RISK MANAGEMENT IN SOUTH AFRICA BEFORE, DURING, AND AFTER THE 2008 GLOBAL FINANCIAL CRISIS: AN APPLICATION TO DIFFERENT SECTORS

Research submitted by

EDEN GROSS

in partial fulfilment of the requirements for the degree of

MASTER OF COMMERCE

specialising in

FINANCE, in the field of INVESTMENT MANAGEMENT

Department of Finance and Tax

University of Cape Town

February 2020

Supervisor:

Associate Professor Ryan Kruger

Abstract

The risk management functions of most financial institutions occupy themselves with the estimation of the value at risk (VaR) of their portfolios as a measure of market risk. Various methods are available to calculate the VaR measure, and this can be done at various degrees of confidence. This study evaluates and analyses the performance of five popular VaR forecasting methods in the South African context, using the closing values of three of the major indices available on the Johannesburg Stock Exchange (JSE), namely the All Share Index (ALSI), the Financials-Industrials Index (FINDI), and the Resources Index (RESI). These three indices are considered based on the findings of prior studies that indicate that not only does decomposing the ALSI into its constituent (the FINDI and the RESI) indices provide a better measurement of market risk on the JSE, but these sub-indices also have different systematic risk exposures which may necessitate different treatments in measuring their risks appropriately. The periods examined surrounded the 2008 global financial crisis in order to allow an evaluation of the impact of varying levels of volatility on the analysis. Overall, the study concludes that the performance of the VaR models examined is similar when assessing the risk of the ALSI and the RESI returns, while they are very different for the FINDI. This conclusion provides crucial insight into the risk management and investment decisions concerning portfolios which are more heavily invested in the FINDI as opposed to the other two, as this study suggests that a blanket treatment to the South African market is incorrect.

UNIVERSITY OF CAPE TOWN

FACULTY OF COMMERCE

Igniting Knowledge and Opportunity

Declaration

COMPULSORY DECLARATION:

1. This dissertation has been submitted to Turnitin (or equivalent similarity and originality checking software) and I confirm that my supervisor has seen my report and any concerns revealed by such have been resolved with my supervisor.
2. I certify that I have received Ethics approval (if applicable) from the Commerce Ethics Committee.
3. This work has not been previously submitted in whole, or in part, for the award of any degree in this or any other university. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works of other people has been attributed, and has been cited and referenced.

Student number	GRSEDE001
Student name	Eden Gross
Signature of Student	Signed by candidate
Date:	09 February 2020

Acknowledgements

First and foremost, I'd like to thank my parents and siblings. Their continuous and endless support throughout this process had been crucial at times.

Second, I'd like to thank my supervisor, Associate Professor Ryan Kruger, for taking me on as a student and for the endless support he has offered me.

Third, a big thank you goes to Associate Professor Francois Toerien, for willing to assist and listen. It is, in part, thanks to him that this dissertation found its direction.

Lastly, to the Department of Finance and Tax, for allowing me to undertake this research, as well as providing me with the tools and the space to complete it, and thank you to Postgraduate Officer Christine Olsen, for the endless support she offers the Department's postgraduate students.

Table of Contents

List of Tables

List of Figures

1. Introduction

The use of value at risk (VaR) as a risk measure metric was the result of a pursuit of safety and the development of the field of risk measurement following Black Monday, or the 1987 American stock crash (Christoffersen, et al., 2001). This, coupled with the advances in both academic literature and industry practice in calculating VaR, has made it the metric of choice for regulatory institutions such as the Basel Committee on Banking Supervision (BCBS) (Basel Committee on Banking Supervision, 2004).

Further updating to regulation, such as that proposed by the BCBS (Basel Committee on Banking Supervision, 2019), has paved the way for various industries to use of any one of many methods to evaluate VaR. However, most institutions opt to use the historical simulation technique due to its ease of implementation and low computation costs (Pritsker, 2006).

The historical simulation method, together with the popular RiskMetrics method, while widely used, failed to predict the 2008 global financial crisis (Lockwood, 2015) – a crisis which did not spare South Africa. This failure then begs the questions: Would another model have managed to predict the crisis better? Do different models perform differently in times of varying volatility in South Africa? And do specific industries within the South African economy display different risk management characteristics when compared to the market as a whole? The examination of the performance of various VaR models before, during, and after the crisis in South Africa then becomes relevant across the South African market, as well as across specific sub-sections of the market, using various models.

While autoregressive models were not yet permitted by financial regulation at the time of the 2008 global financial crisis, it is of interest to test such models and establish whether earlier incorporation of autoregressive models would have provided better predictors to the 2008 global financial crisis in the context of South Africa. On the other hand, it is important to note that the allowance for such models to be used did not necessarily mean that banks opted to use them – as mentioned above, most banks opt to use the historical simulation method and the RiskMetrics model, regardless of the various other methods that are available to them.

In practice, once VaR figures are estimated using the institution's preferred model, the figures are subject to backtesting. The backtesting process is often employed by the risk management division of the institution. Backtesting is the process of evaluating a model by comparing a sample of model-predicted loss values (captured by the VaR metric) to the corresponding experienced loss amounts over a past period (Pérignon, et al., 2008). This allows the institution to evaluate the model's accuracy, and make adjustments to the model (or the reserves held) accordingly.

The international literature which covers the performance evaluation of different VaR models over time usually concludes that GARCH(1,1) models outperform when compared to other models[1], especially the RiskMetrics model. The RiskMetrics model does, however, perform adequately at the 5% confidence level (So & Yu, 2006). McMillan and Kambouroudis (2009) also find that the RiskMetrics method works well when used in conjunction with simple forecasting parameters, and outperforms other methods in Asian markets. These Asian markets were used as the developing markets in the authors' sample, warranting evaluation of the RiskMetrics model in South Africa, as it is also a developing market. Nieto and Ruiz (2016) conclude that the performance of VaR methods is heavily dependent on the period in which they are applied. Hence, an investigation into various sub-periods, forming together a full period, is warranted.

An early observation by Seymour and Polakow (2003) states that the South African market exhibits stock market volatility differently when compared to other markets. The authors found the RiskMetrics method to be superior to the historical simulation method. More recent studies[2], however, find that the generalised autoregressive conditional heteroscedasticity (1,1) (GARCH(1,1)) and exponential GARCH(1,1) (EGARCH(1,1)) models outperform the RiskMetrics and historical simulation models in the South African market. The South African literature often makes use of the Johannesburg Stock Exchange's (JSE's) All Share Index (ALSI) as a proxy to the South African market, as it accounts for 99% of total market capitalisation in the South African market. Hence, it will also be used as the South African market proxy in this study.

This study aims to build on the study of Elenjical, et al., (2016) – a study which investigates the performance of various VaR models in the South African context using the ALSI (for more details, see Section 3.2.). This study aims to break the ALSI down into its sub-indices, thereby differentiating between the risk management treatment applied to each of the sub-indices.

Despite the majority of prior studies employing the ALSI in their tests, van Rensburg (2002) shows that an application of the arbitrage pricing theory model utilising two major sector

[1] See Christoffersen, et al., (2001); Kuester, et al., (2006); So and Yu (2006); and McMillan and Kambouroudis (2009).

[2] See Bonga-Bonga and Mutema (2009); Cifter (2012); and Elenjical, et al., (2016).

indices as the systematic risk factors provides a superior estimate of market risk when compared to a market model using the ALSI as the market index. These sectors are the financials-industrial sector, represented by the Financials-Industrials Index (the FINDI) and the resources sector, represented by the Resources Index (the RESI). van Rensburg finds that the risk exposures of these sector indices are fundamentally different (primarily due to the behaviour of dual-listed shares and mining companies which are exposed to exchange rate risk) and these characteristics are averaged away when the broad ALSI market index is considered. It is, therefore, of interest to evaluate the performance of different VaR models using not only the ALSI but also the FINDI and the RESI in order to better observe their underlying risk behaviour.

A second point of interest is how changing volatility over time impacts the effectiveness and appropriateness of the VaR models selected. To this end, this study examines multiple periods, specifically pre-, during, and post-crisis, as well as the full period combining all sub-periods, to establish which method deals better with the changing volatility experienced as one period changes to the next and, therefore, can perform better across all periods.

The full period examined by this study spans from 4 March 2003 to 31 December 2014 and is divided into three periods, namely the pre-crisis period (4 March 2003 to 31 January 2007), the crisis period (1 February 2007 to 30 September 2009), and the post-crisis period (1 October 2009 to 31 December 2014). The data used are the returns obtained using the ALSI, the FINDI, and the RESI, as traded on the JSE during the period.

While an evaluation of the performance of different VaR evaluation models during periods of changing volatility – specifically examining the periods before, during, and after the global financial crisis – has been performed in the South African context in the past, no study makes use of the two primary sub-indices of the ALSI to evaluate their performance. This novel addition is believed to add further depth to existing research, and to provide useful risk management results to those South African sectors.

The remainder of this study is structured as follows: Chapter 2 introduces the concept of VaR and defines it, while also exploring five different VaR methods to capture market risk; Chapter 3 discusses and analyses the available academic literature on the performance of different VaR models internationally and in the South African context; Chapter 4 then discusses the data and methodology used in this study; while Chapter 5 discusses and analyses the results obtained. Finally, Chapter 6 concludes this study and discusses future research options and the limitations encountered in conducting the study.

2. Value at Risk

This chapter introduces the VaR risk measure as a way to quantify market risk. A definition of VaR is provided before examining various well-established models to calculate VaR.

The increase in the desire of financial institutions to manage their market risks, coupled with increased regulation, has pushed many financial institutions towards the use of VaR as a risk measure model (Jorion, 1996).

VaR provides a portfolio loss figure quoted in absolute terms, capturing the decline in portfolio value due to changes in market prices of the instruments making up the portfolio, which is tied to a period over which the loss may be incurred with a pre-specified degree of certainty (Hendricks, 1995).

A mathematical definition of VaR is as follows: Consider X to be an invertible random loss variable with a cumulative distribution function $F_X(x) = \Pr[X \leq x]$. Moreover, let the cumulative distribution function of X have a left-continuous inverse $F_X^{-1}(y) = \min\{y | F_X(x) \geq y\}$. VaR is then defined to be the value of the left-continuous inverse function at some α-quantile, i.e.,

$$VaR_\alpha(X) \coloneqq F_X^{-1}(\alpha) \tag{1}$$

where α is the degree of certainty mentioned above (Pflug, 2000).

To make sense of the above, consider the following example: A large institution quotes a 1-day 99% VaR figure of R50 million. This quote means that, over any single trading day, the institution has a 99% probability of making a loss which does not exceed R50 million. Alternatively, the quoted VaR figure means that, over any single trading day, the institution has a 1% probability of making a loss which does exceed R50 million.

An advantage of institutions using VaR to quote market risk is that it is easily comparable to other metrics, such as, for example, the institution's total revenue. This eases the shareholders' decision-making process as they can easily comprehend the levels of risk quoted by the institution (Jorion, 1996).

The VaR measure is widely adopted today due to its incorporation in regulation. The BCBS is a regulatory authority which publishes the Basel Accords, a set of recommendations for national regulators to enforce, governing the risk within banks and the banking sector. The BCBS requires banks to quote their VaR figures on a daily basis as a measure of market risk

(Basel Committee on Banking Supervision, 2004). Moreover, the BCBS allows banks to make use of internal models to calculate VaR. The introduction of this regulation has pushed many institutions towards the application of VaR as a means to measure market risk and to provide such for convenience of both management and shareholders. Hence, an examination of the various VaR models available, together with their advantages and disadvantages, is crucial to the understanding of the consequences of the implementation of such models. A discussion of the various models commonly used to measure VaR is presented below, followed by a discussion of the measure's limitations and criticisms.

2.1. Models

Below is a discussion of some of the well-established VaR modelling techniques developed to quantify market risk. A definition of each model is presented together with a brief discussion.

2.1.1. Historical Simulation

As mentioned earlier, being the least computationally intensive method to apply, many institutions find themselves applying the historical simulation method to calculate VaR. The ease of computation, together with the simplicity of the method, makes the historical simulation method the most commonly applied method to calculate VaR.

To compute VaR using the historical simulation method, the closing figures of the profit and loss account for the past year (or approximately 250 trading days) of the firm under examination are collected and sorted to form a histogram. The VaR is then read off the histogram, given the degree of certainty required (Dowd, 2002).

When implementing this model, the firm implicitly assumes that the past distribution of its returns (profit and loss account) is indicative of the distribution of future returns. While this assumption does not cast a parametric distribution to the returns experience (such as the normal distribution), it still assumes that past returns are indicative of future returns, making this method somewhat difficult to justify.

Moreover, the method assumes that the specific values which can be taken on by the random variable, being the firm's daily return, are independent and identically distributed (Pritsker, 2006). This, together with the equally weighted contribution of each day in the period examined, begs the question of whether this method is capable of producing VaR figures which are both adequate and accurate in periods of varying volatility. Hence, this method is suited perfectly to be examined in this study.

2.1.2. Delta-Normal

The application of the delta-normal (commonly referred to as the variance-covariance) method relies on the underlying assumption that a random variable with a well behaved statistical distribution, coupled with the Central Limit Theorem, exhibits a normal distribution in the limit (Dowd, 2002). This means that, given a large database for daily profit and loss account figures, the daily profit and loss account figures can be modelled using a normal distribution, accommodating the application of the normal distribution in the calculation of daily VaR figures.

The normal distribution allows for a simple calculation of VaR. VaR can be calculated as follows:

$$VaR = \mu + z_\alpha \times \sigma \tag{2}$$

where μ is the mean of the distribution; σ is the standard deviation of the distribution; and z_α is the quantile corresponding to the degree of confidence of the VaR figure.

The normality assumption is, in fact, often implied in most VaR calculation methods when it comes to scaling the daily VaR figures that most models calculate. Financial regulation often requires large institutions to quote a 10-day VaR figure, as opposed to a daily VaR figure. This is done by calculating the 1-day figure and scaling it by a factor of $\sqrt{10}$, i.e., the square-root of time rule for the 10-day period. This scaling application, however, is only valid when the returns are identically and independently distributed normal variables (Alexander, 2008). The application of this scaling technique may lead to substantial model risk if applied to returns that are not normally distributed.

A possible explanation for the less frequent use of the delta-normal method (when compared to the historical simulation method) may be the complexity of the implementation of such a method. The starting point in applying the method requires the modeller to first identify the factors which are believed to influence the returns of the institution in question (Linsmeier & Pearson, 2000). As mentioned above, the assumption of identically and independently distributed normality is made when considering such factors, with the further assumption that the combined distribution of those factors is multivariate normal. Further complications may be encountered when calculating the correlation matrix of variables making up the multivariate normal distribution.

The assumption of normality is perhaps the biggest flaw of this method, given the ample literature available[3] as evidence that financial returns are leptokurtic as opposed to normally distributed, as the model assumes. This, once again, may lead to substantial model risk that is inherent within the application of the delta-normal method to estimate VaR.

