“An evaluation and comparison of Value at Risk and Expected Shortfall”

AUTHORS	Tom Burdorf Gary van Vuuren https://orcid.org/0000-0001-6811-0538

ARTICLE INFO	Tom Burdorf and Gary van Vuuren (2018). An evaluation and comparison of Value at Risk and Expected Shortfall. *Investment Management and Financial Innovations*, *15*(4), 17-34. doi:10.21511/imfi.15(4).2018.02
DOI	http://dx.doi.org/10.21511/imfi.15(4).2018.02
RELEASED ON	Friday, 12 October 2018
RECEIVED ON	Thursday, 02 August 2018
ACCEPTED ON	Friday, 21 September 2018

JOURNAL	"Investment Management and Financial Innovations"
ISSN PRINT	1810-4967
ISSN ONLINE	1812-9358
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

BUSINESS PERSPECTIVES

LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10, Sumy,
40022, Ukraine

www.businessperspectives.org

Tom Burdorf (South Africa), **Gary van Vuuren** (South Africa)

AN EVALUATION AND COMPARISON OF VALUE AT RISK AND EXPECTED SHORTFALL

Abstract

As a risk measure, Value at Risk (VaR) is neither sub-additive nor coherent. These drawbacks have coerced regulatory authorities to introduce and mandate Expected Shortfall (ES) as a mainstream regulatory risk management metric. VaR is, however, still needed to estimate the tail conditional expectation (the ES): the average of losses that are greater than the VaR at a significance level. These two risk measures behave quite differently during growth and recession periods in developed and emerging economies. Using equity portfolios assembled from securities of the banking and retail sectors in the UK and South Africa, historical, variance-covariance and Monte Carlo approaches are used to determine VaR (and hence ES). The results are back-tested and compared, and normality assumptions are tested. Key findings are that the results of the variance covariance and the Monte Carlo approach are more consistent in all environments in comparison to the historical outcomes regardless of the equity portfolio regarded. The industries and periods analyzed influenced the accuracy of the risk measures; the different economies did not.

Received on: 2nd of August, 2018
Accepted on: 21st of September, 2018

Keywords market risk, Value at Risk, Expected Shortfall, South Africa, UK, normality, back-testing, historical simulation, variance covariance, Monte Carlo

JEL Classification C6, G2, G3

Tom Burdorf, MSc., School of Economics, University of Cape Town, South Africa.

Gary van Vuuren, Extraordinary Professor, Business School, University of the Free State, South Africa.

INTRODUCTION

Financial institutions are continuously exposed to credit, market, operational, liquidity, reputational and other risks. Although hedging helps to minimize and mitigate some of these risks, the first step towards managing risks is measuring them (BCBS, 1994). The focus of this article is idiosyncratic market risk (i.e. market risk that can be diversified away) along with the metrics which claim to measure it, but the accurate assessment of market risk is non-trivial (Riskmetrics, 1996). Portfolio return volatility weighs upside and downside risks equally, so risk measures such as Value at Risk (VaR) and Expected Shortfall (ES) were introduced to emphasize only downside risk: both measures now constitute part of the market risk regulatory framework of the Basel Committee of Banking Supervision (BCBS, 2016). All qualifying financial institutions must comply with the BCBS rules and must retain sufficient capital reserves to protect them from adverse scenarios with a given level of confidence over a specified period (BCBS, 2016). It is important to correctly estimate these reserves, because regulatory capital does not generate returns, capital retention is costly for institutions (Riskmetrics, 1996).

Calculating market risk has become a pursuit of burgeoning complexity due largely to the increasing degree of global investment and the growing number of interacting securities, which constitute trading book portfolios (Riskmetrics, 1996). The now familiar VaR measure first introduced by JP Morgan 1994 has enjoyed the status of principal market risk metric for several decades (Riskmetrics, 1996). Adopted by the BCBS in 1994, VaR in its various manifestations has been used to estimate the minimum capital required for financial institutions' market risk exposures (BCBS, 1994). VaR measures the maximum potential change in the value of a portfolio of financial instruments with a given probability over a pre-set horizon (Riskmetrics, 1996). Thus, if the random variable X describes potential portfolio profits and losses, with related quantile x_A, A represents a percentage of considered worst cases, i.e. $\alpha = A\% \in (0,1)$. VaR is then usually expressed as the supremum of the worst cases percentage α :

$$VaR_a = -x_a(X) = -\sup\{x | P[X \leq x] < \alpha\}.$$

VaR is measured using one of three approaches (historical, variance-covariance and Monte Carlo simulation), but it does not provide an estimate of the loss severity, should a suitably large loss occur (as determined by the confidence level) – it only provides a measure of the loss frequency (Acerbi & Tasche, 2001). ES estimates the loss severity: it is the probability-weighted average of the losses greater than VaR. It is a superior risk measure, because it is sub-additive and coherent unlike VaR (Acerbi & Tasche, 2001). If the random variable of profits/losses X is continuous, the ES is called tail conditional expectation (TCE): $TCE_a = -E\{X | X \leq x_a\}$. The TCE is the average of losses that are greater than the boundary VaR value at a significance level (Acerbi & Tasche, 2001), α and $TCE_a \geq VaR_a$. For a more general distribution, when the random variable of profits – losses X is discontinuous, the TCE is not sub-additive either and, in this case, the coherent risk measure ES is:

$$ES_a = \frac{1}{a}\int_0^a F^{-1}(p)\,dp,$$

where with $F^{-1}(p)$ being the generalized inverse function of $F(x)$: $F^{-1}(p) = \inf(x | F(x) \geq p)$. Here, $F(x)$ is the distribution function: $F(x) = P[X \leq x]$.

The Expected Shortfall can also be expressed in terms of the TCE and VaR:

$$ES_a = TCE_\alpha + (\lambda - 1)(TCE_\alpha - VaR_a),$$

where $\lambda \equiv \dfrac{P[X \leq x_a]}{\alpha} \geq 1$.

Thus, the following relationship holds $ES_a \geq TCE_a \geq VaR_a$ (Acerbi & Tasche, 2001).

Assuming a normal distribution for the ES, $ES_\alpha = \dfrac{f(VaR_\alpha)}{q}$,

where $f(x) = \dfrac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(-\dfrac{x^2}{2\sigma^2}\right)$

is the probability density function of $N(0, \sigma^2)$. Inserting the probability density function $f(x)$ into the integral leads to (1):

$$ES_\alpha = \int_{-\infty}^{\alpha} x \cdot f(x)\,dx = -\frac{\sigma}{\sqrt{2\pi}} \cdot \exp\left(-\frac{q^2}{2\sigma^2}\right). \quad (1)$$

All three approaches for calculating VaR and each approach's associated ES are calculated and compared for different markets (developing and developed), different time periods (highly volatile and calm) and different market sectors (retail and banks). To ascertain which measure more accurately describes the tail risks, results are back-tested.

http://dx.doi.org/10.21511/imfi.15(4).2018.02

The results of the above are used to answer the questions:

1. Is ES superior to VaR under all market conditions, for different market sectors and for markets operating in different economies?

2. Which method is more consistent?

3. Does the normality assumption of returns hold?

Questions 1 and 2 are answered using the back-testing outcomes prescribed by the BCBS (2016) and ES/VaR ratios and question 3 is dealt with by applying a Jarque-Bera test (Thadewald & Büning, 2004) coupled with a comparison of calculated and actual return volatilities.

The remainder of this paper proceeds as follows: section 1 provides theoretical background, section 2 gives a literature overview, while section 3 outlines the underlying methodology, section 4 presents and discusses the result, and last section concludes.

1. THEORETICAL BACKGROUND

1.1. Approaches for VaR and ES

The Historical Simulation (HS) approach makes no assumption about the return distribution and asserts that tomorrow's returns will behave as they did in the past (a contradiction of the efficient market hypothesis). Unless the historical period selected for simulation covers a turbulent era, losses in the tail region will only comprise a few observations. The VaR measure can thus be volatile (Sharma, 2012).

The Variance-Covariance (VCV) method assumes that the portfolio returns are normally distributed and takes correlations between constituent assets into account (Benninga & Wiener, 1998) and the Monte Carlo (MC) simulation generates correlated, usually but not necessarily normally distributed returns based on historical values to model possible scenarios (Linsmeier & Pearson, 1996). Other different sampling methods exist to calculate VaR (e.g., bootstrap, empirical, maximum likelihood estimation) (Li, Fan, Li, Zhou, Jin, & Liu, 2012).

The ES explores the tail region: all outcomes worse than VaR are averaged. For the MC simulation and HS approaches, outcomes worse than VaR are simply averaged to obtain the ES, while for the VCV method, an analytical solution is used (see Artzner, Delbaen, Eber, & Heath, 1999).

