (possibly due to measurement error), leverage is likely to be a better proxy for cross–sectional risk than firm size, an additional variable like the debt/equity ratio can be the proxy for the risk of common equity to explain the expected common stock returns.

The empirical findings by Stattman (1980), Rosenberg et al (1985) and Chan et al (1991) provides yet another variable in explaining expected stock returns. By forming a strategy of buying stocks with a high ratio of book value of common equity per share to market price per share (BM) and selling stocks with a low book/price ratio, Rosenberg et al. found that the return is positive in 38 of the 54 months they examined. Rather than attribute the positive relationship between average returns on the stocks of the NYSE and the BM ratio to the misspecification of CAPM, they concluded that the market is inefficient. In a broader study that is also motivated by a very limited research relating to the Japanese stock market, Chan et al (1991) examined four predictor variables that have been in existence in a U.S. data. They related the cross-sectional differences in returns on Japanese stocks to earnings yield, firm size, book-to-market equity ratio and cash flow (earnings plus depreciation) yield variables and found that book-to-market equity ratio is the most important variable, statistically and economically. In addition, since cash flow yield is highly correlated with earning yields, it is not surprising that it also has reliable positive impact on expected returns. The study has reinforced the significance of book-to-market equity and identified a better proxy than the earnings yield.

It is clear that there is overwhelming empirical evidence (regardless of whether beta is significant) to conclude that more than one variable is at work in explaining expected stock returns if one chooses to accept that market is rational and efficient. Since the Fama and Macbeth (1973, 1974) study in testing the validity of the CAPM, the pattern of most subsequent studies has been to first discover at least one additional significant variable (besides beta) that is related to average stock returns and then offer an explanation of such proxies. The important issue here is that even if we can exactly identify and narrow down all the important proxies, it is doubtful as to how far we can advance our fundamental understanding in the area of asset pricing theory under this trend of research. A common element that is missing in this research is a theoretical model that pre-specifies a relationship between the underlying factors and stock returns. Without a theory to build upon and to test its correctness, not only do these studies appear ad hoc but they may also create statistical artifacts. In addition, it may also motivate future empirical exercises into a data snooping contest in which the participants' goal is to find something else that is statistically significant.

Other efforts have recently been started toward finding fundamental factor(s) that are proxy by size, and have book-to-market value variables. Growth in earnings/dividends, a potential fundamental factor, has been identified indirectly by the latest studies. For instance, Fama and French (1995), in an attempt to lay down an economic foundation for the empirical relationship between firm size and book-to-market equity and expected stock returns that were observed in their earlier studies, analyzed how these two variables were related to stock earnings and profitability. They reasoned that if stocks were priced rationally, not only must size and BM proxy for sensitivity to common risk factors in returns, but they also must be driven by common factors in shocks to expected earnings that are related to size and BM. To test their hypothesis, annual portfolio returns sorted by size and BM were first regressed on dividends yields and growth in earnings, a proxy for shocks to expected earnings. Then for each market, size and BM factors in stock returns was regressed on the dividends yield and growth variable. Consistent with the prior, the growth variable was statistically significant in explaining portfolio returns and was also significant in relating to the market and size factor. However, BM was found to be weakly related to the growth variable; but the authors speculated that it was caused by noise in measuring growth in earnings.

Further empirical evidence that is consistent with the notion that growth in earnings/dividends may be a missing factor in explaining variation of stock returns was provided by the Jensen et al (1996) study, which found that size and book-to-market equity were significant in explaining average stock returns only in expansive monetary policy periods. According to Myers (1977), a stock's price can be viewed as due to two parts; the present value of assets already in place and the present value of future growth opportunities. Therefore, during periods of monetary expansion where interest rates were low, stock prices of growth firms (low BM and large market equity) relative to value firms (high BM and small market equity) tend to be bided up because of the larger present value derived from growth opportunities. This leads to lower expected returns for growth stocks. On the other hand, during periods of monetary restriction where interest rates were high, the stock prices of growth firms relative to value firms tend to go down, which leads to higher expected returns.

In addition, Harris and Marston (1994) focused their attention on the links among growth, book-to-market equity, and beta. When growth was controlled, the relationship between BM and beta changed from negative to positive. This observation was consistent with rational pricing in which high BM links to high risk (relative distress risk as suggested by Fama and French, 1992), and beta's important role in market pricing. Moreover, when BM was regressed on growth and beta, growth was more significant in explaining BM. They suggest that future works in understanding the economics of BM should incorporate measures of growth. These studies provide evidence linking growth to the anomalies and ultimately, to expected stock returns.

2.3 Modeling using ARCH or GARCH Models

Another group of models that has become popular in modelling asset returns is autoregressive continuous heteroskadastic (ARCH) or generalized autoregressive continuous heteroskedastic (GARCH) models. The basic version of the least squares model assumes that the expected value of all error terms, when squared, is the same at any given point (ie. Homoskedasticity). This assumption is the focus of ARCH/GARCH models. Data in which the variances of the error terms are not equal suffer from heteroskedasticity. ARCH and GARCH models treat heteroskedasticity as a variance to be modelled. As a result, not only are the deficiencies of least squares corrected, but a prediction is made for the variance of each error term based on some computations. This prediction turns out often to be of interest, particularly in applications in finance (Engle, 2001).

One of the popular works in this area is that of French et al (1987). They examined the relations between stocks returns and stock market volatility. They found evidence that the expected market risk premium (the expected return on a stock portfolio minus the Treasury bill yield) is positively related to the predictability of stock returns. There was also evidence that unexpected stock markets are negatively related to the unexpected change in the volatility of stock returns. This negative relationship provides indirect evidence of a positive relation between expected risk premiums and volatility.

Furthermore, other researchers like Bollerslev (1987) also investigated the distribution of speculative price changes and rates of return and found that the data tend to be uncorrelated over time but was characterized by volatile and tranquil periods. He then presented a simple time series model designed to capture this dependence. The model was an extension of the Autoregressive Conditional Heteroskedastic (ARCH) and Generalized ARCH (GARCH) models obtained by allowing for conditionally t-distributed errors. The model could be derived as a simple subordinate stochastic process by including an additive unobservable error term in the conditional variance equation. The descriptive validity of the model was illustrated by Bollerslev (1987) for a set of foreign exchange rates and stock price indices.

Some researchers also have criticized the ARCH and GARCH models. Nelson (1991) revealed that although GARCH models have been applied in modeling the relation between conditional variance and asset risk premium, the models have at least three major drawbacks in asset pricing applications: (i) Researchers beginning with Black (1976) have found a negative correlation between current returns and future returns volatility. GARCH models rule this out by assumptions. (ii) GARCH models impose parameter restrictions that are often violated by estimated coefficients and that may unduly restrict the dynamics of the conditional variance process. (iii) Interpreting whether shocks to conditional variance "persist" or not is difficult in GARCH models, because the usual norms measuring persistence often do not agree. Nelson (1991) then proposed a new ARCH that meets these objections. His method was used to estimate a model of the risk premium on the Center for Research in Security Prices (CRSP) Value–Weighted Market Index from 1962 to 1987.

In cases where a cross section of international stock returns and volatilities were considered for two or more markets, Multivariate GARCH models were employed (Bauwens et al, 2006). For instance, Karolyi (1995) earlier examined the short-run dynamics of returns and volatility for stocks traded on the New York and Toronto Stock Exchanges and found that inferences about the magnitude and persistence of return innovations that originated in one market and transmitted to the other market depended importantly on how the cross-market dynamics in volatility were modelled. Moreover, much weaker cross-market dynamics in returns and volatility prevailed during later sub periods, especially for Canadian stocks, and with shares dually listed in New York.

2.4 Stock Returns Modeling in Developing Economies

Although developed economies like the USA and Great Britain have mature stock markets where most of the above mentioned models work efficiently, emerging markets from developing economies also have effective models. A lot of work has been done on efficient market and econometric models.

- On financial market instability for instance, Kaminsky and Schmukler (2002) observed that rating agencies have been under particular scrutiny lately as promoters of financial excesses, upgrading countries in good times and downgrading them in bad times. Using a panel of emerging economies, Kaminsky and Schmukler (2002) examined whether sovereign ratings affect financial markets. They found that changes in sovereign ratings have an impact on country risk and stock returns. They also found that these changes were transmitted across countries, with neighbour-country effects being more significant. Rating upgrades (downgrades) tends to occur following market rallies (downturns). Countries with more vulnerable economies, as measured by low ratings, were more sensitive to changes in U.S. interest rates.

- Furthermore, Gay (2008) investigated the time-series relationship between stock market index prices/returns and the macroeconomic variables of exchange rate and oil price for Brazil, Russia, India, and China (BRIC) using the Box-Jenkins ARIMA model. Although no significant relationship was found between respective exchange rates and oil price on the stock market index prices of any of the BRIC countries, Gay (2008) concluded that this might be due to the influence of other domestic and international macroeconomic factors on stock market returns, warranting further research.

Also, there was no significant relationship found between present and past stock market returns, suggesting the markets of Brazil, Russia, India, and China exhibit the weak-form of market efficiency.

- Also, Zadorozhna (2009) tested the relationship between stock market variables (indices returns, individual stocks' returns, spreads and trading volumes) and the weather in transition countries of Central and Eastern Europe. In his research, weather was considered to be a proxy for the mood factors that affect decisions of investors and traders. It was hypothesized that investors tend to be more optimistic about the market prospects if the weather is warm and sunny, and are more pessimistic if it is rainy and cloudy. Hence, market players are more predisposed to buy stocks when the weather is fine and sell them when the weather is bad.
- According to Fávero and Belfiore (2011) some investors prefer to use cash flow instead of earnings per share to evaluate stocks' current prices, and they argue that, while the first is not easily manipulated, the same cannot be said for earnings. Based on a sample from Compustat Global, including 3,567 stocks of companies from 35 emerging countries, covering 118 months (1998-2007), totalizing 218,530 observations, their study applied panel data models with different estimators to verify that price to cash flow ratio was more significant to influence returns over time, with more efficient estimators for the fixed effect model.

In most recent times, increasing attention is being paid to the relationship between share prices/returns and the macroeconomic variables by both economists and financial specialists. In the present-day scenario, where there is an increasing integration of the financial markets and implementation of various stock market reforms, the activities in the stock markets and their relationships with the macro economy have assumed significant importance. According to Singh et al (2011), economic agents use information in forming their expectations of future returns for holding stock securities. They examined the casual relationship between index returns and certain crucial macroeconomic variables namely employment rate, exchange rate, gross domestic product (GDP), Inflation and money supply in Taiwan. Their analysis was based on stock portfolios rather than single stocks. In portfolio construction, four criteria were used: Market capitalization, price/earnings ratio (PE) and yield. Empirical findings revealed that exchange rate and GDP seem to affect returns of all portfolios, while inflation rate, exchange rate, and money supply were having negative relationship with returns for portfolios of big and medium companies.

Bayezid (2011) similarly investigated the impact of changes in selected microeconomic and macroeconomic variables on stock returns at Dhaka Stock Exchange. A Multivariate Regression Model computed on standard ordinary least squares formula was used to estimate the relationship. Based on regression coefficients, he found that inflation and foreign remittance have negative influence on industrial production index, and market P/Es, while monthly percent average growth in market capitalization have positive influence on stock returns. All the independent variables could jointly explain 44.48 percent variation in Dhaka Stock Exchnge all share price index. Furthermore, no unidirectional Granger Causality was found between stock prices and all the predictor variables under the study, except one unidirectional causal relation, from stock price and market PEs. In a nut-shell, Bayedzid (2011) concluded that lack of Granger causality between stock price and selected micro and macro variables ultimately reveals the evidence of informationally inefficient market.

Further empirical studies on the performance of the Arbitrage Pricing Theory (APT) in the Nigerian Stock Exchange for the period of 2000 up to 2004 were conducted by Izedonmi and Abdullahi (2011). Three macro-economic variables (inflation, exchange rate and market capitalization) were investigated against 20 sectors of the Nigerian Stock Exchange. Using OLS they observed that there were no significant effects of those variables on the stocks' return in Nigeria. They concluded that their results were broadly consistent with similar studies carried out for most developed and emerging economies.

2.5 Stock Returns Modeling in Ghana

The Ghana Stock Exchange (GSE) was incorporated in July 1989 as a private company limited by guarantee under Ghana's companies' code, 1963. The Exchange however, changed its status to a public company limited by guarantee in April 1994. Trading on the floor of the Exchange commenced in November 1990 (see GSE, About Us).

The Ghana Stock Exchange was set up with the following objectives:

- To provide the facilities and framework to the public for the purchase and sales of bonds, shares and other securities;

- To control the granting of quotations on the securities market in respect of bonds, shares and other securities of any company, corporation, government, municipality, local authority or other corporate body;
- To regulate the dealings of members with their clients and other members;
- To co-ordinate the stock dealing activities of members and facilitate the exchange of information including prices of securities listed for their mutual advantages and for the benefit of their clients;
- To co-operate with associations of stockbrokers and Stock Exchanges in other countries, and to obtain and make available to members information and facilities likely to be useful to them or to their clients.

Since its inception, the GSE's listings have been included in the main index, the GSE All-Share Index. In 1993, the GSE was the sixth best index performing emerging stock market, with a capital appreciation of 116%. In 1994 it was the best index performing stock market among all emerging markets, gaining 124.3% in its index level. 1995's index growth was a disappointing 6.3%, partly because of high inflation and interest rates. Growth of the index for 1997 was 42%, and at the end of 1998 it was 868.35. As of October 2006 the market capitalization of the Ghana Stock Exchange was about 111,500 billon cedis ($11.5 billion). As of December 31 2007, the GSE's market capitalization was 131,633.22 billion cedis (see GSE Publications).

Since its inception, researchers and financial analysts have done a lot of studies on the GSE. Most of the studies hinge around the ARCH/GARCH Models. Sample of the studies are discussed below.

Frimpong and Oteng-Abayie (2006) modelled and forecasted volatility (conditional variance) on the Ghana Stock Exchange using a random walk (RW), GARCH(1,1), EGARCH(1,1), and TGARCH(1,1) models. The unique 'three days a week' Databank Stock Index (DSI) was used to study the dynamics of the Ghana stock market volatility over a 10-year period. The competing volatility models were estimated and their specification and forecast performance compared with each other, using Akaike information criteria (AIC) and log-likelihood information criteria and Brock-Dechert-Scheinkman nonlinearity diagnostic checks. The DSI exhibits the characteristics such as volatility clustering, leptokurtosis and asymmetry effects associated with stock market returns on more advanced stock markets. The random walk hypothesis is rejected for the DSI. Overall, the GARCH (1,1) model outperformed the other models under the assumption that the innovations follow a normal distribution.

A study by Adjasi et al (2008) looked at the relationship between Stock Markets and Foreign Exchange market, and determined whether movements in exchange rates have an effect on stock market in Ghana. The EGARCH model was used in establishing the relationship between exchange rate volatility and stock market volatility. It was found that there is negative relationship between exchange rate volatility and stock market returns – a depreciation in the local currency leads to an increase in stock market returns in the long run. Whereas in the short run it reduces stock market returns. They further found that there was volatility persistence in most of the macroeconomic variables; present rates had an effect on forecast variance of future rate. It was also revealed that an increase (decrease) in trade deficit and expectation in future rise in trade deficit will decrease (increase) stock market volatility. In addition, the consumer price index had a strong relationship with stock market volatility which meant that an increase in consumer price will lead to a rise in stock market volatility. The study also found the presence of leverage effect and volatility shocks in stock returns on the Ghana Stock Exchange.

A recent study by Aliyu (2011) also applied the Generalized Autoregressive Conditional Heteroskedastic (GARCH) model to assess the impact of inflation on stock market returns and volatility using monthly time series data from the two West African countries, that is, Nigeria and Ghana. In addition, the impact of asymmetric shocks was investigated using the QGARCH model developed by Sentana (1995), in both countries. Results for Nigeria show weak support for the hypothesis that bad news exert more adverse effect on stock market volatility than good news of the same magnitude, while a strong opposite case holds for Ghana. Furthermore, inflation rate and its three month average were found to have significant effect on stock market volatility in the two countries. Measures employed towards restraining inflation in the two countries, therefore, would certainly reduce stock market volatility and boost investor confidence.

2.6 Conclusion

In conclusion, researchers to date are searching for the most appropriate capital assets model. Considering the various studies on discussed above, this study seeks to take the microeconomic approach to the subject, by developing a multi-factor asset pricing model that inculcates all the variables which may capture significant variation in stock returns. So far, apart from Beta in the original CAPM, the other significant factors that have been identified include: the earnings to price ratio (Basu, 1977); firm size effect (Banz, 1981 and Chan, 1988); the leverage effect, (that is ratio of debt to equity, Bhandari, 1988); book-to-market value effect (Stattman 1980; Rosenberg et al, 1985; Chan et al, 1991); and growth in earnings (Chien-Ting, 1999 and Basu, 1983). These factors would be examined in detail in this study.

In addition, Treasury-Bill rates were collected from the Central Bank of Ghana whilst other relevant information was collected from annual reports of the individual companies quoted on the Ghana Stock Exchange.

The study period spanned January 2001 to December 2008. Although thirty five equities were listed on the Stock Exchange as at the year 2008, only fourteen equities which had traded consistently within the study period were considered.

3.1.1 Method of Data Collection

The required data were extracted from the sources stated above using the form designed for this purpose. A copy of this form is attached in Appendix 1.

3.1.2 Variables Description

All the variables (dependent and independent) that were considered in this study were derived or computed. The key variables were:

- Returns (R) on stock i, $R_i = \dfrac{SP_i - CP_i + Div_i}{CP_i}$, $i = 1, \ldots, 14$ (3.1)

  where CP, SP, and Div are respectively cost/opening price, selling/closing price, and dividend per share.