2.1.3. Autoregressive Conditional Heteroscedasticity Models

Advances in regulation and computational resources have provided for banks to employ autoregressive models to calculate VaR since the 2008 global financial crisis. Autoregressive models such as autoregressive conditional heteroscedasticity (ARCH), generalised ARCH (GARCH), and exponential GARCH (EGARCH) are often found being employed to calculate VaR. Each of these models is discussed below.

In the application of an ordinary least squares model, the assumption that the square of the expected sum of error terms is constant throughout the data is made. This assumption is given the name of homoscedasticity (Engle, 2001). The opposite is then called heteroscedasticity, which the ARCH family of models is tasked with estimating.

The primary advantage of autoregressive models is their ability to incorporate the daily changes in volatility (Giot & Laurent, 2004), which is key to the application of risk analysis, something that the historical simulation and the delta-normal methods are incapable of doing as efficiently as autoregressive models.

2.1.3.1. ARCH Models

While originally used by Engle (1982) to describe the uncertainty involving inflation in the United Kingdom, ARCH models have been used in many financial and econometric applications whenever the need to track volatility through time arises (Bollerslev, 2007).

An ARCH model is often used to model a time series of the residual terms ε_t, where these residual terms are defined by a time-dependent standard deviation term, σ_t, and a stochastic term z_t, i.e.,

$$\varepsilon_t = \sigma_t \times z_t \tag{3}$$

where z_t is a series of independently distributed normal random variables with zero mean and a standard deviation of 1 (Bollerslev, 2007).

[3] See, for example, Breen, et al., (1989).

The variance of the series at time t is modelled using the following equation:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2, \qquad \alpha_0 > 0, \{\alpha_i\}_{i>0} \geq 0 \tag{4}$$

An ARCH model is often stated together with its order. Engle (1982) defined the ARCH model using Equations (3) and (4) as a q^{th} order ARCH model, denoted by ARCH(q), where q is the length of the ARCH lags.

The procedure to compute the value for q was also presented by Engle (1982). The procedure, in brief, requires the implementation of an autoregressive model of the output data r_t (given any information available up to and including time $t-1$, i.e., given the filtration system of the process up to and including time $t-1$, denoted by $\mathcal{F}_t$). The squared error terms for the autoregressive model are then estimated from the data by regressing those on a linear combination of an initial constant and q lagged values, where q is, as before, the length of the ARCH lag.

The last step, as is in all statistical tests, is the comparison of a test statistic to a critical value (in this case, a Chi-squared with q degrees of freedom) in order to make a conclusion regarding the null hypothesis, that being that no GARCH effects are present in the underlying data.

2.1.3.2. GARCH Models

The GARCH model was developed by Tim Bollerslev in 1986. While the GARCH model is similar to the ARCH model in that it makes use of past squared residuals in equal weights, the contribution of the past squared residual diminishes (although never to zero) as the squared residual moves further into the past (Engle, 2001).

Engle (2001) further states that the prediction of the next period's variance is a combination of inputs gathered from the long-run average variance captured by the model, together with new information revealed in the current period (information that affects both the residuals and the current period's estimated variance).

Following the notation presented by Engle (2001), when applying the GARCH model to financial data, the variance estimated is that of the residuals in the regression $r_t = m_t + \sqrt{h_t}\varepsilon_t$, where r_t is the return on the asset or portfolio in question (i.e., the dependent variable); m_t is the mean of past returns; h is the variance of past returns, making $\sqrt{h}$ the standard deviation of past returns; and ε_t is the error term. The GARCH model of the variance for the next period, h_{t+1}, is then

$$h_{t+1} = \omega + \alpha h_t \varepsilon_t^2 + \beta h_t \tag{5}$$

where the GARCH model parameters ω, α, and β must be estimated.

Moreover, the parameters outlined in Equation (5) must adhere to the following set of conditions:

1. $\alpha > 0$;
2. $\beta > 0$;
3. $\omega > 0$; and
4. $\alpha + \beta < 1$.

The level of long-term average variance of the model is captured by the expression $\sqrt{\omega/(1-\alpha-\beta)}$, making condition 4. above a trivial requirement.

α above is termed the GARCH error parameter, proportionally capturing the model's sensitivity (and, hence, volatility's sensitivity) to market shocks. β is termed the lag parameter, capturing the tenacity of the volatility inherent in the model. Putting the two terms together yields a measure capturing the rate at which the conditional variance in the model converges to its long-term average figure.

GARCH models are often referred to as GARCH(p,q) models, where p represents the number of autoregressive lags used in the model (confusingly notated in this study, this was denoted by q in the ARCH model discussion in Section 2.1.3.1.), and q represents the number of moving average lags used in the GARCH model. The former is often referred to as the ARCH term, while the latter is often referred to as the GARCH term (Engle, 2001). The model described in Equation (5) above is, in fact, a GARCH(1,1) model. The assumption that p and q are greater than zero (while p is allowed to equal to zero) is often made as well (Bollerslev, 1986).

A downfall that is inherent in GARCH models is that they are mean-reverting (Engle, 2001), a quality that may not be necessarily true at all times for financial data.

The GARCH parameters mentioned above can be estimated by maximising the log-likelihood of the model. However, the increase in computing power and the vast number of software available today make the implementation of GARCH models straight forward, providing the user with the GARCH parameters as a form of output of the software employed.

2.1.3.3. EGARCH Models

In the discussion of GARCH models above, it is obvious that the volatility required as an output of the model is non-negative. Intuitively, this makes sense as, when basic financial theory is considered, such as the capital asset pricing model put forward by Sharpe (1964), volatility is known to be at least zero. The GARCH model, however, has made this assumption implicit, by calculating the volatility at time $t+1$, h_{t+1}, as a linear combination of positive values (see Equation (5)). Hence, h_{t+1} itself must be positive.

In 1991, Daniel B. Nelson proposed a new approach to dealing with heteroscedasticity, which overcomes some of the limitation of the GARCH models, as he pointed them out[4].

The proposed model aims to explicitly allow for the asymmetric relationship between financial returns and their volatilities (Bollerslev, 2007), while further imposing the condition that volatility is non-negative by adopting a natural logarithmic function into the GARCH model, thereby resulting in a model which, instead of estimating h_t (or h_t^2), estimates $\ln(h_t^2)$ as some linear combination of variables (Nelson, 1991). The following model, the exponential GARCH (EGARCH) model, was presented:

$$\ln(h_t^2) = \alpha_t + \sum_{k=1}^{\infty} \beta_k g(z_{t-k}), \qquad \beta_1 = 1 \tag{6}$$

where the sequences $\{\alpha_t\}_{t=-\infty}^{\infty}$ and $\{\beta_k\}_{k=1}^{\infty}$ are non-stochastic scalar sequences, such that $\alpha_t, \beta_k \in \mathbb{R}\ \forall\ t, k$. Moreover, $g(z_t) \coloneqq \theta z_t + \gamma(|z_t| - E|z_t|)$, where θ, γ are some of the model coefficients, and $z_t \sim N(0,1)$ (Nelson, 1991).

The model requires no sign restrictions on the model parameters as the natural log of any number can be either positive or negative. Once the left-hand side of Equation (6) is used as the exponent of the base of the natural logarithm (e), the right-hand side is positive, regardless of the signs of the model parameters.

[4] For more details, see Nelson (1991).

Similar to the GARCH model, an EGARCH model is often quoted as an EGARCH(p,q) model, where p and q are similar in interpretation to the interpretation of these variables under the GARCH model (see Section 2.1.3.2.).

2.1.4. RiskMetrics

RiskMetrics was originally published in 1994 by RiskMetrics Group, Inc., a subsidiary of (then) J.P. Morgan & Co. bank (RiskMetrics Group, Inc., 2001).

The RiskMetrics model's methodology assumes that the distribution of logged daily returns (or the balance of the profit and loss account) is conditionally distributed as a normal random variable with a conditional mean at time t of μ_t and a conditional variance at time t of σ_t^2 , i.e., $\log(r_t)\,|\mathcal{F}_t \sim N(\mu_t, \sigma_t^2)$, where $\mathcal{F}_t$ is as defined above. Given the assumption that the process has a conditional mean at time t equal to zero, i.e., $\mu_t = 0$, the following relation holds:

$$\sigma_t^2 = \lambda_{RM}\sigma_{t-1}^2 + (1 - \lambda_{RM})r_{t-1}^2, \qquad \lambda_{RM} \in (0,1) \tag{7}$$

where λ_{RM} is a smoothing parameter; and r_{t-1}^2 is yesterday's squared return, serving as a proxy for true volatility in the market (McMillan & Kambouroudis, 2009).

The RiskMetrics model can also be stated as an exponentially weighted moving average (EWMA) model using backward substitution (González-Rivera, et al., 2007). This yields the following re-statement of the model:

$$\sigma_t^2 = (1 - \lambda_{RM})\sum_{\tau=1}^{\infty} \lambda_{RM}^{\tau-1} r_{t-\tau}^2\,, \qquad \lambda_{RM} \in (0,1) \tag{8}$$

where all variables used are as defined for Equation (7).

The RiskMetrics smoothing parameter, λ_{RM}, is often set by users of the model to equal 0.94 when using daily figures, as it provides the best backtesting results (Pafka & Kondor, 2001), although it should be estimated based on the data available (González-Rivera, et al., 2007).

The model has some great advantages over some of the models already discussed. Some of these advantages are discussed below.

First, the model requires a very small number of inputs, namely the return on any given trading day in the past, the volatility of that trading day, and some smoothing parameter which, as mentioned, may be plugged in as 0.94 for daily return values. The former two variables can be established with ease, while the last one is provided, making the model a simple plug-and-play model to forecast tomorrow's volatility.

Second, the model makes use of a volatility weighting property, as evidenced in the statement of Equation (8). Hence, the model allows for more recent market volatility values to affect forecasted volatility more so than distant past volatility would, something that the historical simulation and the delta-normal methods do not provide for. This allows for forecasting which captures market conditions better and, hence, allows for volatility predictions which are more closely related to today's market.

While the ease of implementation and the simplicity of the model made it a popular choice in the past (McMillan & Kambouroudis, 2009), the more commonly applied method to calculate VaR today is, as mentioned earlier, the historical simulation method. This may be due to the failure of the RiskMetrics model to predict the 2008 global financial crisis, leading institutions, especially banks, to abandon the method altogether.

Lastly, while the method incorporates the functionality of an EWMA model, this does not testify to the method's ability to adequately react to future changes in volatility – it merely implies that volatility updating (given past information) is performed with more emphasis given to more recent (past) information.

2.2. Limitations and Criticisms

Some of the limitations and criticisms regarding the various VaR methods outlined above are discussed in this section. Further discussion is provided in Chapter 3 of this study with reference to some of the specific tests and studies performed in both international literature and the South African markets.

Before any limitations and criticisms are discussed, it is important to note that one of the biggest downfalls of the VaR measure is that it is merely an estimation of market risk (Jorion, 1996). However, most financial institutions used to take VaR figures as a holistic risk measure. Moreover, it is important to highlight that the figure has some notable advantages. These are listed below.

Daníelsson (2002) identifies three main advantages of the VaR measure:

1. It is easily implementable once a valuation method is chosen;

2. It is easily explained, as it captures three key features of risk: a time period, a degree of certainty, and a loss amount; and
3. If all institutions use the measure, investors have an easy measure for comparison, allowing the investor to quantify and rank risk and reward.

Now that the advantages of the measure have been listed, attention can be turned to the measure's shortcomings.

First, the measure is a point estimate. This means that any information around this 'point' is unavailable or uncaptured by the measure. Referring to the example provided earlier in this chapter, recall that, while there is a 1% chance of exceeding a R50 million loss on any given trading day, the estimate does not provide any indication of what loss amount will be incurred, should the unlikely event of a loss exceeding R50 million occur.

Second, the VaR measure is a measure of loss incurred based on profit and loss amounts captured by the profit and loss account. Due to various accounting techniques (such as dealing with accruals, as well as off-balance sheet accounting practices), coupled with management's (mis)judgement, the figures captured by the profit and loss account may be heavily influenced and, therefore, manipulated (Daníelsson, 2002). This manipulation allows institutions, such as large financial institutions, to game the regulatory framework (Lockwood, 2015), leading to disastrous ends, such as the 2008 global financial crisis.

A notable criticism of the VaR measure as a risk measure is the one accusing it of being a non-coherent risk measure. Delbaen (2000) categorises a risk measure as coherent through a formal axiomatic definition, highlighting five key conditions (or elements) that a coherent risk measure must exhibit. Condition number 2 in Delbaen's definition is that of subadditivity. Subadditivity is defined as follows:

Definition 1: A risk measure $\varrho(x)$ associated with the financial activity x is said to be subadditive if, for two financial activities x and y, $\varrho(x+y) \leq \varrho(x) + \varrho(y)$. The result follows due to the diversification effect captured in the combined portfolio of the two positions.

VaR, however, does not meet the definition of a subadditive risk measure, nullifying its qualification as a coherent risk measure (Embrechts, 2000). This can be made clear using an example[5] (which acts as a proof by contradiction).

Consider a portfolio made up of two positions, each in an option offering either cash or nothing at exercise. Both options, Option A and Option B, are out-of-the-money options with a pay-out equal to 1 under the following conditions:

1. Option A offers a pay-out of 1 if the underlying's spot price is greater than some high value H.
2. Option B offers a pay-out of 1 if the underlying's spot price is less than some low value L.

Option A and Option B each carry a premium of P_A and P_B, respectively. Suppose that the risk measure VaR is a risk measure which abides by the subadditivity definition as defined in Definition 1.

Suppose further that the probability that either option ends up in-the-money is 0.75%, i.e., there is a 0.75% probability of pay-out for either option. This also means that there is a 99.25% probability that either option pays out nothing, i.e., maturing out-of-the-money. Hence, the probability of both options expiring worthless is equal to 98.51%, the product of each option expiring worthless.

The above means that, at a degree of confidence of at least 99%, the options will expire worthless. Hence, the 99% VaR for Option A is $-P_A$, and a corresponding VaR figure for Option B is $-P_B$.

Note the perfectly negative correlation of the two options. Hence, the probability of one of the two options ending in-the-money is 1.5%, i.e., the sum of the probabilities of each option ending in-the-money. Hence, one of the two options will be exercised in the lowest 1% of the distribution of the combined portfolio. This implies that the 99% VaR figure for the writer of Option A and Option B is equal to $1 - P_A - P_B$.

However, $VaR(-A - B) = 1 - P_A - P_B > -P_A - P_B = VaR(-A) + VaR(-B)$.

[5] Thank you to Associate Professor Chun-Sung Huang of the University of Cape Town for providing the basis for this example during a master's class at the University.

The equality stated above clearly contradicts the definition of subadditivity made in Definition 1. Hence, VaR does not conform to the subadditivity condition and is, therefore, not a coherent risk measure.