1.2. Coherent risk measures

There are four axioms a coherent risk measure must fulfil. A function $p:\ Z \to \mathbb{R}$, where Z is a set of real-valued random variables, is considered a coherent risk measure if it satisfies the following axioms (Artzner et al., 1999):

a) monotonic:

$$A, B \in Z, R(B) \geq R(A) \Rightarrow p(B) \leq p(A),$$

where Z includes the portfolios A and B. If the returns of portfolio B are always greater or equal to A, then the risk of portfolio B is always smaller or equal to the risk of portfolio A.

b) sub-additive:

$$A, B, A+B \in Z \Rightarrow p(A+B) \leq p(A)+p(B).$$

Combining two portfolios A and B in a third portfolio will result in the new portfolio having less or the same risk as the sub-portfolios added together. This phenomenon is the most important one and known as diversification. VaR violates this axiom. It is possible that VaR for the sum of sub-portfolios is smaller than VaR of the overall portfolio, so losses are underestimated. For example, a regulator should be confident to assume that bank's capital requirements, which depend on the market risk of the institution, consisting of several branches, are not greater than the sum of the individual branch capital requirements. Nonetheless, if a risk measure does not satisfy the axiom of sub-

Source: Authors' calculations.

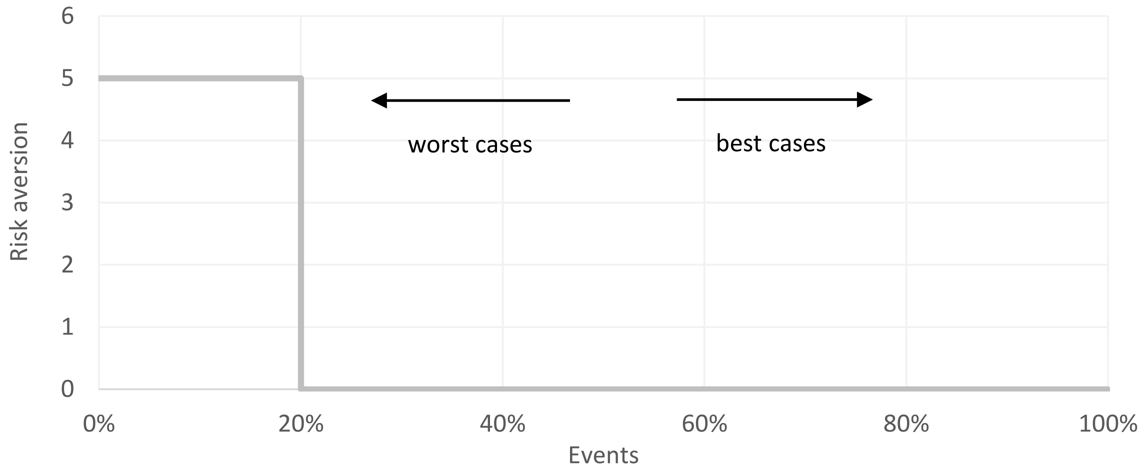

Figure 1. Example of a risk aversion function $\varphi(x)$ for ES

additivity, the total risk of the bank can be much higher (Artzner et al., 1999).

c) positively homogenous:

$$A \in Z,\ g > 0, gA \in Z \Rightarrow p(gA) = gp(A).$$

Positively homogenous states that loss and risk measure both scale the same. For example, if a portfolio position triples $(g = 3)$, then so does the risk.

d) translation invariant:

$$A \in Z,\ g \in \mathbb{R} \Rightarrow p(A + g) = p(A) + g.$$

When the loss distribution moves by a fixed amount g, the axiom implies that the risk measure changes by the same amount.

ES satisfies all four axioms and is, therefore, a coherent risk measure. VaR satisfies all conditions except the axiom of sub-additivity; hence, VaR is not a coherent risk measure (Artzner et al., 1999).

1.3. Risk aversion

Other ways to evaluate coherent risk measures are risk aversion functions (Acerbi, 2001). A coherent risk measure must have a decreasing and positive risk aversion function $\varphi(x)$. In addition,

$$\int_0^1 \varphi(x)\,dx = 1,$$

where $\varphi(x)$ represents the rational attitude towards risk and may be thought of as a function weighing all cases from worst to best.

The risk aversion function for the VaR measure is a spike function, assuming normality (Acerbi, 2001). Thus, it is not an overall declining function (further confirmation that VaR is not a coherent risk measure) (Acerbi, 2001). Alternatively, an example of a risk aversion function for ES that fulfils coherency is shown in Figure 1. For the 20% worst cases, a risk aversion of 5 is assigned. A rational investor may assign their own subjective risk aversion by simply changing the profile of the weight function $\varphi(x)$ (Acerbi, 2001). The only requirements for coherency are that the function is positive, decreasing and normalized to 1 in the interval [0,1]. Within this framework, however, any option for $\varphi(x)$ is a legitimate attitude toward risk (Acerbi, 2001).

2. LITERATURE STUDY

After the definitive introduction to VaR by JP Morgan (1994) in which the development of the metric was introduced and analyzed, Linsmeier and Pearson (1996) explored VaR alternatives such as sensitivity analyses or cash flow at risk. A sensitivity analysis was used to measure the impact of different factors on a dependent variable within a set of assumptions, while cash flow at risk is like VaR, but related to cash flows (Linsmeier & Pearson, 1996). Linsmeier and Pearson's (1996) most significant findings were that the HS and MC simula-

http://dx.doi.org/10.21511/imfi.15(4).2018.02

tion are both able to capture portfolio risks, which included derivatives, while the VCV method did not always provide satisfying results. They further found that any of the three methods can produce misleading results if the recent past was atypical.

VaR was quickly found to not be a coherent risk measure (Artzner et al., 1999). This led Artzner et al. (1999) to introduce ES, then called tail conditional expectation (see equation (1)). However, the definition of ES from Artzner et al. (1999) was insufficient for discontinuous functions, as it did not satisfy the sub-additivity axiom. A more general version of the ES and proof of sub-additivity was then found by Acerbi and Tasche (2001).

Yamai and Yoshiba's (2002) early comparison of VaR and ES under general market conditions analyzed daily logarithmic changes of exchange rates. Their historical data included three established economic markets and 18 emerging markets and they found that VaR and ES do not estimate tail risk accurately in all cases. VaR and ES were both found to underestimate currency risk with fat-tails and high potential for losses. In a further analysis of Yamai and Yoshiba (2002), only Southeast Asian countries (emerging economies except for the Singapore dollar) were examined in which tail dependence was disregarded by both VaR and ES. Their conclusion was that neither risk measure on its own was sufficient: a combined approach of these two methods to analyze financial risk was more sophisticated than either of them alone. In addition, Yamai and Yoshiba (2002) asserted that the profit/loss distribution should be explored from different angles, including tail fatness and asymptotic dependence of the distribution.

Liang and Park (2007) explored risk measures for hedge funds using semi-deviation, VaR, ES, and tail risk, which all measure downside risk. Comparing performances of 1,500 hedge funds, Liang and Park (2007) concluded that skewness and kurtosis of the underlying distribution cannot be ignored. Furthermore, they confirm the findings that Expected Shortfall is superior to VaR for evaluating financial risk precisely when analyzing hedge funds' performances.

Acerbi, Nordio, and Sirtori (2001) compared VaR and ES in a unique non-empirical approach by reviewing classical arguments. A common misconception in the literature is the definition of VaR. VaR is often defined as the maximum potential loss that a portfolio can suffer in the 1% worst cases in a set time period (Riskmetrics, 1994), but Acerbi et al. (2001) assert that a better definition is "VaR is the minimum potential loss that a portfolio can suffer in the 1% worst cases in a set time period". Another definition, which amounts to the same thing, is: "VaR is the maximum potential loss that a portfolio can suffer in the 99% best cases in a set time period" (Acerbi et al., 2001).

Acerbi et al. (2001) also mathematically proved the non-subadditivity of VaR and derived an exact definition of ES, as well as proof of the coherence of this measure. They asserted that ES should replace VaR in risk management, keeping in mind the reliability of approximations and transparency of the underlying hypotheses. They concluded that ES was a solid measure to assess risk with no restrictions on applicability.

Nadarajah, Zhang, and Chan (2013) discuss different estimation methods for ES by providing an overview of the most common approaches to calculate ES. They list 45 different ways to estimate ES and distinguish between 32 underlying distributions. Generally, calculation methods can be categorized into parametric, nonparametric and semiparametric (Nadarajah et al., 2013). Parametric methods assume that the sample data are from a population following a probability distribution. Thus, the method is based on a fixed set of parameters with the most typical being the normal and Student's *t*-distribution. Nonparametric estimation methods do not make any assumptions about the distribution and are often more straightforward to use, even when the use of a parametric approach is applicable. Nonparametric methods have greater robustness than parametric ones and tend to leave less room for improper use and misunderstanding (Nadarajah et al., 2013). According to Nadarajah et al. (2013), the best-known nonparametric method is the HS, but other methods are the filtered HS and the Yamai and Yoshiba's (2002) estimator. Additionally, they recognize two different semiparametric estimations for ES: the heavy-tailed process, and the Necir et al. (2010) estimator.

Yamai and Yoshiba (2002), and Hürlimann (2004) point out two considerable shortcomings of ES: it is inconsistent with right tail risk and a similar level to VaR's accuracy can only be achieved with larger sample sizes.

Miletic, Korenak, and Lutovac (2014) apply VaR methods to the Belgrade Stock Exchange using four equally-weighted stocks from the 15 securities included in the Belgrade – 15 index. VaR is measured using the HS and two parametric methods: one assumes a normal distribution and the other uses the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model with a Student's *t*-distribution. Using six years of return data, they conclude that both methods predict market risk well. Even though these results are limited to the Serbian stock market, Miletic et al. (2014) have shown that VaR can be a good predictor of portfolio risk in a developing country during extraordinary financial events, e.g., the global credit crisis of 2008.

