S Janse van Rensburg*, D Friskin and GD Sharp

A distributional comparison of size-based portfolios on the JSE

ABSTRACT

This paper examines the non-normality of two smaller capitalization and two larger capitalization portfolios, under different weighting schemes. It demonstrates the superior fit from using two-component mixture of normal distributions instead of a single normal distribution. Additionally, the paper analyzes the components of the mixtures in order to contrast the smaller and larger capitalization portfolios. In doing so, it is shown that the portfolios behave similarly during periods of low volatility, but quite differently during periods of high volatility.

1. INTRODUCTION

The assumption of Normality forms the basis of several econometric modeling methods. This assumption is widespread from inferential analysis of linear models to Black and Scholes options pricing. Unfortunately the distribution of equity log returns is highly peaked with heavy-tails and does not fit well with the characteristics of Normally distributed variables.

This paper tests the use of a mixture of two Normal distributions as a suitable model of equity portfolio log returns, against the alternative of a single Normal distribution. The comparison is achieved by fitting both types of distributions to smaller capitalization and larger capitalization portfolios, under equal weighting and logarithmic weighting in terms of market capitalization. Comparison of the distributions of the two smaller capitalization portfolios and the two larger capitalization portfolios further adds to the debate around size-effects, which suggests that the return distribution may be related to firm size or portfolio capitalization.

2. LITERATURE REVIEW

In a study on the stability of equity covariance, Clark and Troskie (2006) found that their sample of thirty three large capitalization JSE-listed shares were, by the Jarque-Bera Test, non-Normal and displayed excessive kurtosis with negligible skewness. Their paper showed that the covariance and correlation between shares changed over time.

The time-varying nature of equity variances and with it the covariance of groups of shares, together with the observed non-Normality, had recently led researchers to avoid the direct modeling of log return distributions by either developing non-parametric models or employing some type of autoregressive conditional heteroskedasticity (ARCH) model.

For example, De Araújo and Maré (2006) applied a non-parametric option pricing technique to calculate the volatility skew of the JSE/FTSE Top 40 index futures options' market. They cited the non-Normality of asset returns as partial motivation for their approach. Samouilhan (2007) used a Component Generalized ARCH model to study the persistence of volatility clustering among different sectors of the JSE and found a material difference between the behavior of sectors' conditional volatility.

Indeed, ARCH-type models proved popular due to their ability to explain volatility clustering. Volatility clustering occurred when large (small) price changes were followed by similarly large (small) changes in the following period. This resulted in periods of sustained high or low volatility, which led to the excessive kurtosis of equity return distributions.

As an alternative to ARCH, Bahng (2004) explained volatility clustering by subdividing a return series and individually modeling each sub-period's distribution. Specifically, Bahng (2004) identified break points, separating sub-periods, on the Swiss Market Index (SMI) using Goldfeld and Quandt's two-period structural break analysis and fitted a Normal distribution to each sub-period. The result, when considering the entire time-series, was a mixture of Normal distributions.

A simpler view, taken from Alexander (2001:297), was that the market always found itself in one of two states, namely a high volatility or low volatility state. The former could be modeled as a Normal distribution with high variance, whereas the latter could be modeled as a Normal distribution with low variance.

*Department of Statistics, Nelson Mandela Metropolitan University. Port elizabeth, Republic of South Africa.
Email: sjansevanrensburg@gmail.com; david.friskin@nmmu.ac.za; gary.sharp@nmmu.ac.za
The authors wish to extend their appreciation to the anonymous referees, the Journal Editor and Dr Catherine Logie whose comments and suggestions improved significantly on an earlier version of this paper. This material is based upon work supported financially by the National Research Foundation, Grant No. TTK2007041600001. Any opinion, findings and conclusions or recommendations expressed in this material are those of the author(s) and therefore the NRF does not accept any liability in regard thereto.