- Returns on risk free investments (*Rf*) will be the 91days Bank of Ghana T-Bill rates.

- The risk component for investing in stock i,

$$\beta_i = \frac{Cov(Rm, R_i)}{\text{var}(Rm)} \quad , \quad i = 1, \dots , 14 \qquad (3.2)$$

where, the returns on the market index at time t is,

$$Rm = \frac{SP_m - CP_m}{CP_m} \quad ;$$

CP_m, and SP_m being the opening, and closing prices of market index respectively. β_i is also the coefficient of the $Rm - Rf$ (Rmf) variable.

- Firm Size (S) of stock i at time t:

$$S_{it} = \ln[ME] = \ln[No \times SP] \qquad (3.3)$$

where, ME, is the market capitalization, No, the number of outstanding shares, and SP, the closing price.

- The book-to-market (BM) equity:

$$\ln[B / M] \qquad (3.4)$$

where, B is the book equity, and M, the market equity.

- The firm's expected growth rate (G):

$$g_t = (E_t / E_{t-5})^{\frac{1}{5}} - 1 \qquad (3.5)$$

where, $E_t = \dfrac{E_{t-1} + E_{t-2} + E_{t-3}}{3}$ is the average earnings between time t-1 and t-3;

$E_t = \dfrac{E_{t-6} + E_{t-7} + E_{t-8}}{3}$ is the average earnings between time t-6 and t-8.

Alternatively:

$$g_t = \left(\frac{E_t - E_{t-1}}{E_{t-1}}\right) \tag{3.6}$$

- Leverage, L (DE), as the ratio of the firms debts, D, to its equity, E is:

$$L = \frac{D}{E} \tag{3.7}$$

- The price to earnings ratio (PE) is simply:

$$\frac{P}{E} \tag{3.8}$$

where, P is the closing price per share and E, the earnings per share.

3.1.3 Problem Encountered in data Collection

For this study, quarterly data for the period of January 2001 to December 2008 was used. In order to obtain better estimates of the coefficients of the explanatory variables of stock returns, use was made of monthly or quarterly data, since longer time periods (eg. annually, biannually, etc) could result in changes over the period examined introducing biases into the estimates. On the other hand, high frequency data such as daily observations covering a relatively short and stable time span could result in the use of very noisy data and thus yield inefficient estimates (Michailidis et al 2006).

Unfortunately, data for most of the explanatory variables incidentally were not available on monthly or quarterly basis but rather annually. Therefore, the annual figures were disaggregated into quarterly data using the best fitted disaggregation technique.

3.2 Methods of Disaggregation

This section provides an adequate disaggregation method for the factors affecting returns on stocks. Some different approaches for disaggregating annual data to quarterly data have been developed in the past years. The two main types of disaggregation described in the literature include:

1. The Plausible method (Linear interpolation; dividing by 4; Lisman and Sandee, 1964)

2. Model – Based

   - Regression (Chow and Lin, 1971; Fernandez, 1981; Litterman, 1983)

   - ARIMA (Guerrero, 1990; Nijman and Palm, 1990; Gudmundsson, 1999)

   - Least Squares (Boot et al, 1967; Stram and Wei, 1986; Jacob et al, 1989; Hodgess and Wei, 1996; Tasdemir, 2008)

In the first group these are methods which divide the annual data into quarterly figures "in a plausible way". These include linear interpolations or simply "dividing by four" where the method of Lisman and Sandee (1964) is a special case.

The second group employs the so called model-based procedures. High correlating time series is used to create the disaggregated series or the wanted disaggregated series is assumed to follow an ARIMA process. Furthermore, the Least Squares tries to minimize the sum of the squared changes of the quarterly values respectively their d-th differences.

Assuming that for an integer T, $\boldsymbol{y}$ is the known T x 1 vector of the annual data for t = 1,...,T, and $\boldsymbol{x}$ is the unknown 4T x 1 vector of the quarterly data, then in case of aggregation we have the following relation:

$$\boldsymbol{y} = \boldsymbol{C}'\boldsymbol{x} \tag{3.9}$$

where $\boldsymbol{C}$ is a 4T x T aggregation matrix.

$$\boldsymbol{C}' = \boldsymbol{I}_T \times \boldsymbol{e}' \tag{3.10}$$

where I_T is the T x T identity matrix and e is e = (1 0 0 0) in the case of stocks if the first quarter is observed otherwise the "1" has to be moved to one of the other positions.

Similarly:

$$\boldsymbol{x} = \boldsymbol{H}'\boldsymbol{y} \tag{3.11}$$

with H being a T x 4T disaggregation matrix as discussed by Kladroba (2005).

3.2.1 The Plausible Methods

a. Dividing by four

It is easy to see that the disaggregation matrix must be:

$$\boldsymbol{H} = \tfrac{1}{4}\,\boldsymbol{C} \tag{3.12}$$

Thus for a given annual figure $\boldsymbol{t}$, we obtain the four equal quarterly figures:

$$x_1 = x_2 = x_3 = x_4 = x = \tfrac{1}{4}\, t \tag{3.13}$$

b. The Procedure of Lisman and Sandee (1964)

In the case of the "dividing by four" method, at the beginning of every year there is a "step" in the disaggregated time series. Lisman and Sandee (1964) wanted to avoid this by building a weighted mean of the quarterly values of the years t_{i-1}, t_i and t_{i+1}. Thus the procedure of Lisman and Sandee includes two steps.

Firstly, divide the annual totals for each year into four equal quarterly figures

$$x_{i1} = x_{i2} = x_{i3} = x_{i4} = x_i = \tfrac{1}{4}\, t_i \tag{3.14}$$

This implies that:

$$\sum_{j=1}^{4} x_{ij} = t_i \tag{3.15}$$

Then, assume the quarterly figures x_{ij} to be a weighted sum of t_{i-1}, t_i and t_{i+1}. The above assumption may be written as:

$$\begin{pmatrix} x_{i1} \\ x_{i2} \\ x_{i3} \\ x_{i4} \end{pmatrix} = \begin{bmatrix} a & e & d \\ b & f & c \\ c & f & b \\ d & e & a \end{bmatrix} \begin{bmatrix} t_{i-1} \\ t_i \\ t_{i+1} \end{bmatrix} \tag{3.16}$$

where the coefficients are unknowns to be determined.

Lisman and Sandee (1964) solved the system above and obtained the values of the coefficients as presented in the matrix below:

$$\operatorname{cov}\left(\hat{x}-x\right)=E\left[\left(\hat{H}'C'u-u\right)\left(\hat{H}'C'u-u\right)^{-1}\right]=\hat{H}'C'VC\hat{H}'-\hat{H}'C'V-VC\hat{H}'+V \quad (3.22)$$

By minimizing this term the unbiased minimum variance estimator is obtained as:

$$\hat{x}=Z\hat{\beta}+\left[C'VC\left(C'VC\right)^{-1}\right]C\hat{u}=Z\hat{\beta}+C\hat{u} \quad (3.23)$$

where $\hat{\beta}$ is the GLS-estimator using the T aggregated data and $\hat{u}$ the corresponding residual vector:

$$\hat{\beta}=\left[Z'C\left(C'VC\right)^{-1}C'Z\right]^{-1}Z'C\left(C'VC\right)^{-1}Cy \quad (3.24)$$

and $$C\hat{u}=y-C'Z\hat{\beta}. \quad (3.25)$$

b. ARIMA Based Models

Assuming that the wanted disaggregated time series follows an ARIMA (p, d, q) – process:

$$\phi(B)(1-B)x_t=\tau(B)\varepsilon_t \quad (3.26)$$

where B is the shift operator and $Bx_t=x_{t-1}$; ε_t is gaussian. In a similar way as in the regression based model above the conditional mean of x is:

$$E(x_t \mid x_1,x_2,\ldots)=E(x_t) \quad (3.27)$$

The unbiased minimum variance estimator becomes:

$$\hat{x}=E(x)+\theta\theta'C\left(C'\theta\theta'C\right)^{-1}\left[y-C'E(x)\right] \quad (3.28)$$

with the estimation error

$$x_t-E(x_t)=\sum_{j=0}^{t-1}\theta_j\varepsilon_{t-j} \quad (3.29)$$

where θ_1, θ_2, ... is the solution of

$$\phi(B)\phi(B)d(B)\tau^{-1}(B)=1 \tag{3.30}$$

c. Minimizing Squared First Difference

We minimize the sum of the squares of the differences between the successive quarterly values, subject to the constraints that during each year the sum of the quarterly totals should equal the yearly total. Mathematically, if there are n years we wish to:

$$\text{minimize} \quad \sum_{i=2}^{4n}(x_i - x_{i-1})^2 \tag{3.31}$$

$$\text{subject to} \quad \sum_{i=4k-3}^{4n} x_i = t_k \qquad (k=1, 2, \ldots, n) \tag{3.32}$$

where x_i is the ith quarterly total and t_k is the given yearly total in year k.

The problem is solved routinely by considering the Lagrangean expression

$$\sum_{i=2}^{4n}(x_i - x_{i-1})^2 - \sum_{k=1}^{n}\lambda_k\left(\sum_{i=4k-3}^{4k} x_i - t_k\right) \tag{3.33}$$

Upon differentiating with respect to x_i (i = 1, 2, …, 4n) and λ_k (k = 1, 2, …,n) and equating the resulting expression to zero we obtain 4n+n equations in 4n+n unknowns of x_i and λ_k.

The problem is to solve the linear system $\begin{bmatrix} B & -J' \\ J & 0 \end{bmatrix}\begin{bmatrix} X \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ t \end{bmatrix}$. (3.34)

The matrix B is a band matrix of order 4nx4n denoted by.

$$B = \begin{bmatrix} 2 & -2 & & & & & & \\ -2 & 4 & -2 & & & & & \\ & -2 & 4 & -2 & & & & \\ & & & & -2 & 4 & -2 & \\ & & & & & -2 & 4 & -2 \\ & & & & & & -2 & 2 \end{bmatrix} \tag{3.35}$$

and J is a matrix of order n x $4n$, which for n = 3 is given by:

$$J = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \tag{3.36}$$

The quarterly figures x_i which satisfy equation (3.34) solve the minimization problem. The solution when n=3 is obtained from:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{bmatrix} = \frac{1}{1836} \begin{bmatrix} 569 & -135 & 25 \\ 525 & -81 & 15 \\ 437 & 27 & -5 \\ 305 & 189 & -35 \\ 129 & 405 & -75 \\ 7 & 513 & -61 \\ -61 & 513 & 7 \\ -75 & 405 & 129 \\ -35 & 189 & 305 \\ -5 & 27 & 437 \\ 15 & -81 & 525 \\ 25 & -135 & 569 \end{bmatrix} \cdot \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} \tag{3.37}$$

The following observations which remain valid when n>3 are made:

1. The sum of the first four x_i terms is identically equal to t_1, etc.

2. If $t_1 = t_2 = t_3 = t$, then all the $x_i = ¼t$.

3. There is symmetry: if the sequence t_1, t_2, t_3, leads to x_1, x_2, ..., x_{12}, then the sequence t_3, t_2, t_1 leads to x_{12}, x_{11}, ..., x_1.

4. If $t_2 - t_1 = t_3 - t_2$, then $x_i - x_{i-1}$ is constant for $i = 5$ *to* 9

d. Minimizing Squared Second Difference

We also minimize the sum of squares of the second difference:

$$\text{minimize} \quad \sum_{i=2}^{4n} (\Delta x_i - \Delta x_{i-1})^2 \tag{3.38}$$

$$\text{subject to} \quad \sum_{i=4k-3}^{4n} x_i = t_k \qquad , (k = 1, 2, \ldots, n) \tag{3.39}$$

$$\text{where} \quad \Delta x_i = x_{i+1} - x_i \tag{3.40}$$

The mathematics here follows the same steps as in the previous case. The problem now is to solve the linear system of 5n equations in 5n unknowns

$$\begin{bmatrix} C & -J' \\ J & 0 \end{bmatrix} \begin{bmatrix} X \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ t \end{bmatrix} \tag{3.41}$$

The only difference between (3.34) and (3.41) is that the matrix B is replaced by the 4n x 4n band matrix

$$C = \begin{bmatrix} 2 & -4 & 2 & & & & & \\ -4 & 10 & -4 & 2 & & & & \\ 2 & -8 & 12 & -8 & 2 & & & \\ & 2 & -8 & 12 & -8 & 2 & & \\ & & 2 & -8 & 12 & -8 & 2 & \\ & & & 2 & -8 & 12 & -8 & 2 \\ & & & & 2 & -4 & 10 & -4 \\ & & & & & 2 & -4 & 2 \end{bmatrix} \quad (3.42)$$

For $n = 3$, the solution is obtained from:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \\ x_8 \\ x_9 \\ x_{10} \\ x_{11} \\ x_{12} \end{bmatrix} = \frac{1}{9280} \begin{bmatrix} 3499 & -1488 & 309 \\ 2697 & -464 & 87 \\ 1911 & 528 & -119 \\ 1173 & 1424 & -277 \\ 531 & 2128 & -339 \\ 49 & 2512 & -241 \\ -241 & 2512 & 49 \\ -339 & 2128 & 531 \\ -277 & 1424 & 1173 \\ -119 & 528 & 1911 \\ 87 & -464 & 2697 \\ 309 & -1488 & 3499 \end{bmatrix} \cdot \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} . \quad (3.43)$$

Again we may verify that:

1. All side conditions are satisfied.

2. If $t_1 = t_2 = t_3 = t$, then all $x_i = \frac{1}{4}t$.

3. There is symmetry.

4. If $t_2 - t_1 = t_3 - t_2$, then the whole curve is a straight line.

3.3 Cross Sectional Models

The objective of this section is to model the relationship between the stock returns and the significant factors that contribute to these returns. The main tools employed in this section were multiple linear regression and distribution fitting.

The first part dealt with fitting general linear models based on normality assumptions for the error terms. The behavior of the error terms was then investigated and their distribution modelled.

3.3.1 Linear Models

As it is widely known, the general linear regression model, estimates the mean of the response variable by using the regression parameters.

The model is of the form

$$y = \beta X + \varepsilon \qquad (3.44)$$

where y is the vector of observed response,

X is the design matrix of predictor variables,

β is the vector of regression parameters and

ε is the vector of random errors.

The random errors are assumed to be independent and normally distributed with a common variance. If these parametric assumptions are valid, then the estimated regression parameters are the best linear unbiased estimates (BLUE).

In matrix form: for an $n \times 1$ response vector y for individual i, equation (3.44) can be written as

$$\underset{n\times 1}{y_i} = \underset{n\times m}{X_{ij}} \underset{m\times 1}{\beta_j} + \underset{n\times 1}{\varepsilon_i} \qquad (3.45)$$

with $i = 1 \ldots n$ individuals and $j = 1 \ldots m$ observations for individual i. Here, y_i is the $n_i \times 1$ dependent variable vector for individual i. X_{ij} is the $n \times m$ covariate matrix for individual i. and ε_i is the $n \times 1$ error vector.

From equation (3.45), taking $j = 7$ for the seven explanatory variables we obtained a linear model:

$$Y_{it} = \beta_0 + \beta_1 X_{1it} + \beta_2 X_{2it} + \beta_3 X_{3it} + \beta_4 X_{4it} + \beta_5 X_{5it} + \beta_6 X_{6it} + \beta_7 X_{7it} + \varepsilon_{it} \quad (3.46)$$

where Y_{it} = return on stock i for time t;

X_{1it} = returns on a risk free investment (T-bill rates) at time t;

X_{2it} = the risk associated with stock i at time t;

X_{3it} = the natural log of the firm size (market capitalization) of stock i at time t;

X_{4it} = the earnings growth of stock i at time t;

X_{5it} = the book to market ratio of stock i at time t;

X_{6it} = the debt to equity ratio of stock i in time t;

X_{7it} = the earning to price ratio of stock i in time t;

$\beta' s$ = regression parameters;

and ε_{it} = regression errors

3.3.2 Non-Normally Distributed Errors

Violations of the assumption of normally distributed errors can threaten the efficiency of estimation (for example, in the case of heavy-tailed error distributions) or can compromise the interpretability of the least-squares fit, which estimates the conditional mean of the response variable as a function of the covariates (for example, in the case of skewed errors).

Least-squares residuals have some different properties from the errors; nevertheless, examining the distribution of the residuals can be informative about the distribution of the errors (Cook and Weisberg, 1982). Furthermore a non-constant error variance (or heteroscedasticity) also threatens the efficiency of least-squares estimation as well as the validity of statistical inference.

3.3.3 Goodness–of–fit Tests

The Kolmogorov-Smirnov and Anderson-Darling tests were the main goodness–of–fit tests used in this research.

a. Kolmogorov–Smirnov Test

This test is used to decide if a sample comes from a hypothesized continuous distribution. It is based on the empirical cumulative distribution function (ECDF). For a random sample $x_1,\ldots,x_n$ from some cumulative distribution function (CDF), *F(x)*. The empirical CDF is denoted by

$$F_n(x) = \frac{1}{n}\left[number\ of\ observatio\ ns \le x\right] \quad (3.47)$$

The Kolmogorov-Smirnov statistic (D) is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function:

$$D = \max_{1 \le i \le n}\left(F(x_i) - \frac{i-1}{n}, \frac{i}{n} - F(x_i)\right) \quad (3.48)$$

Hypothesis Testing

The null and the alternative hypotheses are:

H_0: the data follow the specified distribution;

H_A: the data does not follow the specified distribution.