Wimmerstedt (2015) focused on back-testing ES results to explore its elicitability property. Elicitability is a mathematical concept, which implies that a law invariant risk measure uses a probability distribution to transform into a single-valued point forecast (Brehmer, 2017). Therefore, back-testing is the same as evaluating a forecast. A risk measure is elicitable if a scoring function can be found and minimized. For a scoring function $S(x,y)$, ES is defined by the forecast variable x, while y is the actual observation. Examples of scoring functions are the squared error function $S(x,y)=(x-y)^2$ or the absolute error function $S(x,y)=|x-y|$ (Wimmerstedt, 2015).

Wimmerstedt (2015) uses four different approaches (Fong & Wong, 2008; Righi & Ceretta, 2013; Emmer, Kratz, & Tasche, 2015; Acerbi & Szekely, 2014) to back-test ES. All found functioning back-test methods for ES without requiring elicitability. Two are parametric approaches and two non-parametric (Wimmerstedt, 2015). Fong and Wong (2008) use a saddle point technique to calculate the Probability Density Function (PDF) of the ES and conclude that credit risk of banks' loan portfolio simulated based on a unimodal assumption can in some cases be under-estimated. The Righi and Ceretta's (2013) method is based on a truncated distribution, the distribution of observations below VaR. ES is estimated via the distribution's mean. Righi and Ceretta (2013) further show that there is no need to wait for a whole back-test period using their back-test to spot the inefficiencies of ES outcomes. Emmer et al. (2013) used quantile approximation to back-test ES using approximations of various VaR levels and conclude that ES seems the best risk measure for use in practice, despite the lack of elicitability. Acerbi and Szekely's (2014) non-parametric model uses a defined significance value as a test statistic to evaluate ES results. With this, Acerbi and Szekely (2014) show that ES can in practice be jointly elicited with VaR and that elicitability is influenced by the method selected and not the model used to test results.

Gurrola-Perez and Murphy (2015) introduced a new method to calculate VaR – the filtered HS, in which, though similar to the HS, the return data are modified with more recent market conditions. For example, daily returns can be divided by daily volatilities and then multiplied by the volatility on the day the VaR is measured. Filtering mechanisms modify the properties of the return distribution (unconditional volatility, kurtosis, autocorrelation, and skewness) (Gurrola-Perez & Murphy, 2015). However, the filtering process must be carefully calibrated, as the underlying distribution is changed in non-trivial ways. When applied correctly, however, they found that VaR measured using filtered HS is superior to VaR calculated using HS, because filtered HS models react faster to changing market circumstances compared with HS (Gurrola-Perez & Murphy, 2015).

3. DATA AND METHODS

3.1. Data

The data comprise closing prices of the securities from the London Stock Exchange and the Johannesburg Stock Exchange. Retail and bank portfolios using data from South Africa and the United Kingdom were used. Each of the four portfolios consists of the three biggest retailers or banks by market capitalization in each country as of April 2018. The two industries are compared because banks were directly involved in the financial crisis and the UK banks were heavily exposed to the housing market in the United States (Dimsdale,

http://dx.doi.org/10.21511/imfi.15(4).2018.02

Source: Thomson Reuters.

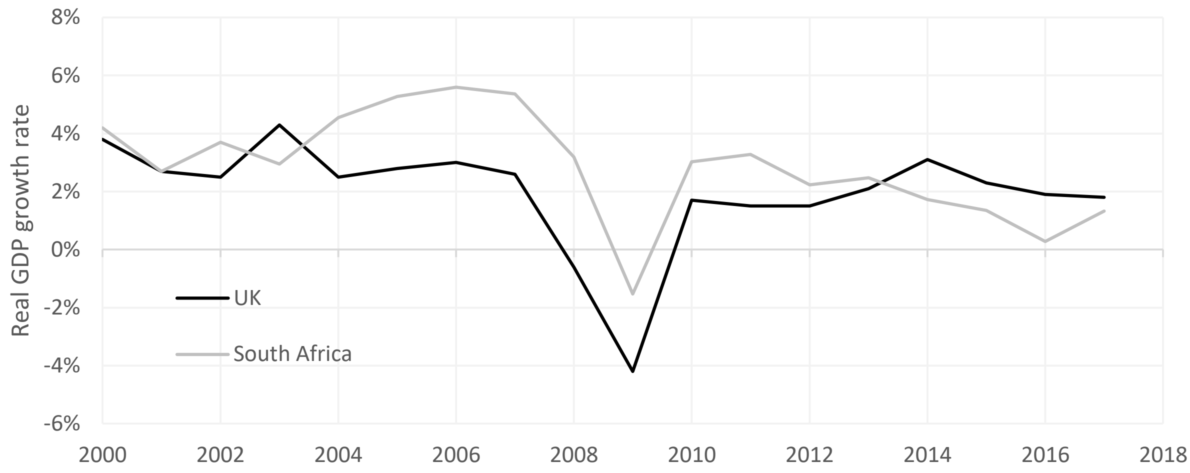

Figure 2. Comparison of South Africa's and the UK's real GDP growth rates from 2000 to 2017

2009). Barclays even bought some of the core assets of the collapsed investment bank Lehmann Brothers (Hughes & MacIntosh, 2008). Retailers are affected by consumer spending which also decreased during and after the crisis (Petev, LSQ-Crest, & Pistaferri, 2012), but since most retailers' products are essentials (groceries, for example), retail stocks should be affected less by the crisis.

All securities are traded in the country's top indexes the JSE Top 40 and the FTSE 100, respectively.

South Africa's retail portfolio comprised Pick 'n Pay, Shoprite, and Woolworths, while the UK's were Tesco, Morrison, and Sainsbury's. The South African bank portfolio comprised First National Bank, Standard Bank, and Nedbank and the UK Barclays, HSBC and Lloyds Bank.

Apart from the financial crisis whose influences could be felt globally, both economies performed differently from 2000 to 2018 as shown in Figures 2 to 4. South Africa's real GDP growth rate increased from 3% in 2003 to 5.3% in only three years, while unemployment decreased by 5% from 31% in 2003 to 26% in 2005. The UK's real GDP growth rate was 4.3% in 2003, dropped to 2.5% in 2009 and increased again to 2.8% in 2005, while the unemployment rate was roughly constant at 5%.

The years from 2003 to 2005 were real growth years for both economies, despite the considerable decrease in real GDP of both countries in 2009. The credit crisis impacted the UK to a greater degree than South Africa (as confirmed by unemployment values in Figure 3). The UK's unemployment rate rose from 5% to 8% over this period, while South

Source: Thomson Reuters.

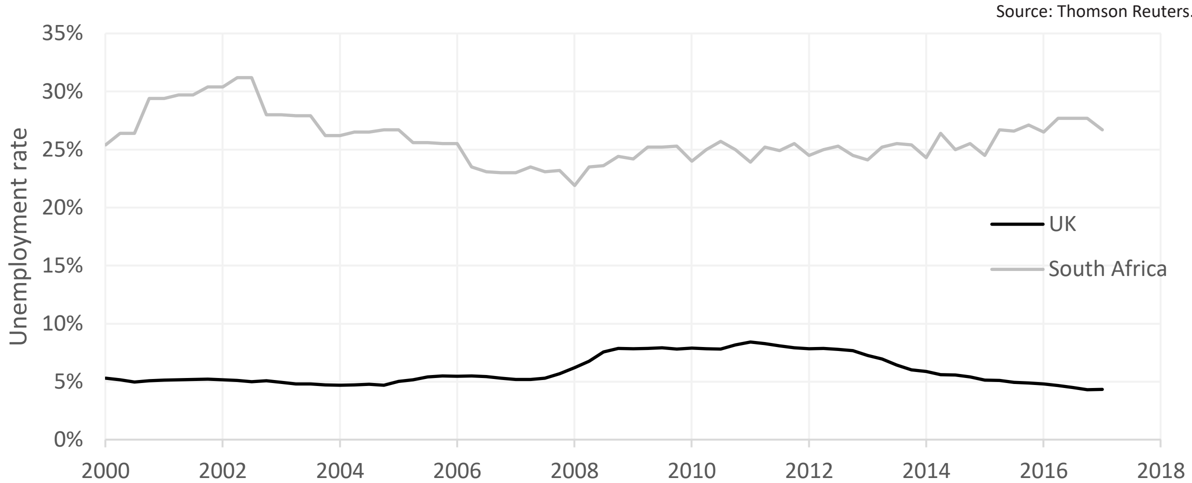

Figure 3. Comparison of South Arica's and the UK's unemployment rates from 2000 to 2017