Chen, Gupta and Troksie (2003) developed a similar model within a simple Capital Asset Pricing Model (CAPM) framework. Assuming that the market proxy followed a mixture of Normal distributions, Chen *et al.* (2003) proved that an individual share's returns would follow a mixture of skew-Normal distributions during a bear or bull market, with differing beta-risk for each market state.

In addition to non-Normality, there remained questions about the effect of firm size on returns. While Cuthbertson and Nitzsche (2004:435) asserted that the small-firm effect had, for the US market, disappeared after the 1980s, several studies still postulated some form of size-effect. Elfakhania and Wei (2003) found evidence of a combined size-price effect on the Canadian market, whereby higher priced small and medium sized firms earned higher returns than large firms. For the US, Zepp (2003) cited studies which supported the claim of a small-firm effect among water utility companies.

Bundoo (2006) used the Fama-French three factor model to show that a significant size and value effect existed on the Stock Exchange of Mauritius (SEM). In South Africa, van Rensburg and Robertson's (2003) style-based two-factor model lent support to the notion of a size-effect. In their model, they identified size as a significant factor in explaining equity returns. The regression coefficient, relating to size, was also negative. This indicated that a decrease in market capitalization led to an increase in returns, implying a small firm effect.

3. RESEARCH METHODOLOGY

This section details the research methodology. First some salient features of the dataset are explained and then the procedure for creating the portfolios, which formed the basis of the study, is specified. Following this return calculations are briefly mentioned, after which distribution fitting and goodness of fit testing are described. The discussion of distribution fitting also serves to give background on the theory of Normal mixture distributions.

3.1 Dataset

The dataset, purchased from Johannesburg Securities Exchange Limited, consisted of monthly price, market capitalization and volume data for companies listed during the period January 1990 till December 2005. There were, in total, 366 companies included in the dataset. Several series had missing observations. These missing values were simply replaced with the last observed record, resulting in 192 observations for any company listed for the entire period of study. Dividends were not considered.

The dataset did not include any delisted firms and no attempt was made to correct for survivorship bias. While Pawley (2006) noted significant survivorship bias within the South African unit trust industry, with an estimated 52,18% survival rate for funds over a fifteen year period, the lack of any specific rates of attrition for smaller and larger capitalization portfolios meant that their findings could not be applied.

Additionally, the Canadian study by Elfakhani and Wei (2003) found that survivorship bias was only significant at the 11% level, indicating that the bias might not be as severe as usually thought. Bundoo (2006) cites further studies which supported the notion that survivorship bias was often overemphasized. The Mauritian study also included nearly all firms listed on the SEM, thus Bundoo's (2006) results did not suffer survivorship bias.

3.2 Portfolio creation

Out of the dataset, two mutually exclusive groups of shares were selected. The smaller capitalization group consisted of the twenty smallest capitalization shares, at the end of January 1990, with price less than R2,50 and positive trading volume. The large capitalization group consisted of the twenty largest capitalization shares, at the end of January 1990, with price greater than R2,50 and positive trading volume.

The price of R2,50 was chosen in order to prevent the inclusion of high-priced shares with a low number of shares in issue into the smaller capitalization portfolio. Due to inflation, R2,50 in the year 1990 was worth about R9,75 in today's money. Shares priced slightly higher than this would violate McLachlan's (2007) popular, though casual, definition that small capitalization shares should be priced below R10,00.

From each group, an equally weighted portfolio was created. Additionally, for the smaller capitalization group a log weighted portfolio was created and for the larger capitalization group an inverse log weighted portfolio was calculated. Weighting was in terms of market capitalization, such that the log weighting assigned more weight to the smallest of the small capitalization shares and the inverse log weighting assigned more weight to the largest of the large capitalization shares. The log and inverse log weighted portfolios were thus more sensitive to size effects than the equally weighted portfolios.

While the usual market capitalization weighting scheme would produce a portfolio sensitive to large-firm effects, assigning weight directly proportional to market capitalization would not amplify any small-firm effect. The log weighting, however, would amplify effects attributable to a lower capitalization. Rather than using the directly proportional market capitalization weighting to amplify large-firm effects, the inverse log weighting scheme was used in this study because it was directly comparable with log weighting.