The lack of subadditivity for the VaR measure is an issue that may be solved by a complex set of equations which allow for their user to calculate a portfolio's VaR figures through adding its components. In other words, while a method exists to calculate the total the VaR figure for a portfolio made up of various sub-portfolios, it is by no means straight forward and grows in complexity with the number of components involved.

Daníelsson (2002) also points out the unreliable assumption revolving the stability of the statistical properties surrounding the computational methods employed to calculate VaR. The assumption entails that the specific set of circumstances affecting the financial data used to calculate the VaR figure will remain unchanged for as long as the model employed is used to generate further VaR figures. This, by definition, does not allow VaR models to deal with changes in these underlying financial conditions with any ease, and a new calibration of the model may be required whenever a new VaR figure is calculated, to account for the changing conditions in the market.

With respect to the specific models discussed above, several observations can be made which serve as disadvantages. These are discussed below.

Concerning the historical simulation and delta-normal methods, it is clear that the models do not employ any time-dependent weighting on the contributions of past data to the forecast. For example, the historical simulation method assumes an equal contribution from each of the trading days which contribute data towards the VaR forecast, suggesting that the most recent contribution is just as indicative of market conditions as the oldest contribution. A similar argument can be made regarding the variance-covariance matrix of correlation coefficients of the different assumed-to-be normally distributed variables. This matrix should be updated as time progresses, to enhance the accuracy of the model and, hence, better reflect changing market conditions.

Moreover, the calibration of the correlation matrix in the variance-covariance method is a challenging task which increases in complexity with the number of variables which are believed to contribute towards the risk in the portfolio. If a firm was to include every single asset which contributes to risk, the matrix would most likely be simply too big for a computer to compute, and the time to attempt such a task would be uneconomic.

Another shortcoming of the delta-normal method is its struggle when it comes to incorporating the effects of options in the portfolio (Linsmeier & Pearson, 2000). This is something that other methods, such as the historical simulation method, may be better at, given the existence of options in the portfolio in the past.

Issues regarding implementation and explanation are often grouped, as a model that is more complicated to implement is often also a model that is more complicated to explain to senior management. This may serve as an explanation as to why various ARCH models are not commonly found to be used in the estimation of VaR (McMillan & Kambouroudis, 2009).

The historical simulation, the delta-normal, and the RiskMetrics methods are unable to deal with what-if scenarios (Linsmeier & Pearson, 2000), a tool that is often used by risk management teams, rendering them less useful than other methods. Moreover, the RiskMetrics method fails to accurately capture the inverse relationship between market volatility and its returns (McMillan & Kambouroudis, 2009).

Regardless of the numerous limitations of the measure, as discussed above, VaR is still widely used in the industry by firms, especially due to mandated regulation (as discussed above), and is often expected to be quoted for shareholders' sake. Hence, VaR is still relevant, and investigations into its calculation and the appropriateness of different VaR calculation methods is appropriate.

3. Literature Review

This chapter begins with a review of the international literature dealing with the performance evaluation of the various VaR methods outlined in Chapter 2, before performing a similar review using South African literature. Where available, a discussion of certain VaR methods' abilities to handle changing volatility, together with a comparison of models on that basis, is provided.

3.1. International Literature Review

The international academic literature examining volatility forecasting in financial data is vast, with some cornerstone studies being evident throughout the literature. One such study is that conducted by Christoffersen, et al., (2001). In their study, the authors compare a GARCH model to the RiskMetrics model, which they consider to be the industry-provided benchmark, to answer two key questions. The more relevant question to this study is that which aims to provide the risk managing function of an institution with a technique to compare two VaR methods, and decide which one is superior.

Using daily returns as their VaR forecast period for the GARCH(1,1) model and the RiskMetrics model with the smoothing parameter, λ_{RM}, being set to the industry-default value of 0.94, Christoffersen, et al., (2001) calculate daily VaR figures by applying both models to daily return data of the Standard and Poor's 500 Index (S&P 500) over 2209 trading days (between November 1985 and October 1994). The authors conclude that the GARCH(1,1) model provides results which are better suited for volatile periods when compared to the results obtained from the RiskMetrics model for the same sample period. It is important to note, however, that the authors quote a conclusion made by Marshall and Siegel (1997), stating notable differences between different RiskMetrics software providers when it comes to the output of VaR figures while using the same parameter and data.

This leads to two key risks when it comes to modelling, which are important to keep in mind whenever an analysis of different modelling techniques and applications is made. The first risk is known as model risk, which arises when the model is applied using incorrect specifications. The second risk, the one identified by Marshall and Siegel (1997), is known as implementation risk. Implementation risk arises when two or more implementations of the same model, using the same specifications, yield different results. In the discussion that follows, these risks are implicitly present, while often not explicitly stated.

The existence of implementation risk and the acknowledgement of such by Christoffersen, et al., (2001) opens the authors' study's conclusions to criticism. By the mere existence of implementation risk, a similar study, using the same parameters and data, may reach a different conclusion regarding the superiority of a GARCH-type model in forecasting VaR when compared to the RiskMetrics model, in particular, and other models, in general.

Another noteworthy study is that performed by Kuester, et al., (2006), in which the authors conduct a comparison of various VaR methods using three decades' worth of daily return data of the NASDAQ Composite Index (an index believed by the authors to adequately mimic the behaviour of a volatile portfolio of financial assets). While the study reveals that poor performance is achieved by most of the VaR models tested (i.e., most models fail both statistical tests and underestimate market risk), some methods (such as the historical simulation method) are more prone than others to exhibit clusters of breaches. This, in turn, suggests that these methods are inadequate in capturing volatility changes and updating their parameters due to the changing circumstances promptly.

Models that incorporate the notion of heteroscedasticity perform better, however, at both predicting acceptable VaR figures and yielding violations which are more often isolated (Kuester, et al., 2006). An application of the GARCH(1,1) model, which may be coupled with some form of forecasting simulations, performs particularly well in the tests performed by Kuester, et al., (2006), supporting the conclusions of Christoffersen, et al., (2001). While the study conducted by Kuester, et al., (2006) does not include the RiskMetrics model in its analysis, it offers a wider selection of models over a longer period, using a different equity market index, suggesting that the superior performance of the GARCH(1,1) model may have some merit across different periods and different markets.

A study that focuses on the comparison of performances of different GARCH-type models (including RiskMetrics, which can be classified as an Integrated GARCH [IGARCH] model (Pafka & Kondor, 2001)), is that by So and Yu (2006). In their study, the authors compare the VaR figures obtained using seven GARCH-type models across twelve different portfolios (comprising of either market indices or foreign exchange rates), employing both long and short positions at varying levels of confidence (namely the 1%, 2.5%, and the 5% levels).

The authors find that the RiskMetrics model proves to be inadequate at both the 1% and the 2.5% levels of confidence when it comes to estimating VaR, but performs tolerably at the 5% level. The GARCH method of estimating market volatility is also found to perform well across the different markets, although the volatility updating property of the ARCH-type family of models is not necessarily required when calculating VaR figures for the foreign exchange rates portfolios, a feature that seems to be necessary when applying VaR models to equity-based portfolios (So & Yu, 2006).

In their analysis of 31 markets across Europe, Asia, and the countries forming the Group of 7 (G7), McMillan and Kambouroudis (2009) build on the research conducted by So and Yu (2006) and address two main questions. Firstly, the authors investigate whether the in-sample literature-claimed superiority of the GARCH model over the RiskMetrics model is also present out-of-sample. Secondly, they investigate the accuracy of the different models using varying levels of confidence. Both of these questions are answered across the various markets included in the authors' study.

When applying the models to simpler forecasting metrics, the RiskMetrics method to evaluate VaR is, in fact, the superior choice (McMillan & Kambouroudis, 2009). The method performs poorly, however, once more complex techniques are applied, leading to a conclusion which supports the discussion made in Section 2.2. above: The RiskMetrics method was used in the period leading to the 2008 global financial crisis due to the simplicity of implementation offered, a characteristic that eases the interaction of the risk management function of an institution with both shareholders and management. The method is, however, inadequate once the simplified assumptions are removed and a more complex application, one that perhaps mimics the economic environment more accurately, is required – a conclusion similar to that of So and Yu (2006).

Moreover, when it comes to applying the models to the more developed of the 31 markets examined (such as the G7 and the European markets), the authors find that an ARCH-type model (specifically, the Asymmetric Power ARCH [APARCH] model) statistically outperforms other models when it comes to volatility prediction. The RiskMetrics model, on the other hand, performs better in volatility prediction in the Asian markets examined in the study, contradicting to some extent the earlier studies performed by Christoffersen, et al., (2001) and Kuester, et al., (2006).

When it comes to the calculation of VaR figures across the different markets, the RiskMetrics model is only adequate across selected levels of significance and markets, while the ARCH-type models (specifically APARCH and GARCH(1,1) models) perform better in almost all situations (McMillan & Kambouroudis, 2009).

Providing further insight into the applications of the GARCH and RiskMetrics models, Daníelsson (2002) concludes that the level of risk is directly proportional to the difficulty encountered when trying to forecast it. This, in turn, highlights that models such as GARCH and RiskMetrics cannot be used at levels of certainty such as 99%, due to their blatant inaccuracy at these ranges. They can, however, be used at levels such as 95%, leading to credible results from both the GARCH and RiskMetrics models (Daníelsson, 2002).

While there are various studies which support the implementation of an autoregressive model to calculate VaR figures, it is not often that a study recommends more basic models, such as the historical simulation method. In their study, Giot and Laurent (2004) evaluate the performance of autoregressive models when it comes to calculating VaR figures for portfolios holding equity indices. While the authors do conclude in favour of an autoregressive model, they also conclude that a model that calculates VaR figures based only on daily return data, such as the historical simulation method, yields satisfactory results.

Contradicting the view of Giot and Laurent (2004) is that of Daníelsson (2002), who states that the VaR figures obtained from the historical simulation method may be misleading, especially in times of high volatility (or crisis). In contrast, studies such as those conducted by Berkowitz and O'Brien (2002) and Pérignon and Smith (2010) conclude that the use of the historical simulation method in calculating VaR (as often employed by large firms) leads to conservative figures. These conservative figures overestimate market risk, rendering them (together with the application of the historical simulation method) useless.

When it comes to the application of the delta-normal method, there exists abundant evidence in the finance literature condemning the assumption that financial data are normally distributed. This assumption, which is the pinnacle in the application of the delta-normal method, results in VaR figures which are contaminated with model risk, more so than most of the other VaR methods discussed above. Moreover, the performance of the delta-normal method in predicting market crises is questionable, as the misspecification of the model underestimates (or, perhaps, overlooks) the likelihood of outliers and, therefore, extreme events (Peiró, 1999).

More relevant to this study are studies that examine the application of VaR models during periods of changing volatility. More specifically, of interest are the studies that analyse the performance of such models before, during, and after the 2008 global financial crisis, as well as those that make the comparison of those three sub-periods to the full period under observation.

A recent study examining the performance of different VaR methods is that of Nieto and Ruiz (2016). The study examines the returns of the S&P 500 Index throughout July 2005 to May 2014 and calculates 99% daily VaR figures using various techniques, including the historical simulation, GARCH(1,1), and EGARCH(1,1) methods, which are of interest in this study. The authors perform various backtesting techniques over two overlapping periods (the first being May 2010 to May 2014, while the second is May 2013 to May 2014). The results of the backtesting techniques are intriguing, as the historical simulation method's results are rejected in both periods, while some GARCH and EGARCH models are not rejected in either period.

The most prominent conclusion of their study is that the performance of different VaR models is heavily dependent on the observation period, as the study concludes that a categorically superior model is not available (Nieto & Ruiz, 2016).

A study whose objectives are more closely aligned with this study is the one presented by Dias (2013). Using three equity indices (NYSE, AMEX, and NASDAQ), the study closely examines several VaR methods by evaluating their performance during both crisis and non-crisis periods, and then also examining the performance of the VaR methods over the full period.

When examining performance over the full period, methods such as the historical simulation method perform best in several tests. This result is attributed by the author to the large number of equities which comprise the portfolio in question (Dias, 2013). This result may also be the result of the historical simulation's method's inability to quickly incorporate changing volatilities. This would result in the method underestimating VaR as the crisis begins, potentially leading to numerous violations, while significantly overestimating VaR once the crisis ends. The combination of periods of underestimation and overestimation, when ignoring the existence of a crisis, results in a balanced method which may perform adequately. Hence, an analysis of the performance of VaR methods without the explicit consideration of crisis periods will lead to misleading conclusions regarding the performance of different VaR methods (Dias, 2013).

3.2. South African Literature Review

Following advances surrounding financial regulation in 1996, the suitability of the RiskMetrics model to the South African market became of interest. Moreover, the model's performance results when compared to the results of more sophisticated volatility updating models were also highlighted in several studies.

One of the earlier studies examining the performance of different VaR methods in the South African market is that of Seymour and Polakow (2003). An early observation made by the authors is that the South African market exhibits different characteristics when it comes to the volatility of its equity market.

The authors conduct a comparison of various VaR evaluation methods, which include the historical simulation method as well as the RiskMetrics method. The study employed a VaR calibration period of 1500 days to calculate VaR figures using the historical simulation method. The historical simulation method, while often exhibiting somewhat adequate results, has exhibited a number of violations that exceeded the expected amount threefold, rendering the method to be of little use in the South African environment (Seymour & Polakow, 2003). This result may suggest that, due to the higher levels of volatility in the South African market when compared to more developed markets (or perhaps due to more frequent changes in volatility regimes, i.e., from little volatility to extreme volatility), the historical simulation method is utterly inadequate as a VaR method for South African institutions holding equity positions.

The study does, however, highlight the need for volatility forecasting in the South African market when it comes to forecasting VaR figures, stipulating that the RiskMetrics method provided the most accurate VaR forecasts (Seymour & Polakow, 2003).

Cifter (2012) builds on the investigation performed by Seymour and Polakow (2003) by examining the performance of a normal mixture GARCH (NM-GARCH) model to those of the RiskMetrics and GARCH models, among others. The study evaluates VaR forecasts using the ALSI from 7 February 2002 to 11 March 2011 (the first 275 trading days of which were used to calibrate the GARCH-type models' parameters).

While the study concludes that the NM-GARCH(1,1) model is superior to all others tested, this model is beyond the scope of this study. The study does provide insight, however, into the superiority of the GARCH(1,1) model over the RiskMetrics model, exhibiting superiority measured by a total of 1 fewer violation (Cifter, 2012).

Another paper that compares the performance of the RiskMetrics method (which is taken as the benchmark method) to several of the GARCH-type methods (GARCH(1,1) and EGARCH(1,1)) in the South African market is that of Bonga-Bonga and Mutema (2009).

The study was published mid-crisis, with its data constituting of ALSI daily return data from 3 January 2005 to 31 October 2008, encompassing data leading up to the 2008 global financial crisis. The study also makes use of a VaR calibration period, but one spanning 250 trading days, as opposed to the 1500 trading days used by Seymour and Polakow (2003). The 250-trading-day period was chosen to comply with regulatory requirements as proposed by the BCBS (Bonga-Bonga & Mutema, 2009).