Source: Thomson Reuters

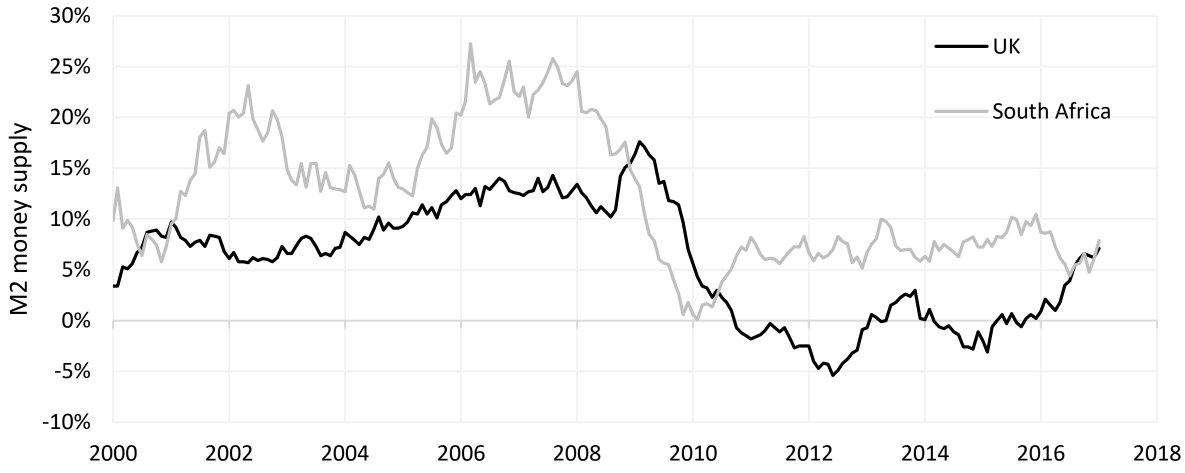

Figure 4. Comparison of South Arica's and the UK's M2 money supply rates from 2000 to 2017

Africa's increased from 23% to 25% (only a 2% absolute increase, and a relative change of 9%, due to the high levels of unemployment in South Africa).

Figure 4 shows the M2 money supply of both economies – a key economic indicator. From 2003 to 2005, the M2 money supply rate grew in both countries. In the financial crisis, the M2 money supply in South Africa was still growing, but the growth reached its lowest level in the aftermath of the credit crisis with a rate of 0.1%. In the UK, the M2 money supply declined during the crisis, falling steadily starting in 2008 and becoming negative at the end of 2010.

3.2. Methods

3.2.1. Jarque-Bera test for normality

Normality of log returns is a chief assumption of the VCV method. Many approaches to test for normality exist, including the Kolmogorov-Smirnov, Shapiro-Wilk, χ^2 and Jarque-Bera tests (Thadewald & Büning, 2004). In this work, the Jarque-Bera test was used. The hypotheses are:

H_0: *Data follow a normal distribution.*

H_1: *Data do not follow a normal distribution.*

The Jarque-Bera test essentially checks whether the sample data skewness and kurtosis match those of a normal distribution. The test statistic is:

$$JB = \frac{n-k+1}{6}\left(S^2 + \frac{1}{4}(C-3)^2\right), \qquad (2)$$

where n is the degrees of freedom, S is sample skewness and C is the sample kurtosis. Skewness is:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{\frac{3}{2}}}$$

and kurtosis:

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2\right)^{2}}.$$

Both values are inserted into the JB equation to get the test statistic. A symmetrical distribution such as the normal distribution has $S = 0$ and $C = 3$. For $C > 3$, fatter tails are indicated.

The JB statistic is asymptotically χ^2 distributed with two degrees of freedom: $JB \sim \chi_2^2$. Assuming a confidence level of 1%, $\chi_2^2 = 9.21$, so, if $JB > \chi_2^2$, the null hypothesis is rejected and vice versa.

http://dx.doi.org/10.21511/imfi.15(4).2018.02

3.2.2. Actual vs. estimated volatility

One-day and 10-day log returns were calculated and a two-sided statistical test was applied. Normality of return data is assumed, so the one-day volatility is scaled by $\sqrt{10}$ to obtain the 10-day volatility. The hypotheses are:

H_0: *The means are the same.*

H_1: *The means are different.*

The null hypothesis is rejected if the $|t| > 2.58$ with a 99% confidence level and not rejected if $|t| < 2.58$.

3.2.3. Back-testing

The BCBS (2016) approach to back-testing allows for a comparison between risk measures. The BCBS back-testing is a forecast and uses one full year of daily return data to estimate the VaR and ES (Wimmerstedt, 2015). Back-testing counts the number of losses that exceed the predicted risk measure. For VaR, these exceedances are (over period t): $e_t = 1$, when $L_t \geq VaR_\alpha(X)$, where L_t represents the loss at time t, $e_t = 1$ means exceedances occurred, while $e_t = 0$ implies no exceptions. Back-testing ES uses the same principles and equations (BCBS, 2016). Exceptions are Bernoulli distributed with probability α. In a $t = 250$ period, exceedances run from $e_1, e_2, ..., e_t$. W is the sum of all independent Bernoulli random variables at probability α (BCBS, 2016):

$$W = \sum_{t=1}^{T} e_t \sim Bin(T, \alpha),$$

for $\alpha = 1\%$, the expected number of exceedances in a 250-day trading year is 3.

The cumulative distribution function is:

$$P(X \leq k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i},$$

where k is the number of exceedances, $n = 250$, while $p = 0.01$.

Calculating the cumulative probability for specific numbers of exceptions is shown in Table 1.

Table 1. BCBS back-testing color table for a 99% confidence level

Source: BCBS (2016).

Zone	Exceedances	Multiplier k	Cumulative probability assuming q* = 0.99
Green	0	3.00	0.0811
	1	3.00	0.2858
	2	3.00	0.5432
	3	3.00	0.7581
	4	3.00	0.8922
Yellow	5	3.40	0.9588
	6	3.50	0.9863
	7	3.65	0.9960
	8	3.75	0.9989
	9	3.85	0.9997
Red	≥ 10	4.00	0.9999

If the number of exceedances > 4, there is a 95% chance that the model is inaccurate. The Basel Committee assigns the colors green, yellow, and red to the numbers of exceedances based on the probabilities. The multiplier column in Table 1 determines the necessary capital reserves of a bank. The market risk capital charge is calculated using multiplier k:

$$MRC_t = \max\left(VaR_{t-1},\ k \cdot \sum_{t=1}^{60} VaR_{t-60}\right), \quad (2)$$

where MRC_t measures the market risk capital at time t. Table 1 shows that multiplier k increases in the yellow zone: bank penalties are manifest in higher capital reserves when they are exposed to higher market risks.

4. RESULTS AND DISCUSSION

4.1. Portfolio performances

Figure 5 shows the development of all four portfolios during the pre-crisis period from 2003 to 2005. In the first quarter of 2003, all four portfolios lost value. After that, the South African retail portfolio increased by more than 230% of its original value, while the South African bank portfolio did not increase in value until the last quarter of 2004, but ended at 172% of its starting value in December 2005.

Both UK portfolios performed worse than the South African ones. Nevertheless, the UK retailer

Source: Thomson Reuters.

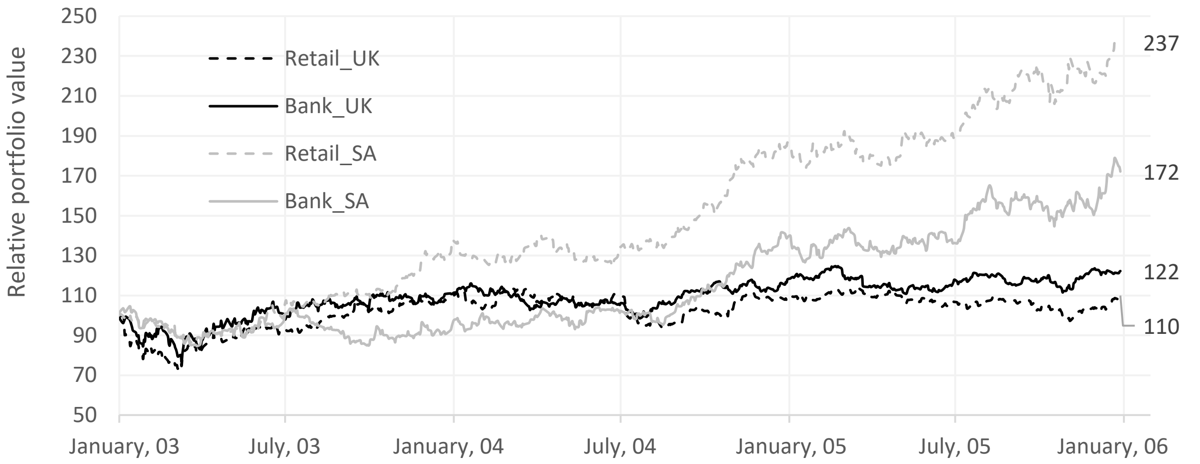

Figure 5. Relative portfolio development from 2003 to 2006. Portfolios rebased to 100 on January 1, 2003

portfolio increased by 22%, while the UK bank portfolio was worth 110% of its original value at the end of 2005. The returns are nominal and hence inflation – considerably higher in South Africa – must be considered.

In the aftermath of the credit crisis, only the South African retail portfolio was worth more by the end of 2010 than its original value in 2008. The portfolio increased by 55%, as shown in Figure 6. All other portfolios lost value over the period. The South African bank portfolio was worth 83% of its original value three years later in December 2010.

The UK retail portfolio decreased by a similar amount and was worth 79% of its starting value at the end of 2010 and the UK bank portfolio more than halved as a result of the financial crisis. An investment in the UK bank portfolio in January 2008 would have only been worth 38% of its initial investment by the end of 2010.