We therefore had the following four portfolios:

1. An equally weighted smaller capitalization portfolio, or EWS-portfolio.

2. An equally weighted larger capitalization portfolio, or EWL-portfolio.

3. A log weighted smaller capitalization portfolio, or LWS-portfolio.

4. An inverse log weighted larger capitalization portfolio, or IWL-portfolio.

From here on, the paper refers to the above portfolios using the stated abbreviations.

3.3 Return calculation

Month-on-month log return series were calculated for each of the four portfolios. Each return series contained 191 monthly observations, starting February 1990.

3.4 Distribution fitting

Both Normal distributions and two-component mixture of Normal distributions were fitted to the four portfolios. While fitting Normal distributions to data via maximum likelihood estimation (MLE) was trivial, the procedure for fitting a mixture of Normal distributions warrants further explanation.

While defining a mixture of Normals as a probability weighted average of Normal distributions, it became clear that the weights as well as multiple means and variances had to be estimated. Relating this to Alexander's two-state model (2001:297), the parameters of the two Normal distributions, which define the mean and variance of the high and low volatility states, and the probability that observations were generated by each Normal distribution, or volatility state, had to be estimated. In the terminology of mixture models, each Normal distribution was a component in a two-component mixture of Normals. The probability associated with observations coming from either state, or component distribution, was the component weight.

Unlike in Bahng (2004), the data was not divided into different sub periods. This meant that observations could not be directly assigned to either component. In other words, there was no observable indicator or label marking observations as coming from one component or the other. The unobservable labels, together with the observed returns, formed a hypothetical complete dataset and MLE parameter estimation was simplified to an incomplete-data problem (McLachlan and Peel 2000:19). This incomplete-data problem necessitated the use of the Expectation Maximization (EM) algorithm.

The EM algorithm consisted of two steps, namely an Expectation-step (E-step) and a Maximization-step (M-step). During the E-step, the unknown labels were estimated by taking the expectation of the complete-data log likelihood given an estimate of the parameters. This expected log likelihood was then maximized, M-step, yielding new parameter estimates. These estimates were substituted back into the complete-data log likelihood and the E-step and M-step were repeated until convergence occurred.

Initial estimates, or starting values, for the EM algorithm were determined using k-means clustering. For the sake of brevity, the details as to the workings of the k-means algorithm is omitted. It functions similarly to the EM algorithm.

3.5 Goodness of fit testing

Goodness of fit, for the fitted Normal and mixture distributions, was determined using a Chi-squared goodness of fit test. Grouping the data into bins, the test evaluated whether the deviation between the observed and expected number of observations per bin was significant.

4. RESULTS

In what follows basic descriptive statistics, the results from distribution fitting and goodness of fit testing are reported.

4.1 Summary statistics

Summary statistics for the four portfolios were recorded in Table 1. In general, the portfolios were highly peaked and slightly skewed to the left. This was in line with the findings of Clark and Troskie (2006). We also noted that the smaller capitalization portfolios had higher mean return, standard deviation and excess kurtosis than the larger capitalization portfolios.

Table 1: Summary statistics for the four portfolios

	EWS-portfolio	EWL-portfolio	LWS-portfolio	IWL-portfolio
Mean	0,01565	0,00475	0,01517	0,004818
Standard Deviation	0,08130	0,07182	0,07348	0,07499
Excess Kurtosis	8,78461	1,49542	5,36520	1,42655
Skewness	-1,44095	-0,66958	-0,20714	-0,61027

4.2 Distribution fitting

Fitting Normal distributions to the four portfolios merely involved setting the mean and variance parameters equal to the sample average and sample variance, respectively.

Taking into account both the small degree of skewness and the complexity of the EM algorithm, it was opted to fit mixtures with common component means and arbitrary variances. These are known as scale mixtures, since only the scale or variance parameter differs between component distributions.

