Modern Finance

Article

Properties of returns and variance and the implications for time series modelling: Evidence from South Africa

Jan Jakub Szczygielski [1,2*], and Chimwemwe Chipeta [3]

* Corresponding author.
1 Assistant Professor, Department of Finance, Kozminski University, ul. Jagiellonska 57/59, Warsaw, 03-301, Poland, jszczygielski@kozmski.edu.pl
2 Department of Financial Management, University of Pretoria, Private Bag x20, Hatfield, Pretoria, 0028, South Africa
3 Associate Professor: Corporate Finance, School of Economic and Business Sciences, University of the Witwatersrand, E-mail: chimwemwe.chipeta@wits.ac.za

Abstract: This paper investigates the properties of South African stock returns and the underlying variance. The investigation into the properties of stock returns and the behaviour of the variance underlying returns is undertaken using model-free approaches and through the application of ARCH/GARCH models. The results indicate that, as with other stock markets, returns on the South African stock market depart from normality and that variance displays evidence of heteroscedasticity, long memory, persistence, and asymmetry. Applying the EGARCH(p,q,m) and IGARCH(p,q) specifications confirms these findings and the application of these models suggests differing characteristics for variance structures underlying the South African stock market. In light of the findings relating to the properties of stock returns and the characteristics of variance and its structure, implications are outlined, and recommendations on how time-series specifications may be estimated are made.

Keywords: Johannesburg Stock Exchange; leverage effect; stock market returns; variance

JEL classification codes: G12

1. Introduction

Stock returns are assumed to conform to a set of *a priori* assumptions – those of *normally, independently and identically distributed (n.i.i.d)* returns - that are crucial for model specification, estimation and inference making (Mandelbrot, 1963; Fama, 1965). However, the literature widely recognizes that these assumptions do not hold in practice, and this has the potential to impact model estimation, notably in the time-series context and when "the great workhorse" - the least-squares methodology - is applied (Engle, 2001: 157; Xiao and Aydemir, 2007). This paper follows Mangani (2007), who investigates the properties of South African stock returns and reports that South African stock returns depart from normality, as evident from the presence of leptokurtosis and skewness in returns, and are not independent. Although non-linearities are not explored by Mangani (2007), the author recognizes that the presence of non-linearities in South African stock returns is an avenue for further research.

The aim of this paper is not only to explore and report upon the characteristics of the return distribution (as in Mangani, 2007) and variance but also to relate the observed results to implications for time-series modelling and to present recommendations as to how these characteristics, which adversely impact model estimation and inference making, may be addressed. As an extension to Mangani (2007), this study focuses on the stationarity, normality, and linearity properties of logarithmic price changes. Consequently, this paper places greater emphasis on non-linearity in stock returns, ARCH

Citation: Szczygielski, J.J., & Chipeta, C. (2023). Properties of returns and variance and the implications for time series modelling. *Modern Finance*, 1(1), 35-55.

Accepting Editor: Adam Zaremba

Received: 4 July 2023
Accepted: 10 August 2022
Published: 11 August 2022

effects, and the leverage effect (the negative correlation between stock returns and volatility). The nature and structure of the variance of South African stock returns is investigated using the Autoregressive Conditional Heteroscedastic/Generalized Autoregressive Heteroscedastic (ARCH/GARCH) modelling framework – an aspect that is not considered by Mangani (2007) in extensive detail. Specifically, this investigation utilizes the asymmetric Exponential GARCH and Integrated GARCH specifications. Most importantly, this paper aims to emphasize that researchers and econometricians should be cognisant of the characteristics of the return distribution and variance structure when applying time-series models to investigate return behaviour, variance structure and return-variable relationships.

From the foregoing discussion, the contribution of this paper is fourfold. First, a comprehensive investigation of the statistical properties of the South African stock market is undertaken. The sample in this study comprises of the market aggregate, represented by the Johannesburg Stock Exchange (JSE) All Share Index, the eight economic sectors comprising the JSE and 26 industrial sectors with a full data history between January 2001 and December 2016. The investigation is therefore conducted at various levels of the JSE. Second, the extensive dataset used takes into account a number of significant events such as the September 11th terrorist attacks and the global financial crisis of 2008. These events are likely to have resulted in periods of prolonged heightened volatility and are likely to be reflected in the returns in the form of volatility clustering, persistence and long memory. Third, this study provides new information and insight into the structure of time-varying volatility and the leverage effect in returns on the JSE and its components. Literature that comprehensively investigates the structure and characteristics of volatility underlying the South African stock market and that takes into account the persistence, long memory and asymmetric aspects of volatility is sparse. This paper considers the leverage effect and the direction of causality - asymmetry in the return volatility relationship - using a model-free and a model-based approach. Bouchaud, Matacz and Potters (2001) state that although the leverage effect has been measured and discussed within ARCH/GARCH literature, its structure has never been quantitatively investigated. Finally, this study is also of relevance to researchers and econometricians who are interested in modelling volatility dynamics *and* deriving linear factor models using financial return series. In the latter case, important examples of such studies are those of Burmeister and Wall (1986) and Berry, Burmeister and McElroy (1988). Although this study relies upon ARCH/GARCH modelling to investigate the properties of variance, Hamilton (2010) suggests that incorporating features of heteroscedasticity into the conditional mean translates into more efficient estimates of the conditional mean. Thus, the characteristics of variance matter not only for stock prices but also in macroeconomic questions (which use data of a lower frequency) and questions that draw upon a combination of finance and macroeconomic theory.

The findings of this paper are that, as expected, returns on the South African stock market and the economic and industrial sectors that comprise the South African stock market depart from normality. Likewise, the assumption of strict independence is not valid; a more accurate description is that independence is a working assumption, given the presence of weak yet statistically significant serial correlation. Furthermore, the results of the analysis point to the widespread presence of ARCH effects and non-linear dependence suggestive of the time varying variance, which necessitates the use of ARCH/GRACH-type models. A notable finding is that a return-volatility leverage effect is present in returns on the JSE All Share Index and a number of economic and industrial sectors, suggesting that the variance underlying the South African stock market is asymmetric in nature. Finally, the application of the ARCH/GARCH framework also suggests that volatility underlying the South African stock market is persistent, has a long-memory and sometimes an infinite memory.

The rest of the paper is organised as follows: Section 2 outlines the literature on the properties and behaviour of stock returns and the structure of volatility. Section 3 elaborates upon the dataset utilized in this study and the methodology employed. Section 4 reports the findings, discusses the implications for econometric modelling and makes recommendations as to how these may be addressed. Section 5 concludes by summarizing and suggesting areas for further research.

2. Literature review

2.1 The properties and behaviour of stock returns and variance

Returns are commonly described as Gaussian or normally distributed with a mean of 0 and variance proportional to the differencing interval, Δt . This implies that the distribution of returns is described by the mean, μ , and innovations from the mean, εt (Mandelbrot, 1963). However, the literature has cast doubt upon the validity of assumptions relating to return behaviour and empirical evidence suggests that the first two moments of the return distribution are not "well-behaved" as implicitly assumed in empirical studies dealing with return-model estimation.

2.1.1 The return distribution

Fama (1965) states that prior to Mandelbrot's (1963) work, the assumption of normality was not widely questioned and according to Officer (1972: 807), the normal distribution was seen as "a good working hypothesis." Mandelbrot (1963) is credited with re-examining the distributional assumptions of stock returns and contends that the normal distribution fails to account for the excess kurtosis and the long tails exhibited by return distributions. In a subsequent paper, Mandelbrot (1967: 396) reiterates his position regarding the high levels of kurtosis observed in financial time series and notes that "Bachalier's assumption that the marginal distribution of L t T (,) (returns) is Gaussian with vanishing expectation might be convenient, but virtually every student of the distribution of prices has commented on their leptokurtic (i.e., very long-tailed) character."

Using data on stocks comprising the Dow Jones Industrial Average (DJIA), Fama (1965) finds that on average, a greater proportion of observations are found to be centred on the mean and a greater number of observations are observed in the tails of the empirical distribution relative to that implied by the normal distribution. Moreover, Fama (1965) reports that the actual level of excess frequency beyond five standard deviations is almost 2000 times greater than that implied by the normal distribution and, in conclusion, states that the normal distribution is not an accurate representation of the return distribution. Brown and Warner (1985), using samples of stocks from the Center for Research in Security Prices (CRSP) database, find that kurtosis is more than double that of the normal distribution and that departures from normality differ according to the frequency of data used. Widespread recognition of the presence of excess kurtosis is more recently acknowledged by Xiao and Aydemir (2007), who state that the level of kurtosis for many studies is above three and by Engle and Patton (2007), who state that it is well-established that return distributions have fat tails and that typical estimates of kurtosis range between four to 50. Mangani (2007) reports that kurtosis is greater than 3 for all return series in his South African sample, these being returns on 42 individual stocks, a portfolio comprising of these stocks and returns on the JSE All Share Index. The author goes on to conclude that there is undisputed evidence of leptokurtosis on the JSE – a finding that provides no support for the normality assumption and that is widely documented in markets. Likewise, Kumar and Dhankar (2010) report kurtosis values in excess of three suggesting that returns in the US stock markets are not normally distributed.