4.2. Correlation

Table 2 shows the correlation between the returns of all securities in this report during the pre-crisis period and the crisis period. The correlation between returns from either a South African bank or retailer and a UK retailer or bank is close to zero in most cases during both periods. Generally, the correlation between South African stocks is higher than the correlation between UK companies (see top left box). While UK retailers' stocks performance remain

Source: Thomson Reuters.

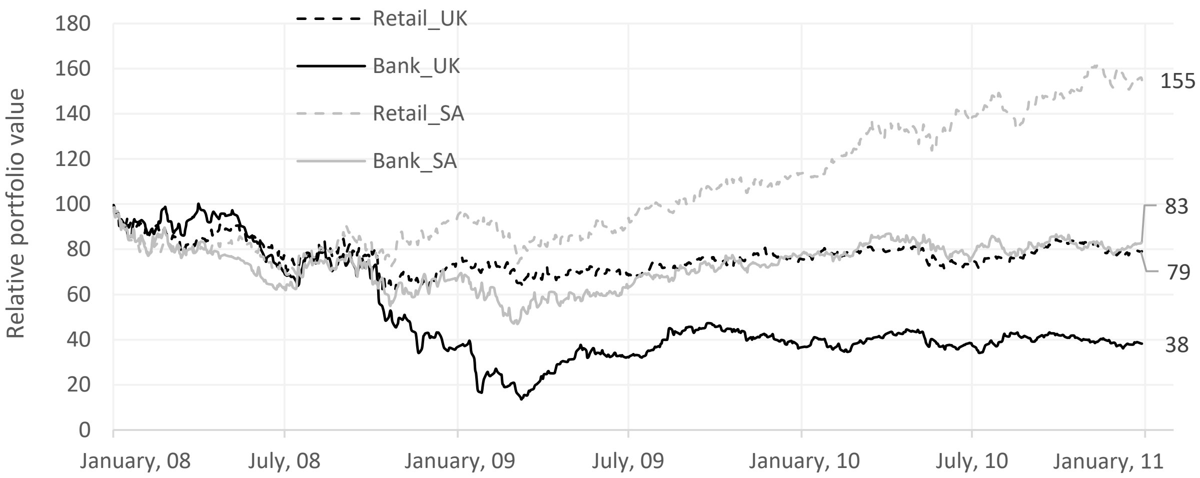

Figure 6. Relative portfolio development from 2008 to 2011. Portfolios rebased to 100 on January 1, 2008

http://dx.doi.org/10.21511/imfi.15(4).2018.02

Table 2. Correlation between all securities (a) from 2003 to 2005 and (b) from 2008 to 2010 using daily returns. Demarcated areas (boxed) indicate SA securities, remainder UK

Source: Authors' calculations.

(a)	PIK	WHL	SHP	FSR	SBAEI	NBKP	SBRY	TSCO	MRW	BARC	HSBA	LLOY
PIK	1	–	–	–	–	–	–	–	–	–	–	–
WHL	0.24	1	–	–	–	–	–	–	–	–	–	–
SHP	0.24	0.18	1	–	–	–	–	–	–	–	–	–
FSR	0.29	0.25	0.14	1	–	–	–	–	–	–	–	–
SBAEI	0.29	0.23	0.15	0.67	1	–	–	–	–	–	–	–
NBKP	0.24	0.23	0.16	0.32	0.30	1	–	–	–	–	–	–
SBRY	0.03	–0.01	–0.01	0.00	0.00	–0.02	1	–	–	–	–	–
TSCO	–0.01	0.02	0.00	0.06	0.04	0.01	0.42	1	–	–	–	–
MRW	0.14	0.00	0.00	0.03	0.00	–0.01	0.40	0.32	1	–	–	–
BARC	–0.06	–0.01	–0.01	–0.01	0.02	0.00	0.36	0.31	0.26	1	–	–
HSBA	–0.03	0.00	–0.03	–0.01	0.01	0.00	0.34	0.29	0.24	0.49	1	–
LLOY	–0.04	–0.04	0.03	0.02	0.05	0.00	0.39	0.35	0.29	0.69	0.50	1
(b)	**PIK**	**WHL**	**SHP**	**FSR**	**SBAEI**	**NBKP**	**SBRY**	**TSCO**	**MRW**	**BARC**	**HSBA**	**LLOY**
PIK	1	–	–	–	–	–	–	–	–	–	–	–
WHL	0.39	1	–	–	–	–	–	–	–	–	–	–
SHP	0.42	0.35	1	–	–	–	–	–	–	–	–	–
FSR	0.39	0.42	0.35	1	–	–	–	–	–	–	–	–
SBAEI	0.42	0.40	0.31	0.70	1	–	–	–	–	–	–	–
NBKP	0.40	0.42	0.39	0.68	0.68	1	–	–	–	–	–	–
SBRY	0.04	0.09	0.08	0.00	0.01	0.06	1	–	–	–	–	–
TSCO	0.04	0.07	0.07	0.07	0.03	0.07	0.59	1	–	–	–	–
MRW	0.03	0.10	0.06	0.01	–0.01	0.01	0.63	0.66	1	–	–	–
BARC	0.02	0.00	0.02	–0.02	–0.10	–0.05	0.32	0.26	0.22	1	–	–
HSBA	0.05	0.02	–0.01	0.08	0.09	–0.05	0.00	0.10	0.06	0.09	1	–
LLOY	–0.05	0.03	0.02	0.02	0.01	0.04	0.04	–0.02	–0.03	0.07	0.02	1

highly correlated with each other, this is not true for UK banks. Perhaps surprisingly, UK banks stocks have low correlations with each other during the credit crisis.

4.3. Normality

The Jarque-Bera test results are shown in Table 3. The test was concluded at a 99% confidence level with two degrees of freedom. Using (2), the JB statistic must be absolute smaller than 9.21 to not reject the null hypothesis, which states normality of returns. The results show that only the returns of FirstRand Bank during the pre-crisis period were normally distributed. The assumption of returns following a normal distribution – a requirement of the variance-covariance approach – is violated in all other cases. Comparison of the JB statistics from both periods indicates that stock returns were 'more normal' in the pre-crisis period than during the crisis, since the JB statistics are smaller for most securities during the growth years.

Analyzing the historical stock price data from Barclays, they show that the bank experienced several days with share price losses greater than 15% and days with gains of more than 20% during the crisis. These unusual events explain the very large number for the JB statistic. In general, South African returns tend to be more normal than UK returns.

4.4. Estimated and actual volatility

The two-sided test findings of calculated and actual returns over a 10-day period can be seen in Table 4. The test was done at a 99% confidence level. The null hypothesis stating that the means are the same can be accepted for a $|z| > 2.58.$ All portfolios during the credit crisis and pre-crisis period fulfil this. If the calculated and actual 10-day means are equal, the normality assumption of portfolio returns also holds. Even though only the share price returns of FirstRand Bank followed a normal distribution, all multi-asset portfolios follow a normal distribution. The least differences of

Table 3. JB test results during (a) 2003–2005 and (b) 2008–2010, 99% confidence level, daily returns. Demarcated areas are for SA, remainder UK

Source: Authors' calculations.

(a)	PIK	WHL	SHP	FSR	SBAEI	NBKP	SBRY	TSCO	MRW	BARC	HSBA	LLOY
Skewness	–0.14	0.23	–0.08	0.02	0.13	–0.07	–0.28	0.01	–0.23	0.69	0.16	0.17
Excess k	1.06	1.61	1.73	0.35	0.82	1.86	2.49	2.56	18.24	4.00	0.99	4.71
Obs	751	751	751	751	751	751	758	758	758	758	758	758
JB stat	40	138	162	1	19	200	498	529	191,787	2,085	34	3,303
Normal	NO	NO	NO	YES	NO	NO	NO	NO	NO	NO	NO	NO
(b)	**PIK**	**WHL**	**SHP**	**FSR**	**SBAEI**	**NBKP**	**SBRY**	**TSCO**	**MRW**	**BARC**	**HSBA**	**LLOY**
Skewness	0.39	–0.07	0.17	–0.11	0.06	0.10	–0.41	0.27	–0.03	1.34	–0.28	–0.75
Excess k	2.87	1.26	1.17	1.63	2.39	2.48	6.12	4.08	2.84	20.49	8.79	15.20
Obs	750	750	750	750	750	750	758	758	758	758	758	758
JB stat	758	63	54	136	425	476	7,245	2,159	721	272,089	21,429	110,886
Normal	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

means are observed for the South African retailer portfolio during the financial crisis and the UK bank portfolio during the pre-crisis period.

4.5. Risk measure comparison

4.5.1. General trends

Tables 5 through 7 show that the VaR and ES are generally higher in crisis periods: an expected result. While the economy was growing in the pre-crisis period, VaR and ES were smaller for the UK bank portfolio than for the South African one. However, the VaR and ES for the UK retail portfolio were slightly larger than the South African retail portfolio during the pre-crisis period, regardless of the approach used. During the crisis, VaR and ES were higher for the UK companies than for the South African ones.

4.5.2. VaR results

VaR is similar during both economic periods for both industries in South Africa using the HS and VCV method. In the UK, this differs. UK banks' VaR during 2008 and 2009 is significantly higher using the HS compared to the VCV method (Tables 5 and 6). Otherwise, UK VaR results using HS and VCV are similar. MC simulation in Table 7 shows similar outcomes as the VCV method for all VaR measures in both industries, periods and countries. The VaR HS is thus also higher in 2008 and 2009 than the VaR produced by the MC simulation for the UK banking portfolio. Overall, the VCV and MC outcomes are similar, while they both differ using the HS.