GNU R (R Development Core Team 2007), together with the package mixtools (Young, Elmore, Hettmansperger, Hunter, Thomas and Xuan 2007), were used to fit two-component scale mixture of Normal distributions to the return series. The results, summarized in Table 2, revealed several interesting patterns. The weight associated with the higher variance component was lower for the smaller capitalization portfolios than it was for the larger capitalization portfolios. Indeed, the weighting for the EWS-portfolio's and LWS-portfolio's higher variance component was over 4,6 and 3,6 times smaller than that of the EWL-portfolio and IWL-portfolio, respectively. The high-variance component standard deviation was, however, greater for the smaller capitalization portfolios. Specifically, the higher variance component's standard deviation for the EWS-portfolio and LWS-portfolio was about 2,32 and 1,72 times larger than that of EWL-portfolio and IWL-portfolio respectively. Lastly, the standard deviations of the lower-variance components were, for all four portfolios, of a similar order of magnitude, with the smaller capitalization portfolios displaying slightly higher variance.

The above claims, concerning component variances were tested, using a standard F-test. Table 3 showed that there was insufficient statistical evidence to conclude that the variances of the low variance components differed between the smaller and larger capitalization portfolios. There was, however, a difference between the high variance components' variances. This was tested further, with Table 4 showing that the smaller capitalization portfolios' higher variance components had significantly larger variances than that of the larger capitalization portfolios' higher variance component.

In terms of a two-state model, the above suggested that the portfolios were exposed to a similar degree of risk when in a low-volatility state. During the high-volatility state, the smaller capitalization portfolios experienced greater volatility than the larger capitalization portfolios. The larger capitalization portfolios were, however, in the high volatility state more often than the smaller capitalization portfolios.

Table 2: Estimated parameters from fitting scale mixtures of two Normal distributions

Name	$weight_1$	$weight_2$	mean	sd_1	sd_2
EWS-portfolio	0,95595	0,04405	0,01565	0,06158	0,25878
EWL-portfolio	0,79400	0,20600	0,00475	0,05656	0,11215
LWS-portfolio	0,93082	0,06918	0,01517	0,05496	0,19233
IWL-portfolio	0,75021	0,24979	0,00482	0,05726	0,11203

Table 3: F-test for equality of variances between different portfolio component variances

Component	Portfolios	F-statistic	p-value
1	EWS-EWL	1,18539	0,28055
	LWS-IWL	0,92128	0,60338
2	EWS-EWL	5,32430	0,00053
	LWS-IWL	2,94731	0,00819

Table 4: One-sided F-test where the alternative hypothesis claims that the smaller capitalization portfolio's high volatility component variance exceeds the larger capitalization portfolio's high volatility component variance

Portfolios	F-statistic	p-value
EWS-EWL	5,32430	0,00026
LWS-IWL	2,94731	0,00410

4.3 Goodness of fit testing

The Chi-squared test required that each series was divided into bins. Intervals for the bins were determined by averaging the equal-probability intervals for the scale mixtures and Normal distributions.
Table 5 demonstrated the effectiveness of the two-component scale mixtures relative to the single Normal distribution. For all four portfolios, the two-component scale mixture produced a better fit than the Normal distribution. The scale mixtures also performed better for the smaller capitalization portfolios than for the larger capitalization portfolios.

Table 5: Chi-squared test results from fitting Normal and two-component scale mixtures to the portfolios. Note that smaller test statistic values, Q, and larger p-values indicate better fit

Name	Normal		Mixture	
	Q	p-value	*Q*	p-value
EWS-portfolio	32,20324	0,07400	18,76930	0,53686
EWL-portfolio	20,03965	0,58056	17,50778	0,61980
LWS-portfolio	24,07894	0,34310	13,00909	0,87699
IWL-portfolio	22,68781	0,41953	19,90873	0,46365

5. CONCLUSIONS AND RECOMMENDATIONS

The study demonstrated the effectiveness of Normal scale mixtures as a model for equity portfolio returns. Mixture models not only yielded a superior fit to the Normal distribution, but also provided insight into the behavior of returns within a two-state model. These insights pointed to possible differences in the risk associated with smaller and larger capitalization portfolios. We could only hypothesize that there was some differential-information effect, whereby large firms were subject to greater scrutiny from the media which led to larger capitalization portfolios switching states more often. This, however, requires more research.