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, D, is greater than the critical value Dc, where Dc is obtained from a Chi-Squared table.

b. Anderson–Darling Test

The Anderson-Darling procedure is a general test to compare the fit of an observed cumulative distribution function to an expected cumulative distribution function. This test gives more weight to the tails than the Kolmogorov-Smirnov test.

The Anderson-Darling statistic (A^2) is defined as:

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\cdot\left[\ln F(x_i) + \ln(1 - F(x_i))\right] \quad (3.49)$$

Hypothesis Testing

The null and the alternative hypotheses are:

H_0: the data follow the specified distribution;

H_A: the data does not follow the specified distribution.

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, A^2, is greater than the critical value Dc, where Dc is obtained from a Chi-squared table.

Note that in both tests, the smaller the test statistic value the better the fit.

3.3.4 Validation Methods

Bootstrapping, jack-knifing and cross–validation are three superficially similar statistical techniques that involve reusing or re-sampling data. In each case a single sample of observations is considered as many samples with the same estimation process being applied to each of them. However, the purposes of this reuse of the samples are quite different for each method. In summary:

- Bootstrapping is a method for evaluating the variance of an estimator.
- Jack-knifing is a method for reducing the bias of an estimator, and evaluating the variance of an estimator.
- Cross validation is a method for evaluating the error involved in making predictions.

Since we wanted to reduce the biases in our estimates we applied the jack-knifing method in this research.

The Jack-knifing Method

The jack-knifing method is a more orderly version of the bootstrap. Instead of generating a set of random samples from $X_1, \ldots, X_n$, we generate n samples of size $(n-1)$ by leaving out one observation at a time. The steps involved in this method are:

1. Observe a sample $X = \{X_1, \ldots, X_n\}$.
2. Compute $\hat{\theta}(X)$ a function of the data which estimates some parameter θ of the model.
3. For $i = 1$ up to n
   - generate a jack-knife sample $X^{-i} = \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots\}$ by leaving out the i^{th} observation.
   - Calculate $\hat{\theta}_{-i}$ by applying the estimation process to the jack-knife sample.
4. Calculate
   - the jack-knifed estimate $\hat{\theta}_{\bullet} = \frac{1}{n}\sum_{i=1}^{n}\hat{\theta}_{-i}$.
   - and the jack knife estimate of variance $\frac{n-1}{n}\sum_{i=1}^{n}\left(\hat{\theta}_{-i} - \hat{\theta}_{\bullet}\right)$.

3.4 Incorporating Randomness

An objective of this study is to examine the behaviour of the responses (returns) over time. This behaviour is best modeled as longitudinal analysis. The defining feature of any longitudinal model is its ability to track changes over time within subjects (individual equities) and changes over time between groups (market). Longitudinal data analyses further have the characteristics of nonlinear growth patterns and the errors within exhibit heteroscedasticity and dependence. Consequently, longitudinal data analysis is more complicated than that of cross-sectional data analysis.

Special methods of statistical analysis are needed for longitudinal data because the set of measurements on the response tends to be correlated. Measurements on the same subject taken over narrow time interval tend to be more highly correlated than measurements taken wide apart in time, and the variances of longitudinal data often change with time. These potential patterns of correlation and variation may combine to produce a complicated covariance structure. This covariance structure must be taken into account to draw valid statistical inferences. Therefore, standard regression and analysis of variance (ANOVA) models may produce invalid results because two of the parametric assumptions (independent observations and equal variances) may not be valid. Concequently, we fitted General Linear Mixed Model (GLMM) would be fitted to the stock returns data collected for this work.

3.4.1 General Linear Mixed Models (GLMM)

The general linear mixed model is an extension of the general linear model (GLM). The standard linear regression model, which is used in the GLM procedure, models the mean of the response variable by using the regression parameters. The random errors are assumed to be independent and normally distributed with a common variance. If these parametric assumptions are valid, then the estimated regression parameters are the best linear unbiased estimates (BLUE).

The GLMM extends the GLM by the addition of random effect parameters and by allowing a more flexible specification of the covariance matrix of the random errors. For example, general linear mixed models allow for both correlated error terms and error terms with heterogeneous variances. The name mixed model indicates that the model contains both fixed–effect parameters and random-effect parameters.

The general linear mixed model is of the form

$$y = \beta X + bZ + \varepsilon \qquad (3.50)$$

y is the vector of observed response.

X is the design matrix of predictor variables.

β is the (fixed) vector of regression (fixed–effect) parameters.

Z is the design matrix of random variables.

b is the vector of random effect parameters. It represents parameters that are allowed to vary over subjects. It also represents subject–specific regression coefficients that reflect the natural heterogeneity in the population with $\mathrm{Cov}(b) = \mathrm{G}$.

ε represents within–subject variation. It is not required to be independent and homogeneous. It also has a covariance matrix (R) that is block diagonal with each block corresponding to a subject.

The variable effects are either fixed or random depending on how the levels of the variables that appear in the study are selected.

In matrx form, for an $n_i \times 1$ response vector y for an individual i, (i=1,...,N) equation (3.50) can be written as:

$$\underset{n_i \times 1}{y_i} = \underset{n_i \times m}{X_i} \; \underset{m \times 1}{\beta} + \underset{n_i \times p}{Z_i} \; \underset{p \times 1}{b_i} + \underset{n_i \times 1}{\varepsilon_i} \qquad (3.51)$$

for $j = 1 \ldots n_i$ observations on each individual i. Here, y_i is the $n_i \times 1$ dependent variable vector for individual i. X_i is the $n_i \times m$ covariate matrix for individual i. β is the $m \times 1$ vector of fixed regression parameters. Z_i is the $n_i \times p$ design matrix for the random effects, b_i is the $p \times 1$ vector of random individual effects, and ε_i is the $n_i \times 1$ error vector.

A GLMM has the following assumptions:

1. Random effects and error terms are normally distributed with means zero.
2. Random effects and error terms are independent of each other.
3. The relationship between the response variable and predictor variables is linear.
4. The variance-covariance matrices for random effects and error terms are block diagonal with each block corresponding to a subject.

With these assumptions, by appropriately defining the model matrices for fixed and random effects, and the covariance structures for the random effects and the error terms, one can perform numerous mixed model analyses. It is important to note that the residuals of the general linear models are cauchy distributed hence the need to incorporate that into the mixed models being developed. Specifically, GLMM with cauchy distrbuted error terms is the one developed in this work.

3.4.2 Random Effect Models

Now for a simple linear regression model of say,

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \varepsilon_{ij} \tag{3.52}$$

which represents the regression of the outcome variable y on the independent variable time (denoted *t)*, the change across time is the same for all individuals since the model parameters (β_0, the intercept or initial level, and β_i, the linear change across time) do not vary across individuals. For this reason, it is useful to add individual–specific effects into the model that will account for the data dependency and describe differential time trends for different individuals.

An extension of the regression model given in equation (3.52) to allow for the influence of each individual on their repeated outcomes is provided by

$$y_{ij} = \beta_0 + \beta_1 t_{ij} + \upsilon_{0i} + \varepsilon_{ij} \tag{3.53}$$

where υ_{0i} represents the influence of an individual *i* on repeated observations of the same. In a hierarchical or multilevel form (Goldstein, 1995; Raudenbush and Bryk, 2002). Equation (3.53) is partitioned into the following within-subjects (or level–1) model:

$$y_{ij} = b_{0i} + b_{1i}t_{ij} + \varepsilon_{ij} \tag{3.54}$$

and between-subjects (or level-2) model:

$$b_{0i} = \beta_0 + \upsilon_{0i}$$

$$b_{1i} = \beta_1 \tag{3.55}$$

Here, the level-1 model indicates that an individual, *i's,* response at time j is influenced by his/her initial level b_{0i} and time trend, or slope, b_{1i}. The level-2 model indicates that an individual, *i's,* initial level is determined by the initial level of the population β_0 plus a unique contribution for that individual υ_{0i}. Thus, each individual has his/her own distinct initial level. Conversely, the present model indicates that the slope of each individual is the same; each is equal to the population slope β_0.

In the model given by equation (3.53) the errors ε_{ij} are assumed to be normally and conditionally independently distributed in the population with zero mean and common variance σ^2. Conditional independence here means conditional on the random individual-specific effects υ_{0i} which contributes an additional variance σ_{υ}^2. This model is sometimes called a random-intercept model, with each υ_{0i} indicating how an individual i deviates from the population trend. Finally, the random intercept model implies a compound symmetry assumption for the variances and covariances of the longitudinal data. That is, both the variances and covariances across time are assumed to be the same, thus

$$V(y_{ij}) = \sigma^2 + \sigma_{\upsilon}^2$$

$$Cov(y_{ij}, y_{ij'}) = \sigma^2. \tag{3.56}$$

The random intercept model is sometimes criticized for its simplicity, in that, individuals differ in their time trends and the compound symmetry assumption is usually untenable for most longitudinal data. For these reasons, one needs a more realistic model that allows both the intercept and time slope to vary by individuals, such that the level-1 model will remain as before in equation (3.55) but the level-2 model is augmented as:

$$b_{0i} = \beta_0 + \upsilon_{0i}$$

$$b_{1i} = \beta_1 + \upsilon_{1i} \qquad (3.57)$$

β_0 is the overall population intercept, β_1 is the overall population slope, υ_{0i} is the intercept deviation for subject i, υ_{1i} is the slope deviation for subject i. ε_{ij} are now assumed to be normally and conditionally independently distributed in the population with zero mean and common variance σ^2. Conditional independence here similarly means conditional on the random individual-specific effects υ_{0i} and υ_{1i}. This model is often called the random intercept and slope model (Hedeker and Gibbons, 2006). With the two individual specific-effects, the population distribution of intercept and slope deviations are assumed to be bivariate normal $\mathrm{N}(0, \Sigma_{\upsilon})$, with the random-effects variance–covariance matrix being:

$$\Sigma_{\upsilon} = \begin{bmatrix} \sigma^2_{\upsilon_0} & \sigma_{\upsilon_0\upsilon_1} \\ \sigma_{\upsilon_0\upsilon_1} & \sigma^2_{\upsilon_1} \end{bmatrix}. \qquad (3.58)$$

Several variance covariance structures that are used to model these effects, including the compound symmetry, would be discussed in Section 3.4.5

3.4.3 Estimations in Mixed Models

The variance–covariance matrix of the observations in the case of mixed models involves the covariance structure of the random effects, denoted **G** and the covariance structure of the random errors, denoted **R**. Ordinary least squares is no longer the best method of estimation because the distributional assumptions regarding the random error terms are too restrictive. In other words, the parameter estimates are no longer the best linear unbiased estimates. The generalized least square (GLS) estimate of β is

$$\hat{\beta} = (X' \hat{V}^{-1} X')^{-} X' \hat{V}^{-1} y \qquad (3.59)$$

where $\hat{V} = ZGZ' + R$

GLS requires the knowledge of **G** and **R** and the goal is to find a reasonable estimate for **G** and **R**.

Two likelihood–based methods: maximum likelihood (ML) and restricted maximum likelihood (REML) can be used to estimate the parameters in **G** and **R**. The difference between ML and REML is with the construction of their likelihood functions. However, the two methods are asymptotically equivalent and often give very similar results. The distinction between ML and REML becomes important only when the number of fixed effects is relatively large. In that case, the comparisons unequivocally favour REML.

First, REML copes much more effectively with strong correlations among the responses for the subjects than does ML. Second, REML estimates do not have the downward bias that ML estimates have because REML estimators take into account the degrees of freedom from the fixed effects in the model. Finally, REML estimators are less sensitive to outliers in the data than ML estimators. In fact, when the estimates do vary substantially, they favour REML (Diggle et al, 1994).

There is also the non iterative MIVQUE0 method, which performs minimum variance quadratic unbiased estimation of the covariance parameters. However, Swallow and Monahan (1984) present simulation evidence favouring REML and ML over MIVQUE0. MIVQUE0 is generally not recommended except for situations when the iterative REML and ML methods fail to converge and it is necessary to obtain parameter estimates from a fitted model.

3.4.4 Covariance Structures

The validity of the statistical inference of the general linear mixed model depends upon the covariance structure selected for **R**. Therefore, a large amount of time spent on building the model is spent on choosing a reasonable covariance structure for **R**. The practical knowledge is that if you choose a structure that is too simple, you risk increasing the Type I error rate. On the other hand, if the structure is too complex, you sacrifice power and efficiency.

To be discussed now; the various covariance structures found in most of the literature. It is important to note here that all the structures discussed below are symmetric.

a. The simplest covariance structure is the independent or variance component (VC) model, where the within–subject error correlation is zero. Therefore:

$$\mathrm{VC} = \begin{pmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{pmatrix}. \qquad (3.60)$$

For the between–subject errors, the simple covariance structure may be a reasonable assumption. However, for the within–subject errors, the simple covariance structure may be a reasonable choice if the repeated measurements occurred at long enough intervals so that the correlation is effectively zero relative to other variation.

b. The covariance structure with the simplest correlation model is the compound symmetry (CS) structure, given by:

$$\mathrm{CS} = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \rho \\ & 1 & \rho & \rho \\ & & 1 & \rho \\ & & & 1 \end{pmatrix}. \qquad (3.61)$$

It assumes that the correlation (ρ) is constant regardless of the distance between the time points. This is the assumption that univariate ANOVA makes, but it is usually not a reasonable choice in longitudinal data analysis. However, this covariance structure may be reasonable when the repeated measurements are not obtained over time.

c. The unstructured (UN) covariance structure is parameterized directly in terms of variances and covariances where the observations for each pair of times have their own unique correlations. Thus:

$$\mathrm{UN} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ & & \sigma_3^2 & \sigma_{34} \\ & & & \sigma_4^2 \end{pmatrix}. \tag{3.62}$$

The variances are constrained to be nonnegative and the covariances are unconstrained. This is the covariance structure used in multivariate ANOVA.

The correlation coefficient for row 1 column 2 is

$$\rho_{12} = \frac{\sigma_{12}}{\sqrt{\sigma_1^2 * \sigma_2^2}}. \tag{3.63}$$

There are two potential problems with using the unstructured covariance. First, it requires the estimation of a large number of variance and covariance parameters. This can lead to severe computational problems, especially with unbalanced data. Second, it does not exploit the existence of trends in variances and covariances over time, and this can result in erratic patterns of standard error estimates (Littell et al, 1998). If a simpler covariance structure is a reasonable alternative, then the unstructured covariance structure wastes a great deal of information, which would adversely affect efficiency and power.

d. The first–order autoregressive (AR(1)) covariance structure takes into account a common trend in longitudinal data; the correlation between observations is a function of the number of time points apart:

$$\mathrm{AR}(1) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \rho^3 \\ & 1 & \rho & \rho^2 \\ & & 1 & \rho \\ & & & 1 \end{pmatrix}. \qquad (3.64)$$

In this structure, the correlation between adjacent observations is ρ, regardless of whether the pair of observations is the 1st and 2nd pair, the 2nd and 3rd pair, and so on. The correlation is ρ^2 for any pair of observations 2 units apart, and ρ^d for any pair of observations d units apart. Notice that the AR(1) model only requires estimates for just two parameters, σ^2 and ρ, whereas the unstructured models require estimates for $\frac{(1+T)T}{2}$ parameters (where T is the number of time points).

The assumption in the AR(1) model is that the longitudinal data is equally spaced (Littell et al, 1996). This means that the distance between time 1 and 2 is the same as time 2 and 3, time 3 and 4, and so on. The AR(1) structure also assumes that the correlation structure does not change appreciably over time (Littell et al, 2002).

e. The Toeplitz (TOEP) covariance structure is similar to the AR(1) covariance structure in that the pairs of observations separated by a common distance share the same correlation. However, observations d units apart have correlation ρ_d instead of ρ^d. The Toeplitz structure requires the estimation of T parameters instead of just two parameters:

$$\text{TOEP} = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ & 1 & \rho_1 & \rho_2 \\ & & 1 & \rho_1 \\ & & & 1 \end{pmatrix}. \quad (3.65)$$

One can also specify banded Toeplitz structure in which one specifies the number of time points apart for which the measurements were still correlated. For example, a TOEP(3), thus, Toeplitz with three bands structure would indicate that measurements are correlated if they are three or fewer time points apart. If they are four or more time points apart, the correlation is zero.

As with the AR(1) structure, the Toeplitz structure assumes that the observations are equally spaced and the correlation structure does not change appreciably over time (Littell et al, 2002).

f. Covariance structures that allow for unequal spacing are the spatial power (SP) covariance structures:

$$\text{SP} = \sigma^2 \begin{pmatrix} 1 & \rho^{|t_1-t_2|} & \rho^{|t_1-t_3|} & \rho^{|t_1-t_4|} \\ & 1 & \rho^{|t_2-t_3|} & \rho^{|t_2-t_4|} \\ & & 1 & \rho^{|t_3-t_4|} \\ & & & 1 \end{pmatrix}. \quad (3.66)$$

These structures are mainly used in geo-statistical models, but they are very useful for unequally spaced longitudinal measurements where the correlations decline with time. The connection between geo-statistics and longitudinal data is that the unequally spaced data can be viewed as a spatial process in one dimension (Littell et al, 1998).