4.5.3. ES results

The VCV and MC VaR results for ES are similar, while HS generates significantly higher ES measures during the credit crisis for the UK and South African bank portfolios. For example, the one-day ES for the UK bank portfolio in 2009 is 16.8%. That means it is estimated that the portfolio loses almost 17% of its value in a single day in the 1% worst events when the HS is used. The MC simu-

Table 4. Results of a two-sided means test at a 99% confidence level (daily returns)

Source: Authors' calculations.

	2003–2005				2008–2010			
	SA		UK		SA		UK	
	Bank	Retail	Bank	Retail	Bank	Retail	Bank	Retail
μ_1	0.80%	1.21%	0.35%	0.22%	–0.01%	0.70%	–0.86%	–0.16%
μ_2	0.71%	1.08%	0.33%	0.25%	0.04%	0.68%	–0.73%	–0.08%
σ_1	3.99%	3.60%	3.68%	3.71%	6.84%	4.92%	8.98%	5.33%
σ_2	3.90%	3.28%	3.31%	3.36%	5.45%	4.36%	11.84%	4.01%
n_1	751	751	758	758	750	750	758	758
n_2	742	742	749	749	741	741	749	749
z-score	0.46	0.75	0.10	–0.16	–0.16	0.08	–0.25	–0.33
Means	Same	Same	Same	Same	Same	Same	Same	Same

http://dx.doi.org/10.21511/imfi.15(4).2018.02

Table 5. Historical simulation results of 99% VaR and ES (daily returns)

Source: Authors' calculations.

Share sector	Performance measure	Pre-crisis			Credit crisis		
		2003	2004	2005	2008	2009	2010
South Africa							
Bank	VaR	2.77%	2.25%	2.63%	6.50%	5.30%	3.64%
	ES	3.34%	2.83%	2.77%	8.89%	5.86%	4.66%
Retail	VaR	2.66%	2.46%	2.35%	4.23%	3.26%	2.53%
	ES	3.58%	2.99%	2.53%	4.94%	4.38%	2.95%
UK							
Bank	VaR	3.65%	2.09%	1.68%	7.88%	15.39%	3.47%
	ES	4.06%	2.35%	2.64%	9.09%	16.84%	3.87%
Retail	VaR	3.50%	3.24%	1.77%	5.55%	3.34%	2.36%
	ES	5.66%	3.70%	2.41%	7.96%	3.86%	2.82%

Table 6. Variance-covariance method results of 99% VaR and ES (daily returns)

Source: Authors' calculations.

Share sector	Performance measure	Pre-crisis			Credit crisis		
		2003	2004	2005	2008	2009	2010
South Africa							
Bank	VaR	3.01%	2.50%	3.04%	6.32%	5.07%	3.29%
	ES	3.43%	3.05%	3.59%	7.10%	5.91%	3.81%
Retail	VaR	2.82%	2.27%	2.46%	4.60%	3.35%	2.38%
	ES	3.39%	2.74%	2.94%	5.26%	3.95%	2.88%
UK							
Bank	VaR	3.80%	2.01%	1.69%	6.29%	8.95%	3.71%
	ES	4.42%	2.34%	1.95%	6.81%	10.34%	4.27%
Retail	VaR	3.59%	2.37%	1.85%	5.53%	3.28%	2.31%
	ES	4.17%	2.71%	2.14%	6.25%	3.78%	2.66%

Table 7. Monte Carlo simulation results of 99% VaR and ES (daily returns)

Source: Authors' calculations.

Share sector	Performance measure	Pre-crisis			Credit crisis		
		2003	2004	2005	2008	2009	2010
South Africa							
Bank	VaR	2.87%	2.48%	2.86%	5.78%	4.84%	3.29%
	ES	3.37%	2.85%	3.41%	6.73%	5.41%	3.63%
Retail	VaR	2.72%	2.25%	2.49%	4.62%	3.53%	2.20%
	ES	3.07%	2.52%	2.78%	5.20%	4.22%	2.51%
UK							
Bank	VaR	3.67%	1.99%	1.63%	5.72%	8.49%	3.60%
	ES	4.15%	2.40%	1.85%	6.75%	9.86%	4.18%
Retail	VaR	3.64%	2.04%	1.56%	5.50%	2.92%	3.34%
	ES	4.07%	2.31%	1.83%	6.33%	3.43%	3.83%

lation estimates that only 9.86% of the portfolio's value is lost, while the VCV method produces an ES of 10.34%. Using the HS to estimate market risk in this case would require much higher capital reserves by banks than the other two approaches.

4.5.4. ES/VaR ratios

Examining the ratio of ES/VaR is instructive. This ratio differs across methods, countries, industries and time periods.

Source: Authors' calculations.

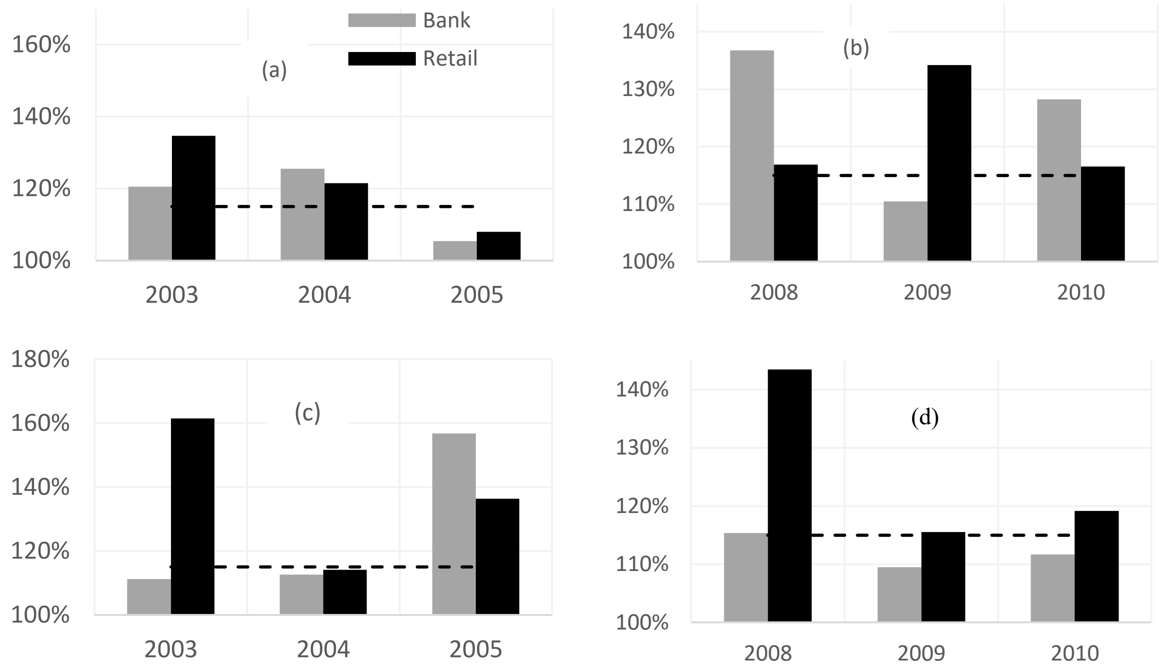

Figure 7. Historical simulation ES/VaR ratios
for (a) SA 03-05, (b) SA 08-10, (c) UK 03-05 and (d) UK 08-10

Source: Authors' calculations.

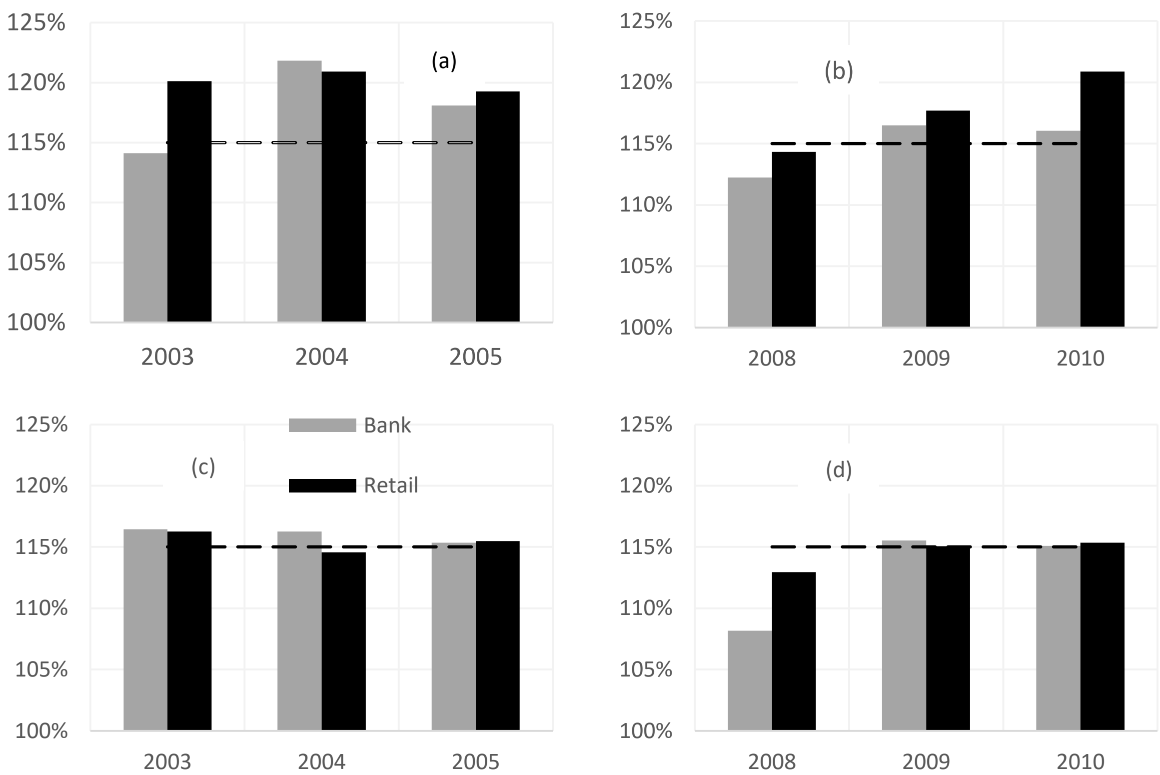

Figure 8. Variance-covariance ES/VaR ratios
for (a) SA 03-05, (b) SA 08-10, (c) UK 03-05 and (d) UK 08-10