The study further highlights the importance of asymmetric volatility in the South African market, stating that both the GARCH(1,1) and EGARCH(1,1) models resulted in superior VaR forecasts when compared to the benchmark used, i.e., the RiskMetrics model (Bonga-Bonga & Mutema, 2009). The authors do acknowledge, however, the existence of financial instability towards the end of the observation period, due to the unfolding of the 2008 global financial crisis.

In a more comprehensive study of nine emerging markets (including South Africa) benchmarked against the S&P 500 Index, Thupayagale (2010) obtains VaR figures for each of the ten markets using daily return data from 1 January 1998 to 31 January 2010. The study employs a calibration period of 1256 trading days (approximately five years) for the estimation of the GARCH-type models' parameters.

The study concludes that a GARCH-type model (specifically, the fractionally integrated GARCH [FIGARCH] model) outperforms all other models examined with varying degrees of confidence in the South African market, while the EGARCH model outperforms in the American market when calibrated using the regulatory parameters proposed by the BCBS (Thupayagale, 2010). The study further emphasises the key role and relevance of the degree of volatility in an emerging market, such as South Africa, together with the ability to capture volatility movements throughout the observation period. Models that account for volatility updating are found to be superior in the South African market more often than not.

More recently, a study building on from Thupayagale (2010) by undertaking the task set out in this study, namely investigating the performance of various VaR models over periods exhibiting different volatility levels with relation to the 2008 global financial crisis, is that of Elenjical, et al., (2016).

The authors point out, as did Dias (2013) above, that the use of long observation periods in this kind of studies may lead to conclusions surrounding some models which lack any form of robustness in reality (Elenjical, et al., 2016). Hence, the division of the full period into sub-periods is warranted to provide various risk management operations with clarity surrounding the performance of the various VaR models discussed.

Using ALSI daily return data for the period from 3 January 2005 to 17 August 2012, the authors separate the full period into three sub-periods: The Pre-Crisis period (3 January 2005 to 31 January 2007, 520 trading days), the Subprime Mortgage Crisis period (1 February 207 to 30 September 2009, 665 trading days), and the Post-Crisis period (1 October 2009 to 17 August 2012, 722 trading days). They then proceed to evaluate thirteen models, with the GARCH-type family of models' parameters calibrated using five years' worth of data as a calibration period, as standard in the literature, with the parameters being updated every 60 days.

The results of the study conclude that, when examining the full period, GARCH-type models (specifically, the fractionally integrated EGARCH [FIEGARCH] model) outperform other models in the South African market. When the various sub-periods were examined, it was noted that the EGARCH(1,1) model also offered more accurate results when forecasting VaR figures for long positions (Elenjical, et al., 2016).

The literature discussed above provides various opinions regarding the applicability and performance of the various VaR models across markets, as well as volatility levels. The autoregressive family of models seems to outperform other models in most scenarios, even in South Africa, as noted by Elenjical, et al., (2016), using ALSI daily return data over the periods.

As mentioned above, the South African market has been empirically shown by van Rensburg (2002) to be adequately predicted using a two-factor arbitrage pricing theory model. van Rensburg shows that the use of the ALSI as the market proxy in the capital asset pricing model (CAPM) is flawed, as the South African market is not mean-variance efficient once investors are permitted to invest off-shore. This nullifies the application of the CAPM and an application of the arbitrage pricing theory model is then undertaken.

van Rensburg (2002) concludes that a two-factor model is sufficient in describing the South African market's systematic risk, and reaffirms this conclusion by regressing the returns on each of the other minor sectors in the market on the two-factor model consisting of the RESI and the FINDI. The results reaffirm the existence of a dichotomy that exists in the South African market as the regressed returns are affected by either of the two sub-indices identified. Therefore, the two-factor model consisting of the FINDI and the RESI is a sufficient arbitrage pricing theory model for describing the systematic risk in the South African market, and it is a model that provides a superior account of results and greater detail when compared to the CAPM using the ALSI.

Hence, an application of the analysis into the performance of various VaR models in the South African market can be applied to both sub-indices and compared to the performance of these same models when making use of the ALSI, thereby expanding on the research performed by Elenjical, et al., (2016).

4. Data and Methodology

This section explores the data used and how the data points were divided into the three periods of interest, namely the pre-crisis, the crisis, and the post-crisis periods, for each of the three indices, namely the ALSI, the RESI, and the FINDI. The methodology employed to prepare the data for the analysis discussed in Chapter 5, involving the application of the five methods to calculate VaR figures (as outlined in Chapter 2) and then backtested using Kupiec's proportion of failure test (see below), is also discussed in this chapter.

The data used in this study were the closing values of the JSE's ALSI, along with its two primary sub-indices, as identified by van Rensburg (2002), namely the FINDI and the RESI. The indices' values were collected from the DataStream database for every trading day in the period described below.

The full period examined in this study spanned from 4 March 2003 to 31 December 2014 and was divided as follows:

1. The crisis period was considered to take place from 1 February 2007 to 30 September 2009 (Elenjical, et al., 2016), totalling to 695 trading days.
2. The pre-crisis period was considered to take place from 4 March 2003 to 31 January 2007, totalling to 1022 trading days.
3. The post-crisis period was considered to take place from 1 October 2009 to 31 December 2014, totalling to 1370 trading days.

Data for the calibration period, the period spanning 2 March 1998 to 3 March 2003, were also collected for the three indices. The length of the calibration period is in line with the literature and the period is used to calibrate the parameters of the various VaR models used in this study.

Expanding the research undertaken by Elenjical, et al., (2016), this study examines and evaluates the performance of different VaR models using the ALSI, and further breaks down the South African market into its two primary sub-indices, the RESI and the FINDI. It is of interest to test whether the VaR models discussed in Chapter 2 differ in their performances when a comparison is made between the ALSI and its sub-indices.

The ALSI was used as the proxy for the South African equity market as it captures 99% of the South African equity market's capitalisation. Further, as initially outlined by van Rensburg (2002), and confirmed by numerous studies, the returns earned on the South African equities market can be better captured by applying a two-factor arbitrage pricing theory model. The two factors being the FINDI and the RESI, two of the sub-indices of the ALSI representing the financials-industrials sector and the resources sector in South Africa, respectively.

The evaluation of the VaR prediction models on both the ALSI as well as its sub-indices is believed to provide depth into the management of risk in the South African context. As the two sub-indices are prominent sectors of the economy (as evidenced by South African equity returns being captured by these two variables in an arbitrage pricing theory application), it is of interest to investigate whether the various VaR models outlined in Chapter 2 perform similarly for the sub-indices as they do for the ALSI.

It is also important to note that the pre-crisis and the post-crisis periods are longer than those employed by Elenjical, et al., (2016). It is believed that the longer period will allow for better parameter estimation when employing the various VaR models and, hence, better forecasted VaR figures.

Using the closing values for the three indices for the full period, as well as the calibration period, returns on each index were calculated using the following formula:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) \tag{9}$$

where r_t is the daily return on the index examined at time t; P_t is the price (or level) of the index examined at time t; and P_{t-1} is the previous day's index price (or level). Hence, the series of returns r_t for the period represents the logged returns for the entire period, which can then be separated into the various sub-period under consideration. In this study, the terms 'logged returns', 'log returns', and 'returns' are used interchangeably and refer to the returns on the three indices calculated by using Equation (9).

A breach of the VaR reserves was recorded when the negative returns obtained on any signle trading day exceeded the value of the reserves held, as calculated by the VaR measure. This means that as the VaR figure was calculated daily, these figures were then compared daily to the returns achieved. If the negative return was greater than the VaR reserve in absolute terms, a breach was recorded.

Before any analysis can begin, outliers in the data were identified. Winsorisation was performed on any outlier in the pre-crisis and post-crisis periods. Winsorisation was not performed on the crisis period, as outliers are expected in such a high volatility period, and such outliers are crucial for the study conducted.

Table 1: Descriptive Statistics and Winsorisation Information

Index	*Number of Observations*	*Mean*	*Standard Deviation*	*Number of Winsorised Outliers*	*Number of Crisis Outliers*
ALSI	*3087*	*0.05754%*	*1.21256%*	*60*	*105*
RESI	*3087*	*0.02881%*	*1.81005%*	*55*	*102*
FINDI	*3087*	*0.07062%*	*1.01878%*	*77*	*92*

Note: This table reports the descriptive statistics for three major indices of the Johannesburg Stock Exchange (JSE), namely the All Share Index (ALSI), the Resources Index (RESI), and the Financials-Industrials Index (FINDI). The data were collected from the DataStream database for the period starting 4 March 2003 and ending 31 December 2014. The number of observations for each index was calculated by counting the number of closing prices of each index, as downloaded from DataStream. The number of outliers was calculated, whereby an outlier was identified as any data point whose value either exceeds the mean of the data plus two standard deviations or is less than the mean of the data less two standard deviations. Outliers in the pre-crisis and post-crisis periods were Winsorised, where their values were replaced with the boundary to which they were closest too, while crisis period outliers were not Winsorised, as these outliers are crucial to the evaluation of the value-at-risk models examined in this study.

Using Winsorisation, together with the 2 standard deviation rule, any data point whose value fell above the mean of the data plus two standard deviations, or below the mean less two standard deviations, had its value replaced with the closer of the two boundaries. Table 1 above shows the mean and standard deviation of the data for each of the three indices, together with the number of outliers which were Winsorised, and the number of outliers identified in the crisis period, which were not Winsorised.

For the sake of both completeness and robustness, all of the VaR models examined in this study were also applied to the raw data, i.e., the data without any Winsorisation applied to it, at the 1% significance level. The use of the raw data had no material effect on the top-performing models across different periods.

The use of the VaR models examined in this study required the use of a calibration period when employing the various models, as mentioned above. This calibration period allows for the calibration of the various parameters employed by the different VaR models and, hence, allows for the calculation of the forecasted VaR figures for the period under examination. Therefore, a period of 1306 trading days, or approximately five years, prior to the start of each of the three periods outlined above (i.e., the pre-crisis, the crisis, and the post-crisis periods) was included in the data as a calibration period. This calibration period is consistent with prior literature (see Chapter 3 and Elenjical, et al., 2016).

Table 2: Periods Examined

Period	*Period Start*	*Period End*	*Number of Observations*
Pre-Crisis	*4 March 2003*	*31 January 2007*	*1022*
Crisis	*1 February 2007*	*30 September 2009*	*695*
Post-Crisis	*1 October 2009*	*31 December 2014*	*1370*
Full Period	*4 March 2003*	*31 December 2014*	*3087*

Note: This table reports the periods used in this study. The pre-crisis period is taken as the period from 4 March 2003 to 31 January 2007, with a 5-year calibration period starting on 2 March 1998. The crisis period is taken as the period from 1 February 2007 to 30 September 2009, with a 5-year calibration period starting on 1 February 2002. The post-crisis period is taken as the period from 1 October 2009 to 31 December 2014, with a 5-year calibration period starting on 1 October 2005. The full period is taken as the period from 4 March 2003 to 31 December 2014, with a 5-year calibration period starting on 2 March 1998. The 5-year calibration period is consistent with the literature and is used to calibrate some parameters which are used in the different value-at-risk models. The data was collected from the DataStream database for the Johannesburg Stock Exchange's All Share Index, Resources Index, and the Financials-Industrials Index for the periods covering each period as well as each period's calibration period.

Table 2 above outlines the sub-periods. As mentioned earlier, a calibration period was used to calibrate the model prior to any forecasting period. While it is possible to have separate calibration periods for each of the sub-periods, each starting 1306 trading prior to the first closing value of the index in question, the literature suggests that a rolling calibration period allows for more accurate calibration. Hence, all periods are calibrated by first examining the calibration period for the pre-crisis period, namely the period spanning 2 March 1998 to 3 March 2003, with a rolling period of 1306 trading days. This means that as the model forecasted VaR one day further into the future, the calibration period for that forecast moved sequentially with the forecasting output. The daily returns obtained for the calibration period were Winsorised using the same criteria as the pre-crisis and post-crisis periods' data.

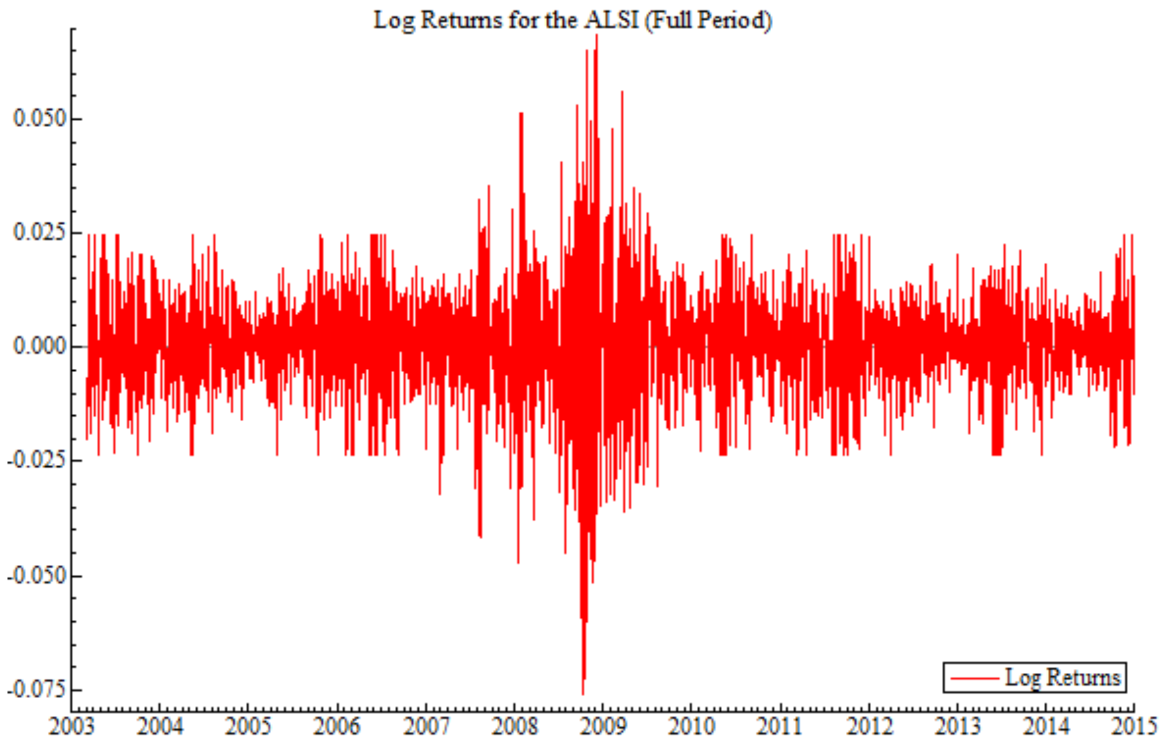
As a rolling calibration period was used, the calibration periods overlap with previous periods of interest, e.g., the calibration period of the post-crisis period may have included data from the pre-crisis period, the crisis period, and the post-crisis period. This is expected and is of use in this study, as these overlapping periods assist in providing insight into the performance of the different VaR models as return volatility changes.