Peiró (1999) suggests that the assumption of symmetry implies that upside and downside risks are considered equally by investors. The author argues that while high levels of kurtosis are a well-recognized feature of return distributions, less consideration is given to the symmetry of the distribution as it is considered to be less important. This is problematic given that leptokurtosis is usually accompanied by asymmetry. Mangani

(2007) finds that the sample skewness parameter is statistically insignificant in six out of the 42 individual stocks considered. For both the portfolio and returns on the JSE All Share Index, the skewness parameter is found to be significant. Returns on the portfolio and JSE All Share Index are negatively skewed, whereas returns on 24 of the individual stocks are found to be negatively skewed. Similarly, Kumar and Dhankar (2010) find that returns on the S&P500 are negatively skewed, suggesting that there is a high probability that skewness extends to individual returns series. Similar results are reported in Lim, Luo and Kim (2013) for the S&P500, the Dow Jones Industrial Average (DJIA) and the New York Stock Exchange (NYSE). In a study of the volatilities of leading US and Eurozone market indices, Ning, Xu and Wirjanto (2015) find that the return series of these indices depart significantly for the assumptions of normality. Returns on U.S. indices, namely the S&P 500, the DJIA, the Nasdaq 100 and the Euro Stoxx 50 index exhibit kurtosis coefficients of over 3 and positive skewness. Findings that return distributions are characterized by both leptokurtosis and skewness pose a challenge to the assumption of normality.

According to Cont (2001), it is a well-known fact that there is no significant linear correlation in returns. Therefore, the independence assumption assumes that the serial correlation function of returns decays rapidly to zero. The absence of (linear) serial correlation is often cited as evidence in favour of the efficient market hypothesis (Cont, 2001). Campbell, Lo and MacKinlay (1997) suggest that if the serial correlation function is equal to zero, returns are serially uncorrelated and mutually independent. Returns are assumed to show little or no linear serial correlation and if serial correlation is present, it is short-lived. Furthermore, Independence can further be defined from two perspectives. The first relates to statistical independence in returns. The second relates to whether investors can use the knowledge of past returns to increase expected profits (Fama, 1965; Mandelbrot, 1967). Kendall and Hill (1953: 11) in an early analysis of the properties of returns find that the pattern of events in price series is less systematic than generally accepted. Changes from one period to another behave almost like a "wandering series" implying that subsequent returns follow a random walk and are independent. Kendall and Hill (1953) first report findings for the Chicago Wheat Series. The series follows a random walk with changes from one period to the other appearing to be independent, thus making serial correlation unlikely and mostly negligible for this series. An analysis of serial correlation in British Industrial Share Prices yields similar results; for the most part, changes in prices are independent and where dependence is observed, it is too low to exploit for predictive purposes. Mangani (2007), on the basis of an examination of the serial correlation structure of return series in his sample, fails to find statistically significant serial correlation in 24 of the 42 return series for individual stocks. However, returns on the JSE All Share Index and the stock portfolio exhibit statistically significant serial correlation at the first order. Mangani (2007) concludes that the results suggest that the assumption of independence is invalid.

Baur, Dimpfl and Jung (2012) adopt a quantile regression approach to examine serial correlation of daily, weekly and monthly return series for stocks comprising the Dow Jones Stoxx 600 index. The authors find evidence of serial correlation in the lower quantiles, while negative dependence with past stock price returns is observed in the upper quantiles. The middle quantiles exhibit weak or no serial correlation with past returns. Lim, Luo and Kim (2013) re-examine the issue of return predictability for three US stock indices using rolling window estimation of time-varying variability of the predictability of returns. The authors show that periods of serial correlation are characterised by exogenous shocks arising from of events such as the global financial crisis. Specifically, the authors conjecture that investors' misreaction to news during such periods causes significant return serial correlation. Likewise, Kinnunen (2013) shows that predictability in Russian aggregate market returns is largely conditional on information flow; serial correlation tends to increase (reduce) during periods of low (high) information flow.

In contrast to the assumption of normality, which is widely rejected, the independence assumption continues to be debated. Given these findings, it is difficult to conclusively pronounce upon the validity of the assumption of independence. The optimal approach may be to investigate the independence assumption on a "case-by-case" basis – as is the case in this paper.

Mandelbrot (1967) defines stationarity as the non-variation in the sample moments. To test for stationarity in the mean of South African stock returns, Mangani (2007) applies the DickeyFuller (DF) and Augmented Dickey-Fuller (ADF) tests. For the DF and ADF tests, the hypothesis of a unit root in the return series for the individual stocks, the portfolio and the JSE All Share Index is rejected in all instances and the author concludes that JSE stock return series are of a stationary nature. While these findings provide insight into the stationarity of the mean, they do not provide insight into the stationarity of variance. Given the possibility of nonstationarity of the variance, the validity of the assumption of identically distributed returns remains questionable; it can be argued that return distributions are stationary in the mean but not in the variance. This is an aspect that is investigated by this paper.

2.1.2 The Characteristics of volatility

An investigation of volatility yields further insight into the behaviour of returns and the return distribution and similarly to stock returns, return volatility is characterized by a number of stylized facts, namely volatility clustering, persistence, leverage effects and mean reversion.

The phenomena of volatility clustering, although not referred to by that name at the time, is acknowledged early on by Mandelbrot (1963: 418) who notes that "large changes (in prices) tend to be followed by large changes – of either sign – (and) small changes tend to be followed by small changes." Volatility clustering implies that volatility exhibits alternating periods of tranquillity and heightened amplitude suggesting that fluctuations in returns are lumped together (Poon, 2005; Chan and Cryer, 2008). The presence of volatility clustering is further evidence in favour of the proposition that variance is of a time-varying nature (Jacobsen and Dannenburg, 2003). Engle (2001) states that time-varying variance is easily observed and an examination of a time-series plot of returns is all that is required to establish whether volatility clustering is present or not. Engle (2001) demonstrates this by reporting plots of DJIA and NASDAQ returns (see Engle 2001; Figure 1). The amplitude of returns is shown to vary over time; the amplitude is greater around initial observations and declines towards the middle and increases greatly towards the end of the sample period. This is an observable example of volatility clustering, also referred to as the "ARCH effect," often cited as an explanation for leptokurtosis (Akgiray, 1989; Engle, 2001). Engle (2001) interprets the variance as the risk level of returns and volatility clustering implies that certain time periods are riskier than others. Notably, these riskier times are not random and are serially correlated. In Ning, Xu and Wirjanto (2015), the asymmetric nature of volatility clustering is examined using highfrequency data for several stock market indices from the US and the Eurozone. The authors find evidence of non-linear and asymmetric volatility clustering. Further, the authors report that, in most of the return series, volatility clustering is highly persistent.

The concepts of volatility clustering and volatility persistence are closely related to the extent that some authors do not make an explicit distinction between these two phenomena (see Engle and Patton, 2007). In fact, Niu and Wang (2013) show that volatility clustering may lead to long-range dependence of the volatility time series. Thus, volatility clustering implies volatility persistence; if extended periods are characterized by greater variability in returns and other periods by lower variability, then this suggests that variability must be persistent to create identifiable periods of greater and lower volatility. Perhaps a more fitting term for persistence is "long memory". Whereas volatility clustering implies that extended periods of volatility arise from the clustering of news or the clustering of information arrivals, the persistence or longmemory property of volatility implies that a single shock will have an impact upon future volatility in periods to come (Engle, 2004; Engle, Focardi and Fabozzi, 2008; McMillan and Ruiz, 2009). McMillan and Ruiz (2009) state that the standard approach to examining the longmemory property in time series is to examine the sample serial correlation function for nonlinear transformations of returns. Whereas non-linear serial correlation of any length is a symptom of volatility clustering, in the context of memory, what is of interest is how long it takes for a shock to die out. If it takes the sample serial correlation function an extended period of time to decline to zero, the process exhibits long memory. In other words, levels of heightened volatility persist and shocks do not die out immediately.

A considerable amount of literature has examined the issue of long memory in stock price returns and the evidence appears to be mixed. Using parametric and semiparametric estimation techniques, Henry (2002) finds evidence of long memory in four out of nine international stock market indices. Likewise, Chkili, Aloui and Nguyen (2012) examine conditional volatility in stock returns and exchange rates for three European stock markets, and find evidence of long memory in the conditional variance of the time series of returns for all the sampled markets. Long-run dependence has also been investigated in emerging markets. For instance, Bhattacharya and Bhattacharya (2012) examine this issue in the stock market indices of 10 emerging markets and confirm evidence of long memory in the return series of these markets. However, Kasman, Turgutlu and Ayhan (2009) document conflicting results for Central European stock market data. The authors find weak evidence of long-run dependence in returns for Hungary and the Czech Republic and strong evidence for Slovakia. In contrast, empirical evidence for North African markets seems to suggest that long-run dependence is present in these markets. For example, Boubaker and Makram (2012) show that the North African stock markets exhibit long memory in both the returns and the volatility of returns. Evidence of long memory in the return series for these markets contradicts the proposition of the weak-form market efficiency. In a related study, Anagnostidis and Emmanouilides (2015) explore nonlinear dependence in high-frequency data from the Athens Exchange Composite Index. Consistent with the Mixture-of-Distribution hypothesis, the authors find evidence of long memory in the volatility process. This, taken together with the other studies, suggests that not all market exhibit long-run dependence although this is a characteristic of volatility.

[6] The NASDAQ, as it is commonly referred to today, stands for the National Association of Securities Dealers Automated Quotation.