http://dx.doi.org/10.21511/imfi.15(4).2018.02

Source: Authors' calculations.

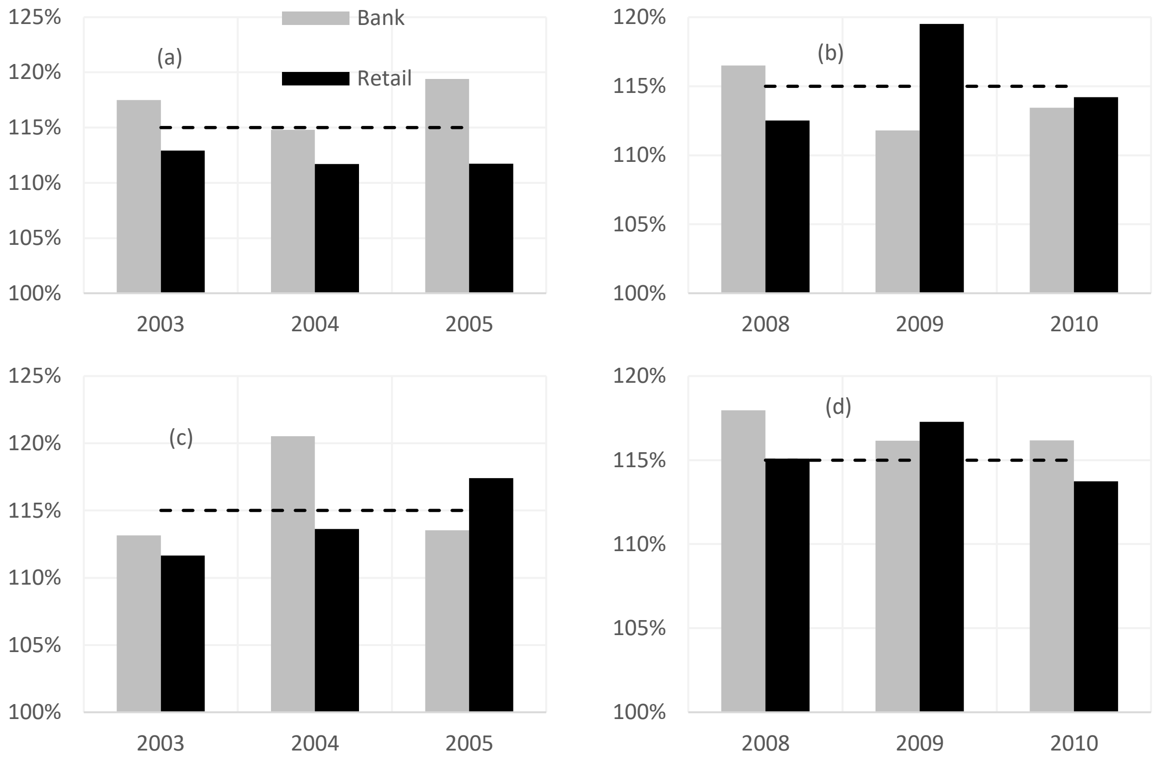

Figure 9. Monte Carlo simulation ES/VaR ratios for (a) SA 03-05, (b) SA 08-10, (c) UK 03-05 and (d) UK 08-10

The comparison for these can be seen in Figures 7-9. The normal ratio of ES/VaR (assuming a normal distribution of returns) is 1.15 (see equation (1)). This ratio is shown in each graph as a dashed black line for comparison. For the HS (Figure 7), the ratios are volatile regardless of the industry, country or period: the average ES/VaR ratio is 1.24. For example, the UK retail ES in 2003 is 1.62 times greater than VaR. SA banks' ES in 2008 is 1.37 times greater than VaR. In comparison, using the VCV approach (Figure 8), all ratios are closer to 1.15. Here, the highest ratio is 1.22 (Figure 8a; SA bank, 2004), and the lowest 1.08 (Figure 8d; UK bank, 2008) with an average ratio of 1.16. The least volatile ratios occur using the MC approach: the highest ratio being 1.21 (Figure 9b; SA bank, 2009) and the lowest 1.12 (Figure 9c; UK bank, 2004). This is further indicated by the average ratio of the MC outcomes, which equal the normal ratio of 1.15.

MC results show that the difference between ES and VaR is often smaller than 15%. This is the case when the bars do not reach the dashed normal ratio line, which occurs 50% of the time (12 out of 24 results) (Figure 9). This implies that VaR and ES values are similar for MC. It therefore matters less which risk measure (VaR or ES) is superior, because the differences between VaR and ES are relatively small using the MC approach. Basel III capital requirements for the MC method are similar, then, for financial institutions regardless of the risk measure used.

4.5.5. Back-testing results

VaR and ES for the HS and VCV methods were back-tested using the BCBS rules (testing MC results is not necessary, since returns are simulated). Back-testing results for HS and VCV methods are shown in Tables 8 and 9. If the risk measure is flagged yellow, the multiplier is greater than three. Thus, banks must hold higher capital reserves (see Table 1 and equation (2)). In addition, yellow implies that the model used may not be accurate, but to be certain, more information is needed. As expected, the risk measure VaR is flagged yellow for both bank portfolios during the credit crisis (Table 8). It is surprising that HS VaR also flags yellow during a pre-crisis period in 2005 for South African retailers and banks. The VCV VaR

Table 8. BCBS back-testing results for VaR at a 99% confidence level

Source: Authors' calculations.

South Africa		2003–2004	2004–2005	2008–2009	2009–2010
Bank	VCV	Green	Green	Green	Yellow
	HS	Green	Yellow	Green	Yellow
Retail	VCV	Green	Green	Green	Green
	HS	Green	Yellow	Green	Green
UK		**2003–2004**	**2004–2005**	**2008–2009**	**2009–2010**
Bank	VCV	Green	Green	Yellow	Green
	HS	Green	Green	Yellow	Green
Retail	VCV	Green	Green	Green	Green
	HS	Green	Green	Green	Green

Table 9. BCBS back-testing results for ES at a 99% confidence level

Source: Authors' calculations.

South Africa		2003–2004	2004–2005	2008–2009	2009–2010
Bank	VCV	Green	Green	Green	Green
	HS	Green	Green	Green	Green
Retail	VCV	Green	Green	Green	Green
	HS	Green	Green	Green	Green
UK		**2003–2004**	**2004–2005**	**2008–2009**	**2009–2010**
Bank	VCV	Green	Green	Yellow	Green
	HS	Green	Green	Yellow	Green
Retail	VCV	Green	Green	Green	Green
	HS	Green	Green	Green	Green

results in these periods, however, do not require higher capital reserves: the VaR model works accurately, since exceedances are < 5 (marked green in Table 8). Overall, VaR measures flag yellow in four cases and only twice for ES. The models are least accurate in 2009 when UK banks' share prices more than halved due to considerable US housing market exposure. Financial institutions that use ES and estimated using the VCV VaR method would have calculated far more accurate risk measures in comparison to those banks, which used HS VaR.

CONCLUSION AND LIMITATIONS

A developed and an emerging economy, different industries and two periods of contrasting economic growth were considered for evaluating VaR (in its three common manifestations) and ES risk measures in this paper.

ES and VaR were found to be considerably higher during recession periods. The MC simulation and VCV method are more accurate (in terms of estimating exceedances) than the HS approach, especially in times of recessions and the ratio of ES/VaR using the VCV method and the MC simulation is more consistent than the HS method.

The assumption of normality for single stock returns, a requirement of the VCV approach, has been demonstrated to be largely untrue, as it has often been shown in the literature (e.g., Richardson & Smith, 1993; Sheikh & Qiao, 2009). Nevertheless, for the portfolios of stocks used in this work, this is a reasonable approximation. No statistically significant differences are found using the different risk measures whether applied to a developed or emerging economy.

 http://dx.doi.org/10.21511/imfi.15(4).2018.02

Differences arise between industries. Banks were directly exposed to the crisis initiated by the collapse of the US housing market, while retailers were only affected indirectly due to diminished consumer spending. VaR and ES are generally higher for the banking industry in both countries during the crisis period. Risk models were also inaccurate for banks – in line with Linsmeier and Pearson (1996) who found that if the recent past is atypical, all three estimation models are flawed.