Considering the details outlined above, while the full observation period is considered to be that starting 4 March 2003 and ending 31 December 2014, data were collected for the period starting 2 March 1998 and ending 31 December 2014.

Figures 1 to 3 below graphically depict the logged returns experienced over the full period for the ALSI, the FINDI, and the RESI, respectively. As mentioned above, the log returns for the pre-crisis and post-crisis periods were Winsorised, while the log-returns for the crisis period were not.

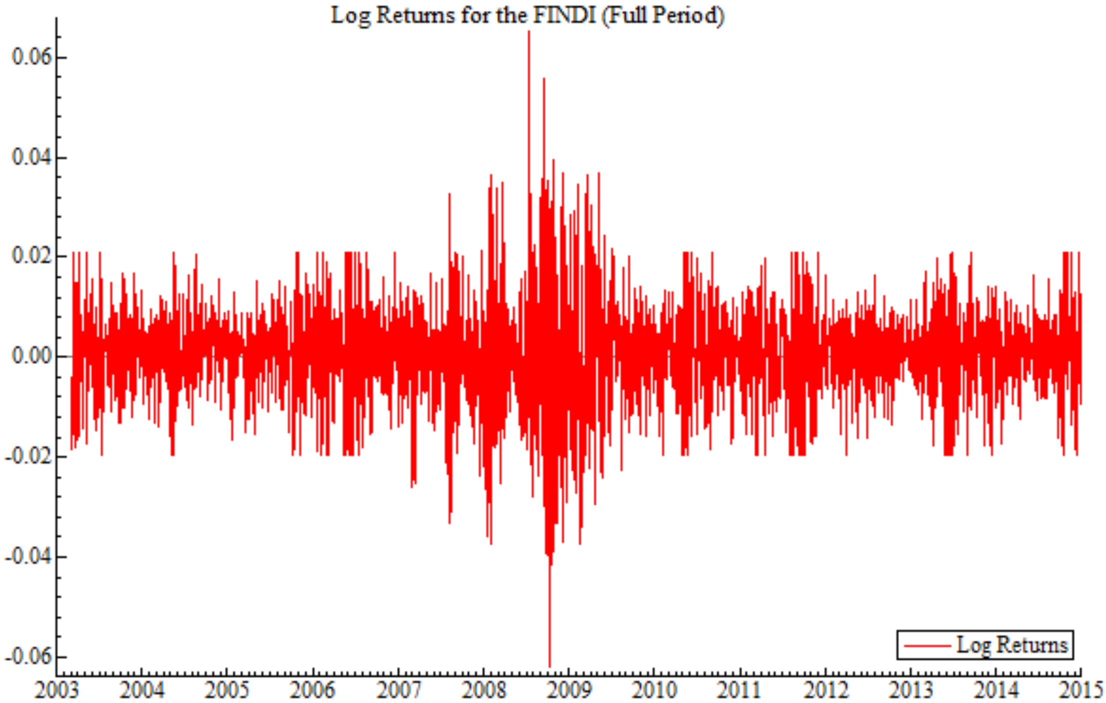
Each figure graphically highlights the highly volatile market conditions experienced during the crisis period. This is evident by the large spikes in each of the three diagrams, found during the crisis period. This is especially evident between 2008 and 2009, leading to the conclusion that the effects of the 2008 global financial crisis were only really felt in the South African market in the second half of 2008, about 18 months after the start of the crisis period identified by Elenjical, et al., (2016).

Figure 1: Daily log returns for the ALSI for the full period



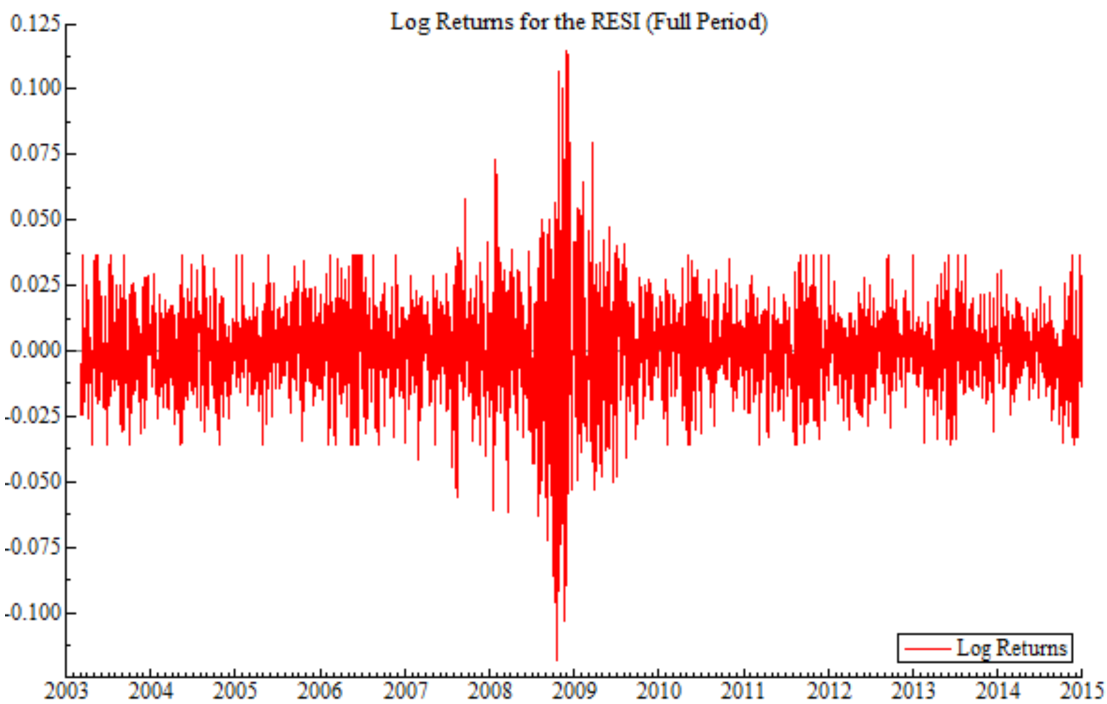
Note: This figure graphically depicts the logged daily returns of the South African All Share Index (ALSI) for the full period under consideration in this study, 4 March 2003 to 31 December 2014. The graph was obtained using data from the DataStream database and graphed using the OxMetrics 7 software.

Figure 2: Daily log returns for the FINDI for the full period



Note: This figure graphically depicts the logged daily returns of the South African Financials-Industrials Index (FINDI) for the full period under consideration in this study, 4 March 2003 to 31 December 2014. The graph was obtained using data from the DataStream database and graphed using the OxMetrics 7 software.

Figure 3: Daily log returns for the RESI for the full period



Note: This figure graphically depicts the logged daily returns of the South African Resources Index (RESI) for the full period under consideration in this study, 4 March 2003 to 31 December 2014. The graph was obtained using data from the DataStream database and graphed using the OxMetrics 7 software.

The daily log returns calculated for each of the three indices were used as input for the OxMetrics 7 software, which was used as the software of choice for the analysis performed in this study. The software employed maximum log-likelihood estimation to determine the parameters of the model. The models applied using the OxMetrics software included the GARCH(1,1), the EGARCH(1,1), and the RiskMetrics models. Each model was applied to each of the datasets (i.e., each time series of returns of each of the three indices) using a Gaussian distribution twice: Once producing a series of forecasted VaR figures for the index examined using a 95% significance level, and a second time using the same returns series to produce a series of forecasted VaR figures for the index examined using a 99% confidence interval. The RiskMetrics method was calibrated using the default λ_{RM} value of 0.94.

The historical simulation and the delta-normal methods were computed using Microsoft Excel. Similar to the GARCH, EGARCH, and RiskMetrics models implemented by the OxMetrics software, the historical simulation and the delta-normal methods also made use of a calibration period of equal length, i.e., 1306 trading days were used to calibrate each model. The parameters for each of the two models were updated daily.

The historical simulation method outputted daily VaR forecasts using the rolling calibration period as its historical period. The daily forecasts were made using either 95% or 99% as the degree of confidence with the function PERCENTILE in Microsoft Excel.

The delta-normal method also outputted daily VaR forecasts using the same rolling calibration period. While also using either 95% or 99% as the degree of confidence, the product of the functions NORMSINV and STDEV was used to calculate the forecasts in Microsoft Excel.

The VaR forecasts achieved for all models (i.e., all combinations of VaR model, significance level, index, and period), along with their breaches, was tested for statistical significance using Kupiec's proportion of failure test.

Kupiec's proportion of failure test was introduced by Paul Kupiec in 1995. Kupiec assumes that the breaches recorded for any model are randomly distributed as a binomial random variable with the parameters $N + 1$ trading days and probability p.

The test is one of comparing probabilities under the two statistical hypotheses. The null hypothesis of the test states that the likelihood of breaches experienced in the model, p, is equal to the likelihood observed during the backtesting procedure, i.e., $p = \hat{p}$, where $\hat{p}$ is the ratio of the number of observed violations, denoted by x, to 1 less than the total number of trading days examined under the model, i.e., N. The alternative hypothesis under the proportion of failure test assumes that p is that likelihood decided upon by the modeller, often 0.01 (Kupiec, 1995).

Kupiec (1995) then formulates the proportion of failure test as the ratio of the likelihood of a breach under the alternative hypothesis to likelihood of a breach under the null hypothesis, denoted by Λ, which is calculated as follows:

$$\Lambda = \frac{p^{N-x} \times (1-p)^x}{(1-\hat{p})^{N-x} \times \hat{p}^x} \tag{10}$$

A manipulation of Equation (10) allows for a more intuitive use in inference procedures. This manipulation is presented in Equation (11) below, where λ is used to denote $-2\ln(\Lambda)$, and q is used to denote $1-p$.

$$\lambda = 2\ln\left(\left(\frac{N-x}{pN}\right)^{N-x} \times \left(\frac{x}{Nq}\right)^x\right) \tag{11}$$

λ is then used as the test statistic of the proportion of failure test. It is compared to the critical value obtained for a Chi-squared random variable with 1 degree of freedom (Lehmann & Romano, 2005), say c. If $\lambda > c$, the null hypothesis of the proportion of failure test is rejected, implying that the model under examination is inaccurate and the number of breaches experienced is, in fact, false.

An alternative method to determining the outcome of the proportion of failure test is to construct a confidence interval for the number of violations. A confidence level of $(1-\alpha_{CL})\%$ must be determined for each level of confidence (i.e., either 95% or 99%).

By substituting either 1% or 5% for the value of α_{CL}, a confidence interval can be set using the following two equations:

$$\Pr[X < l] \leq \frac{\alpha_{CL}}{2} \tag{12}$$

$$\Pr[u < X] \leq \frac{\alpha_{CL}}{2} \tag{13}$$

where l is the lower bound of the confidence interval; u is the upper bound of the same interval; and X is the random variable taking on the value of x violations in the observed model under the null hypothesis.

While recalling the fact that λ in Equation (11) is distributed as a Chi-squared with 1 degree of freedom, by equating the critical value corresponding to each of Equations (12) and (13) to

Equation (11), the values of l and u can be determined, thereby leading to the confidence interval $[\lfloor l \rfloor, \lceil u \rceil]$. Note that the confidence interval was made slightly larger using the floor function for l and the ceiling function for u to allow for an integer confidence interval. This is discussed further and computed below.

Using Equations (11), (12), and (13), the boundaries of the confidence intervals were obtained for each of the sub-periods. Since all models (again, all combinations) were forecasted using either 95% or 99% as the degree of confidence, the critical values of the Chi-squared test (with 1 degree of freedom) are common to all tests. These are 3.841459 and 6.634897, respectively.

Table 3: Kupiec's Proportion of Failure Test Confidence Intervals

Sub-Period	$\alpha = 1\%$	$\alpha = 5\%$
Pre-Crisis Period	[3,20]	[38,66]
Crisis Period	[1,15]	[24,47]
Post-Crisis Period	[5,25]	[53,85]
Full Period	[17,47]	[131,179]

Note: This table reports the confidence intervals attained for each of the four sub-periods under examination. The length of the sub-crisis, crisis, post-crisis, and full periods are 1022, 695, 1370, and 3087 trading days, respectively. The critical values of the χ_1^2 for the 95% and 99% degrees of confidence are 3.841459 and 6.634897, respectively.

Combining the χ_1^2 critical values with the lengths of the sub-periods yielded confidence intervals for each of the four sub-periods (the pre-crisis period, the crisis period, the post-crisis period, and the full period), as displayed in Table 3 above.

The null hypothesis of the proportion of failure test was rejected if the number of observed violations did not fall within the confidence interval applicable to the degree of confidence. This will be used in the analysis of the study.

5. Results and Analysis

This chapter analyses the results obtained using the data and methodology outlined in Chapter 4, as applied to the various VaR models outlined in Chapter 2. More importantly, an application of Kupiec's proportion of failure backtesting technique, as outlined in Chapter 4, is performed in this chapter.

Throughout this chapter, tables and figures relating to models' output at the 99% degree of confidence are displayed, while similar tables and figures for the 95% degree of confidence are included in the appendices. The analysis and discussion of all figures, however, are performed in this chapter.

Five models were applied at two varying degrees of confidence, for each of the three indices. Hence, 30 models were effectively tested. All autoregressive models used the normal distribution as the underlying distribution of returns. VaR predictions were forecasted using a rolling calibration period of 1306 trading days, with the first forecast performed for 4 March 2003 and the last forecasted performed for 31 December 2014.

Each sub-period first discusses some of the descriptive statistics of the period, together with some brief analysis of the performance of the various VaR models when evaluated solely based on the number of breaches observed. Then, each section discusses the statistical significance of the violations experienced in each of the models by employing Kupiec's proportion of failure test.

5.1. The Pre-Crisis Period

This section discusses and analyses the results of the VaR forecasting models for the three indices under examination for the pre-crisis period. The pre-crisis period took place from 4 March 2003 to 31 January 2007, spanning a total of 1022 trading days.

Table 4 below depicts some of the descriptive statistics observed from the data for the pre-crisis period. As can be seen, when evaluating the performance of a model based on its number of breaches, it is of no surprise that the historical simulation method outperforms all others, across all indices and degrees of confidence, achieving 0 breaches in the pre-crisis period. This can be attributed to the calibration period involved in the forecasting procedure being relatively stable, or maybe even a booming period, leading to few violations. This implies that the historical simulation method forecasts loss values which are too high, rendering it useless in South Africa. This is in line with the conclusion made by Seymour and Polakow (2003).

The next best model, again evaluated using solely the number of breaches, is the delta-normal method. This method seems to outperform most other models (excluding the historical simulation method) across the three indices and both degrees of confidence.

It is interesting to note that the relatively less sophisticated methods seem to outperform the more sophisticated autoregressive VaR models during the pre-crisis period, with a relatively calm calibration period.