3. Data and methodology

3.1. Data

The data sample consists of monthly returns on the FTSE/JSE All Share Index (henceforth JSE All Share Index) and the FTSE/JSE Africa economic- and industrial-sector indices with a full history over the January 2001 to December 2016 period.7 The data is sourced from the IRESS Expert Database and month-end total return (returns adjusted for dividends) data is used. The use of these indices makes it possible to conduct a comprehensive and broad investigation of the return distribution and volatility of the South African stock market using 192 months of data for the JSE as represented by the FTSE/JSE All Share Index comprising of eight economic sectors which, in turn, comprise of 26 industrial sectors. The economic group and industrial sector indices considered in the study are listed in Table 1.

As the sample spans the period from January 2001 to December 2016, this paper traces the growth of the economic and industrial sectors constituting the South African stock market over this period. The sample period coincides with a number of significant events and coincides with certain events that are not taken into account in Mangani's (2007) sample. Some of these events are the aftermath of the bursting of the Dot-com bubble in 2000, an unprecedented terrorist attack on the twin towers of the World Trade

Centre in New York in 2001, the sub-prime mortgage crisis of 2008 and its aftermath, growing trade liberalization and the emergence and consolidation of the economic clout of the "Asian Tigers" in the form of India and China.

Table 1. Economic and industrial sector FTSE/JSE All-Africa series indices

Economic Sector Index	Industrial Sector Index
1.Basic Materials	1.1. Chemicals
	1.2. Forestry & Paper
	1.3. Industrial Metals
	1.4. Mining
2.Industrials	2.1. Construction & Materials
	2.2. General Industrials
	2.3. Electronic & Electrical Equipment
	2.4. Industrial Engineering
	2.5. Industrial Transport
	2.6. Support Services
3.Consumer Goods	3.1. Automobiles & Parts
	3.2. Beverages
	3.3. Food Producers
4.Health Care	4.1. Health Care Equipment & Services
	4.2. Pharmaceuticals & Biotechnology
5.Consumer Services	5.1. Food & Drug Retailers
	5.2. General Retailers
	5.3. Media
	5.4. Travel & Leisure
6. Telecommunication	6.1. Fixed-line Telecommunications
7. Financials	7.1. Banks
	7.2. Non-life Insurance
	7.3. Life Insurance
	7.4. General Financial
	7.5. Equity Investment Instruments
8. Technology	8.1. Software & Computer Services

Notes: Economic and industrial sector classification based upon the FTSE/JSE Global Classification system.

Formally, the returns used are continuously compounded total returns – the natural logarithm of monthly total returns over the sample period:

$$r_{it} = ln\ S_{it} - lnS_{it-1} \tag{1}$$

where r_{it} is the total return on index *i* at time *t*, and Sit is the level of index i at time t. Excess total returns (henceforth referred to as returns), r_{it} , are obtained by subtracting the risk-free rate, as measured by the yield on the R186 government bond, from the logarithm of total returns in equation (1).

3.2. Methodology

Preliminary analysis is conducted on each return series and the mean, standard deviation, kurtosis and skewness are reported for each series in Table 2. To formally test whether each series conforms to the normality assumption, the Jarque-Bera (JB) test, which assumes that a normal distribution is characterized by a skewness (S) coefficient of zero and a kurtosis (K) coefficient of three is applied to test the joint hypothesis that S = 0 and K = 3 (Cryer and Chan, 2008). As outliers are likely to bias normality tests towards a rejection of the normality assumption, box plots are used to identify outliers and far (extreme) outliers are excluded for the purposes of testing the normality assumption

(Hodge and Austin, 2004; Poon, 2005; Agung, 2009).[10] A rejection of the null hypothesis implies that a return series is not normally distributed.

To investigate the assumption of (statistical) independence, the approach of Fama (1965) and Campbell, Lo and MacKinlay (1997) is adopted in the form of the serial-correlation model. Serial-correlation coefficients not only provide insight into whether the assumption of independence holds, but also reveal the magnitude of dependence (Fama, 1965). The assumption of independence is further investigated using Ljung-Box Q-statistics (henceforth Q-statistics). Unlike serial-correlation coefficients, which indicate the level of serial correlation at individual lags, the Q-statistic indicates whether serial-correlation coefficients up to a certain order are jointly equal to zero (Campbell, Lo and MacKinlay, 1997; Gujarati, 2003). Individual serial-correlation coefficients for the first five orders are reported for each series together with the Q-statistics for the first five and 10 orders.

Using the Q-statistic to test whether serial coefficients are jointly equal to zero complements the serial-correlation model; while individual correlation coefficients may be statistically significant, jointly they may be equal to zero suggesting a negligible level of dependence (see Fama, 1965). An analysis of the serial-correlation structure of stock returns also provides preliminary insight into the validity of the assumption of identically distributed returns; if a time series is white noise, then the series is mostly likely stationary. As in Mangani (2007), the more formal Augmented Dickey-Fuller (ADF) unit root test is employed to test the stationarity of each series. The null hypothesis implies that the series has a unit root and is non-stationary and a rejection of the null hypothesis implies that the series is stationary (Gujarati, 2003). While returns are most likely to be stationary in the mean, results of the ADF test for returns are reported for comprehensiveness (see Sadorsky, 2001; Sadorsky and Henriques, 2001).

To investigate the characteristics of volatility underlying the economic and industrial sector return series, a (econometric) model-free approach in the form of Q-statistics is used to determine whether squared returns – a proxy for volatility – are serially correlated. Statistically significant Q-statistics indicate that the ARCH effect is present in the return series (Poon, 2005; Cryer and Chan, 2008). Following Engle (2001), a Q-statistic for the first 15 serial-correlation coefficients of squared returns is reported for each series. The presence of ARCH effects implies that the ARCH/GARCH framework is appropriate for modelling and analyzing the return-generating process of South African stock returns (see Elyasiani and Mansur, 1998). zvarying variance and indicate that the ARCH effect is present in the residuals. The approach undertaken in this study is to determine whether the residual terms are conditionally heteroscedastic. If residuals are conditionally heteroscedastic, then it can be argued that the residuals reflect volatility clustering and time-varying variance in returns. Residual series are generated for testing by applying an AR(1) model to the return series (see Akgiray, 1989):

$$R_{it} = \alpha = b_1 R_{it-1} + \varepsilon_{it} \quad (2)$$

where R_{it} is the return on index i at time t and R_{it-1} is the autoregressive term. Tests are conducted to determine whether ARCH(1), (5) and (10) effects are present in the residuals, ε_{it} .

Another feature of volatility, aside from the presence of time variation and ARCH effects, is the asymmetric relationship between returns and volatility – the leverage effect – which has been cited as an explanation for negatively skewed return distributions (Black, 1976; Bouchaud, Matacz and Potters, 2001). While the direction of the causality of the effect is questioned, the presence of leverage effects may be established by considering the correlation between squared returns representative of volatility and past returns, or alternatively, by considering the correlation between returns and prior volatility, depending upon the hypothesized direction of causality:

$$L1(\tau) = corr(R^2_{it}, R_{it-\tau}) \quad (3)$$

$L2(\tau) = corr(R_{it}, R^2_{it-\tau})$ (4)

where the proxy for volatility in the form of squared returns is denoted by R^2_{it}, returns are denoted by R_{it} and τ is the lag order. If the correlation function, $L1(\tau)$, starts from a negative value and decays to zero, negative (positive) return)s result in increases (decreases) in volatility (equation 3). Of course, increases (decreases) in volatility may result in decreases (increases) in returns (equation 4). Although equations (3) and (4) represent the same approach to testing for the leverage effect, they differentiate between the direction of the casual relationship (see Bouchaud, Matacz & Potters, 2001).

To investigate the persistence of volatility, the approach of Engle and Patton (2007) and McMillan and Ruiz (2009) is also followed whereby ARCH/GARCH (see discussion relating to ARCH/GARCH models that follows) specifications are used to investigate volatility persistence and mean reversion (or the lack thereof). Although 192 monthly observation are used in this study, there is precedent for using a sample of this length in terms of the (lower) number of observations and the (lower) frequency in ARCH/GARCH modelling, especially in applications for which the conditional mean is of interest (see Sadorsky & Henriques, 2001; Hamilton, 2010). The sum of the coefficients of the conditional variance specification, $\alpha_i + \beta_j$ - the persistence parameter, κ - is less than unity if unconditional variance is finite, implying that volatility reverts to a mean level. The closer κ is to 1, the longer it takes for volatility to revert to its mean. In investigating the duration of shocks, the volatility half-life is estimated as follows (McMillan & Ruiz, 2009):

$\xi = log(0.5)/log(x),$ (5)

where ξ is the volatility half-life in months and κ is the persistence parameter. The volatility half-life is a measure of the amount of time that it takes volatility to move halfway back to its unconditional mean following a shock (assuming that volatility is mean-reverting).