The spatial power structure provides a direct generalization of the AR(1) structure for equally spaced data. In fact, the AR(1) structure is more efficient when you have equal spacing. Only two parameters are estimated (σ^2 and ρ).

g. A frequently used covariance structure for unequally spaced measurements is the spatial gaussian (SG) structure:

$$\text{SP} = \sigma^2 \begin{pmatrix} 1 & e^{\left(\frac{-|t_1-t_2|^2}{\rho^2}\right)} & e^{\left(\frac{-|t_1-t_3|^2}{\rho^2}\right)} & e^{\left(\frac{-|t_1-t_4|^2}{\rho^2}\right)} \\ & 1 & e^{\left(\frac{-|t_2-t_3|^2}{\rho^2}\right)} & e^{\left(\frac{-|t_2-t_4|^2}{\rho^2}\right)} \\ & & 1 & e^{\left(\frac{-|t_3-t_4|^2}{\rho^2}\right)} \\ & & & 1 \end{pmatrix}. \quad (3.67)$$

The difference between the spatial covariance structures is the assumptions made on how the correlation between the error terms decreases as the length of the time interval increases.

3.4.5 Covariance Structure Selection Criteria

To determine which correlation function is the best fit for the stock returns of this sudy, the intraclass correlation (ICC) is used for the various covariance structures described above. The intraclass correlation represents the proportion of (unexplained) variation in the dependent variable that is due to subjects. Here, "unexplained" refers to the variation not explained by the fixed effects of the model: group, time, and group by time. As the intraclass correlation approaches zero one can conclude that there is little correlation among the repeated observations over time.

The numeric formula for the ICC is:

$$ICC = \frac{\sigma_{\upsilon_0}^2}{\sigma_{\upsilon_0}^2 + \sigma^2} \tag{3.68}$$

where σ^2 is the residual variance and $\sigma_{\upsilon_0}^2$ is the sum of covariance estimates.

3.4.6 Degrees of Freedom Method

In a simulation study performed by Guerin and Stroup (2000), the Kenward–Roger (KR) degrees of freedom adjustment was shown to be superior, or at worst equal, to the Satterthwaite and other options. Their study revealed that, for complex covariance structures, the Type I error rate inflation was extremely severe unless the KR adjustment was used. Guerin and Stroup (2000) strongly recommend the KR adjustment as the standard operating procedure for longitudinal models.

3.5 Conclusion

The various mathematical and statistical techniques which were employed in the research have been discussed in this chapter. These techniques mainly included: the methods applied in data collection; methods of disaggregation of annual data into quarterly figures; the General Linear Model; and the General Linear Mixed Models. These techniques were then used for emperical analyses and the results discussed in the next chapter.

CHAPTER 4

RESULTS AND ANALYSES

4.0 Introduction

Presented in this chapter are the analyses of the empirical results using the various techniques described in the preceding Chapter.

The chapter begins with exploratory analysis. The purpose of this analysis is to organize/integrate various visual/graphical techniques to explore the stock returns data collected. This would help to identify trends at both group and individual levels, the time points where important changes occur, and unusual subjects/outliers. It would also help in selecting suitable statistical models and suggesting possible within-error variance.

Furthermore, the linear regression analyses would be performed where R (returns on equity) was regressed on Rf, Rmf, S, BM, G, DE and PE as described in Section 3.1.2 for each of the fourteen equities and also for their combined portfolio.

Analysis of variance tests would also be conducted where the p–values (F-probabilities) and adjusted R^2 values (% variance accounted for) were recorded.

4.1 Empirical Results

a. Descriptive Statistics

Descriptive statistics for the stock return variable (R) were generated using SAS. The results are as given in Table 4.1

Table 4.1: Basic Statistics for Stock Returns (R)

N	504	**Sum Weights**	504
Mean	13.8414985	**Sum Observations**	6976.11523
Std Deviation	24.8502237	**Variance**	617.53362
Skewness	2.02295891	**Kurtosis**	7.67539142
Uncorrected SS	407179.299	**Corrected SS**	310619.411
Coeff Variation	179.534201	**Std Error Mean**	1.10691694
Median	7.70254	**Range**	207.50000
Mode	0.00000	**Interquartile Range**	20.45076

From Table 4.1, the distribution of stock returns is not Normal since the values of mean, median and mode are not the same. The high standard deviation indicates a wide range of scores and generally speaking, less representative the mean becomes. The distribution is positvely skewed and leptokurtic.

A histogram for the Stock Returns (R) was also drawn in SAS as given in Figure 4.1

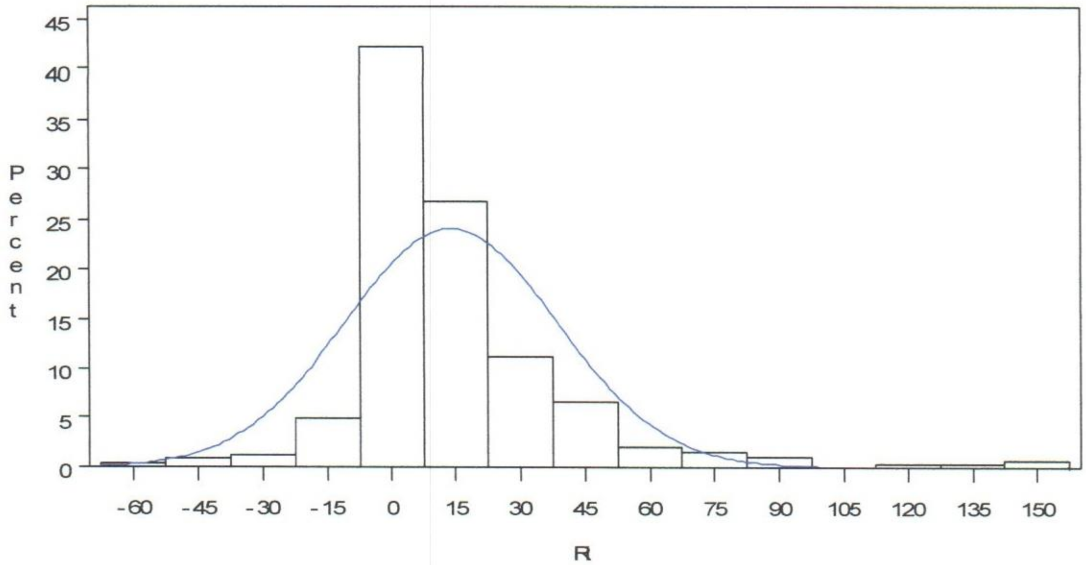

Figure 4.1: Histogram for Stock Returns

From Figure 4.1 the distribution of **R (Stock Returns)** appears to be skewed to the right and slightly peaked.

b. Profile Plots

Individual profile plots were generated for the various equity returns with an average trend line using 60 as the smoothing factor. The result is shown in Figure 4.2.

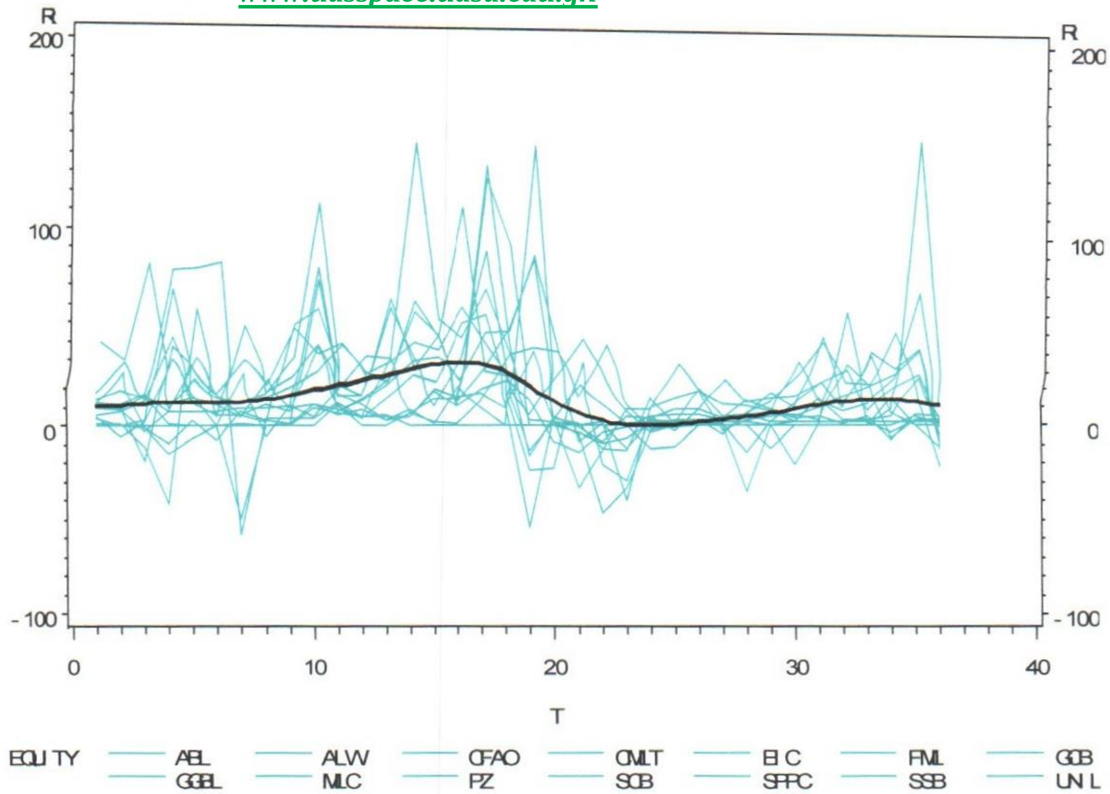

Figure 4.2: Individual Profiles of Ghana Stock Returns

From Figure 4.2, the Ghana stock returns, on the average, appears to increase gently within the first quarter of the study period (Jan 2001 – Dec 2002) and then increased slightly in the second quarter of the study period (Jan 2003 – Dec 2004). It then decreased slightly for about one and half years and rose slowly again to the end of the study period (Dec 2008).

The variability of Ghana stock returns at the beginning of the study seems low but then rose to its highest point around half way through the study period. It then dropped to its lowest and increased again to the end.

Table 4.3: Disaggregated figures of GSE All Share Index Rates (%)

Year	Quarter	Annual Rates	Quarterly Rates	Dividing by 4	Lisman Sandee	First Difference	Second Difference
2nd	1st	15.74	3.36	3.93		5.24	6.26
3rd	2nd		7.48	3.93		4.92	4.98
4th	3rd		4.61	3.93		4.27	3.76
	4th		0.29	3.93		3.31	2.73
2002	1st	11.02.	4.81	2.76	2.63	1.97	1.67
	2nd		3.69	2.76	1.47	2.03	2.11
	3rd		2.53	2.76	2.18	2.78	2.96
	4th		-0.01	2.76	4.74	4.23	4.27
2003	1st	40.26	6.49	10.07	6.53	6.14	4.28
	2nd		20.20	10.07	7.58	8.28	7.98
	3rd		7.11	10.07	10.59	11.13	11.95
	4th		6.46	10.07	15.55	14.69	16.40
2004	1st	107.27	17.80	26.82	22.51	23.95	23.28
	2nd		26.83	26.82	28.62	26.8	26.08
	3rd		17.20	26.82	29.84	28.24	28.26
	4th		45.43	26.82	26.26	28.27	29.65
2005	1st	80.27	59.43	20.07	24.42	25.11	25.58
	2nd		24.37	20.07	24.51	22.26	22.64
	3rd		-0.68	20.07	22.11	18.64	18.55
	4th		-2.85	20.07	11.23	14.26	13.49
2006	1st	-33.26	-5.07	-8.31		0.67	4.90
	2nd		-9.16	-8.31		-7.05	-3.68
	3rd		-16.79	-8.31		-12.20	-12.70
	4th		-2.24	-8.31		-14.78	-21.88
Kolmogorov – Smirnov's Test			**D (p-value)**	**0.2083 (0.6749)**	**0.3958 (0.0813)**	**0.2083 (0.686)**	**0.2083 (0.686)**
Mean Sq. Error						***3775.3***	***3525.6***

The Kolmogorov-Smirnov's test results were the same for the two least squared models so the mean squared errors were employed to arrive at the best fit. The "second difference" method again performed better.

Since the "second difference" model was consistently the best performers in the two tests conducted above, it is hereby recommended as the most adequate disaggregation method for analysts studying the Ghanaian Security System.

4.3 Cross Sectional Analysis

Cross Sectional Analysis was performed to identify the variables (in Section 3.1.2) that contributed significantly to the Stock Returns (R). Firstly, linear regression model with the constant term and the main effects of the variates (a linear regression with constant coefficient) were fitted to the stock returns of the selected equities and their combined portfolio. In the second test, the constant term was omitted from the model and the main effects (a linear regression without constant coefficient) maintained. The third test was conducted with the main effects as well as 1 level of interactions among the main effects (to check for possible associations within the explanatory variables).

The results of the three tests conducted are summarized in the Tables 4.4 to 4.6.

Table 4.4: Results of Linear Regression with Constant Coefficients (C)

Equity	F-pr	Adjusted R^2 (%)	Sig. fitted terms with t-pr ≤ 0.05	Sig. fitted terms with 0.05 < t-pr ≤ 0.10
ABL	0.826	0.0	-	-
ALW	0.289	5.6	Rmf	Rf
CFAO	0.080	17.7	PE, Rf	DE
CMLT	0.038	23.4	DE	-
EIC	0.001	42.3	Rmf, Rf	-
FML	0.124	14.0	Rmf	-
GCB	0.002	41.6	Rmf, S, C	DE
GGBL	<.001	44.7	Rmf	Rf
MLC	<.001	56.3	Rmf, Rf, DE, PE	BM, S
PZ	0.010	32.2	G, BM	-
SCB	0.260	6.8	Rmf, PE	S
SPPC	0.039	23.1	S, Rf, C	DE
SSB	<.001	44.7	Rmf	-
UNIL	0.004	37.1	-	DE, PE
Portfolio	**< .001**	**20.7**	**C, Rf, Rmf, S**	-

The model for the tests in Table 4.4 is in the form:

Response variate: R

Fitted terms: Constant + Rf + Rmf + Si + BM + G + DE + PE

(*See Appendix 2A for more details*)

From Table 4.4, the Adjusted R^2 values were generally low (less than 50%) indicating that the model does not fit the data perfectly. Furthermore, all the seven explanatory variables appeared significant at varying levels for the equities. Finally, the constant term (C) was significant in only three out of fifteen tests.

Table 4.5: Results of Linear Regression without Constant Coefficients

Equity	F-pr	Adjusted R^2 (%)	Sig. fitted terms with t-pr ≤ 0.05	Sig. fitted terms with 0.05 < t-pr ≤ 0.10
ABL	0.663	0.0	-	S
ALW	0.033	8.8	Rmf, Rf	-
CFAO	0.003	19.2	PE, Rf, S	DE
CMLT	<.001	25.7	DE, S	-
EIC	0.001	43.0	Rmf, Rf	-
FML	<.001	15.1	Rmf	Rf
GCB	<.001	33.1	Rmf, Rf	-
GGBL	<.001	46.6	Rmf, Rf	-
MLC	<.001	52.7	Rmf, Rf	PE
PZ	<.001	34.1	G, S, BM	Rf
SCB	<.001	0.8	Rmf	-
SPPC	0.065	12.4	DE, BM, S	Rf
SSB	<.001	45.6	Rmf	-
UNIL	<.001	39.1	S, DE	PE, Rmf
Portfolio	**<.001**	**18.7**	**Rf, Rmf, S, DE**	**-**

The model in Table 4.5 is in the form:

Response variate: R

Fitted terms: Rf + Rmf + S + BM + G + DE + PE

(*See Appendix 2B for more details*)

The F-pr (model rejection) values from Table 4.5 were significantly lower (better) than that of Table 4.4. This indicates that, there are no fixed returns for investments in Ghana but returns are directly related to one variable or the other. The variable Rmf, and/or Rf seem to influence most equities in Ghana, except CMLT and ABL.

Table 4.6: Results of Linear Regression with one level of interaction among the Main Effects

Equity	F- pr	Adjusted R^2	Sig. fitted terms with t-pr ≤ 0.05	Sig. fitted terms with 0.05<t-pr ≤ 0.10
ABL	0.307	24.5	S, G	BM, PE
ALW	<.001	90.0	-	Rf, BM
CFAO	0.029	58.7	-	Rmf
CMLT	0.066	42.3	-	-
EIC	0.165	18.1	-	-
FML	0.018	58.3	BM	-
GCB	0.045	47.7	-	-
GGBL	0.014	65.6	DE, PE, S, Rmf	BM
MLC	<.001	93.5	-	DE
PZ	0.077	39.9	-	-
SCB	0.009	47.6	Rmf	G
SPPC	0.013	72.9	Rmf, Rf, DE	-
SSB	0.059	46.6	-	-
UNIL	0.031	39.4	-	-
Portfolio	**<.001**	**27.4**	**Rf, Rmf, S, BM**	-

The model for stock returns in Table 4.6 will be:

Response variate: R

Fitted terms: Rf + Rmf + S + BM + G + DE + PE + Rf.Rmf + Rf.S + Rmf.S + Rf.BM + Rmf.BM + S.BM + Rf.G + Rmf.G + S.G + BM.G + Rf.DE

+ Rmf.DE + S.DE + BM.DE + G.DE + Rf.PE + Rmf.PE + S.PE + BM.PE + G.PE + DE.PE

(*See Appendix 2C for more details*)

From Tables 4.4, 4.5 and 4.6, the model with interactive effect (Table 4.6) explained the variations in stock returns best (from the Adjusted R^2 values) even though no variables in themselves influenced stock returns significantly.

Generally, for ABL, the stock returns were not based on any of the covariates. Indeed the linear regressions could not explain the returns significantly. Also, all the seven explanatory variables appeared significant at varying levels for the various equities. This phenomenon is good for risk diversification and portfolio selection and formation. Finally, the adjusted R^2 values were generally low (especially in Tables 4.4 and 4.5) and that called for investigation of the distribution of the errors (residuals) in the next section.