Table 4: Violations Observed for the Various Value-at-Risk Models at both 1% and 5% Significance Levels (Pre-Crisis Period)

Index		***ALSI***		***FINDI***		***RESI***	
Model	***Confidence Level***	***(1) Observed Violations***	***(2) (1)/1022***	***(1)***	***(2)***	***(1)***	***(2)***
GARCH(1,1)	*1%*	*8*	*0.7828%*	*9*	*0.8806%*	*12*	*1.1742%*
	5%	*40*	*3.9139%*	*43*	*4.2074%*	*49*	*4.7945%*
EGARCH(1,1)	*1%*	*11*	*1.0763%*	*16*	*1.5656%*	*15*	*1.4677%*
	5%	*37*	*3.6204%*	*41*	*4.0117%*	*47*	*4.5988%*
RiskMetrics	*1%*	*12*	*1.1742%*	*17*	*1.6634%*	*16*	*1.5656%*
	5%	*50*	*4.8924%*	*52*	*5.0881%*	*53*	*5.1859%*
Historical Simulation	*1%*	*0*	*0.0000%*	*0*	*0.0000%*	*0*	*0.0000%*
	5%	*34*	*3.3268%*	*30*	*2.9354%*	*49*	*4.7945%*
Delta-Normal	*1%*	*9*	*0.8806%*	*4*	*0.3914%*	*11*	*1.0763%*
	5%	*34*	*3.3268%*	*42*	*4.1096%*	*44*	*4.3053%*

Note: This table reports the pre-crisis period descriptive statistics for the application of various value-at-risk (VaR) forecasting models at varying levels of confidence, to three indices listed on the Johannesburg Stock Exchange, namely the All Share Index (ALSI), the Financials-Industrials Index (FINDI), and the Resources Index (RESI). The number of violations was then recorded [Columns (1)], where a violation is considered a log return achieved on the index that's lower than the forecasted VaR figure. The proportion of violations out of the total number of observations are shown in the respective Columns (2). The total number of observations is 1022.

Tables 5 to 7 below, together with Tables 20 to 22 in Appendix 1, depict Kupiec's proportion of failure test for the ALSI, the FINDI, and the RESI for the pre-crisis period. The confidence interval was taken from Table 3 above (see Chapter 4). Tables 5 to 7 show the results of the proportion of failure test at the 1% significance level, while Tables 20 to 22 show the results corresponding to the 5% significance level.

Kupiec's proportion of failure test yields the conclusion that the historical simulation method inaccurately forecasts VaR figures as it is an inaccurate model for all three indices at the 1% confidence level, a conclusion in line with other South African studies. The method yields no violations at all, leading to the conclusion that its VaR estimates are excessively high when compared to actual losses incurred during the period. This can also be seen as the number of violations, zero, is less than the lower bound of the confidence interval for the pre-crisis period at the 1% confidence level.

At the 5% significance level, on the other hand, the historical simulation method is only rejected for the ALSI and the FINDI, while not rejected for the RESI. Moreover, the EGARCH(1,1) method and the delta-normal method are also rejected as accurate models for the ALSI during the pre-crisis period at the 5% significance level.

Table 5: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (ALSI, Pre-Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	8	$8 \in [3,20]$. *Cannot reject* H_0
EGARCH(1,1)	11	$11 \in [3,20]$. *Cannot reject* H_0
RiskMetrics	12	$12 \in [3,20]$. *Cannot reject* H_0
Historical Simulation	0	$0 \notin [3,20]$. *Reject* H_0
Delta-Normal	9	$9 \in [3,20]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [3,20] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 6: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (FINDI, Pre-Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	9	$9 \in [3,20]$. *Cannot reject* H_0
EGARCH(1,1)	16	$16 \in [3,20]$. *Cannot reject* H_0
RiskMetrics	17	$17 \in [3,20]$. *Cannot reject* H_0
Historical Simulation	0	$0 \notin [3,20]$. *Reject* H_0
Delta-Normal	4	$4 \in [3,20]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [3,20] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 7: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (RESI, Pre-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	12	$12 \in [3,20]$. *Cannot reject* H_0
EGARCH(1,1)	15	$15 \in [3,20]$. *Cannot reject* H_0
RiskMetrics	16	$16 \in [3,20]$. *Cannot reject* H_0
Historical Simulation	0	$0 \notin [3,20]$. *Reject* H_0
Delta-Normal	11	$11 \in [3,20]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [3,20] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

The rejection of the historical simulation method as an accurate VaR model for the pre-crisis period at both significance levels contradicts the conclusions made above when only the number of violations were considered. This highlights the importance of conducting meaningful statistical tests to confirm any conclusions made by simply looking at the figures provided. The rejection of the method does, however, support the conclusions of earlier studies, such as Seymour and Polakow (2003).

Since the historical simulation was rejected, the next best model for the pre-crisis period, as identified in the discussion above, is the delta-normal method. Although the model was rejected for the ALSI at the 5% significance level, its VaR forecasting results were still more accurate and statistically meaningful for the three indices across both confidence levels.

5.2. The Crisis Period

This section discusses and analyses the results of the VaR forecasting models for the three indices under examination for the crisis period. The crisis period took place from 1 February 2007 to 30 September 2009, spanning a total of 695 trading days.

Table 8 below depicts some of the descriptive statistics observed from the data for the crisis period. The descriptive statistics reveal that every single model fails to maintain a breach level below its prescribed degree of certainty. This is not surprising given that the calibration period begins by examining the periods before the pre-crisis period, and then proceeds to roll over to the pre-crisis period. Both of these periods are relatively stable when it comes to the volatility of returns, leading many models to be calibrated with data that is not an accurate reflection of the true market conditions.

In complete contrast to the pre-crisis period, the less sophisticated historical simulation and delta-normal methods excessively underperform across all three indices, with their violations being sometimes double those of the more sophisticated autoregressive models. This is not a surprise considering the nature of the calibration of these models.

Table 8: Violations Observed for the Various Value-at-Risk Models at both 1% and 5% Significance Levels (Crisis Period)

Index		***ALSI***		***FINDI***		***RESI***	
Model	***Confidence Level***	***(1) Observed Violations***	***(2) (1)/695***	***(1)***	***(2)***	***(1)***	***(2)***
GARCH(1,1)	*1%*	*13*	*1.8705%*	*12*	*1.7266%*	*15*	*2.1583%*
	5%	*50*	*7.1942%*	*45*	*6.4748%*	*52*	*7.4820%*
EGARCH(1,1)	*1%*	*17*	*2.4460%*	*5*	*0.7194%*	*18*	*2.5899%*
	5%	*59*	*8.4892%*	*34*	*4.8921%*	*55*	*7.9137%*
RiskMetrics	*1%*	*12*	*1.7266%*	*10*	*1.4388%*	*14*	*2.0144%*
	5%	*44*	*6.3309%*	*38*	*5.4676%*	*47*	*6.7626%*
Historical Simulation	*1%*	*32*	*4.6043%*	*33*	*4.7482%*	*32*	*4.6043%*
	5%	*88*	*12.6619%*	*78*	*11.2230%*	*49*	*11.3669%*
Delta-Normal	*1%*	*44*	*6.3309%*	*38*	*5.4676%*	*39*	*5.6115%*
	5%	*87*	*12.5180%*	*83*	*11.9242%*	*76*	*10.9353%*

Note: This table reports the crisis period descriptive statistics for the application of various value-at-risk (VaR) forecasting models at varying levels of confidence, to three indices listed on the Johannesburg Stock Exchange, namely the All Share Index (ALSI), the Financials-Industrials Index (FINDI), and the Resources Index (RESI). The number of violations was then recorded [Columns (1)], where a violation is considered a log return achieved on the index that's lower than the forecasted VaR figure. The proportion of violations out of the total number of observations are shown in the respective Columns (2). The total number of observations is 695.

The models that outperform the others based on the number of violations are the RiskMetrics model for the ALSI and the RESI, while the EGARCH(1,1) model seems to be the superior model for the FINDI. These results are in line with some of the international literature, such as So and Yu (2006) for the RiskMetrics model.

An interesting observation is the outperformance of the RiskMetrics model using the ALSI contradicts the findings of Elenjical, et al., (2016). This, however, may be due to the different methodology applied in this study, namely the length of the sub-periods examined. An even more interesting observation is the fact that a different model outperforms when evaluating FINDI returns, suggesting that the FINDI may behave differently to the ALSI, while the RESI's behaviour may be more closely aligned with the ALSI when it comes to risk management.

Tables 9 to 11 below, together with Tables 23 to 25 in Appendix 2, depict Kupiec's proportion of failure test for the ALSI, the FINDI, and the RESI for the crisis period. The confidence interval was again taken from Table 3 above. Tables 9 to 11 show the results of the proportion of failure test at the 1% significance level, while Tables 23 to 25 show the results corresponding to the 5% significance level.

The proportion of failure test provides the following conclusions regarding rejected models:

1. At the 1% level of confidence,
    a. the EGARCH(1,1) model, the historical simulation model, and the delta-normal model are all inaccurate models when evaluated using the ALSI and the RESI.
    b. the historical simulation model and the delta-normal model are inaccurate models when evaluated using the FINDI.
2. At the 5% level of confidence,
    a. the GARCH(1,1) model, the EGARCH(1,1) model, the historical simulation model, and the delta-normal model are all inaccurate models when evaluated using the ALSI and the RESI.
    b. the historical simulation model and the delta-normal model are inaccurate models when evaluated using the FINDI.

The conclusions provided by Kupiec's proportion of failure test once again provide statistically significant detail about the performance of the various models employed. It seems that the performances of models at both levels of confidence are similar for both the ALSI and the RESI during the crisis period, while very different for the FINDI. This is an interesting observation, suggesting that the FINDI reacts differently to the market volatility than the ALSI.

This further supports the analysis performed earlier in this section, considering solely the number of violations experienced in this period. The RESI behaved more closely to the ALSI rather than the FINDI, a conclusion that is supported by the proportion of failure test.

The RiskMetrics model has not been rejected by any model across all indices and levels of confidence. This conclusion is in line with international literature, stating that it is well-suited for developing markets (McMillan & Kambouroudis, 2009). The model was identified in the discussion above to be best suited for the ALSI and the RESI, but it seems to prove to be only statistically adequate for the FINDI during the crisis period. The GARCH(1,1) model is a close second, with only one more violation when compared to the RiskMetrics model for almost all of its runs. The EGARCH(1,1) model is also a statistically adequate model for the FINDI during the crisis period.

Table 9: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (ALSI, Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	13	$13 \in [1,15]$. *Cannot reject* H_0
EGARCH(1,1)	17	$17 \notin [1,15]$. *Reject* H_0
RiskMetrics	12	$12 \in [1,15]$. *Cannot reject* H_0
Historical Simulation	32	$32 \notin [1,15]$. *Reject* H_0
Delta-Normal	44	$44 \notin [1,15]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the crisis period. The number of observed violations must fall outside the confidence interval $[1,15]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 10: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (FINDI, Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	12	$12 \in [1,15]$. *Cannot reject* H_0
EGARCH(1,1)	5	$5 \in [1,15]$. *Cannot reject* H_0
RiskMetrics	10	$10 \in [1,15]$. *Cannot reject* H_0
Historical Simulation	33	$33 \notin [1,15]$. *Reject* H_0
Delta-Normal	38	$38 \notin [1,15]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the crisis period. The number of observed violations must fall outside the confidence interval $[1,15]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 11: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (RESI, Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	15	$15 \in [1,15]$. *Cannot reject* H_0
EGARCH(1,1)	18	$18 \notin [1,15]$. *Reject* H_0
RiskMetrics	14	$14 \in [1,15]$. *Cannot reject* H_0
Historical Simulation	32	$32 \notin [1,15]$. *Reject* H_0
Delta-Normal	39	$39 \notin [1,15]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the crisis period. The number of observed violations must fall outside the confidence interval $[1,15]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

5.3. The Post-Crisis Period

This section discusses and analyses the results of the VaR forecasting models for the three indices under examination for the post-crisis period. The post-crisis period took place from 1 October 2009 to 31 December 2014, spanning a total of 1370 trading days.

Table 12 below depicts some of the descriptive statistics observed from the data for the post-crisis period. The analysis for the post-crisis period is slightly different from the analysis performed for the pre-crisis and the crisis periods. This is because the volatile conditions captured in the calibration period of this period will most likely lead to models with very few breaches. This is not due to the accuracy of the prediction model, but rather because the models overestimate the VaR reserves required, leading to little to no violations. This point is clearly observed when examining the performance of the less sophisticated methods, but specifically, the historical simulation method, achieving no violations for two out of the three indices examined at the 1% significance level, and almost a fifth of the number of violations of some of the other models at the 5% significance level.

Due to the nature of the autoregressive models, it is expected that they will outperform in this period. This should be clearer when examining the statistical significance of the violations observed using Kupiec's proportion of failure test further down. However, if the historical simulation and delta-normal methods are excluded, the EGARCH(1,1) model outperforms the RiskMetrics and GARCH(1,1) model across all three indices and two degrees of confidence. This observation is in line with the findings of Elenjical, et al., (2016), as well as other studies (both South African and international), although it is interesting to note the lack of heterogeneity observed in the models' performance when compared to the crisis period.

The more sophisticated methods seem to exhibit a proportion of violations that is closer to the level of significance of the various models evaluated when compared to the less sophisticated methods. This can be seen in Table 12 below. The RiskMetrics method, however, exceeds the 5% significance level across all three indices, achieving a proportion of violations as high as 6.5693% for the RESI. This suggests a large number of violations experienced with this method, even though the calibration period should result in estimates that are too high.

Table 12: Violations Observed for the Various Value-at-Risk Models at both 1% and 5% Significance Levels (Post-Crisis Period)

Index		***ALSI***		***FINDI***		***RESI***	
Model	***Confidence Level***	***(1) Observed Violations***	***(2) (1)/1370***	***(1)***	***(2)***	***(1)***	***(2)***
GARCH(1,1)	*1%*	*14*	*1.0219%*	*16*	*1.1679%*	*14*	*1.0219%*
	5%	*70*	*5.1095%*	*64*	*4.6715%*	*71*	*5.1825%*
EGARCH(1,1)	*1%*	*14*	*1.0219%*	*13*	*0.9489%*	*11*	*0.8029%*
	5%	*55*	*4.0146%*	*60*	*4.3796%*	*66*	*4.8175%*
RiskMetrics	*1%*	*19*	*1.3869%*	*22*	*1.6058%*	*23*	*1.6788%*
	5%	*87*	*6.3504%*	*80*	*5.8394%*	*90*	*6.5693%*
Historical Simulation	*1%*	*0*	*0.0000%*	*0*	*0.0000%*	*2*	*0.1460%*
	5%	*13*	*0.9489%*	*30*	*2.1898%*	*24*	*1.7518%*
Delta-Normal	*1%*	*4*	*0.2920%*	*3*	*0.2190%*	*5*	*0.3650%*
	5%	*12*	*0.8759%*	*32*	*2.3358%*	*19*	*1.3869%*

Note: This table reports the post-crisis period descriptive statistics for the application of various value-at-risk (VaR) forecasting models at varying levels of confidence, to three indices listed on the Johannesburg Stock Exchange, namely the All Share Index (ALSI), the Financials-Industrials Index (FINDI), and the Resources Index (RESI). The number of violations was then recorded [Columns (1)], where a violation is considered a log return achieved on the index that's lower than the forecasted VaR figure. The proportion of violations out of the total number of observations are shown in the respective Columns (2). The total number of observations is 1370.