The two ARCH/GARCH specifications that are applied to investigate the implications of the statistical properties of returns and the behaviour of volatility for modelling South African return series are Engle and Bollerslev's IGARCH(*p,q*) and Nelson's (1991) EGARCH(*p,q*). The EGARCH(*p,q,m*) model is the first choice model and has a number of attractive properties. Firstly, it may be argued that this specification nests Engle's (1982) short-memory ARCH model and Bollerslev's (1986) long-memory GARCH model as it incorporates both the ARCH and GARCH parameters (see Dowd, 2005). Secondly, the specification permits for asymmetries in volatility and therefore nests a test for the presence of the leverage effect. In the EGARCH(*p,q,m*) specification, m denotes the asymmetry parameter. Finally, it does not impose positivity restrictions on the ARCH and GARCH coefficients (Nelson and Cao, 1992; Francq, Wintenberger and Zakoïan, 2013). The IGARCH(*p,q*) specification is treated as a second choice model and is applied if the ARCH and GARCH parameters of the EGARCH(*p,q*) specification are equal to or greater than unity ($\alpha_i + \beta_j \geq 1$) implying that variance is nonstationary and that the EGARCH(*p,q,m*) specification is misspecified for that particular series (Zivot, 2009).

4. Results

4.1. The properties of returns and variance

Table 2 reports the results of the analysis of the distributional properties of the return series in the sample.

The kurtosis coefficient for the returns on the JSE All Share Index is above 3 at 3.631. For the economic sector return series, all series exhibit levels of kurtosis greater than 3 although the null hypothesis of normally distributed returns is not rejected for two series; consumer services and telecommunications. For the industrial sector indices, 25 out of 26 series exhibit a kurtosis above 3. Xiao and Aydemir (2007) and Engle and Patton (2007) state that it is common to find that levels of kurtosis in financial return series are above 3 and the results in Table 2 attest to this. The presence of widespread leptokurtosis also suggests that variance is non-stationary (Akgiray, 1989).

Return distributions are also asymmetric; the average level of skewness for the economic sector and industrial sector return series is -0.397 and -0.244 respectively, suggesting that negatively skewed distributions are more prevalent than positively skewed distributions. Nevertheless, isolated instances of positive skewness are observed for returns on industrial metals and mining, industrial engineering, food producers, pharmaceuticals and biotechnology and nonlife insurance industrial sector indices. The remainder are characterized by negative skewness. In line with these findings, returns on the JSE All Share Index also exhibit negative skewness.

Table 2. Distributional properties

	Obs.	Mean	Std Dev	Skewness	Kurtosis	JB Test Statistic
JSE All Share Index	192	0.005	0.048	-0.305	3.631	6.156**
Economic Sector Index						
1. Basic Materials	192	0.003	0.069	-0.505	4.020	16.482***
2. Industrials	192	0.005	0.049	-0.551	3.941	16.793***
3. Consumer Goods	190	0.009	0.057	-0.277	3.687	6.167**
4. Health Care	192	0.009	0.053	-0.445	3.308	7.089**
5. Consumer Services	192	0.011	0.055	-0.360	3.049	4.155
6. Telecommunication	192	0.004	0.074	-0.141	3.209	0.989
7. Financials	192	0.004	0.048	-0.349	3.611	6.890**
8. Technology	190	0.001	0.086	-0.547	4.961	39.921***
Industrial Sector Index						
1.1.Chemicals	191	0.007	0.053	-0.503	4.237	20.230***
1.2. Forestry & Paper	191	0.006	0.083	-0.327	3.018	3.410
1.3. Ind. Metals & Mining	190	0.007	0.110	0.220	4.276	14.422***
1.4. Mining	192	0.001	0.081	-0.225	3.506	3.700
2.1. Const. & Materials	192	0.002	0.070	-0.495	3.876	13.994***
2.2. General Industrials	192	0.009	0.051	-0.514	4.071	17.637***
2.3. Elec. & Elec. Equip.	192	0.001	0.055	-0.460	4.297	20.230
2.4. Indust. Engineering	190	0.013	0.060	0.144	4.547	19.615***
2.5. Indust. Transp.	192	0.004	0.065	-0.591	4.245	23.581***
2.6. Support Services	192	0.000	0.056	-0.468	3.370	8.087**
3.1. Automobiles & Parts	191	0.001	0.085	-0.095	3.887	6.548**
3.2. Beverages	192	0.008	0.060	-0.117	4.132	10.692***
3.3. Food Producers	192	0.009	0.044	0.090	3.115	1.655
4.1. Health Care Eq & Serv.	192	0.013	0.059	-0.128	3.939	7.576**
4.2. Pharm & Biotech.	192	0.013	0.073	0.068	3.486	2.036
5.1. Food & Drug Retailers	192	0.010	0.058	-0.218	3.622	4.608***
5.2. General Retailers	192	0.008	0.066	-0.220	3.008	1.554
5.3. Media	191	0.017	0.081	-0.293	2.941	2.757
5.4. Travel & Leisure	192	0.005	0.055	-0.629	4.173	23.679***
6.1. Fixed-line Telecoms	192	0.003	0.092	-0.416	3.545	7.899**
7.1. Banks	192	0.005	0.062	-0.061	3.277	0.735
7.2. Non-life Insurance	191	0.006	0.055	0.196	3.149	1.394
7.3. Life Insurance	192	0.002	0.059	-0.483	4.140	17.859***
7.4. General Financial	192	0.003	0.064	-0.192	4.402	16.897***
7.5. Equity Investment Inst.	192	0.005	0.044	-0.195	4.023	9.584***
8.1. Software & Comp. Serv.	189	0.002	0.088	-0.428	4.817	31.773***

Notes: *** Indicates statistical significance at the 1 percent level of significance. ** Indicates statistical significance at the 5 percent level of significance. * Indicates statistical significance at the 10 percent level of significance. Obs. refers to the number of observations. Numbers below 192 indicate that outliers have been purposefully omitted.

In contrast to the findings of Simkowitz and Beedles (1980: 10) who state that “securities display a habitual tendency to positive skewness,” returns on South African economic sectors and industrial sector indices display a habitual tendency towards negative skewness. On the basis of the JB test, the null hypothesis of normally distributed returns is rejected for the JSE All Share Index, six out of the eight economic sector indices and for 17 out of the 26 industrial sector indices. These findings are in line with those of Mangani (2007) who finds that the returns on the JSE All Share Index and a portfolio of individual stocks exhibit negative skewness and levels of kurtosis in excess of 3. For individual stocks, 25 out of 42 return series exhibit negative skewness and all series exhibit excess kurtosis.

Table 3. Serial correlation structure and ADF test statistics

	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	Q(5)	Q(10)	ADF Test
JSE All Share Index	-0.028	0.063	0.099	0.050	-0.080	4.623	7.237	-13.991***
Economic Sector Index								
1. Basic Materials	-0.042	0.163*	0.068	0.091	-0.144*	12.254**	15.897	-14.326***
2. Industrials	0.062	0.015	-0.009	-0.057	0.011	1.492	4.724	-12.810***
3. Consumer Goods	-0.071	0.033	0.195*	-0.135	0.050	12.811**	19.955**	-14.803***
4. Health Care	0.059	0.044	-0.034	0.124	0.130	7.694	12.029	-12.845***
5. Consumer Services	0.036	0.061	-0.109	-0.119	0.139	10.007*	11.716	-12.823***
6. Telecommunication	0.043	-0.012	0.127	-0.053	0.045	4.510	6.107	-13.548***
7. Financials	0.034	-0.022	-0.055	0.098	0.031	2.991	9.705	-13.372***
8. Technology	0.179*	0.040	0.044	0.095	0.101	10.792*	20.101**	-11.531***
Industrial Sector Index								
1.1.Chemicals	-0.049	0.027	-0.037	0.164*	-0.055	7.027	9.413	-14.231***
1.2. Forestry & Paper	-0.007	-0.137	0.149*	0.190*	-0.071	16.235***	18.851**	-13.884***
1.3. Ind. Metals & Mining	0.035	0.242*	-0.004	0.218*	-0.012	21.127***	23.098**	-7.501***
1.4. Mining	-0.038	0.139*	0.046	0.037	-0.142*	8.808	12.073	-14.269***
2.1. Const. & Materials	0.177*	0.099	-0.061	0.064	0.151	14.058**	20.855**	-11.528***
2.2. General Industrials	-0.031	-0.048	0.016	-0.146*	-0.032	5.1252	11.575	-13.802***
2.3. Elec. & Elec. Equip.	0.147*	-0.090	0.091	0.032	0.023	7.749	18.490**	-11.889***
2.4. Indust. Engineering	0.152*	0.007	0.005	0.140*	0.165*	13.860**	14.592	-11.621***
2.5. Indust. Transp.	-0.000	0.106	-0.093	0.115	0.059	7.2055	12.418	-13.741***
2.6. Support Services	0.057	0.107	-0.087	-0.004	0.004	4.3767	7.605	-13.050***
3.1. Automobiles & Parts	0.117	-0.027	-0.052	0.083	0.084	6.095	18.430**	-12.292***
3.2. Beverages	-0.050	-0.11	-0.048	-0.008	-0.053	1.539	7.309	-14.313***
3.3. Food Producers	0.057	-0.055	-0.018	-0.034	-0.053	2.077	2.722	-12.881***
4.1. Health Care Eq & Serv.	0.050	0.021	0.026	0.165*	0.186*	13.039**	15.203	-12.609***
4.2. Pharm & Biotech.	-0.014	-0.045	-0.011	0.103	-0.071	3.559	6.636	-13.638***
5.1. Food & Drug Retailers	-0.032	-0.025	-0.046	-0.102	0.048	3.228	5.106	-14.396***
5.2. General Retailers	0.105	-0.023	-0.003	-0.099	0.040	4.520	5.832	-12.273***
5.3. Media	0.025	-0.059	-0.141*	-0.054	0.215*	14.533**	24.733***	-13.223***
5.4. Travel & Leisure	0.103	0.033	-0.01	0.125	0.023	5.513	8.339	-12.365***
6.1. Fixed-line Telecoms	0.114	0.076	0.073	-0.022	0.162*	10.014*	15.840	-12.579***
7.1. Banks	-0.070	-0.012	-0.093	0.038	-0.040	3.297	6.794	-14.780***
7.2. Non-life Insurance	0.019	-0.161*	-0.136	-0.009	0.051	9.367*	18.812**	-13.453***
7.3. Life Insurance	0.013	0.013	0.019	0.074	0.114	3.813	9.041	-13.660***
7.4. General Financial	0.132*	-0.053	0.031	0.066	0.132*	8.493	13.637	-12.106***
7.5. Equity Investment Inst.	-0.091	0.135*	0.027	0.057	0.036	6.226	15.726	-15.312***
8.1. Software & Comp. Serv.	0.189*	0.045	0.036	0.081	0.094	10.690*	21.747**	-11.407***