4.4 Distribution of Residuals

The residuals observed from the general linear model (Portfolio of the 14 equities) were fitted to various probability distribution functions to investigate the best fit. Initially, probability-probability (P-P) plots, histograms and cumulative density fucnctions (CDF) curves were fitted to the residuals for visual accessments. In all 36 probability distribution functions (PDF) were considered. Since the residuals contained some negative values, only functions which had domains ranging from negative infinity to positive infinity were considered. Finally, the fitted distributions were fitted with the residuals using Kolmogorov–Smirnov, Anderson–Darling and the Chi–squared goodness of fit tests. The software used in the tests was the *Easy Fit 5.5 Professional by Mathwave Technologies.*

4.4.1 Exploratory Analysis of Residuals

a. The P-P plots show graphs of the empirical (residuals) CDF values plotted against the theoretical CDF values. It is used to determine how well a specific distribution fits the observed data. This plot will be approximately linear if the specified theoretical distribution is the correct model. *Easy Fit* displays the reference diagonal line along which the graph points should fall. Figure 4.3 shows the P-P plots of nine out of the 36 distributions that were considered.

b. In the probability density plot, the empirical PDF is displayed as a histogram consisting of equal-width vertical bars (bins), each representing the number of sample data values (falling into the corresponding interval), divided by the total number of data points. The theoretical PDF is displayed as a continuous curve properly scaled depending on the number of intervals. Figure 4.4 shows the PDF plots for nine out of the 36 fitted distributions.

c. Finally the cumulative distribution plots showing the empirical CDFs and the fitted theoretical CDFs are displayed in Figure 4.5.

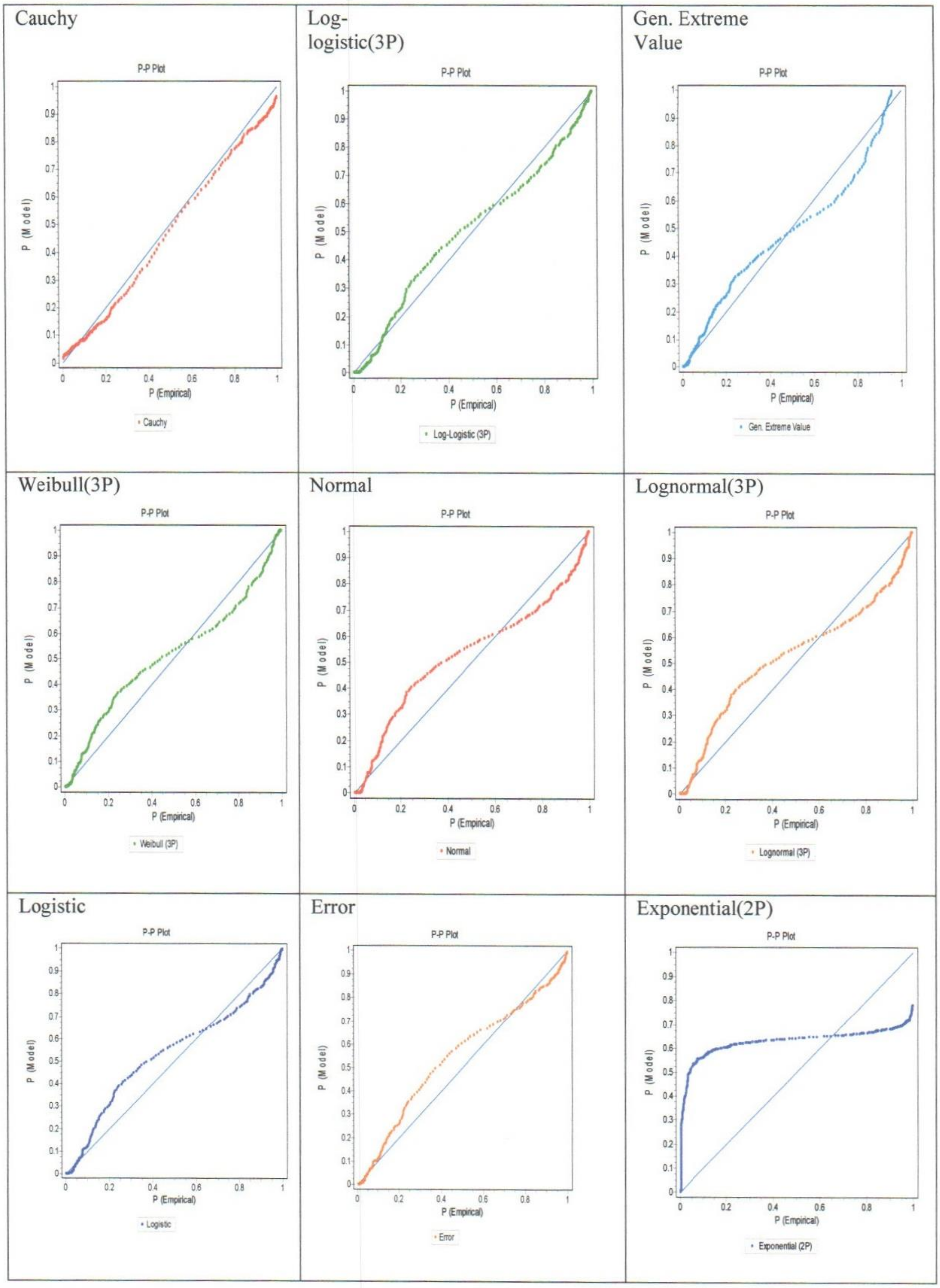

Figure 4.3: Probability–Probability Plots of Residuals

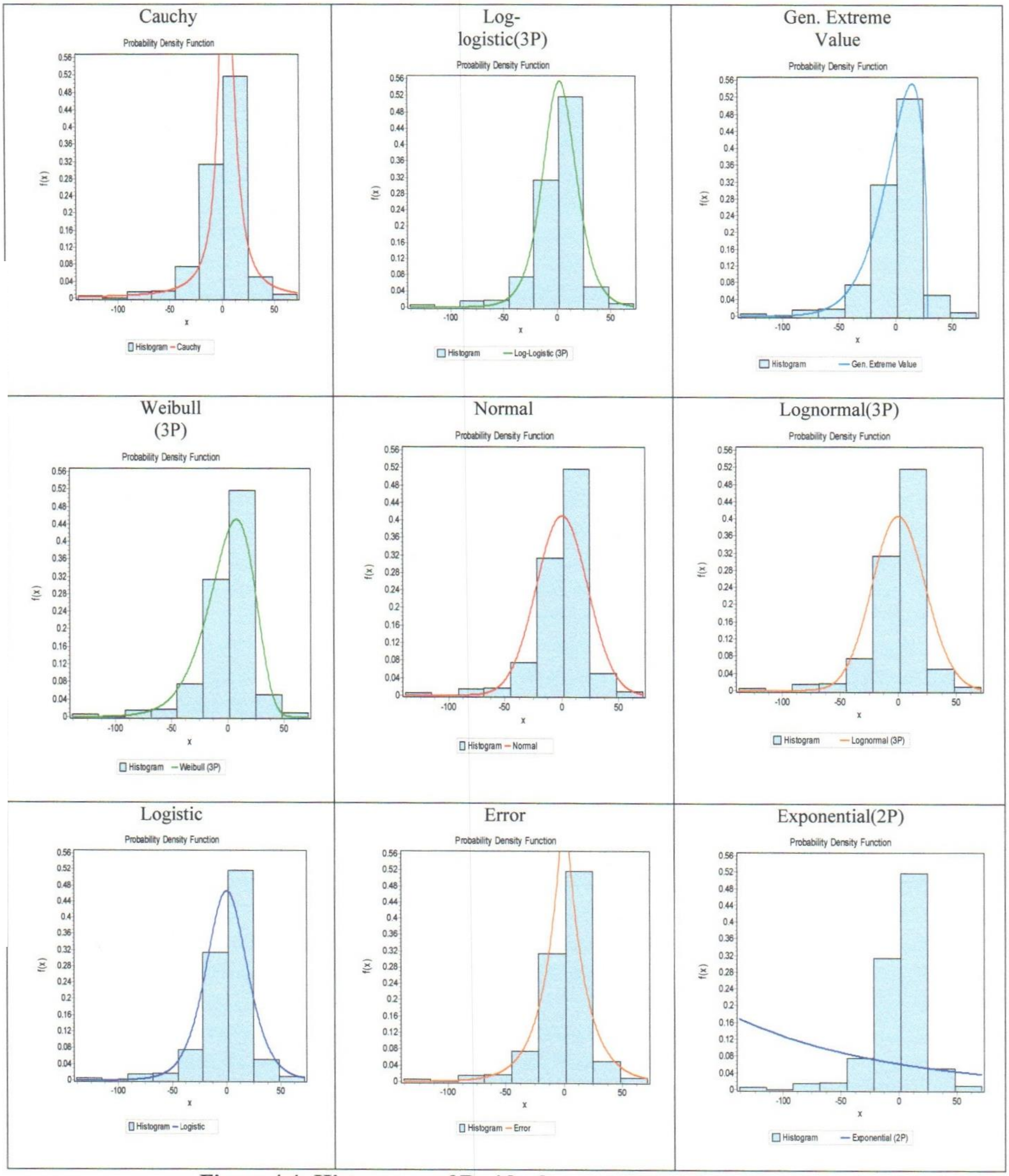

Figure 4.4: Histograms of Residuals

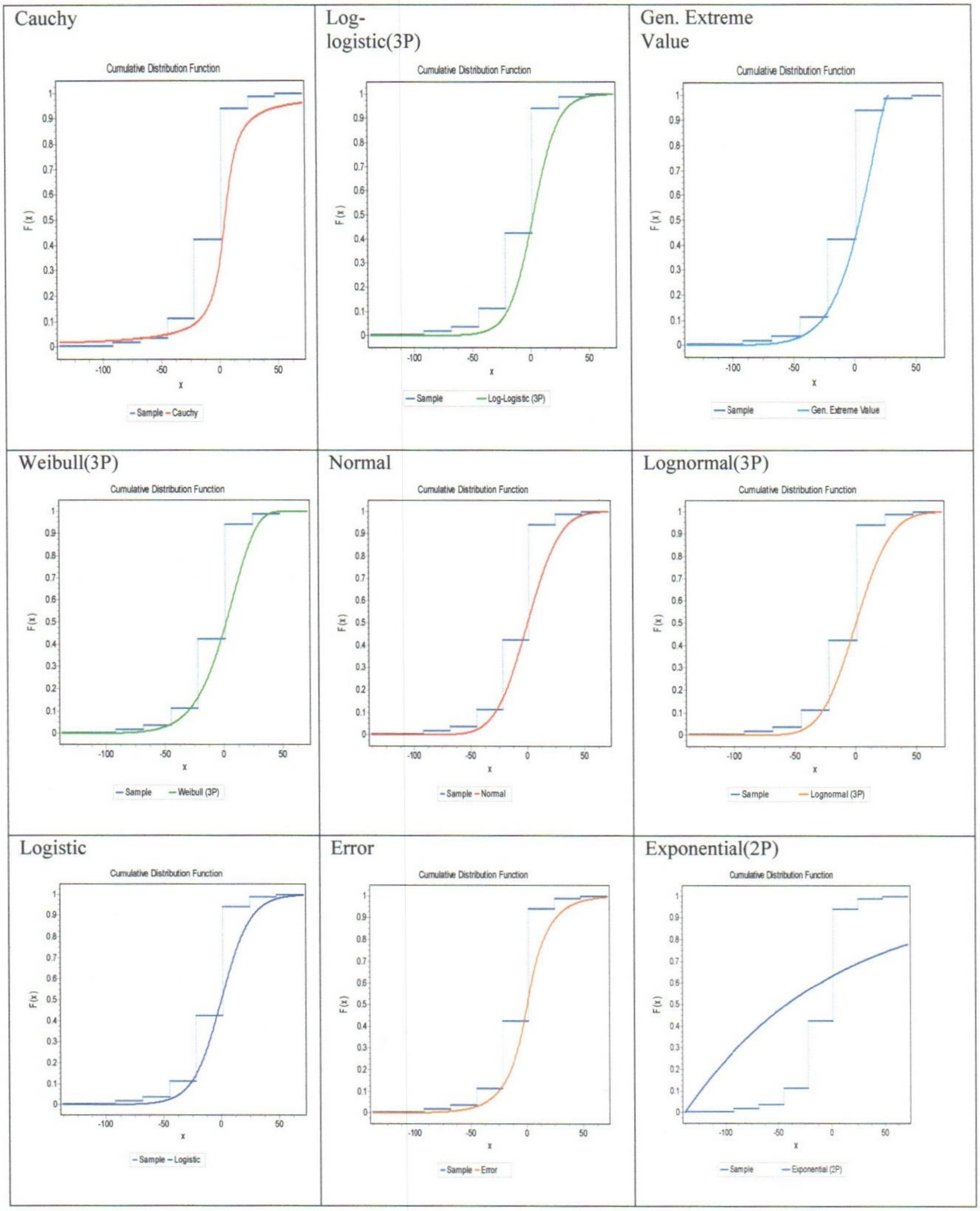

Figure 4.5: Cumulative Density Functions of Residuals

4.4.2 Comments on the Exploratory Analysis

It is quite clear from the P-P plots in Figure 4.3 that the best fitted distribution (closest to the diagonal line) of the stock returns residuals is the Cauchy distribution and not the presumed normal distribution. With the histograms and the CDF plots in Figures 4.4 and 4.5 respectively, much distinctions can not be made visually on the various functions, except for the fact that the Exponential (2P) distribution could not fit the data well.

4.4.3 Goodness of Fit Tests

The goodness of fit (GOF) test measures the compatibility of a random sample with a theoretical probability distribution function. In other words, these tests show how well the distribution you selected fits to your data.

The fitted results of our residuals are presented in Table 4.7 and distributional parameters for the fittings are given in Table 4.8.

Table 4.7a: Summary of Goodness of Fit Tests

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Cauchy	0.06199	1	4.603	1	19.159	1
2	Log-Logistic (3P)	0.0786	2	7.9545	2	63.981	2
3	Gen. Extreme Value	0.11257	3	89.1	27	N/A	
4	Johnson SU	0.11273	4	8.7613	3	68.115	3
5	Gumbel Min	0.11348	5	11.884	4	99.899	6
6	Weibull (3P)	0.12345	6	13.336	7	105.15	8
7	Kumaraswamy	0.12463	7	13.602	8	105.47	9
8	Laplace	0.12644	8	12.357	5	78.731	5
9	Error	0.12644	9	12.357	6	78.731	4
10	Beta	0.13257	10	15.875	10	118.57	11
11	Hypersecant	0.13845	11	13.993	9	102.3	7
12	Logistic	0.14843	12	16.132	11	112.81	10

Table 4.7b: Summary of Goodness of Fit Tests (continuation)

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
13	Lognormal (3P)	0.15597	13	21.09	16	153.26	17
14	Pearson 6 (4P)	0.15657	14	20.376	13	143.87	13
15	Fatigue Life (3P)	0.15839	15	20.743	14	154.49	18
16	Dagum (4P)	0.15944	16	20.274	12	134.5	12
17	Normal	0.16057	17	20.875	15	150.6	14
18	Inv. Gaussian (3P)	0.16149	18	21.093	17	150.63	15
19	Error Function	0.16632	19	21.547	18	150.93	16
20	Gen. Gamma (4P)	0.17153	20	23.342	19	164.68	19
21	Erlang (3P)	0.1717	21	24.929	20	188.48	20
22	Gen. Pareto	0.17301	22	179.94	33	N/A	
23	Gamma (3P)	0.17965	23	26.577	21	195.5	22
24	Chi-Squared (2P)	0.18449	24	29.177	23	227.55	23
25	Pearson 5 (3P)	0.18766	25	27.035	22	194.36	21
26	Uniform	0.18908	26	125.05	30	N/A	
29	Frechet (3P)	0.23412	29	76.573	26	N/A	
31	Power Function	0.38231	31	106.19	29	906.17	27
32	Rayleigh (2P)	0.39746	32	129.43	31	1406.0	28
33	Exponential (2P)	0.47384	33	174.05	32	2753.6	30
34	Student's t	0.49495	34	584.15	36	1479.7	29
35	Levy (2P)	0.58386	35	205.18	34	4298.2	31
36	Burr (4P)	0.60959	36	247.36	35	770.98	26

Since the Cauchy distribution ranked first in all the three goodness of fit tests it would be necessary to model the residuals of the Ghana stock returns with it.