Tables 13 to 15 below, together with Tables 26 to 28 in Appendix 3, depict Kupiec's proportion of failure test for the ALSI, the FINDI, and the RESI for the post-crisis period. The confidence interval was again taken from Table 3 above. Tables 13 to 15 show the results of the proportion of failure test at the 1% significance level, while Tables 26 to 28 show the results corresponding to the 5% significance level.

The proportion of failure test provides the following conclusions regarding rejected models:

1. At the 1% significance level,
    a. the historical simulation model is rejected as an accurate model across all three indices.
    b. the delta-normal model is rejected as an accurate model when evaluated using ALSI and FINDI returns.

2. At the 5% significance level,
    a. The historical simulation model and the delta-normal model are rejected as accurate models across all indices.
    b. The RiskMetrics model is rejected as an accurate model when evaluated using ALSI and RESI returns.

As expected, the historical simulation method and the delta-normal method are concluded to be inaccurate. This is unsurprising due to the calibration nature of the models, where the data used to calibrate the two models include the highly-volatile crisis period, leading to higher VaR forecasts and excess capital held as reserves.

Moreover, the RiskMetrics method is also concluded to be inaccurate, supporting the discussion made above surrounding its excessive proportion of violations when compared to the significance level. It is interesting, however, that the method is not rejected for the FINDI, once again suggesting that the RESI and the ALSI are closely related, and the FINDI reacts differently to risk management models.

Further to the discussion above, since the historical simulation model and the delta-normal model were excluded by Kupiec's proportion of failure test, it is concluded that the EGARCH(1,1) model is the superior model during the post-crisis period across all indices and levels of certainty. The GARCH(1,1) model is also a good fit for the ALSI. These conclusions are in line with the literature, as well as Elenjical, et al., (2016).

Table 13: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (ALSI, Post-Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	14	$14 \in [5,25]$. *Cannot reject* H_0
EGARCH(1,1)	14	$14 \in [5,25]$. *Cannot reject* H_0
RiskMetrics	19	$19 \in [5,25]$. *Cannot reject* H_0
Historical Simulation	0	$0 \notin [5,25]$. *Reject* H_0
Delta-Normal	4	$4 \notin [5,25]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the post-crisis period. The number of observed violations must fall outside the confidence interval $[5,25]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 14: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (FINDI, Post-Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	16	$16 \in [5,25]$. *Cannot reject* H_0
EGARCH(1,1)	13	$13 \in [5,25]$. *Cannot reject* H_0
RiskMetrics	22	$22 \in [5,25]$. *Cannot reject* H_0
Historical Simulation	0	$0 \notin [5,25]$. *Reject* H_0
Delta-Normal	3	$3 \notin [5,25]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the post-crisis period. The number of observed violations must fall outside the confidence interval $[5,25]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 15: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (RESI, Post-Crisis Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	14	$14 \in [5,25]$. *Cannot reject* H_0
EGARCH(1,1)	11	$11 \in [5,25]$. *Cannot reject* H_0
RiskMetrics	23	$23 \in [5,25]$. *Cannot reject* H_0
Historical Simulation	2	$2 \notin [5,25]$. *Reject* H_0
Delta-Normal	5	$5 \in [5,25]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the post-crisis period. The number of observed violations must fall outside the confidence interval $[5,25]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

5.4. The Full Period

This section discusses and analyses the results of the VaR forecasting models for the three indices under examination for the full period. The full period took place from 4 March 2003 to 31 December 2014, spanning a total of 3087 trading days.

Table 16 below depicts some of the descriptive statistics observed from the data for the full period. As can be seen, no model achieves a proportion of violations that is less than 1% for the 99% VaR predicted values for any of the three indices, while 8 models achieve less than 5% violations for the 95% VaR predicted values.

If, once again, a comparison is made based solely on the number of violations observed, the results are somewhat mixed. Overall, however, the historical simulation method and the delta-normal method seem to offer the best performance, with the exception of the FINDI, where the EGARCH(1,1) model seems to be the superior model. The two former models' superior performance is in line with Dias (2013). The latter leads, once again, to the conclusion that the FINDI may behave differently to the ALSI and the RESI.

As pointed out by Dias (2013), the fact that the less sophisticated models may offer better performances could be attributed to their averaging effects over the three sub-periods of the full period, namely the pre-crisis period, the crisis period, and the post-crisis period. The very few (or considerably lower number of) violations experienced by the historical simulation method and the delta-normal method during the pre-crisis and post-crisis periods balance out the very high number of violations experienced over the crisis period, therefore leaving a number that's relatively low. This does not mean that the methods are superior in their performance throughout, but rather that their averaged performance is seemingly better, most likely due to the clustering of violations experienced by the historical simulation method during the crisis period, and the lack of many violations in the other two sub-periods.

Table 16: Violations Observed for the Various Value-at-Risk Models at both 1% and 5% Significance Levels (Full Period)

Index		***ALSI***		***FINDI***		***RESI***	
Model	***Confidence Level***	***(1) Observed Violations***	***(2) (1)/3087***	***(1)***	***(2)***	***(1)***	***(2)***
GARCH(1,1)	*1%*	*35*	*1.1338%*	*37*	*1.1986%*	*41*	*1.3282%*
	5%	*160*	*5.1830%*	*152*	*4.9239%*	*172*	*5.5718%*
EGARCH(1,1)	*1%*	*42*	*1.3605%*	*34*	*1.1014%*	*44*	*1.4253%*
	5%	*151*	*4.8915%*	*135*	*4.3732%*	*168*	*5.4422%*
RiskMetrics	*1%*	*43*	*1.3929%*	*49*	*1.15873%*	*53*	*1.7169%*
	5%	*181*	*5.8633%*	*170*	*5.5070%*	*190*	*6.1548%*
Historical Simulation	*1%*	*32*	*1.0366%*	*33*	*1.0690%*	*34*	*1.1014%*
	5%	*135*	*4.3732%*	*138*	*4.4704%*	*152*	*4.9239%*
Delta-Normal	*1%*	*57*	*1.8465%*	*45*	*1.4577%*	*55*	*1.7817%*
	5%	*133*	*4.3084%*	*157*	*5.0858%*	*139*	*4.5028%*

Note: This table reports the full period descriptive statistics for the application of various value-at-risk (VaR) forecasting models at varying levels of confidence, to three indices listed on the Johannesburg Stock Exchange, namely the All Share Index (ALSI), the Financials-Industrials Index (FINDI), and the Resources Index (RESI). The number of violations was then recorded [Columns (1)], where a violation is considered a log return achieved on the index that's lower than the forecasted VaR figure. The proportion of violations out of the total number of observations are shown in the respective Columns (2). The total number of observations is 3087.

Tables 17 to 19 below, together with Tables 29 to 31 in Appendix 4, depict Kupiec's proportion of failure test for the ALSI, the FINDI, and the RESI for the full period. The confidence interval was again taken from Table 3 above. Tables 17 to 19 show the results of the proportion of failure test at the 1% significance level, while Tables 29 to 31 show the results corresponding to the 5% significance level.

The proportion of failure test provides the following conclusions regarding rejected models:

1. At the 1% significance level,
    a. the delta-normal method can be rejected as an inaccurate model for the ALSI.
    b. the RiskMetrics model can be rejected as an inaccurate model for the FINDI and the RESI.
2. At the 5% significance level,
    a. the RiskMetrics model can be rejected as an inaccurate model for the ALSI and the RESI.

Due to the rejection of the delta-normal method, as well as the discussion above regarding the averaging nature of the historical simulation method, the GARCH(1,1) model and the EGARCH(1,1) model seem to be the superior models in modelling and forecasting VaR figures over the full period, with the EGARCH(1,1) model performing better for the FINDI, while the GARCH(1,1) model performing better for the ALSI and the RESI, further highlighting the similarities in their performances.

The conclusions made above through the use of Kupiec's proportion of failure test are the core reason for this study. It seems that risk management differs between the ALSI and its sub-index, the FINDI, while displaying similarities between itself and the RESI over all of the different periods examined in this study. The RiskMetrics model is evidently a statistically adequate model for the ALSI over the full period under examination, while it is definitely not the case for the FINDI and the RESI, at the 1% significance level, while the model is rejected for the ALSI and the RESI, but not for the FINDI, at the 5% significance level. Overall, the EGARCH(1,1) and GARCH(1,1) models provide statistically adequate results for all three indices over the full period, a conclusion that is in line with those of Elenjical, et al., (2016), as well as other studies.

Table 17: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (ALSI, Full Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	*35*	$35 \in [17{,}47]$. *Cannot reject* H_0
EGARCH(1,1)	*42*	$42 \in [17{,}47]$. *Cannot reject* H_0
RiskMetrics	*43*	$43 \in [17{,}47]$. *Cannot reject* H_0
Historical Simulation	*32*	$32 \in [17{,}47]$. *Cannot reject* H_0
Delta-Normal	*57*	$57 \notin [17{,}47]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the full period. The number of observed violations must fall outside the confidence interval $[17{,}47]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 18: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (FINDI, Full Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	*37*	$37 \in [17{,}47]$. *Cannot reject* H_0
EGARCH(1,1)	*34*	$34 \in [17{,}47]$. *Cannot reject* H_0
RiskMetrics	*49*	$49 \notin [17{,}47]$. *Reject* H_0
Historical Simulation	*33*	$33 \in [17{,}47]$. *Cannot reject* H_0
Delta-Normal	*45*	$45 \in [17{,}47]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the full period. The number of observed violations must fall outside the confidence interval $[17{,}47]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 19: Results: Kupiec's Proportion of Failure Test at the 1% Significance Level (RESI, Full Period)

Model	***Observed Violations***	***Conclusion***
GARCH(1,1)	*41*	$41 \in [17{,}47]$. *Cannot reject* H_0
EGARCH(1,1)	*44*	$44 \in [17{,}47]$. *Cannot reject* H_0
RiskMetrics	*53*	$53 \notin [17{,}47]$. *Reject* H_0
Historical Simulation	*33*	$33 \in [17{,}47]$. *Cannot reject* H_0
Delta-Normal	*45*	$45 \in [17{,}47]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 1% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the full period. The number of observed violations must fall outside the confidence interval $[17{,}47]$ to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

6. Conclusion

In conclusion, this study was conducted using data from 2 March 1998 to 31 December 2014 to explore the performance of various VaR forecasting models before, during, and after the 2008 global financial crisis, as well as throughout the full period. The data used were those of the major index on the Johannesburg Stock Exchange, the ALSI, together with its two primary sub-indices, the FINDI and the RESI. The study aimed to investigate whether the performance of the VaR methods differed between the ALSI and its sub-indices, thereby expanding on prior studies.

The study divided the data of each of the three indices into three sub-periods, as well as the full period. The pre-crisis period took place from 4 March 2003 to 31 January 2007; the crisis period took place from 1 February 2007 to 31 September 2009; and the post-crisis period took place from 1 October 2009 to 31 December 2014. A historical period from 2 March 1998 to 3 March 2003 was used as a calibration period. The data for the pre-crisis and post-crisis periods for each of the three indices were Winsorised, while the crisis period data were not.

The data were then used as input when using the OxMetrics software in order to calibrate and produce forecasts for the autoregressive models investigated in this study, namely the GARCH(1,1) model, the EGARCH(1,1) model, and the RiskMetrics model with a lambda value of 0.94. The data were also used as input in Microsoft Excel, to produce VaR forecasts for the historical simulation method and the delta-normal method. These five methods were calibrated at 1% and 5% significance levels across all three indices, therefore producing 30 models in total.

The VaR forecasts were then analysed and a backtesting technique (Kupiec's proportion of failure test) was applied in order to test the statistical significance of each of the model and, therefore, its accuracy. The conclusions reached are as follows:

1. For the pre-crisis period, the delta-normal method was the superior method across all three models and the two levels of significance.
2. For the crisis period,
    a. the RiskMetrics method was the superior method across all three models and the two levels of significance with the GARCH(1,1) model, while the EGARCH(1,1) model was also concluded to be adequate for the FINDI. These results are consistent with the various literature concerning the performance of

the RiskMetrics model in developing markets and the performance of the GARCH(1,1) and EGARCH(1,1) models in general.

   b. the RESI was found to exhibit similar results to the ALSI when it comes to the performance of VaR models.

3. For the post-crisis period,
   a. the historical simulation method and the delta-normal method were found to be statistically inaccurate methods to forecast VaR.
   b. the RESI and the ALSI once again showed similar performances, especially when employing the RiskMetrics method.
   c. the EGARCH(1,1) method was the superior method across all three indices and the two levels of significance, while the GARCH(1,1) method was a good fit when using the ALSI. These conclusions are, once again, in line with the literature.
4. For the full period,
   a. the averaging tendencies of the historical simulation method were magnified and became evident, together with its violation-clustering characteristics. This result is consistent with the literature.
   b. the EGARCH(1,1) method was the superior method when using the FINDI across both levels of significance.
   c. the GARCH(1,1) method was the superior method when using the ALSI and the RESI across both levels of significance.
   d. the analysis suggests that the RESI exhibits similar results to the ALSI when it comes to the performance of VaR models once again.

The study shows a number of key findings. First, it reaffirms the conclusion made by Nieto and Ruiz (2016), which states that the performance of the various VaR methods is heavily dependent on the period examined. This means that institutions employing these VaR methods for the purpose of risk management must ensure that the period used to forecast VaR has similar characteristics to the period believed to begin with the forecasts. The underlying market conditions play a crucial role in VaR forecasting, even when employing autoregressive models.

Second, the study reveals similarities between the ALSI and its sub-index, the RESI, while also revealing dissimilarities between the former two and the FINDI. This is an important aspect when it comes to risk management as well as investment management in South Africa in general. The dissimilarities firstly highlight that caution should be taken when evaluating portfolios which are more heavily invested in the FINDI, as it reacts differently to the VaR models which South African firms would often apply to the market (i.e., the ALSI), without making any distinction. Secondly, the dissimilarities may prove some investment diversification benefits, a question which may be expanded on in future research.

The limitations encountered when performing this study mainly came down to the availability of data (no data was available for all three indices prior to March 1998), as well as the sole use of the normal distribution as the underlying distribution of returns. Further research into the effect of a change in distribution would be of interest, especially when employing a skewed Student's t distribution. Moreover, the RiskMetrics model parameter, λ_{RM}, could be evaluated from the data, as opposed to being set to the industry standard value of 0.94. The inclusion of more models in the investigation may also be of use, models such as extreme value theory, for example, as they have been used in the South African context in past literature. Further investigations should explore several backtesting techniques, as opposed to just Kupiec's proportion of failure test, as employed in this study. An example of one such technique is Christoffersen's test which deals with both clusters and breaches of VaR values.