Notes: For the Q-statistics and ADF test, *** indicates statistical significance at the 1 percent level of significance, ** indicates statistical significance at the 5 percent level of significance and * indicates statistical significance at the 10 percent level of significance. For individual correlation coefficients * indicates statistical significance at the 5 percent level of significance. The ADF test applied here assumes a strict random walk process. Lag selection is based upon the SIC. The results of the ADF test are validated by the Phillips-Perron (PP) test as in Sadorsky and Henriques (2001). The PP supports the conclusions of the ADF test for all return series. See Table A1.1 in Appendix 1.4. Outliers not excluded.

The results presented in this paper support Mangani's (2007) results; departures from normality in the form of excess kurtosis, skewness or both are widely observed in return series comprising the South African stock market. The normality assumption does not appear to be a valid working hypothesis for the South African stock market.

The results in Table 3 indicate that returns on the JSE All Share Index are uncorrelated and the Q-statistics for the first five and 10 orders are statistically insignificant, suggesting that the assumption of (linear) independence holds for the JSE All Share Index return series. For the economic and industrial sector indices, one and five serial correlation coefficients are statistically significant at the first order respectively although isolated instances of statistically significant serial correlation are observed at higher orders. While most are of a small magnitude, the largest higher-order statistically significant serial-correlation coefficients for the economic and industrial sector indices are observed for returns on the consumer goods economic sector index at 0.195 (ρ_3) and for the mining and metals industrial sector index at 0.242 (ρ_2).It is questionable whether this level of correlation is significant from a practical perspective; it is most likely too low to exploit for trading purposes (see Fama, 1965). The null hypothesis that the first five and 10 serial-correlation coefficients are jointly equal to zero is rejected in four and two instances for the economic sector series respectively. The null hypothesis is rejected in nine and eight instances respectively for the industrial sector return series. In instances where Q-statistics are statistically significant, an analysis of serialcorrelation coefficients suggests that the respective Q-statistics may be biased upwards by large individual serial-correlation coefficients. Unlike Mangani (2007), who concludes that returns are far from independent on a basis of a finding that 16 of the 42 individual stock return series are serially correlated, the results presented here are more moderate; while overall the assumption of independence holds, violations of this assumption occur in isolated instances. Even in instances where the assumption of statistical independence is violated, prior returns are unlikely to be useful in predicting price changes, given the low and isolated instances of significant correlation. Therefore, independence remains a working assumption, as opposed to a purely statistical one, for the South African stock market.

Finally, results of the ADF test suggest that the presence of a unit root may be rejected for returns on the JSE All Share Index, the economic and industrial sector indices. This implies that the return series constituting the sample are stationary as expected. This finding is in agreement with Mangani (2007) who states that the unit-root hypothesis is rejected for all return series in the sample.

The results in Table 4 show that returns on the JSE All Share Index exhibit non-linear dependence, suggestive of time-varying variance and ARCH effects and that returns on most economic sectors and industrial sector indices are characterized by some form of non-linear dependence in returns or ARCH effects in the residuals of the AR(1) model. Statistically significant non-linear dependence in returns or statistically significant ARCH effects in the residuals are detected in seven out of the eight economic sectors and in 21 of the 26 industrial sectors.

Higher-order ARCH(5) and ARCH(10) effects are more prevalent than ARCH(1) effects; four economic sectors and five industrial sectors exhibit ARCH(1) effects in comparison to six economic sectors and 12 industrial sectors, which exhibit ARCH(10) effects. The frequency of statistically significant ARCH(5) effects is approximately equal to that of ARCH(1) effects in the economic sector series and the frequency of significant ARCH(5) effects for industrial sector return series exceeds that of ARCH(1) effects. While Mangani (2007) suggests that returns on the JSE may be modelled as ARCH-type processes, this aspect is not investigated further. In contrast, the results presented here indicate that returns on the South African stock market should indeed be seen and modelled as ARCH-type processes.

Table 4. ARCH effects and the leverage effect

	$Q^2(15)$	ARCH(1)	ARCH(5)	ARCH(10)	$L_1(\tau)$	$L_2(\tau)$
JSE All Share Index	62.586***	7.073***	3.009**	5.500***	-0.280***	-0.029
Economic Sector Index						
1. Basic Materials	39.135***	6.338**	5.226***	3.655***	-0.146**	-0.138*
2. Industrials	27.900**	0.629	1.674	22.159**	-0.033	0.099
3. Consumer Goods	67.927***	0.096	24.560***	42.482***	-0.152**	-0.039
4. Health Care	33.097***	0.028	2.172*	2.048**	-0.0107	-0.029
5. Consumer Services	10.586	0.000	0.731	0.921	-0.087	0.001
6. Telecommunication	25.321**	4.366**	1.668	1.451	-0.042	0.065
7. Financials	58.927***	3.783*	6.850***	4.705***	-0.252***	0.070
8. Technology	68.331***	2.815*	6.838	3.742	-0.246***	-0.135*
Industrial Sector Index						
1.1.Chemicals	31.817***	4.398**	1.605	1.486	-0.151**	-0.100
1.2. Forestry & Paper	18.851	0.000	2.729**	1.406	-0.062	0.018
1.3. Ind. Metals & Mining	19.926	0.815	3.153***	1.556	-0.086	0.030
1.4. Mining	23.918*	1.885	2.263*	2.066**	-0.079	-0.018
2.1. Const. & Materials	28.242**	11.057***	2.788**	2.226**	-0.145**	-0.087
2.2. General Industrials	17.621	0.264	0.626	1.734*	0.006	0.030
2.3. Elec. & Elec. Equip.	36.162***	0.554	5.654***	3.182***	0.015	-0.059
2.4. Indust. Engineering	32.322***	0.089	3.903***	2.250**	-0.141*	-0.033
2.5. Indust. Transp.	24.142*	0.304	2.086*	1.488	-0.101	0.008
2.6. Support Services	20.821	0.720	0.746	0.557	-0.114	0.000
3.1. Automobiles & Parts	6.472	0.012	0.171	0.340	-0.138*	-0.039
3.2. Beverages	4.6764	0.199	0.145	0.443	-0.064	-0.025
3.3. Food Producers	12.447	0.001	2.061*	1.412	-0.085	-0.028
4.1. Health Care Eq & Serv.	21.844	0.243	2.601**	1.582	0.017	-0.010
4.2. Pharm & Biotech.	37.322***	0.174	0.288	2.558***	-0.053	-0.003
5.1. Food & Drug Retailers	13.679	0.064	0.400	0.370	-0.002	0.015
5.2. General Retailers	8.831	0.009	0.896	0.975	-0.053	0.007
5.3. Media	29.528**	1.812	1.501	4.217***	-0.117	0.015
5.4. Travel & Leisure	18.228	1.854	2.183*	1.446	-0.123*	0.018
6.1. Fixed-line Telecoms	20.983	0.200	1.769	1.714*	0.017	0.107
7.1. Banks	43.946***	2.794*	6.024***	3.236***	-0.228**	-0.018
7.2. Non-life Insurance	19.205	0.364	1.973*	1.176	0.105	-0.127*
7.3. Life Insurance	51.191***	0.002	6.063***	4.532***	-0.153**	-0.014
7.4. General Financial	32.439***	3.637*	3.111***	1.695*	-0.162**	0.150*
7.5. Equity Investment Inst.	9.299	0.236	0.260	1.065	-0.131*	0.055
8.1. Software & Comp. Serv.	77.734***	2.923*	6.530***	4.633***	-0.254***	-0.144**

Notes: *** Indicates statistical significance at the 1 percent level of significance. ** Indicates statistical significance at the 5 percent level of significance. * Indicates statistical significance at the 10 percent level of significance. Leverage effect established by testing whether the correlation between squared returns and lagged returns (squared returns) is statistically significant. The lag order is 1 .The results of the ARCH LM test for ARCH effects at the 1st order are validated by the White test. While the results of the ARCH LM and White tests are for the most part consistent, the results of the White test suggest that the ARCH LM test may (slightly) understate the presence of ARCH effects.