Table 4.8: Parameters For the Fitted Distributions

#	Distribution	Parameters
3	Cauchy	σ=7.8045 μ=4.5821
7	Error	k=1.0 σ=22.633 μ=0.33469
8	Error Function	h=0.03124
9	Exponential (2P)	λ=0.00724 γ=-137.79
13	Gen. Extreme Value	k=-0.7042 σ=21.848 μ=-2.475
24	Log-Logistic (3P)	α=1.3372E+9 β=1.4006E+10 γ=-1.4006E+10
25	Logistic	σ=12.478 μ=0.33469
26	Lognormal (3P)	σ=0.01745 μ=7.1772 γ=-1308.5
27	Normal	σ=22.633 μ=0.33469
28	Pearson 5 (3P)	α=824.06 β=5.6408E+5 γ=-685.84
35	Uniform	a=-38.866 b=39.535
36	Weibull (3P)	α=15.533 β=295.12 γ=-285.79

4.4.4 Determination of the Residual Distribution

A random variable x is Cauchy distributed with parameters $\sigma > 0$ (continuous scale parameter) and μ (continuous location parameter), where $-\infty < x < \infty$ has a probability density function:

$$f(x) = \left(\pi\sigma\left(1 + \left(\frac{x-\mu}{\sigma}\right)^2\right)\right)^{-1} \tag{4.1}$$

It has a cumulative distribution function

$$F(x) = \frac{1}{\pi}\arctan\left(\frac{x-\mu}{\sigma}\right) + 0.5 \tag{4.2}$$

Equation (4.1) could be written as

$$f(x) = \frac{\sigma}{\pi\left(\sigma^2 + (x-\mu)^2\right)} \tag{4.3}$$

Substituting the empirical estimates of the Cauchy parameters from Table 4.8 into equation (4.3) we obtain

$$f(x) = \frac{7.8045}{\pi\left(7.8045^2 + (x-4.5821)^2\right)} \tag{4.4}$$

Let $a = \frac{7.8045}{\pi}$, $b = 7.8045$, and $c = 4.5821$, then equation (4.4) will be simplified as

$$f(x) = \frac{a}{b^2 + (x-c)^2} \tag{4.5}$$

Since a, b and c are known constants, we conclude that the returns on stock i at time t will have a residual density function

$$f(x) = \frac{a}{b^2 + (x-c)^2}, \quad -\infty < x < \infty \tag{4.6}$$

4.5.5 Validation of Test Results

In order to preserve unbiasedness of our estimated Cauchy location and scale parameters, they are validated with the jack-knifing method as described in Section 3.3.4.

The results obtained by the jack-knifing procedure are as tabulated in the Table 4.9:

Table 4.9: Jack-knife validated results of Cauchy Parameters

Deleted (knifed) Equity	**Scale (Sigma)**	**Location (Mu)**
	7.8045	**4.5812**
ABL	7.8944	3.8697
ALW	7.6304	4.7762
CFAO	7.6597	4.0348
CMLT	7.7615	4.7048
EIC	8.0759	4.7484
FML	7.8907	4.7659
GCB	7.6194	4.7908
GGBL	7.8547	4.8961
MLC	7.9786	4.0457
PZ	7.7123	4.4915
SCB	7.6658	5.1600
SPPC	7.9097	4.2514
SSB	7.7540	4.8255
UNIL	7.7582	4.9365

The mean and standard error of the estimates in Table 4.9 were:

	Sigma	**Mu**
Mean	**7.7947**	**4.5812**
S.E	0.1375	0.3896

Thus the parameter estimates were good, since they fell within the expected ranges:

$$7.5712 \le \sigma \le 7.9322 \quad \text{and} \quad 4.1916 \le \mu \le 4.9709 .$$

4.5 Longitudinal Analyses

The four main procedures undertaken here were:

1. evaluating and selecting the best variance covariance structure,
2. fitting the random effects model,
3. fitting the mixed model and
4. incorporating Cauchy distribution into the model.

4.5.1 Evaluation of Covariance Structure

The various covariance structures described in Sections 3.5.5 and 3.5.6 were used to fit random intercept and slope models in SAS. The intra-class correlations coefficients (ICC) were then computed and the Akaike information criteria (AIC) values recorded. The results of the two tests are given in Tables 4.10 and 4.11.

Table 4.10: Intra Class Correlation Coefficients for Fitted Covariance Strctures

COVARIANCE STRUCTURES		INTRA CLASS CORRELATION COEF (%)
1.	Variance Componenet (VC)	1.74
2.	Unstructured (UN)	No convergence
3.	Compound Symmetry (CS)	12.28
4.	Autoregressive [AR(1)]	13.36
5.	Banded Toeplitz (2) [TOEP(2)]	28.63
6.	TOEP(3)	33.97
7.	TOEP(4)	37.52
8.	TOEP(5)	48.94
9.	TOEP(6)	50.95
10.	TOEP(7)	42.41
11.	ARMA(1,1)	12.99
12.	Spatial Power (SP)	13.36
13.	Spatial Gaussian (SG)	12.06

The % ICC values in Table 4.10 indicate the percentage of the variation in the data was explained by the various models. Thus, the bigger the ICC value the better. Therefore, the best model was TOEP(6), followed by TOEP(5), TOEP(7), TOEP(4), and so on.

Table 4.11: Akaike information criteria (AIC) for the fitted Covariance Structures

Covariance Structres		AIC for Random Intercept and Slope
1.	VC	4667.0
2.	UN	-
3.	CS	4669.0
4.	AR(1)	4614.7
5.	TOEP(2)	4622.2
6.	TOEP(3)	4620.5
7.	TOEP(4)	4611.1
8.	TOEP(5)	4608.7
9.	TOEP(6)	4610.2
10.	TOEP(7)	4607.2
11.	ARMA(1,1)	4611.8
12.	SP	4614.7
13.	SG	4621.7

In Table 4.11, the best model had the lowest AIC value. Therefore, the best model was TOEP(7), followed by TOEP(5), TOEP(6), TOEP(4), and so on.

Toeplitz (5) is thus finally chosen as the best covariance structure for the Ghana Stock returns data, due to its consistent optimum results.

4.5.2 Fitting the Random Effects

The unstructured banded Toeplitz (5) covariance structure was used to fit the random time and intercept model in SAS. The REML estimation method was used. The results are presented in Table 4.12.

Table 4.12: Results for the Random Estimates

Covariance Parameter Estimates

Cov Parm	Subject	Estimate
UN(1,1)	Equity	3.72E-16
UN(2,1)	Equity	-0.4311
UN(2,2)	Equity	0
TOEP(2)	Equity	114.10
TOEP(3)	Equity	42.0512
TOEP(4)	Equity	62.8503
TOEP(5)	Equity	24.7392
Residual		510.06

Row	Effect	Equity	Col1	Col2
1	Intercept		3.72E-16	-0.4311
2	T		-0.4311	0

From Table 4.12, the estimates for random intercept and slope by the stock returns (as explained at Section 3.4.2) are 3.72E-16 and -0.4311 respectively.

4.5.3 Evaluating Fixed Effects

The banded Toeplitz (5) covariance structure was used to fit the full model with all of the main effects, the time by main effect interactions, and the quadratic and cubic effects for time. The ML estimation method was used. The outputs are displayed in Table 4.13.

Table 4.13: Initial Results of the Fixed Effects Model

Solution for Fixed Effects

Effect	Estimate	Error	DF	t Value	Pr > |t|
Intercept	-79.8148	30.5603	121	-2.61	0.0101
T	4.6036	2.7660	203	1.66	0.0976
Rf	0.7689	0.5394	368	1.43	0.1548
Rmf	0.9133	0.2870	425	3.18	0.0016
S	6.1466	1.6516	73.3	3.72	0.0004
BM	0.000378	1.8281	99.7	0.00	0.9998
G	0.02423	0.01363	464	1.78	0.0760
PE	-0.07657	0.1086	288	-0.71	0.4813
DE	3.3251	6.5572	135	0.51	0.6129
T*Rf	-0.02204	0.03534	319	-0.62	0.5333
T*Rmf	-0.01849	0.01432	428	-1.29	0.1972
T*S	-0.1774	0.07935	87.4	-2.23	0.0280
T*BM	-0.05925	0.09574	160	-0.62	0.5369
T*G	-0.00075	0.000474	449	-1.57	0.1169
T*PE	0.003855	0.004146	328	0.93	0.3532
T*DE	-0.1417	0.2193	132	-0.65	0.5195
T*T	-0.1810	0.1393	185	-1.30	0.1954
T*T*T	0.003693	0.002533	179	1.46	0.1466

It is to be observed that at a significance level of 0.05 only S, Rmf and T*S were of significant effects. Therefore, the effect which was least significant in the model was first removed. Although in normal practice all the main effects are saved, BM had a p-value of 0.9998 which is simply unacceptable.

The initial AIC value at this point was 4562.7. We always checked the AIC values were always checked and the poor effects eliminated, to ensure that a better (smaller) value was obtained. The elimination process was continued until the best results was obtained.

Table 4.14: Final Solution for Fixed Effects

Fit Statistics

-2 Log Likelihood	4523.1
AIC (smaller is better)	4549.1
AICC (smaller is better)	4549.9
BIC (smaller is better)	4557.5

Solution for Fixed Effects

Effect	Estimate	Error	DF	t Value	Pr > |t|
Intercept	-42.5137	12.9538	80.4	-3.28	0.0015
Rf	0.7893	0.2003	212	3.94	0.0001
Rmf	0.9052	0.1523	224	5.94	<.0001
S	3.8037	0.8348	73.4	4.56	<.0001
G	0.02584	0.01362	462	1.90	0.0583
Rmf*T	-0.01563	0.006806	303	-2.30	0.0224
S*T	-0.04301	0.01324	133	-3.25	0.0015
G*T	-0.00080	0.000473	452	-1.69	0.0910

The results show that the factors that contribute significantly are T-bill rates (Rf), equity's risk factor (Rmf), size of equity (Si) and the growth in earnings (G). All these factors had positive relationships to the stock returns. Finally, time (T) itself did not have a significant relationship but had a negative interactive effect with Rmf, Si and slightly G on the Ghana stock returns.

4.5.4 Incorporating Cauchy Distribution into the Model

From Section 4.4 it is clear that the distribution of the residuals of the Ghana stock returns was Cauchy and not Normal. This inconsistency needed to be resolved. Two main approaches were adopted to do so.

Firstly, the method of Generalized Linear Mixed Models (GLMM) which is a distribution free model could be used. Secondly, after obtaining the results based on normal distribution assumptions, SAS programming software allows the incorporation of the residual function into the original algorithm to re-evaluate the coefficients of the estimated parameters.

The table below shows the results of our reevaluated estimates.

(The estimates as given in Table 4.14 above were: Intercept (beta0) = - 42.52; Rf (beta1) = 0.79; Rmf (beta2) = 0.91; S (beta3) = 3.81; G (beta4) = 0.026; Rmf*T (beta5) = - 0.016; S*T (beta6) = -0.043; G*T (beta7) = -0.0008. Also the estimates for the Cauchy parameters were; scale (sigma) = 7.8045 and location (theta) = 4.5821)

Table 4.15: SAS Reevaluated Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Gradient
beta0	-26.3358	3.9519	504	-6.66	<.0001	0.05	0.000052
beta1	0.3671	0.07962	504	4.61	<.0001	0.05	0.001568
beta2	0.4133	0.07659	504	5.40	<.0001	0.05	-0.00201
beta3	2.4290	0.2669	504	9.10	<.0001	0.05	0.000795
beta4	0.009871	0.005327	504	1.85	0.0645	0.05	-0.03099
beta5	-0.00424	0.003152	504	-1.34	0.1796	0.05	0.018292
beta6	-0.02661	0.005461	504	-4.87	<.0001	0.05	-0.01435
beta7	-0.00032	0.000169	504	-1.87	0.0618	0.05	-0.85276
Sigma	6.3798	0.4179	504	15.27	<.0001	0.05	-0.00001
Theta	4.5821	0	504	Infty	<.0001	0.05	0

Since the p–value for beta5 was not acceptable at $\alpha = 0.05$ it was removed from the model and the test performed again. The final optimum results are given in Table 4.16.

Table 4.16: Final Reevaluated Estimates

Parameter	Estimate	Standard Error	DF	t Value	Pr > |t|	Alpha	Gradient
beta0	-26.6462	3.9518	504	-6.74	<.0001	0.05	-0.00026
beta1	0.3403	0.07591	504	4.48	<.0001	0.05	-0.00815
beta2	0.3306	0.03914	504	8.45	<.0001	0.05	0.006343
beta3	2.3631	0.2637	504	8.96	<.0001	0.05	-0.00382
beta4	0.01042	0.005180	504	2.01	0.0448	0.05	-0.03
beta6	-0.02210	0.004474	504	-4.94	<.0001	0.05	-0.01272
beta7	-0.00034	0.000164	504	-2.05	0.0408	0.05	-1.25624
Sigma	6.4105	0.4159	504	15.41	<.0001	0.05	0.000299
Theta	4.5821	0	504	Infty	<.0001	0.05	0

The p – values in Table 4.16 indicate that all the variables are significant at $\alpha = 0.05$ which concludes the computations.

4.5.5 Proposed Linear Mixed Model

The proposed mixed model for the returns on equity (company, stock) i listed at the Ghana Stock Exchange based on quarterly data collected from fourteen equities from the period January 2000 to December 2008 is:

$$Ri_{ij} = \beta_0 + \beta_1 Rf + \beta_2 Rmf + \beta_3 Si + \beta_4 G + (\beta_5 Si + \beta_6 G)t_{ij} + b_{0i} + b_{1i} + \varepsilon_{ij}$$

with $i = 1, \ldots, 14$ are the equities and $j = 1, \ldots, 36$ representing the time period.

The distributions are as follows:

$$\hat{\beta} \sim N(\beta,\ \Sigma_k).$$

Thus the β estimates are multivariate normally distributed with mean β and a variance - covariance structure Σ.

$$\hat{b}_i \sim N\left(0,\ \sigma_{|i-j|} + 1^{1(|i-j|<5)}\right).$$

Thus the b_i estimates are multivariate normally distributed with mean zero and a banded Toeplitz (5) covariance structure.

$$\varepsilon_{ij} \sim Cauchy(4.5821,\ 6.4105).$$

Thus the residuals are Cauchy distributed with location parameter 4.5821 and scale parameter 6.4105.

4.6 Conclusion

Various empirical analyses were performed in this Chapter and the results obtained included: an adequate disaggregation method for financial data; significant explanatory variables for Ghana stock returns; a consistent probability distribution function for the residuals of Ghana stock returns; and a proposed linear mixed model for the returns on equities listed at the Ghana Stock Exchange.

A detailed summary and conclusions are presented in the next Chapter.

CHAPTER FIVE

SUMMARY AND CONCLUSIONS

5.0 Introduction

We shall now discuss the final empirical findings of our research and make recommendations for financial analysts, investors and further researchers in this chapter.

5.1 Summary

- **Disaggregation of Stock Data**

Although this task was not originally part of the objectives, it became an essential part before proceeding. Many procedures have been described in the literature but they needed to be tested for and the most adequate one to be selected for the purpose of this research.

The model based procedures, specifically the ***least square method of second differencing*** was found to be the most adequate disaggregation method for breaking down annual data into quarterly figures. This was evident from the Kolmogorov–Smirnov test and also the least mean squared error results which were obtained when we analyzed Bank of Ghana Treasury bill rates and the returns on the All Share Index.

- **Cross Sectional Analyses**

The two main objectives were to identify the economic variables that contribute significantly to the expected returns of individual equities listed on the Ghana Stock Exchange and they develop a general linear model for the Ghanaian market. These objectives were achieved by performing cross sectional analyses.

Although thirty-five equities were listed at the Ghana Stock Exchange as of December 2008, only fourteen were sampled for this research, since these were all the equities which had traded consistently throughout the study period of January 2000 to December 2008. The qualified equities were: ABL, ALW, CFAO, CMLT, EIC, FML, GCB, GGBL, MLC, PZ, SCB, SPPC, SSB and UNIL.

Some economic factors were found in the literature to contribute significantly to stock returns (R). They include: return on a risk free investment (Rf), systemic risk component (Rmf), size of equity (S), book to market ratio of equity (BM), growth in earnings of equity (G), price to equity ratio (PE) and the debt to equity ratio (DE). So it became necessary to find out which of these factors affected the returns on the equities and also the market as a whole. The results were summarized in tables 4.4 and 4.5. For ABL, only it's size (S) affected it slightly indicating that any time more ABL shares were floated or the stock price increased, there was an increase in its returns, S being the product of number of outstanding share and the price per share. Furthermore, as the size decreased the returns on ABL stock decreased. It is important to note here that there might be other significant factors which may not have been included in this research.

The significant factors for ALW, EIC, FML and GGBL were Rf and Rmf. These follow the capital asset pricing model (CAPM) and strictly indicate that the CAPM should not be ignored totally. This means investors have to consider the movements of both the riskless rates (T-Bills) as well as the markets systematic risk.

The returns on investing in CFAO depended significantly on Rf, S, PE, and DE at various levels. The investor therefore should pay attention to the movements in Treasury bill rates, the firm's size, the price to earnings ratio, which gave the proportion of dividends to the stock price and finally the debt equity ratio, which represented the proportion of the firm's liabilities to assets, in order to achieve an optimum returns.

In the case of the returns of CMLT, the significant factors were DE, and S.

PZ got G, BM, S and Rf as the significantly contributing factors to its returns. This meant that the movements of: growth of earnings, which is the firm's percentage rate of change of dividends; the equities ratio of book to market value; the firm's size; and Treasury bill rates should be considered for optimum investments.

MLC got the highest number of significant factors. Apart from G, all the other factors contributed significantly. The significant factors for the other equities were given as follows: SCB had Rmf, PE and S; SPPC got S, Rf and DE; SSB had only Rmf; and finally UNIL got PE, DE, S and Rmf.

In the case of the Ghanaian market, when all the equities were combined into a portfolio, the significant factors for the returns were Rf, Rmf, S and DE. Hence the general linear model proposed for the returns on investments in the Ghanaian Market is:

$$Ri_{ij} = \beta_0 + \beta_1 Rf + \beta_2 Rmf + \beta_3 Si + \beta_4 DE + \varepsilon_{ij} \quad (5.1)$$

with $i = 1, \ldots, 14$ being the equities and $j = 1, \ldots, 36$ representing the time period.

- **Distribution Fitting Analyses**

The third objective for this research was to describe the behaviour of the residuals of stock returns. This was achieved by fitting an adequate probability distribution function to the residuals. This analysis became necessary because, although the central limit theorem in probability theory justified the approximation of large sampled statistics to the normal distribution, in controlled experiments, using the specified distribution of the residuals eliminated biased regression coefficients.