References

Acerbi, C. & Tasche, D., 2002. On the Coherence of Expected Shortfall. *Journal of Banking & Finance,* Volume 26, pp. 1487-1503.

Alexander, C., 2008. *Market Risk Analysis IV: Value-at-Risk Models.* Chichester: John Wiley & Sons, Ltd.

Basel Committee on Banking Supervision, 1996. *Amendment to the Capital Accord to Incorporate Market Risks,* Basel: Bank for International Settlements.

Basel Committee on Banking Supervision, 2004. *International Covergence of Capital Measurement and Capital Standards: A Revised Framework,* Basel: Bank for International Settlements.

Basel Committee on Banking Supervision, 2016. *Minimum Capital Requirements for Market Risk,* Basel: Bank for International Settlements.

Basel Committee on Banking Supervision, 2017. *Basel III: Finalising Post-Crisis Reforms,* Basel: Bank for International Settlements.

Basel Committee on Banking Supervision, 2019. *Internal Models Approach: Backtesting and P&L Attribution Test Requirements,* Basel: Bank for International Settlements.

Berkowitz, J. & O'Brien, J., 2002. How Accurate are Value-at-Risk Models at Commercial Banks?. *The Journal of Finance,* 57(3), pp. 1093-1111.

Bollerslev, T., 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* Issue 307-327, p. 31.

Bollerslev, T., 2007. *Glossary to ARCH (GARCH),* s.l.: Duke University.

Bonga-Bonga, L. & Mutema, G., 2009. Volatility Forecasting and Value-at-Risk Estimation in Emerging Markets: The Case of the Stock Market Index Portfolio in South Africa. *South African Journal of Economic and Management Sciences,* Volume 4, pp. 401-411.

Breen, W., Glosten, L. R. & Jagannathan, R., 1989. Economic Significance of Predictable Variations in Stock Index Returns. *The Journal of Finance ,,* 44(5), pp. 1177-1189.

Christoffersen, P., Hahn, J. & Inoue, A., 2001. Testing and Comparing Value-at-Risk Measures. *Journal of Empirical Finance,* Volume 8, pp. 325-342.

Cifter, A., 2012. Volatility Forecasting with Asymmetric Normal Mixture GARCH Model: Evidence from South Africa. *Romanian Journal of Economic Forecasting,* Volume 2, pp. 127-142.

Daníelsson, J., 2002. The Emperor has no Clothes: Limits to Risk Modelling. *Journal of Banking & Finance,* Volume 26, pp. 1273-1296.

Delbaen, F., 2000. Coherent Risk Measures. In: *Blätter DGVFM.* Zurich: Springer-Verlag, pp. 733-739.

Dias, A., 2013. Market Capitalization and Value-at-Risk. *Journal of Banking & Finance,* Volume 37, pp. 5248-5260.

Dowd, K., 2002. *Measuring Market Risk.* Chichester: John Wiley & Sons, Ltd.

Elenjical, T., Mwangi, P., Panulo, B. & Huang, C.-S., 2016. A Comparative Cross-Regime Analysis on the Performance of GARCH-based Value-at-Risk Models: Evidence from the Johannesburg Stock Exchange. *Risk Management,* 18(2/3), pp. 89-110.

Embrechts, P., 2000. *Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool,* Zurich: ETHZ (Swiss Federal Institute of Technology, Zurich).

Engle, R., 1982. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica,* Volume 50, pp. 897-1007.

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

Giot, P. & Laurent, S., 2004. Modelling Daily Value-at-Risk using Realized Volatility and ARCH Type Models. *Journal of Empirical Finance,* Volume 11, pp. 379-398.

González-Rivera, G., Lee, T.-H. & Yoldas, E., 2007. Optimality of the RiskMetrics VaR Model. *Finance Research Letters,* Volume 4, pp. 137-145.

Hendricks, D., 1995. *Evaluation of Value-at-Risk Models Using Historical Data,* New York: The Federal Reserve Bank of New York.

Jorion, P., 1996. Risk2: Measuring the Risk in Value at Risk. *Financial Analysts Journal,* 52(6), pp. 47-56.

Kuester, K., Mittnik, S. & Paolella, M. S., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics,* 4(1), pp. 53-89.

Kupiec, P., 1995. Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives,* 3(2), pp. 73-84.

Lehmann, E. L. & Romano, J. P., 2005. *Testing Statistical Hypotheses.* 3rd ed. New York City: Springer.

Linsmeier, T. J. & Pearson, N. D., 2000. Value at Risk. *Financial Analyst Journal,* 56(2), pp. 47-67.

Lockwood, E., 2015. Predicting the Unpredictable: Value-at-Risk, Performativity, and the Politics of Financial Uncertainty. *Review of International Political Economy,* 22(4), pp. 719-756.

Marshall, C. & Siegel, M., 1997. Value at Risk: Implementing a Risk Measurement Standard. *Journal of Derivatives,* Volume 4, pp. 91-111.

McAleer, M. & da Veiga, B., 2008. Forecasting Value-at-Risk with a Parsimonious Portfolio Spillover GARCH (PS-GARCH) Model. *Journal of Forecasting,* Volume 27, pp. 1-19.

McMillan, D. G. & Kambouroudis, D., 2009. Are RiskMetrics Forecasts Good Enough? Evidence from 31 Stock Markets. *Inernational Review of Financial Analysis,* Volume 18, pp. 117-124.

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

Nieto, M. R. & Ruiz, E., 2016. Frontier in VaR Forecasting and Backtesting. *International Journal of Forecasting,* Volume 32, pp. 475-501.

Pafka, S. & Kondor, I., 2001. Evaluating the RiskMetrics Methodology in Measuring Volatility and Value-at-Risk in Financial Markets. *Physica A: Statistical Mechanics and its Applications,* 299(1-2), pp. 305-310.

Peiró, A., 1999. Skewness in Financial Returns. *Journal of Banking & Finance,* Volume 23, pp. 847-862.

Pérignon, C., Deng, Z. Y. & Wang, Z. J., 2008. Do Banks Overestimate their Value-at-Risk?. *Journal of Banking & Finance,* Volume 32, pp. 783-794.

Pérignon, C. & Smith, D. R., 2010. The Level and Quality of Value-at-Risk Disclosed by COmmercial Banks. *Journal of Banking & Finance,* Volume 34, pp. 362-377.

Pflug, G. C., 2000. Some Remarks on the Value-at-Risk and the Conditional Value-at-Risk. In: S. Uryasev, ed. *Probabilistic Constrained Optimization: Methodology and Applications.* Boston: Springer, pp. 272-282.

Pritsker, M., 2006. The Hidden Dangers of Historical Simulation. *Journal of Banking & Finance,* Volume 30, pp. 561-582.

RiskMetrics Group, Inc., 2001. *Return to RiskMetrics: The Evolution of a Standard,* New York: RiskMetrics Group, Inc..

Seymour, A. J. & Polakow, D. A., 2003. A Coupling of Extreme-Value Theory and Volatility Updating with Value-at-Risk Estimation in Emerging Markets: A South African Test. *Multinational Finance Journal,* 7(1&2), pp. 3-23.

Sharma, M., 2012. Evaluation of Basel III Revision of Quantitative Standards for Implementation of Internal Models for Market Risk. *IIMB Management Review,* Volume 24, p. 2340244.

Sharpe, W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance,* 19(3), pp. 425-442.

So, M. K. P. & Yu, P. L. H., 2006. Empirical Analysis of GARCH Models in Value at Risk Estimation. *International Financial Markets, Institutions & Money,* Volume 16, pp. 180-197.

Thupayagale, P., 2010. Evaluation of GARCH-based Models in Value-at-Risk Estimation: Evidence from Emerging Equity Markets. *Investment Analysts Journal,* 39(72), pp. 13-29.

van Rensburg, P., 2002. Market Segmentation on the Johannesburg Stock Exchange II. *Journal of Studies in Economics and Econometrics,* 26(1), pp. 83-100.

Appendices

Appendix 1: Kupiec's Proportion of Failure Test at 5% Significance Level (Pre-Crisis Period)

Table 20: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (ALSI, Pre-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*40*	$40 \in [38,66]$. *Cannot reject* H_0
EGARCH(1,1)	*37*	$37 \notin [38,66]$. *Reject* H_0
RiskMetrics	*50*	$50 \in [38,66]$. *Cannot reject* H_0
Historical Simulation	*34*	$34 \notin [38,66]$. *Reject* H_0
Delta-Normal	*34*	$34 \notin [38,66]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [38,66] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 21: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (FINDI, PRE-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*43*	$43 \in [38,66]$. *Cannot reject* H_0
EGARCH(1,1)	*41*	$41 \in [38,66]$. *Cannot reject* H_0
RiskMetrics	*52*	$52 \in [38,66]$. *Cannot reject* H_0
Historical Simulation	*30*	$30 \notin [38,66]$. *Reject* H_0
Delta-Normal	*42*	$42 \in [38,66]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [38,66] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 22: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (RESI, Pre-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*49*	$49 \in [38,66]$. *Cannot reject* H_0
EGARCH(1,1)	*47*	$47 \in [38,66]$. *Cannot reject* H_0
RiskMetrics	*53*	$53 \in [38,66]$. *Cannot reject* H_0
Historical Simulation	*49*	$49 \in [38,66]$. *Cannot reject* H_0
Delta-Normal	*44*	$44 \in [38,66]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the pre-crisis period. The number of observed violations must fall outside the confidence interval [38,66] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Appendix 2: Kupiec's Proportion of Failure Test at 5% Significance Level (Crisis Period)

Table 23: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (ALSI, Crisis Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	50	$50 \notin [24,47]$. *Reject* H_0
EGARCH(1,1)	59	$59 \notin [24,47]$. *Reject* H_0
RiskMetrics	44	$44 \in [24,47]$. *Cannot reject* H_0
Historical Simulation	88	$88 \notin [24,47]$. *Reject* H_0
Delta-Normal	87	$87 \notin [24,47]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the crisis period. The number of observed violations must fall outside the confidence interval [24,47] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 24: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (FINDI, Crisis Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	45	$45 \in [24,47]$. *Cannot reject* H_0
EGARCH(1,1)	34	$34 \in [24,47]$. *Cannot reject* H_0
RiskMetrics	38	$38 \in [24,47]$. *Cannot reject* H_0
Historical Simulation	78	$78 \notin [24,47]$. *Reject* H_0
Delta-Normal	83	$83 \notin [24,47]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the crisis period. The number of observed violations must fall outside the confidence interval [24,47] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 25: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (RESI, Crisis Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	52	$52 \notin [24,47]$. *Reject* H_0
EGARCH(1,1)	55	$55 \notin [24,47]$. *Reject* H_0
RiskMetrics	47	$47 \in [24,47]$. *Cannot reject* H_0
Historical Simulation	79	$79 \notin [24,47]$. *Reject* H_0
Delta-Normal	76	$76 \notin [24,47]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the crisis period. The number of observed violations must fall outside the confidence interval [24,47] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Appendix 3: Kupiec's Proportion of Failure Test at 5% Significance Level (Post-Crisis Period)

Table 26: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (ALSI, Post-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*70*	$70 \in [53,85]$. *Cannot reject* H_0
EGARCH(1,1)	*55*	$55 \in [53,85]$. *Cannot reject* H_0
RiskMetrics	*87*	$87 \notin [53,85]$. *Reject* H_0
Historical Simulation	*13*	$13 \notin [53,85]$. *Reject* H_0
Delta-Normal	*12*	$12 \notin [53,85]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the post-crisis period. The number of observed violations must fall outside the confidence interval [53,85] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 27: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (FINDI, Post-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*64*	$64 \in [53,85]$. *Cannot reject* H_0
EGARCH(1,1)	*60*	$60 \in [53,85]$. *Cannot reject* H_0
RiskMetrics	*80*	$80 \in [53,85]$. *Cannot reject* H_0
Historical Simulation	*30*	$30 \notin [53,85]$. *Reject* H_0
Delta-Normal	*32*	$32 \notin [53,85]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the post-crisis period. The number of observed violations must fall outside the confidence interval [53,85] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 28: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (RESI, Post-Crisis Period)

Model	*Observed Violations*	*Conclusion*
GARCH(1,1)	*71*	$71 \in [53,85]$. *Cannot reject* H_0
EGARCH(1,1)	*66*	$66 \in [53,85]$. *Cannot reject* H_0
RiskMetrics	*90*	$90 \notin [53,85]$. *Reject* H_0
Historical Simulation	*24*	$24 \notin [53,85]$. *Reject* H_0
Delta-Normal	*19*	$19 \notin [53,85]$. *Reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the post-crisis period. The number of observed violations must fall outside the confidence interval [53,85] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Appendix 4: Kupiec's Proportion of Failure Test at 5% Significance Level (Full Period)

Table 29: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (ALSI, Full Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	*160*	$160 \in [131{,}179]$. *Cannot reject* H_0
EGARCH(1,1)	*151*	$151 \in [131{,}179]$. *Cannot reject* H_0
RiskMetrics	*181*	$181 \notin [131{,}179]$. *Reject* H_0
Historical Simulation	*135*	$135 \in [131{,}179]$. *Cannot reject* H_0
Delta-Normal	*133*	$133 \in [131{,}179]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's All Share Index (ALSI) for the full period. The number of observed violations must fall outside the confidence interval [131,179] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 30: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (FINDI, Full Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	*152*	$152 \in [131{,}179]$. *Cannot reject* H_0
EGARCH(1,1)	*135*	$135 \in [131{,}179]$. *Cannot reject* H_0
RiskMetrics	*170*	$170 \in [131{,}179]$. *Cannot reject* H_0
Historical Simulation	*138*	$138 \in [131{,}179]$. *Cannot reject* H_0
Delta-Normal	*157*	$157 \in [131{,}179]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Financials-Industrials Index (FINDI) for the full period. The number of observed violations must fall outside the confidence interval [131,179] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.

Table 31: Results: Kupiec's Proportion of Failure Test at the 5% Significance Level (RESI, Full Period)

Model	Observed Violations	Conclusion
GARCH(1,1)	*172*	$172 \in [131{,}179]$. *Cannot reject* H_0
EGARCH(1,1)	*168*	$168 \in [131{,}179]$. *Cannot reject* H_0
RiskMetrics	*190*	$190 \notin [131{,}179]$. *Reject* H_0
Historical Simulation	*152*	$152 \in [131{,}179]$. *Cannot reject* H_0
Delta-Normal	*139*	$139 \in [131{,}179]$. *Cannot reject* H_0

Note: This table reports the number of observed violations in each of the value-at-risk (VaR) forecasting models tested at the 5% significance level using the Johannesburg Stock Exchange's Resources Index (RESI) for the full period. The number of observed violations must fall outside the confidence interval [131,179] to reject the null hypothesis of Kupiec's proportion of failure test. The null hypothesis states that the number of true violations of the model is equal to the observed number of violations.