Finally, there is evidence of an infrequent leverage effect as evident from casual returnvolatility relationships. The return-volatility correlation coefficient, $L1(\tau)$, is significant for the JSE All Share Index and for four of the eight economic sectors and 10 industrial sectors. A finding of a statistically significant volatility-return correlation coefficients, L2(τ) , is limited to two economic groups and three industrial sectors. This suggests that the direction of the leverage effect, where present, is predominantly from returns to volatility – volatility increases following a decrease in stock prices. This is in line with the directionality suggested by Bouchaud, Matacz and Potters (2001) who report that correlation is between past returns and future volatility, implying that a decrease in returns results in an increase in volatility.

4.2 ARCH/GARCH modelling

The results in Table 5 are for the EGARCH(*p*,*q*,*m*) or IGARCH(*p*,*q*) models fitted to returns on the JSE All Share Index, the economic and the industrial sectors. An EGARCH(1,1,1) with a conditional general error-distribution (G.E.D) is fitted to returns on the JSE All Share Index, suggesting that the variance underlying returns on the JSE All Share Index is asymmetric and has a long memory, as suggested by the statistically significant ARCH and GARCH coefficients. The presence of a leverage effect in JSE All Share Index returns is indicated by a significant coefficient of asymmetry, $\boldsymbol{\gamma}$, of -0.327 and the half-life, ζ, is 9.137 months suggesting that shocks to the conditional variance persist for an extended period of time and that it takes 9.137 months for half the shock to decay. The persistence parameter, κ , of 0.927 also suggests that shocks are highly persistent (Engle, 2001; McMillan and Ruiz, 2009). It is not uncommon to observe high-levels of persistence in national markets; Ding, Granger and Engle (1993) report that shocks to variance for the NYSE and German stock market (DAX) take an exceptionally long period of time to subside.

Most of the conditional variance structures underlying returns on the economic sectors are described by the EGARCH(1,1,1) model, with the exception of the industrials, health care and telecommunications industrial sectors, which are described by IGARCH(1,1) specifications. The IGARCH(1,1) is an infinite-variance model suggesting that any shocks to the variance of returns on these economic sectors remain important for all forecast horizons (and hence no half-life is reported as shocks never die out) and that conditional variance does not converge to unconditional variance – it is non-mean reverting (Xiao and Aydemir, 2007; Kirchgässner and Wolters, 2007; Kang, Kang and Yoon, 2009). The GARCH parameters are statistically significant for each EGARCH(*p*,*q*,*m*) model that is applied to the respective economic sectors, suggesting the conditional variance generally exhibits long memory for these sectors but that it not explosive (non-stationary). Notably, the ARCH parameters are statistically significant for the consumer goods and technology economic sectors suggesting that these series also exhibit lower order ARCH effects and volatility-clustering (see discussion below).

The sum of the coefficient of the EGARCH(2,1,1) model fitted to the basic material return series suggests that the variance underlying the returns on this series has long finite-memory; this is attested to by a κ of 0.783 and a half-life of 2.837 months. The variance of returns on the financials economic sector also appears to have a long half-life (longer than that of the basic materials economic sector) of 2.976 months and shocks are persistent as evident from a κ of 0.792. All sectors to which the EGARCH(*p*,*q*,*m*) specifications are applied are characterized by conditional asymmetry, as evident from statistically significant asymmetry coefficients all five economic sectors. Of the series that are described by EGARCH(*p*,*q*,*m*) specifications, the consumer services economic sector has the shortest half-life of 1.102 months. As in McMillan and Ruiz (2009), the variance of the economic group series exhibits varying levels of finite persistence although the industrials, health care and telecommunications sectors require the use of the IGARCH(*p*,*q*) suggesting that the conditional variance underlying these sectors is nonstationary and that shocks never die out. The normal distribution is the most appropriate distribution for six economic sectors although the basic materials and industrials sectors are characterized by the student's t distribution suggesting that the normal distribution may not always be the best approximation of the distribution of the residuals.

The conditional variance for the industrial sectors is most appropriately described by an EGARCH(1,1,1) specification (16 sectors). ARCH/GARCH modelling does not appear to be appropriate for the general industrials sector series. This sector does not exhibit non-linear dependence in the residuals and displays a weak ARCH(10) effect. Consequently, the ARCH and GARCH coefficients are statistically insignificant and the model appears to be misspecified. However, all series that are described by an EGARCH(1,1,1) specification exhibit significant GARCH coefficients whereas a number of sectors exhibit statistically significant ARCH and GARCH coefficients. Examples of sectors that exhibit significant GARCH coefficients but not ARCH coefficients are the mining, industrial transportation and the support services industrial sectors. Examples of series that exhibit statistically significant ARCH and GARCH coefficients are the construction and materials, food and drug and the media industrial sectors. A closer (unreported) analysis of non-linear dependence and ARCH effects in the residuals of the AR(1) model in equation (2) for series for which only the GARCH coefficient and for which both ARCH and GARCH coefficients are statistically significant is undertaken. This analysis indicates that series for which only GARCH coefficients are significant are characterized by higher order non-linear dependence and higher order ARCH effects. Series for which both ARCH and GARCH coefficients are significant, are characterized by both lower and higher order non-linear dependence and ARCH effects.

The widespread statistical significance of the GARCH coefficients also suggests that the conditional variance of the industrial sectors is characterized by long-memory and the long-run persistence of shocks. Furthermore, the widespread statistical significance of the coefficient of asymmetry, for i.e. mining, industrial transport, support services and other industrial sectors, suggests that the variance underlying a substantial number of South African industrial sectors is conditionally asymmetric in character. A total of 10 industrial sectors exhibit evidence of conditional asymmetry. The series with the greatest (finite) half-life is the chemical industrial sectors with ξ equal to 106.653 suggesting that the conditional variance underlying this sector may potentially also be described more adequately by an IGARCH(p,q) specification (see Franke, Härdle and Hafner,, 2011). The persistence parameter for this sector is 0.994. The series with the second highest finite half-life of 4.090 months is the forestry and paper sector. The conditional variance for this series is described by an EGARCH(1,2,1) model with a persistence parameter of 0.844. The higher number of GARCH parameters (q=2) supports a finding of a longer half-life and suggests that this series is characterized by a longer memory relative to the other series described by the EGARCH(p,q,m) specification. The remaining nine sectors, i.e. industrial metals and mining, construction and materials, electronic and electrical equipment ,etc, are described by the IGARCH(p,q) model which indicates the presence of the IGARCH effect - the infinite persistence of shocks to variance. This suggests that a number of industrial sectors that comprise the South African stock market have an explosive variance and that shocks persist infinitely – a characteristic that would not be captured by a finite variance model that assumes mean reversion.

The conditional error distribution fitted to 11 of the series is the normal distribution, generally a useful general approximation for conditional errors. However, the most appropriate distribution for eight industrial sectors is the generalized error distribution and for seven industrial sectors, the most appropriate conditional error distribution is the student's t distribution suggesting that the normal distribution is not always the best approximation of the residuals. This further suggests that other distributions, such as the student's t and generalized error distribution, should be considered.

Table 5. ARCH/GARCH modelling

	ω	α_1	α_2	β_1	β_2	γ	Model	Dist.	AIC	κ	ξ
JSE All Share Index	-1.739	0.469***		0.784***		-0.327***	EGARCH(1,1,1)	G.E.D	-3.386	0.927	9.137
Economic Sector Index											
1. Basic Materials	-2.206	-0.096	0.469**	0.653***	-	-0.243***	EGARCH(2,1,1)	Student's t	-2.563	0.783	2.837
2. Industrials	-	0.507***	-	0.943***	-	-	IGARCH(1,1)	Student's t	-3.213	1.000	-
3. Consumer Goods	-0.170	-0.200***	-	0.945***	-	-0.197***	EGARCH(1,1,1)	Normal	-2.998	0.547	1.150
4. Health Care	-	0.110***	-	0.889***	-	-	IGARCH(1,1)	Normal	-3.040	1.000	-
5. Consumer Services	-1.122	-0.093	-	0.796***	-	-0.170**	EGARCH(1,1,1)	Normal	-2.949	0.533	1.102
6. Telecommunication	-	0.0886**	-	0.911***	-	-	IGARCH(1,1)	Normal	-2.411	1.000	-
7. Financials	-1.267	0.175	-	0.821***	-	-0.204**	EGARCH(1,1,1)	Normal	-3.325	0.792	2.976
8. Technology	-1.01	-0.150***	-	0.961***	-	-0.130***	EGARCH(1,1,1)	Normalt	-2.338	0.681	1.8012. 837

Table 5. ARCH/GARCH modelling (continued)