With the aid of the statistical software *Easy Fit 5.5 Professional by Mathwave Technologies*, the residuals obtained (equation (5.1)) were fitted to 36 probability distrbution functions. Exploratory analyses (P-P plots, histograms and CDF curves) were done and goodness of fit tests (Kolmogorov-Smirnov test, Anderson-Darlings test, Chi-squared test) perfomed. In all the analyses and tests performed, the ***cauchy distribution*** ranked first. The cauchy distribution parameter estimates were then validated using the jack-knifing procedure.

- **Longitudinal Analyses**

As a final objective, the changes of the stock returns over time were modelled. Longitudinal models have been found to be the most adequate in this respect, since it modelled both fixed and random effects.

In selecting the variance-covariance structure for the random effects the intraclass correlation coefficients (ICC) and the Akaike information criteria (AIC) for various structures in the literature were evaluated and it was found that the ***Banded Toeplitz (5)*** was the most adequate for the Ghana stock returns. The mixed model was then fitted and it was found that, the factors that contributed significantly to Ghana stock returns were Treasury bill rates (Rf), systematic risk factor (Rmf), size of equity (S) and the growth in earnings (G). All these factors had positive relationships to the stock returns. Furthermore, time (T) itself did not have a significant relationship but had a negative interactive effect with Rmf, S and slightly with G on the Ghana stock returns. Finally the significant parameter estimates were re-evaluated by incorporating Cauchy distribution as the residual distribution to obtain the final unbiased results. The final parsimonious model for the Ghana Stock Returns was:

$$Ri_{ij} = \beta_0 + \beta_1 Rf + \beta_2 Rmf + \beta_3 Si + \beta_4 G + \left(\beta_5 Si + \beta_6 G\right)t_{ij} + b_{0i} + b_{1i} + \varepsilon_{ij}$$

with $i = 1, \ldots, 14$ being the equities and $j = 1, \ldots, 36$ representing the time period and with distributions as specified in Section 4.5.5.

5.2 Conclusion

5.2.1 Contribution to Knowledge

A method for disaggregating annual into quarterly stock data has been identified. Furthermore, a general linear model for monitoring movement of returns has been developed. These have positive implications for national economic planning.

5.2.2 Recommendations

On the basis of the findings noted in the preceding Sections, the following recommendations are made:

- Whenever the need arises that quarterly figures are required for any economic study (eg. GDP, Inflation, etc) especially in Ghana and only annual data was available, then the least square method of disaggregation with second differencing is highly recommended.

- Investors in Ghana Stock Exchange are advised to consider the movements in the following economic factors for optimum returns (% profit) in their investments: the Treasury bill rates, systemic risk factors, the size of firms and the growth of earnings.

- In the case of individual equities, investors are to look out for the movements in the various economic factors that contributed significantly to each firm as discussed in Section 5.1.

- Financial analysts and researchers are advised to incorporate the Cauchy distribution as the distribution of the residuals with scale parameter 7.8045 and location 4.5812 instead of the normal distribution, any time the response variable happens to be the returns on investment.

- The results of ABL indicate that there might be some other significant explanatory variables. This issue is open for further research.

5.2.3 Future Work

The low Adjusted R^2 values obtained in Tables 4.4 and 4.5 call for further investigation of other significant factors that may contribute to the returns of stock.

In order to compare the behaviour of stock returns in various developing economies, Spatio-Temporal (involving space and time) models of stock returns in emerging markets could be done.

REFERENCES

Adjasi, C, Harvey S. K. and Agyapong D. 2008, "Effect of Exchange Rate Volatility on the Ghana Stock Exchange." *African Journal of Accounting, Economics, Finance and Banking Research*. v. 3(2), pp. 28-47.

Arnott, R.D. 1993. "The CAPM and equity management, asset allocation, and performance measurement in The CAPM controversy: Policy and strategy implications for investment management". *ICFA Continuing Education*, edited by D.R. Harrington and R.A. Korajczyk. AIMR, New York, pp.16-24.

Aliyu, S. U. R. 2011, "Does Inflation has an Impact on Stock Return and Volatility? Evidence from Nigeria and Ghana" *MPRA Paper 20091*.

Banz, R. 1981, "The Relationship between Return and Market Value of Common Stocks," *Journal of Financial Economics*, v.9, pp. 3-18.

Basu, S. 1977, 'Investment Performance of Common Stocks in Relation to their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, v.32, pp.663-682.

Basu, S. 1983, "The Relationship Between Earnings Yield, Market Value and Return for NYSE Common Stocks: Further Evidence," *Journal of Financial Economics*, v.l2, pp.129-156.

Bauwens, L., Laurent S. and Rombouts J. V. K. 2006, "Multivariate GARCH Models: A Survey." *Journal of Applied Economics*. v. 21, pp. 79-109.

Bayezid Ali, M 2011, "Impact of Micro and Macroeconomic variables on Emerging Stock Market Returns: A Case on Dhaka Stock Exchange (DSE)," *Interdisciplinary Journal of Research in Business*, v.1, Issue 5, pp 8-16.

Bhandari, L.C. 1988, "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence," *Journal of Finance* v. 43, pp. 507-528.

Black, F. 1976, "Studies of Stock Market Volatility Changes." *1976 Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.

Black, F., Jensen M. and Scholes M. 1972, *The Capital Asset Pricing Model: Some Empirical Test and Studies in the Theory of Capital Markets*, ed., Michael C. Jensen. New York, Praeger Publishers.

Bollerslev, T. 1987, "A Conditionally Heteroskedastic Time Series Model for Speculative Price and Rates of Return." *The Review of Economics and Statistics*, v. 69(3), pp. 542-547.

Boot J., Feibes W. and Lisman J. 1967, "Further Methods of Derivation of Quarterly Figures from Annual data". *Applied Statistics*, v. 16, pp. 65 – 75.

Chan, K. C. 1988, "On the Contrarian Investment Strategy." *Journal of Business*, v. 61, pp. 147-163.

Chan, L., Hamao Y. and Lakonishok J. 1991, "Fundamentals and Stock Returns in Japan," *Journal of Finance*, v. 46, pp. 1739-1789.

Chien-Ting, L. 1999, *Misspecification of CAPM: Implication for size and book-to-market effects*. PhD Dissertation, Texas Tech University, Texas.

Chow, G. and Lin A. 1971, "Best Linear Unbiased Interpolation, Distribution, and Extrapolation of Time Series by Related Series": *Review of Economics and Statistics* 53, S. 373 – 375

Cook, R. D. and Weisberg S. 1982, *Residuals and Influence in Regression*, Chapman and Hall Ltd., London

Diggle, P.J., Liang, K., and Zeger, S.L. 1994, *Analysis of Longitudinal Data*, Oxford University Press, New York.

Engle, R. 2001, "GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics." *Journal of Econometric Perspectives*, v. 15(4), pp. 157-168.

Fama, E.F., and Macbeth J. 1973,"Risk, Return, Equilibrium: Empirical Tests,"' *Journal of Political Economy*, pp. 607-636.

Fama, E.F., and Macbeth J. 1974,"Test of the Multiperiod Two-Parameter Model." *Journal of financial Economics*, v.l, pp. 43-66.

Fama, E.F., and French K. R. 1992,"The Cross-Section of Expected Stock Returns," *Journal of Finance*, v. 47, pp. 427-465.

Fama, E.F., and French K. R. 1995,"Size and Book-to-Market Factors in Earnings and Returns," *Journal of Finance*, v. 50, pp. 131-155.

Fávero, L. P. and Belfiore P. 2011, "Cash Flow, Earnings Ratio and Stock Returns in Emerging Global Regions: Evidence from Longitudinal Data," *Global Economy and Finance Journal*, v. 4, pp. 32-43.

Fernandez, R. 1981, "A Methodological Note on the Estimation of Time Series": *Review of Economics and Statistics,* v. 63, pp. 471 – 476

French, K. R., Schwert G. W. and Stambaugh R. F. 1987, "Expected Stock Returns and Volatility." *Journal of Financial Economics*, v. 19, pp. 3-29.

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

Gay, R. D. J. 2008, "Effect of Macroeconomic variables on Stock Market Returns for four Emerging Economies: Brazil, Russia, India and China," *International Business & Economics Research Journal*, v. 7, no. 3, pp. 1-8.

GSE About Us, ---- *http://www.gse.com.gh/index1.php?linkid=1*

GSE Publications,---- *http://en.wikipedia.org/wiki/Ghana_Stock_Exchange*

Goldstein, H. 1995, *Multilevel Statistical Models*, 2nd edition. Halstead Press, New York.

Gudmundsson, G. 1999, "Disaggregation of Annual Flow Data with Multiplicative Trends": *Journal of Forecasting*, v. 18, pp. 33 – 37

Guerin, L., and Stroup, W.W. 2000, "A Simulation Study to Evaluate PROC MIXED Analysis of Repeated Measures Data," *Proceedings of the Twelfth Annual Kansas State University Conference on Applied Statistics in Agriculture*, April 30 – May 2, 2000.

Guerrero, V. 1990, "Temporal Disaggregation of time Series: An ARIMA-based Approach": *International Statistical Review* 58, S. 29 – 46

Harrington, D.R. 1987, *Modern portfolio theory: the Capital Asset Pricing Model and Arbitrage Pricing Theory: A user's guide*. 2nd edition. Prentice-Hall, Englewood Cliffs, New Jersey.

Harris, R.S., and Marston, F.C. 1994,"Value versus Growth Stocks: Book-to-Market, Growth and Beta," *Financial Analysts Journal*. pp. 18-24.

Hedeker, D and Gibbons R. P. 2006, *Longitudinal Data Analyses*, John Wiley & Sons, Inc.

Hodgess, E. M. and Wei W. W. S. 1996, "Temporal Disaggregation of Time Series", *M. Ahsanullah (Hrsg.), Applied Statistical Science I*, Commack, New York

Izedonmi, P. F. and Abdullahi I. B. 2011, "The Effect of Macroeconomic Factors on the Nigerian Stock Returns: A Sectoral Approach." *Global Journal of Management and Business Research*. v. 11(7)

Jacobs, J., Kroonenberg S. and Wansbeek T. 1989, "Dividing by 4: An efficient Algorithm for the Optimal Disaggregation of Annual Data into Quarterly Data", *CCSO Series* No. 7, Rijksuniversiteit Groningen

Jensen, G. R., Mercer J. M. and Johnson R. R. 1996, "Business Conditions, Monetary Policy, and Expected Security Returns." *Journal of Financial Economics*, v. 40, pp. 213-237.

Jones, C.P. 1998, *Investments: Analysis and Management*, 6th edition, Wiley, New York.

Kaminsky, G. and Schmukler S. 2002, "Emerging Markets Instability: Do Sovereign Ratings affect Country Risk and Stock Returns?" *World Bank Policy Research Working Paper,* No. 2678

Karnosky, D.S. 1993. "Global investment in the CAPM framework, in The CAPM controversy: Policy and strategy implications for investment management," *ICFA Continuing Education,* edited by D.R. Harrington and R.A. Korajczyk, AIMR, New York, pp.56-61.

Karolyi, G. A. 1995, "A Multivariate GARCH Model of International Transmissions of Stock Returns and Volatility: The Case of the United States and Canada." *Journal of Business and Economic Statistics*, v. 13(1), pp. 11-25.

Keogh, W. J. 1994. *The stability of beta and the usability of the Capital Asset Pricing Model in the South African context*, M.Econ Dissertation, University of the Orange Free State, Bloemfontein.

Kladroba, A. 2005, "The Temporary Disaggregation of Times series" *Duskussionsbeitrag* Nr. 148 Uni. Duisburg-Essen.

Laubscher, E.R. 2001, *Capital Market Theories and Pricing Models: Evaluation and Consolidation of the available body of knowledge*, M.Com Dissertation, University of South Africa, Pretoria.

Leland, H.E. 1999, "Beyond mean-variance: Performance measurement in a nonsymmetrical world". *Financial Analysts Journal*. January/February 1999, Vol. 55, No. 6, pp.27-35.

Levy, M., Levy, H. and Solomon, S. 2000, *Microscopic Simulation of Financial Markets – from Investor Behavior to Market Phenomena*. Academic Press, San Diego, California.

Lintner, J. 1965, "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, v. 47, pp. 13-37.

Lisman J. and Sandee J. 1964, "Derivation of Quarterly Figures from Annual Data". *Applied Statistics*, 13, s. 87 – 90.

Littell, R.C., Freund, R.J. and Spector, P.C. 2002, *SAS System for Linear Models*, 4th Edition, SAS Institute Inc., Cary, NC

Littell, R.C., Henry, P.R., and Ammerman, C.B. 1998, "Statistical Analysis of Repeated Measures Data Using SAS Procedures," *Journal of Animal Science*, v. 76, pp. 1216-1231.

Littell, R.C., Milliken, G.A., Stroup, W.W., and Wolfinger, R.D. 1996, *SAS Systems for Mixed Models*, SAS Institute Inc., Cary, NC

Litterman, R. 1983, "A Random Walk, Markov Model for the Distribution of Time Series," *Journal of Business & Economic Statistics*, v. 1, pp. 169–173.

Markowitz, H.M. 1952, "Portfolio Selection and Foundation of Portfolio Theory." *Journal of Finance*, pp.469-477.

Michailidis G., Tsopoglou S., Papanastasiou D. and Mariola E 2006. "Testing the Capital Asset Pricing Model (CAPM): The case of the Emerging Greek Security Market". *International Research Journal of Finance and Economics*. Issue 4, pp. 78-91.

Mossin, J. 1966, "Equilibrium in a Capital Asset Market," *Econometrica*, pp. 768-783.

Moyer, R.C., McGuigan, J.R. and Kretlow, W.J. 2001, *Contemporary Financial Management*. 8th edition, South-Western Cincinnati, Ohio.

Myers, S.C. 1977, "Determinants of Corporate Borrowing," *Journal of Financial Economics*, v. 5, pp. 147-175.

Nelson, D. B. 1991, "Conditional Heteroskedasticity in Asset Returns: A New Approach." *Econometrica*, v. 59(2), pp. 347-370.

Nijman, T. and Palm F. 1990, "Predictive Accuracy Gain from Disaggregate Sampling in ARIMA Models," *Journal of Business & Statistics*, v. 8, pp. 405-415

Pike, R. and Neale B. 1996, "*Corporate Finance and Investment–Decisions and Strategies*," 2nd ed., Prentice-Hall, Hemel Hempstead, Hertfordshire.

Radcliffe, R.C. 1997, "*Investment: Concepts, Analysis and Strategy*," 5th edition, Addison-Wesley, Reading, Massachusetts.

Raudenbush, S. W. and Bryk A. S. 2002, "*Hierarchical Linear Models*," 2nd ed. Sage, Thousand Oaks, CA

Reilly, F.K. & Brown, K.C. 1997, "*Investment Analysis and Portfolio Management*," 5th edition, Dryden, Orlando, Florida.

Roll, R. 1977, "A Critique of the Asset Pricing Theory's Test." *Journal of Financial Economics*, pp. 129-176.

Roll, R., and Ross S. A. 1994,"On the Cross-Sectional Relationship between Expected Returns and Betas," *Journal of Finance*, v. 49, pp. 101-121.

Rosenberg, B., K. Reid, and R. Lanstein 1985,"Persuasive Evidence of Market Inefficiency," *Journal of Portfolio Management*, v. 11., pp. 9-17.

Ross, S.A. 1976, "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, v. 13, pp. 3341-360.

Ross, S. A. 1993. "Is beta useful in The CAPM controversy?: Policy and strategy implications for investment management," *ICFA Continuing Education*, edited by D.R. Harrington and R.A. Korajczyk, AIMR, New York, pp.11-15.

Sentata, E. 1995, "Quadratic ARCH Models," *Review of Economic Studies*, v. 62, pp. 639-661.

Sharpe, W. F. 1964, "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, v. 19, pp. 425-442.

Sharpe, W. F. 1985, *Investments*, 3rd edition, Prentice-Hall, Englewood Cliffs, New Jersey.

Singh, T., S. Mehta and M. S. Varsha 2011, "Macroeconomic Factors and Stock Returns: Evidence from Taiwan," *Journal of Economics and International Finance,* v. 2(4), pp. 217-227.

Stattman, D. 1980, "Book Values and Stock Returns," *The Chicago MBA: A Journal of Selected Papers*, v. 4, pp. 25-45.

Stram, D. and Wei W. 1986, "A Methodological Note on the Disaggregation of Time Series Totals," *Journal of Time Series Analysis,* v.7, pp. 293 - 302

Swallow, W. H. and Monahan J. F. 1984, "Monte Carlo Comparison of ANOVA, MIVQUE, REML, and ML Estimators of Variance Components," *Technometrics*, v. 28, pp. 47-57.

Tasdemir, M. 2008, "Estimating Monthly GDP for Turkey by State-Space Approach," *International Research Journal of Finance and Economics*, Issue 17, pp. 144-151.

Van Rhijn, H. J. P. 1994, *The Capital Asset Pricing Model for Financial Decision-Making under South African Conditions*, DBA Thesis, University of Pretoria, Pretoria.

Ward, M. 2000, "The CAPM in an options pricing framework". *Investment Analysts Journal,* v. 52, pp. 35-44.

Watson, D. and A. Head 1998, "Corporate Finance – Principles & Practice". *Financial Times,* Pitman, London.

Zadorozhna, O. 2009, *Does weather affect Stock returns across Emerging Markets?* MA Thesis, Kyiv School of Economics, Ukraine.