ES was shown to be a superior risk measure (Liang & Park, 2007). Nevertheless, this is only true to the extent of which VaR measure is used. Financial institutions should indicate both measures employed. Both risk measures are insufficient for evaluating all potential portfolio risks (as demonstrated by Yamai & Yoshiba, 2002). Extraordinary market events, as witnessed during the crisis, are exceedingly difficult to predict. This is emphasized by back-testing results, which resulted in exceedances > 4 (flagged yellow in Tables 8 and 9) using the VCV and HS methods for the UK banking sector from 2008 to 2010. UK banks experienced substantial losses during the crisis, and the risk measures were unable to adapt them sufficiently quickly. Overall, ES performs better than VaR, and both VCV and MC simulation approaches provide more consistent results.

The research conducted is limited to one developed and one emerging economy, which are compared with each other. Also, only specific industries are compared and therefore the results might differ for other industry sectors or economies. Lastly, the time periods regarded might not reflect other past or future time periods. However, because both recession and growth timeframes were studied it is not unlikely that future periods of these economic conditions will lead to similar outcomes regarding the risk measures.

Future research could involve additional testing of risk measures under extreme market conditions. Also, back-testing or validating of test results are only in the early stages of development and can be researched in more detail. Finally, much research focuses on comparing VaR and ES, to ascertain which risk measure is superior (in terms of back-test exceedance numbers), future focus could shift to assessing and comparing various methods to calculate both risk measures. For example, the BCBS back-test is not a good method to validate VaR and ES using the MC simulation approach. Since it does not make sense to back-test simulated returns because they have not actually occurred in the past. The development of a universal back-test could be the next step for research in this field. This would allow a reliable comparison of all methods: currently determining which of the three methods is best under all market conditions is complex and can be contradictory.

New methods to calculate more accurate (and easier to apply) VaR and ES measurements could be developed. If policymakers take the initiative and decide on one framework and one risk measure, the risk measures would become much more comparable, allowing differences and similarities between them to be studied more easily. The current most suitable method for this seems to be the MC method because it is more consistent than the HS and no false assumption is necessary, as for the VCV approach.

REFERENCES

1. Acerbi, & Tasche (2003). Expected shortfall: a natural coherent alternative to Value at Risk. *Economic Notes, 31*(2), 379-388. https://doi.org/10.1111/1468-0300.00091
2. Acerbi (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking and Finance, 26*(7), 1505-1518. https://doi.org/10.1016/S0378-4266(02)00281-9
3. Acerbi, C., Nordio, C., & Sirtori, C. (2001). *Expected shortfall as a tool for financial risk management* (Working paper). Retrieved from https://arxiv.org/pdf/cond-mat/0102304.pdf
4. Acerbi, C., & Szekely, B. (2014). Back-testing expected shortfall. *Risk Magazine, December 2014*. Retrieved from https://www.msci.com/documents/10199/22aa9922-f874-4060-b77a-0f0e267a489b
5. Artzner, P., Delbaen, F., Eber, J., & Heath, D. (1999). Coherent Measure of Risk. *Mathematical Finance, 9*(3), 203-228. https://doi.org/10.1111/1467-9965.00068
6. Basel Committee on Banking Supervision (BCBS) (2016). *Minimum capital requirements*

for market risk. Retrieved from https://www.bis.org/bcbs/publ/d352.pdf

7. Basel Committee on Banking Supervision (BCBS) (1994). *Risk management guidelines for derivatives*. Retrieved from https://www.bis.org/publ/bcbsc211.pdf

8. Benninga, S., & Wiener, Z. (1998). Value at Risk (VaR). *Mathematica in Education and Research, 7*(4), 1-8. Retrieved from http://simon-benninga.com/wiener/MiER74.pdf

9. Brehmer, J. (2017). *Elicitability and its Application in Risk Management* (Thesis: University of Mannheim). Retrieved from https://arxiv.org/pdf/1707.09604.pdf

10. Dimsdale, N. (2009). The Financial Crisis of 2007–9 and the British Experience. *Oxonomics, 4*(1), 1-9. Retrieved from https://doi.org/10.1111/j.1752-5209.2009.00026.x

11. Emmer, S., Kratz, M., & Tasche, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *Journal of Risk, 18*(2), 31-60. Retrieved from https://arxiv.org/abs/1312.1645

12. Hughes, C., & MacIntosh, J. (2008, September 17). Barclays strikes deal for Lehman spoils. *Financial Times*. Retrieved from https://www.ft.com/content/5c9dcc26-83f1-11dd-bf00-000077b07658,

13. Gurrola-Perez, P., & Murphy, D. (2015). Filtered historical simulation Value-at-Risk models and their competitors. *Bank of England Working Paper, 525,* 1-33. Retrieved from https://www.bankofengland.co.uk/working-paper/2015/filtered-historical-simulation-value-at-risk-models-and-their-competitors

14. Hürlimann, W. (2004). Distortion risk measure and economic capital. *North American Actuarial Journal, 8*(1), 86-95. https://doi.org/10.1080/10920277.2004.10596130

15. JP Morgan (1996). Riskmetrics Technical Document, JP Morgan/Reuters (Fourth Edition). Retrieved from https://www.msci.com/documents/10199/5915b101-4206-4ba0-aee2-3449d5c7e95a

16. Li, H., Fan, X., Li, Y., Zhou, Y., Jin, Z., & Liu, Z. (2012). *Approaches to VaR*. Project: Stanford University. Retrieved from https://web.stanford.edu/class/msande444/2012/MS&E444_2012_Group2a.pdf

17. Liang, B., & Park, H. (2007). Risk measures for hedge funds: a cross-sectional approach. *European Financial Management, 13*(2), 333-370. https://doi.org/10.1111/j.1468-036X.2006.00357.x

18. Linsmeier, T., & Pearson, N. (1996). *Risk Measurement: An Introduction to Value at Risk* (Working Paper, 96-04). Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=7875

19. Miletic, S., Korenak, B., & Lutovac, M. (2014). Application of VaR (Value at Risk) method on Belgrade Stock Exchange (BSE) optimal portfolio. *Faculty of Business Economics and Entrepreneurship, 1*(2), 142-157. Retrieved from https://scindeks-clanci.ceon.rs/data/pdf/2217-9739/2014/2217-97391402142M.pdf

20. Nadarajah, S., Zhang, B., & Chan, S. (2013). Estimation methods for expected shortfall. *Quantitative Finance, 14*(2), 271-291. https://doi.org/10.1080/14697688.2013.816767

21. Necir, A., Rassoul, A., & Zitikis, R. (2010). Estimating the conditional tail expectation in the case of heavy-tailed losses. *Journal of Probability and Statistics, 2010,* 1-17. https://doi.org/10.1155/2010/596839

22. Richardson, M., & Smith, T. (1993). A test for multivariate normality in stock returns. *The Journal of Business, 66*(2), 295-321. https://doi.org/10.1086/296605

23. Righi, M., & Ceretta, P. (2013). Individual and flexible expected shortfall back-testing. *Journal of Risk Model Validation, 7*(3), 3-20. https://doi.org/10.21314/jrmv.2013.108

24. Petev, I., LSQ-Crest, & Pistaferri, L. (2012). *Consumption in the Great Recession*. Stanford: The Russell Sage Foundation and The Stanford Center on Poverty and Inequality. Retrieved from https://inequality.stanford.edu/sites/default/files/Consumption_fact_sheet.pdf

25. Sharma, M. (2012). What is wrong with quantitative standard deviation? *Research Journal of Finance and Accounting, 3*(4), 16-25.

26. Sheihk, A., & Qiao, H. (2009). *Non-normality of Market Returns*. JP Morgan. Retrieved from https://www.jpmorgan.com/cm/BlobServer/AM_Non-normality_of_market_returns.pdf?blobcol=urldata&blobtable=MungoBlobs&blobkey=id&blobwhere=1158543264199&blobheader=application%2Fpd

27. Thadewald, T., & Büning, H. (2004). Jarque-Bera test and its competitors for testing normality: A power comparison. *Diskussionsbeiträge, 2004/9*. Freie Universität Berlin, Fachbereich Wirtschaftswissenschaft, Berlin. Retrieved from http://hdl.handle.net/10419/49919

28. Wimmerstedt, L. (2015). *Backtesting Expected Shortfall: the design and implementation of different back-test* (Thesis: KTH School of Engineering). Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A848996&dswid=318

29. Fong, T., & Wong, C. (2008). *Stress testing banks' credit risk using mixture vector autoregressive models* (Working paper). Hong Kong Monetary Authority. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1326833.

30. Yamai, Y., & Yoshiba, T. (2002). Comparative analyses of expected shortfall and Value-at-Risk: their estimation error, decomposition, and optimization. *Monetary and Economic Studies, 20*(1), 87-121. Retrieved from https://econpapers.repec.org/article/imeimemes/v_3a20_3ay_3a2002_3ai_3a1_3ap_3a87-121.htm

http://dx.doi.org/10.21511/imfi.15(4).2018.02