Industrial Sector Index											
1.1.Chemicals	-0.768**	0.224*	-	0.901***	-	-0.131	EGARCH(1,1,1)	G.E.D	-3.059	0.994	106.653
1.2. Forestry & Paper	-0.932	0.199*	-	0.006	0.841***	-0.201**	EGARCH(1,2,1)	Normal	-2.063	0.844	4.090
1.3. Ind. Metals & M.	-	0.114***	-	0.886***	-	-	IGARCH(1,1)	Student's t	-1.500	1.000	-
1.4. Mining	-1.170	0.199	-	0.802***	-	-0.168**	EGARCH(1,1,1)	Normal	-2.190	0.833	3.803
2.1. Const. & Materials		0.066***	-	0.934***	-	-	IGARCH(1,1)	G.E.D	-2.507	1.00	-
2.2. General Industrials	-7.381	-0.188	-	-0.259	-	0.043	EGARCH(1,1,1)	G.E.D	-3.086	-	-
2.3. Elec. & Elec. Equip.	-	0.005	-	0.995***	-	-	IGARCH(1,1)	Normal	-2.953	1.000	-
2.4. Indust. Engineering	-	0.063***	-	0.937***	-	-	IGARCH(1,1)	G.E.D	-2.619	1.000	-
2.5. Indust. Transp.	-0.634	0.042	-	0.894***	-	-0.160***	EGARCJ(1,1,1)	Normal	-2.689	0.775	2.723
2.6. Support Services	-1.219	-0.038	-	0.785***	-	-0.150*	EGARCH(1,1,1)	Normal	-2.914	0.597	1.342
3.1. Automobiles & Parts	-0.574*	-0.030	-	0.881***	-	-0.089	EGRACH(1,1,1)	Student's t	-2.061	0.762	2.554
3.2. Beverages	-1.120	0.129	-	0.820***	-	-0.187	EGARCH(1,1,1)	G.E.D	-2.793	0.761	2.541
3.3. Food Producers	-2.556	0.052	-	0.598*	-	-0.154	EGARCH(1,1,1)	Normal	-3.362	0.497	0.991
4.1. Health Care Eq &		0.122***	-	0.878***	-	-	IGARCH(1,1)	G.E.D	-2.871	1.000	-
4.2. Pharm & Biotech.		0.064***	-	0.936***	-	-	IGARCH(1,1)	Normal	-2.375	1.000	-
5.1. Food & Drug	-0.164***	-0.182***	-	0.947***	-	-0.065***	EGARCH(1,1,1)	G.E.D	-2.866	0.700	1.943
5.2. General Retailers	-0.958*	-0.157	-	0.805***	-	-0.151**	EGARCH(1,1,1)	Normal	-2.599	0.497	0.990
5.3. Media	-0.218***	-0.219***	-	0.920***	-	-0.129***	EGARCH(1,1,1)	G.E.D	-2.197	0.572	1.242
5.4. Travel & Leisure	-1.113	0.120	-	0.828***	-	-0.136	EGARCH(1,1,1)	Student's t	-2.990	0.812	3.332
6.1. Fixed-line Telecoms		0.059***	-	0.941***	-		EGARCH(1,1,1)	Student's t	-1.951	1.000	-
7.1. Banks	-1.251**	0.092	-	0.793***	-	-0.229**	EGARCH(1,1,1)	Normal	-2.771	0.656	1.646
7.2. Non-life Insurance	-	0.003	-	0.997***	-	-	IGARCH(1,1)	Student's t	-2.841	1.000	-
7.3. Life Insurance		0.112***	-	0.888***	-	-	IGARCH(1,1)	Student's t	-2.830	1.000	-
7.4. General Financial		0.120***	-	0.879***	-	-	IGARCH(1,1)	Normal	-2.777	1.000	-
7.5. Equity Investment	-0.947**	-0.037	-	0.846***	-	-0.143***	EGARCH(1,1,1)	Normal	-3.418	0.666	1.708
8.1. Software & Comp.	-0.078*	-0.141***	-	0.966***	-	-0.115***	EGARCH(1,1,1)	Student's t	-2.274	0.711	2.028

Notes: *** Indicates statistical significance at the 1 percent level of significance. ** Indicates statistical significance at the 5 percent level of significance. * Indicates statistical significance at the 10 percent level of significance.

4.3 Implications

The results of this paper point towards a number of important implications for the modelling of financial return series. An immediate observation is that South African return series are not “well-behaved”; returns exhibit excess kurtosis, are generally negatively skewed and are therefore not normally distributed, as evident from the results in Table 2. In light of the literature, this is to be expected. Departures from non-normality are in themselves not problematic in model estimation. However, Wong and Bian (2003) argue that the characteristics the return series may carry over into the residuals and impact hypothesis tests. For example, Roll (1992) and Ford (2003) suggests that if the residuals of model are not normally distributed, then t-tests and F-tests based upon estimated standard errors used in inference making will be misleading. Furthermore, leptokurtosis may be indicative of nonstationary variance and which calls for the application of an appropriate econometric methodology that is able to account for this non-stationarity (Engle, 2001; Jacobsen and Dannenburg, 2003). Additionally, when using parametric methods that require the specification of a conditional error distribution, such as the ARCH/GARCH framework - which may be used to model the very same time-varying variance that leptokurtosis is indicative of - the misspecification of the conditional error distribution will lead to a misspecification of the log-likelihood function and will result in inconsistent parameter estimates (Herwartz, 2004). Although the misspecification of the conditional error distribution may be overcome by using quasi-maximum likelihood estimation techniques, robustness will come at the cost of a loss in efficiency (Fan, Qi and Xiu, 2014).

The results in Table 3 show that return series exhibit a limited degree of dependence as exemplified by significant serial correlation. The presence of serial correlation has the potential to create spurious relationships although it is questionable whether the levels of serial correlation in the present sample are problematic. Nevertheless, if conventional time-series regression techniques are applied, the recommendation is that consideration is given to the use of unexpected components of the return series (Poon and Taylor, 1991). This may be done by pre-whitening series prior to model estimation (see Priestley, 1996). Also, Studenmund (2011) suggests that the presence of serial correlation will lead to a bias in the standard errors which will result in misleading inferences. The results in Table 4 point towards the presence of nonlinear dependence, which is associated with the presence of heteroscedasticity and is suggestive of time-varying variance. The presence of conditional heteroscedasticity is confirmed by the ARCH LM test that is applied at the 1, 5 and 10th orders. If the least-squares methodology is applied in the presence of heteroscedasticity, coefficient estimates will no longer be efficient and variance estimators will be biased, resulting in misleading statistical inferences (Gujarati, 2003). In the presence of heteroscedasticity and serial correlation, specifications may still be estimated using the least-squares methodology but with robust standard errors such as NeweyWest (1987) HAC (heteroscedasticity and serial-correlation consistent) standard errors (see Andersen, Bollerslev, Diebold and Vega, 2003). Alternatively, the ARCH/GARCH framework, which treats heteroscedasticity as variance to be modelled may be applied (Engle, 2001).

Table 4 also suggests that South African stock returns exhibit a leverage effect and the leverage effect may be viewed as a characteristic of the variance underlying South African return series. Furthermore, considered together, the results in Table 4 and Table 5 point towards the presence of asymmetry, long memory and high levels of persistence in the variance and in some instance explosive variance – characteristics that may be captured by the ARCH/GARCH framework. In Table 5, this is exemplified by the widespread appropriateness of the EGARCH(p,q,m) model that is in essence a long-memory model which captures asymmetry and the appropriateness of the IGARCH(p,q) specification that model non-stationary variance. These characteristic have the potential impact model coefficient estimates. Evidence that the structure of conditional variance has an impact on coefficient estimates is provided by Bera, Bubnys and Park (1988), Hamilton (2010) and Armitage and Brzeszczynski (2011). Specifically, Hamilton (2010) suggests that the failure to consider the structure of the variance will result in misleading inferences and that incorporating the observed features of heteroscedasticity will yield substantially more efficient coefficient estimates. Finally, the results in Table 5 suggest that although the normal distribution may be appropriate in describing the conditional error distribution for a substantial number of industrial sector series, other distributions should be considered if a parametric modelling approach is applied.

5. Conclusion

This paper investigates the properties of South African stock returns and the structure of the variance underlying the return series using a comprehensive sample. The evidence points towards widespread departures from normality, in line with observations made for other markets. The evidence also suggests that variance is non-constant, heteroscedastic, asymmetric and characterized by long memory and the persistence of shocks. The structure of the variance ranges between relatively simple as suggested EGARCH(p,q,m) specifications that do not reflect the presence of asymmetry to relatively complex as suggested by the presence of asymmetry and IGARCH effects. It is recommended that the structure of variance is taken into consideration when modelling the mean in a time-series context and a failure to do so may result in misleading inferences. Furthermore, if relying upon a parametric approach when modelling return behavior, consideration should be given to the conditional error distribution. In line with the initial results and accompanying analysis for Table 2 and Table 4, it appears that non-normality, asymmetry and long memory are also reflected by the EGARCH(p,q,m) and IGACRH(p,q) specifications and selected conditional error distributions.

Supplementary materials: Additional appendices are available from the authors.

Author Contributions: Both authors contributed equally to the research.

Funding: This research received no external funding.

Data Availability Statement: The data in this research comes from widely accessible databases.

Conflicts of Interest: The authors declare no conflict of interest.

References

Agung, I. G. N. (2009). Time series analysis using Eviews. John Wiley and Sons.

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: evidence and forecasts. *Journal of Business*, 62(1), 55–88.

Anagnostidis, P., & Emmanouilides, C. J. (2015). Nonlinearity in high-frequency stock returns: evidence from the Athens Stock Exchange. *Physica A: Statistical Mechanics and its Applications*, 421, 473–487.

Andersen, T.G., Bollerslev, T., Diebold, F.X., & Vega, C. (2003). Micro effects of macro announcements: real-time price discovery in foreign exchange. *The American Economic Review*, 93(1), 38–62.

Armitage, S., & Brzeszczynski, J. (2011). Heteroscedasticity and interval effects in estimating beta: UK evidence. *Applied Financial Economics*, 21(20), 1525–1538.