APPENDIX A

DATA EXTRACTION FORMS

1. Form (1)

			ACCRA BREWERY LTD. (ABL)						
Year	Quarter	GSE	Market Ca	Net Asset	Div	Ear	Price	No. share	Dis Div
		736.16			0	0.000409	0.0458		
2000	Q1	760.86	2494500	1450672	0.001	0.000409	0.05	49890000	0.000229
	Q2	817.79	2694060	1450672		0.000409	0.054	49890000	0.000271
	Q3	855.51	3118125	1450672		0.002018	0.0625	49890000	0.000271
	Q4	857.98	5238450	1450672		0.001211	0.063	83150000	0.000229
2001	Q1	899.26	5238450	3523400	0	0.001211	0.063	83150000	0
	Q2	932.47	5238450	3523400		0.001211	0.063	83150000	0
	Q3	956.04	5321600	3523400		0.000606	0.032	1.66E+08	0
	Q4	955.95	5321600	3523400		0.000298	0.032	1.66E+08	0
2002	Q1	1018.02	5371490	3794900	0.001	0.000298	0.0323	1.66E+08	0.000186
	Q2	1223.69	5986800	3794900		0.003045	0.036	1.66E+08	0.00023
	Q3	1310.67	6319400	3794900		0.003045	0.038	1.66E+08	0.000272
	Q4	1395.31	6818300	3794900		0.00175	0.041	1.66E+08	0.000312
2003	Q1	1643.71	6984600	4338400	0.0015	0.00175	0.042	1.66E+08	0.000346
	Q2	2084.72	7583280	4338400		0.004826	0.0456	1.66E+08	0.000372
	Q3	2443.34	7649800	4338400		0.004826	0.046	1.66E+08	0.000388
	Q4	3553.42	9179760	4338400		0.005714	0.0552	1.66E+08	0.000393
2004	Q1	5665.04	10061150	9439700	0.0015	0.007092	0.0605	1.66E+08	0.00039
	Q2	7045.4	15091725	9439700		0.00547	0.0605	2.49E+08	0.000381
	Q3	6997.79	36918600	9439700		0.002533	0.148	2.49E+08	0.00037
	Q4	6798.46	36918600	9439700		0.002964	0.148	2.49E+08	0.000358
2005	Q1	6453.84	36918600	10439700	0.0015	0.003646	0.148	2.49E+08	0.00043
	Q4	4769.02	32428500	10439700		-0.0087	0.13	2.49E+08	0.000289
2006	Q1	4764.07	28686750	10192600	0	-0.005	0.115	2.49E+08	0
	Q2	4833.33	28686750	10192600		-0.00099	0.115	2.49E+08	0
	Q3	4943.45	28686750	10192600		-0.00678	0.115	2.49E+08	0
	Q4	5006.02	28686750	10192600		-0.00265	0.115	2.49E+08	0
2007	Q1	5092.25	28686750	10530000	0.001	-0.002	0.115	2.49E+08	0.000138
	Q2	5294.58	28686750	10530000		0.0014	0.115	2.49E+08	0.000206
	Q3	5676.77	28686750	10530000		-0.0006	0.115	2.49E+08	0.000284
	Q4	6599.77	28686750	10530000		-0.0041	0.115	2.49E+08	0.000372
2008	Q1	7848.14	28686750	11167000	0.0025	0.0025	0.115	2.49E+08	0.000469
	Q2	10346.3	28686750	11167000		0.0036	0.115	2.49E+08	0.000572
	Q3	10890.8	29934000	11167000		-0.0856	0.12	2.49E+08	0.000676
	Q4	10431.64	29934000	11167000		-0.0856	0.12	2.49E+08	0.000782

2. Forms (2)

				ACCRA BREWERY LTD. (ABL)							
Year	Quarter	R	Rf	Rm	S	BM	G	DE	PE	Rmf	T
2000	Q1	9.72	8.55	3.36	14.73	0.58	0.00	1.18	122.25	-5.19	1
	Q2	8.50	8.99	7.48	14.81	0.54	0.00	1.18	132.03	-1.51	2
	Q3	16.20	11.03	4.61	14.95	0.47	393.40	1.18	30.97	-6.42	3
	Q4	1.20	10.48	0.29	15.47	0.28	-39.99	1.18	52.02	-10.19	4
2001	Q1	0.40	10.61	4.81	15.47	0.67	0.00	0.08	52.02	-5.79	5
	Q2	0.40	11.48	3.69	15.47	0.67	0.00	0.08	52.02	-7.79	6
	Q3	-49.21	10.78	2.53	15.49	0.66	-49.96	0.08	52.81	-8.25	7
	Q4	0.00	8.16	-0.01	15.49	0.66	-50.83	0.08	107.38	-8.17	8
2002	Q1	1.72	5.73	6.49	15.50	0.71	0.00	0.13	108.39	0.76	9
	Q2	12.23	5.97	20.20	15.61	0.63	921.81	0.13	11.82	14.24	10
	Q3	6.25	6.41	7.11	15.66	0.60	0.00	0.13	12.48	0.70	11
	Q4	8.55	6.54	6.46	15.74	0.56	-42.53	0.13	23.43	-0.08	12
2003	Q1	3.35	6.81	17.80	15.76	0.62	0.00	0.28	24.00	10.99	13
	Q2	9.46	8.31	26.83	15.84	0.57	175.77	0.28	9.45	18.52	14
	Q3	1.70	7.60	17.20	15.85	0.57	0.00	0.28	9.53	9.60	15
	Q4	20.82	5.65	45.43	16.03	0.47	18.40	0.28	9.66	39.79	16
2004	Q1	10.28	4.41	59.43	16.12	0.94	24.12	0.07	8.53	55.01	17
	Q2	0.62	4.36	24.37	16.53	0.63	-22.87	0.07	11.06	20.01	18
	Q3	145.25	4.25	-0.68	17.42	0.26	-53.69	0.07	58.43	-4.92	19
	Q4	0.25	4.27	-2.85	17.42	0.26	17.02	0.07	49.93	-7.12	20
2005	Q1	0.25	4.29	-5.07	17.42	0.28	23.01	0.08	40.59	-9.36	21
	Q2	0.25	4.26	-9.16	17.42	0.28	51.10	0.08	26.87	-13.42	22
	Q3	0.25	3.75	-16.79	17.42	0.28	-118.33	0.08	-146.53	-20.54	23
	Q4	-11.91	3.15	-2.24	17.29	0.32	761.78	0.08	-14.94	-5.39	24
2006	Q1	-11.54	2.67	-0.10	17.17	0.36	-42.53	0.05	-22.99	-2.78	25
	Q2	0.00	2.43	1.45	17.17	0.36	-80.19	0.05	-116.04	-0.98	26
	Q3	0.00	2.56	2.28	17.17	0.36	583.96	0.05	-16.97	-0.29	27
	Q4	0.00	2.57	1.27	17.17	0.36	-60.95	0.05	-43.45	-1.31	28
2007	Q1	0.22	2.44	1.72	17.17	0.37	-24.44	0.05	-57.50	-0.72	29
	Q2	0.22	2.40	3.97	17.17	0.37	-170.00	0.05	82.14	1.57	30
	Q3	0.22	2.44	7.22	17.17	0.37	-142.86	0.05	-191.67	4.78	31
	Q4	0.22	2.62	16.26	17.17	0.37	583.33	0.05	-28.05	13.64	32
2008	Q1	0.54	2.70	18.92	17.17	0.39	-160.98	0.08	46.00	16.22	33
	Q2	0.54	3.40	31.83	17.17	0.39	44.00	0.08	31.94	28.43	34
	Q3	4.89	5.53	5.26	17.21	0.37	-2477.78	0.08	-1.40	-0.27	35
	Q4	0.52	6.17	-4.22	17.21	0.37	0.00	0.08	-1.40	-10.38	36

APPENDIX B

GenStats RESULTS FOR CROSS - SECTIONAL ANALYSES

B1. Linear Regression with Constanat Coefficients

Response variate: R
Fitted terms: Constant, Rf, Rmf, S, BM, G, DE, PE

1. ABL

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	2686.	383.7	0.50	0.826
Residual	28	21456.	766.3		
Total	35	24141.	689.8		

Residual variance exceeds variance of response variate
Standard error of observations is estimated to be 27.7

*** Estimates of parameters ***

	estimate	s.e.	t(28)	t pr.
Constant	64.	401.	0.16	0.874
Rf	0.40	1.06	0.37	0.711
Rmf	0.447	0.438	1.02	0.317
S	-1.8	21.5	-0.08	0.933
BM	-69.0	82.8	-0.83	0.412
G	-0.0005	0.0101	-0.05	0.963
DE	-2.6	29.6	-0.09	0.932
PE	0.0956	0.0837	1.14	0.263

2. ALW

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	9695.	1385.	1.29	0.289
Residual	28	29964.	1070.		
Total	35	39659.	1133.		

Percentage variance accounted for 5.6
Standard error of observations is estimated to be 32.7

*** Estimates of parameters ***

	estimate	s.e.	t(28)	t pr.
Constant	39.	504.	0.08	0.939
Rf	1.720	0.909	1.89	0.069
Rmf	0.959	0.381	2.52	0.018
S	-2.8	28.3	-0.10	0.923
BM	-4.0	26.5	-0.15	0.882
G	0.0089	0.0183	0.49	0.631
DE	0.02	5.75	0.00	0.997
PE	0.079	0.109	0.72	0.477

3. CFAO

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	2167.	309.5	2.08	0.080
Residual	28	4174.	149.1		
Total	35	6341.	181.2		

Percentage variance accounted for 17.7
Standard error of observations is estimated to be 12.2

*** Estimates of parameters ***

	estimate	s.e.	t(28)	t pr.
Constant	-37.2	53.7	-0.69	0.494
Rf	1.571	0.554	2.84	0.008
Rmf	0.216	0.196	1.10	0.280
S	-1.01	3.25	-0.31	0.759
BM	2.60	2.68	0.97	0.342
G	0.0112	0.0185	0.60	0.551
DE	32.9	19.1	1.72	0.096
PE	2.384	0.694	3.44	0.002

4. Port

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	61023.	8717.5	17.32	<.001
Residual	496	249597.	503.2		
Total	503	310619.	617.5		

Percentage variance accounted for 18.5
Standard error of observations is estimated to be 22.4

*** Estimates of parameters ***

	estimate	s.e.	t(496)	t pr.
Constant	-42.1	11.2	-3.75	<.001
Rf	2.104	0.422	4.98	<.001
Rmf	0.6652	0.0683	9.74	<.001
S	2.625	0.615	4.27	<.001
BM	1.54	1.20	1.28	0.200
G	0.00509	0.00489	1.04	0.299
DE	-0.618	0.933	-0.66	0.508
PE	0.0320	0.0328	0.98	0.329

B2. Linear Regression without Constant Coefficients

Response variate: R

Fitted terms: Rf, Rmf, S, BM, G, DE, PE

1. ABL

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	3685.	526.5	0.71	0.663
Residual	29	21475.	740.5		
Total	36	25160.	698.9		

Residual variance exceeds variance of response variate
Standard error of observations is estimated to be 27.2

*** Estimates of parameters ***

	estimate	s.e.	t(29)	t pr.
Rf	0.463	0.955	0.48	0.632
Rmf	0.430	0.418	1.03	0.312
S	1.614	0.872	1.85	0.074
BM	-58.0	45.3	-1.28	0.210
G	0.00000	0.00947	0.00	1.000
DE	1.2	17.9	0.07	0.948
PE	0.0963	0.0822	1.17	0.251

2. ALW

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	18787.	2684.	2.60	0.033
Residual	29	29970.	1033.		
Total	36	48757.	1354.		

Percentage variance accounted for 8.8
Standard error of observations is estimated to be 32.1

*** Estimates of parameters ***

	estimate	s.e.	t(29)	t pr.
Rf	1.756	0.766	2.29	0.029
Rmf	0.956	0.372	2.57	0.016
S	-0.59	1.04	-0.57	0.574
BM	-2.1	11.6	-0.18	0.855
G	0.0090	0.0179	0.50	0.620
DE	0.10	5.57	0.02	0.986
PE	0.080	0.107	0.75	0.462

3. CFAO

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	4120.	588.6	4.02	0.003
Residual	29	4246.	146.4		
Total	36	8366.	232.4		

Percentage variance accounted for 19.2
Standard error of observations is estimated to be 12.1

*** Estimates of parameters ***

	estimate	s.e.	t(29)	t pr.
Rf	1.329	0.426	3.12	0.004
Rmf	0.208	0.194	1.08	0.291
S	-3.14	1.04	-3.03	0.005
BM	1.71	2.34	0.73	0.471
G	0.0144	0.0178	0.81	0.423
DE	33.1	18.9	1.75	0.091
PE	2.353	0.686	3.43	0.002

4. Port

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	7	150496.	21499.5	41.63	<.001
Residual	497	256683.	516.5		
Total	504	407179.	807.9		

Percentage variance accounted for 16.4
Standard error of observations is estimated to be 22.7

*** Estimates of parameters ***

	estimate	s.e.	t(497)	t pr.
Rf	1.197	0.351	3.41	<.001
Rmf	0.6458	0.0690	9.36	<.001
S	0.388	0.154	2.52	0.012
BM	0.35	1.17	0.30	0.765
G	0.00593	0.00495	1.20	0.231
DE	-1.796	0.890	-2.02	0.044
PE	0.0243	0.0331	0.73	0.464

B3. Linear Regression with One Level of Interaction

```
Response variate: R
     Fitted terms: Rf + Rmf + S + BM + G + DE + PE + Rf.Rmf +
                   Rf.S + Rmf.S + Rf.BM + Rmf.BM + S.BM +
                   Rf.G + Rmf.G + S.G + BM.G + Rf.DE + Rmf.DE +
                   S.DE + BM.DE + G.DE + Rf.PE + Rmf.PE + S.PE +
                   BM.PE + G.PE + DE.PE
                   (FACTORIAL limit for expansion of formula = 2)
```

1. ABL

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	28	20992.	749.7	1.44	0.307
Residual	8	4168.	521.1		
Total	36	25160.	698.9		

Percentage variance accounted for 25.7
Standard error of observations is estimated to be 22.6

*** Estimates of parameters ***

	estimate	s.e.	t(8)	t pr.
Rf	96.	191.	0.50	0.629
Rmf	344.	133.	2.58	0.032
S	21.7	11.0	1.98	0.083
BM	5494.	3401.	1.62	0.145
G	-39.7	15.4	-2.57	0.033
DE	-12395.	9132.	-1.36	0.212
PE	100.8	52.9	1.91	0.093

2. ALW

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	28	47850.5	1708.9	15.08	<.001
Residual	8	906.5	113.3		
Total	36	48756.9	1354.4		

Percentage variance accounted for 90.0
Standard error of observations is estimated to be 10.6

*** Estimates of parameters ***

	estimate	s.e.	t(8)	t pr.
Rf	-119.5	62.7	-1.90	0.093
Rmf	28.5	26.0	1.10	0.304
S	8.07	6.11	1.32	0.223
BM	-2625.	1398.	-1.88	0.097
G	-11.3	22.9	-0.49	0.635
DE	3674.	2017.	1.82	0.106
PE	-78.1	58.2	-1.34	0.216

3. CFAO

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	28	7767.6	277.41	3.71	0.029
Residual	8	598.5	74.82		
Total	36	8366.1	232.39		

Percentage variance accounted for 58.7
Standard error of observations is estimated to be 8.65

*** Estimates of parameters ***

	estimate	s.e.	t(8)	t pr.
Rf	-19.1	16.7	-1.14	0.285
Rmf	-22.6	11.5	-1.96	0.086
S	-5.30	4.88	-1.08	0.310
BM	-345.	188.	-1.84	0.103
G	-1.17	1.38	-0.84	0.423
DE	1171.	1139.	1.03	0.334
PE	-72.8	46.0	-1.58	0.152

4. Port

*** Summary of analysis ***

	d.f.	s.s.	m.s.	v.r.	F pr.
Regression	28	180240.	6437.1	13.50	<.001
Residual	476	226939.	476.8		
Total	504	407179.	807.9		

Percentage variance accounted for 22.8
Standard error of observations is estimated to be 21.8

*** Estimates of parameters ***

	estimate	s.e.	t(476)	t pr.
Rf	-5.11	1.69	-3.02	0.003
Rmf	-3.772	0.851	-4.43	<.001
S	-0.091	0.238	-0.38	0.701
BM	-1.7	10.9	-0.16	0.877
G	-0.0521	0.0700	-0.74	0.457
DE	2.99	8.10	0.37	0.712
PE	0.621	0.439	1.41	0.158
Rf.Rmf	0.1257	0.0455	2.76	0.006
Rf.S	0.486	0.114	4.26	<.001
Rmf.S	0.2286	0.0465	4.92	<.001
Rf.BM	0.100	0.434	0.23	0.817
Rmf.BM	0.191	0.113	1.69	0.091
S.BM	0.125	0.650	0.19	0.848
Rf.G	0.00521	0.00251	2.08	0.038
Rmf.G	-0.000312	0.000594	-0.52	0.600
S.G	0.00217	0.00407	0.53	0.595
BM.G	0.00662	0.00728	0.91	0.364
Rf.DE	-0.039	0.792	-0.05	0.961
Rmf.DE	0.138	0.102	1.35	0.177
S.DE	-0.127	0.513	-0.25	0.805
BM.DE	1.48	2.88	0.52	0.606
G.DE	0.00242	0.00275	0.88	0.379
Rf.PE	-0.0431	0.0197	-2.19	0.029
DE.PE	-0.0219	0.0155	-1.41	0.158