Bachalier, L.J.B.A. (1914). Le Jeu, la Chance, et le Hasard. E. Flammarion.

Barreto, H., & Howland, F.M. (2006). *Introductory econometrics: Using Monte Carlo simulation with Microsoft Excel*. Cambridge University Press.

Baur, D.G., Dimpfl, T., & Jung, R.C. (2012). Stock return autocorrelations revisited: a quantile regression approach. *Journal of Empirical Finance*, 19(2), 254–265.

Bera, A., Bubnys, E., & Park, H. (1988). Conditional heteroscedasticity in the market model and efficient estimates of betas. *Financial Review*, 23(2), 201–214.

Berry, M.A., Burmeister, E., & McElroy, M.B. (1988). Sorting out risks using known apt factors. *Financial Analysts Journal*, 44(2), 29–42.

Bhattacharya, S. N., & Bhattacharya, M. (2012). Long memory in stock returns: A study of emerging markets. *Iranian Journal of Management Studies*, 5(2), 67–88.

Black, F. (1976). Studies of stock market volatility changes. Paper presented at the 1976 Meetings of the American Statistical Association, Business and Economic Statistics Section.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3), 307–327.

Boubaker, A., & Makram, B. (2012). Modelling heavy tails and double long memory in North African stock market returns. *The Journal of North African Studies*, 17(2), 195–214.

Bouchaud, J-P., Matacz, A., & Potters, M. (2001). Leverage effects in financial markets: the retarded volatility model. *Physical Review Letters*, 87(22), 1–4.

Box, G.E., & Pierce, D.A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American statistical Association*, 65(332), 1509–1526.

Brown, S.J., & Warner, J.B. (1985). Using daily stock returns: the case of event studies. *Journal of Financial Economics*, 14(1), 3–31.

Campbell, J.Y., Lo, A.W., & MacKinlay, A.C. (1997). *The econometrics of financial markets*. Princeton University Press.

Chkili, W., Aloui, C., & Nguyen, D.K. (2012). Asymmetric effects and long memory in dynamic volatility relationships between stock returns and exchange rates. *Journal of International Financial Markets, Institutions and Money*, 22(4), 738–757.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2), 223–236.

Cryer, J.D., & Chan, K-S. (2008). *Time series analysis with applications in R (2nd ed).* Springer.

Desjardins, J. (2016). All of the world's stock exchanges. Retrieved from http://money.visualcapitalist.com/all-of-theworlds-stock-exchanges-by-size/ on September 6, 2017.

Desjardins, J. (2017). The 20 largest stock exchange in the world. Retrieved from http://www.visualcapitalist.com/20-largest-stock-exchanges-world/ on September 6, 2017.

Ding, Z., Granger, C.W.J., & Engle, R.F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1), 83–106.

Franke, J., Härdle, W.K., & Hafner, C.M. (2011). *Statistics of Financial Markets. An Introduction (4th ed).* Springer.

Ford, G.S. (2003). Daily stock returns, non-normality and hypothesis testing. Retrieved September 6, 2017, from http://www.aestudies.com/library/daily.pdf

Giannopolous, K. (2000). Measuring volatility. In M. Lore & L. Borodovsky (Eds.), *The Professional's Handbook of Financial Risk Management (pp. 42–74).* Butterworth-Heinemann.

Gujarati, D.N. (2003). *Basic Econometrics (4th ed.)* McGraw-Hill/Irwin.

Hamilton, J.D. (2010). Macroeconomics and ARCH. In T. Bollerslev, J. Russell, & M. Watson (Eds.), *Volatility and Time Series Econometrics: Essays in Honor of Robert Engle (pp. 79-96).* Oxford University Press.

Henry, O.T. (2002). Long memory in stock returns: some international evidence. *Applied Financial Economics*, 12, 725–729.

Herwartz, H. (2004). Conditional heteroscedasticity. In H. Lütkepohl & M. Krätzig (Eds.), *Applied Time Series Econometrics* (pp. 197–220). Cambridge University Press.

Hodge, V.J., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, 22 (2), 85–126.

Jacobsen, B., & Dannenburg, D. (2003). Volatility clustering in monthly stock returns. *Journal of Empirical Finance*, 10 (4), 479 – 503.

Johannesburg Stock Exchange. (2017). Ground rules for the management of the FTSE/JSE Africa Index Series. Retrieved September 6, 2017, from https://www.jse.co.za/content/JSERulesPoliciesandRegulationItems/FTSE%20JSE%20Ground%20Rules%20.pdf

Kang, S.H., Kang, S.M., & Yoon, S.M. (2009). Forecasting volatility of crude oil markets. *Energy Economics*, 31 (1), 119–125.

Kasman, S., Turgutlu, E., & Ayhan, A.D. (2009). Long memory in stock returns: evidence from the major emerging central European stock markets. *Applied Economics Letters*, 16(17), 1763–1768.

Kendall, M.G., & Hill, A.B. (1953). The analysis of economic time-series-part I: prices. *Journal of the Royal Statistical Society: Series A (General),* 116 (1), 11–34.

Kinnunen, J. (2013). Dynamic return predictability in the Russian stock market. *Emerging Markets Review*, 15(June 2013), 107–121.

Kirchgässner, G., & Wolters, J. (2007). *Introduction to Modern Time Series Analysis*. Springer.

Kumar, R. & Dhankar, R.S. (2010). Empirical analysis of conditional heteroscedasticity in time series of stock returns and asymmetric effect on volatility. *Global Business Review*, 11 (1), 21–33.

Lim, K.P., Luo, W. & Kim, J.H. (2013). Are US stock index returns predictable? Evidence from automatic autocorrelation-based tests. *Applied Economics*, 45 (8), 953–962.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business*, 36 (4), 394– 419.

Mandelbrot, B. (1967). The variation of some other speculative prices. *Journal of Business*, 40 (4), 393–413.

Mangani, R. (2007). Distributional properties of JSE prices and returns. *Investment Analysts Journal*, 36(66), 57–72.

McMillan, D.G., & Ruiz, I. (2009). Volatility persistence, long memory and time-varying unconditional mean: Evidence from 10 equity indices. *Quarterly Review of Economics and Finance*, 49 (2), 578–595.

Nel, W.S. (2011). The application of the Capital Asset Pricing Model: A South African perspective. *African Journal of Business Management*, 5(13), 5336–5347.

Nelson, D.B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 347–370.

Nelson, D.B. (1992). Inequality constraints in the univariate GARCH model. *Journal of Business and Economic Statistics*, 10 (2), 229-235.

Newey, W. K., & West, K. D. (1987). Hypothesis testing with efficient method of moments estimation. *International Economic Review*, 28(3), 777–787.

Ning, C., Xu, D., & Wirjanto, T. S. (2015). Is volatility clustering of asset returns asymmetric? *Journal of Banking and Finance*, 52, 62–76.

Niu, H., & Wang, J. (2013). Volatility clustering and long memory of financial time series and financial price model. *Digital Signal Processing*, 23(2), 489–498.

Officer, R. R. (1972). The distribution of stock returns. *Journal of the American Statistical Association*, 67(340), 807-812.

Peiró, A. (1999). Skewness in financial returns. *Journal of Banking and Finance*, 23(6), 847–862.

Poon, S., & Taylor, S. J. (1991). Macroeconomic factors and the UK stock market. *Journal of Business Finance and Accounting*, 8(5), 619–636.

Poon, S-H. (2005). *A Practical Guide to Forecasting Financial Market Volatility*. John Wiley and Sons, Inc.

Priestley, R. (1996). The arbitrage pricing theory, macroeconomic and financial factors, and expectation generating processes. *Journal of Banking and Finance*, 20 (5), 869–890.

PWC. (2015). Valuation methodology survey 2014/2015. Retrieved September 6, 2017, from https://www.pwc.co.za/en/assets/pdf/valuation-methodology-survey-2015.pdf

Roll, R. (1992). Industrial structure and the comparative behavior of international stock market indices. *Journal of Finance*, 47(1), 3–41.

Sadorsky, P. (2001). Risk factors in stock returns of Canadian oil and gas companies. *Energy Economics*, 23(1), 17–28.

Sadorsky, P., & Henriques, I. (2001). Multifactor risk and the stock returns of Canadian paper and forest products companies. *Forest Policy and Economics*, 3(3–4), 199–208.

Simkowitz, M. A., & Beedles, W. L. (1980). Asymmetric stable distributed security returns. *Journal of the American Statistical Association*, 75(370), 306–312.

Studenmund, A. H. (2011). *Using Econometrics: A Practical Guide* (6th ed.). Pearson Education Limited.

Wong, W. K., & Bian, G. (2000). Robust estimation in capital asset pricing model. *Journal of Applied Mathematics and Decision Sciences*, 4(1), 65–82.

Xiao, L., & Aydemir, A. (2007). Volatility modeling and forecasting in finance. In J. Knight & S. Satchell (Eds.), *Forecasting Volatility in the Financial Markets* (3rd ed., pp. 1–45). Elsevier Ltd.

Zivot, E. (2009). Practical issues in the analysis of univariate GARCH models. In T. G. Anderson, R. A. Davis, J-P. Kreiß, & T. V. Mikosch (Eds.), *Handbook of Financial Time Series* (pp. 113-155). Springer.