Having shown that the variances of the high volatility states differ significantly lends support to the claim that the distributions of the two portfolios are different. An area of research that can be expanded on is a test for comparing both mixture distributions simultaneously.

We must additionally note some shortcomings in the study. These include the inability to correct for survivorship bias and the exclusion of dividends in the returns calculation. Were we able to include dividends and a correction for survivorship bias, we would expect the larger capitalization portfolio to become positively skewed, due to higher dividends, while the smaller capitalization portfolio would most likely become negatively skewed, due to higher levels of attrition. Such return distributions could be modeled using two-component Normal mixtures with arbitrary means and variances.

In summary, scale mixtures provide a simple model able to capture the subtle differences in the risk structures of size-based portfolios. These differences, as well as the inclusion of survivorship bias and dividends, provide for future avenues of research.

REFERENCES

Alexander C. 2001. *Market models: A guide to financial data analysis*. Chichester: Wiley.

Bahng J. 2004. Structural breaks and the normality of stock returns. *Swiss Journal of Economics and Statistics*, 127(2):207-227.

Bundoo SK. 2006. An investigation of the size and value premium on the stock exchange of Mauritius. *The African Finance Journal,* 8(1):14-25.

Chen JT, Gupta AK, Troskie CG. 2003. The distribution of stock returns when the market is up. *Communications in Statistics - Theory and Methods*, 32(8):1541-1558.

Clark AE and Troskie CG. 2006. The stability of South African stock returns. *The African Finance Journal,* 8(2):66-77.

Cuthbertson K and Nitzsche D. 2004. *Quantitative financial economics: stocks, bonds and foreign exchange*. Chichester: Wiley.

De Araújo M and Maré E. 2006. Examining the volatility skew in the South African equity market using risk-neutral historical distributions. *The Investment Analysts Journal*, 64:15-19.

Elfakhani S and Wei J. 2003. The survivorship bias, share price effect, and small firm effect in Canadian markets. *Review of Financial Economics*, 12(4):397-411.

McLachlan G and Peel D. 2000. *Finite Mixture Models*. USA: Wiley.

McLachlan, K. 2007. *Welcome to Small Caps*. Available: http://smallcaps.co.za/index.php?option=com_content &task=view&id=44&Itemid=28. Last accessed 21 December 2007.

Pawley MG. 2006. The impact of survivorship bias on South African unit trust performance: 1972-2004 *The Investment Analysts Journal*, 64:21-26.

R Development Core Team. 2007. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Samouilhan NL. 2007. The persistence of SA equity volatility: A component ARCH perspective. *Journal for Studies in Economics and Econometrics*, 31(1):99-118.

Van Rensburg P and Robertson M. 2003. Style characteristics and the cross-section of JSE returns. *The Investment Analysts Journal*, 57:7-16.

Young D, Elmore R, Hettmansperger T, Hunter D, Thomas H and Xuan F. 2007. *Mixtools: Tools for mixture models*. R package version 0.2.0.5.

Zepp T. 2003. Utility stocks and the size effect—revisited. *The Quarterly Review of Economics and Finance*, 43(3):578-582.

Appendix 1: List of companies in portfolios

List of companies used for constructing the portfolios.

Table 6: List of companies included in the smaller capitalization portfolios

Adcorp Hldgs Ltd Ord	Goodhope Diam Kim Ltd
Aflease Gold Ltd	Invicta Holdings Ltd
Basil Read Hldgs Ltd	Jasco Electronics Hldgs
Cashbuild Ltd	Matodzi Resources Ltd
Ceramic Industries Ltd	Merafe Resources Ltd
Combined Motor Hldgs Ltd	Nu World Holdings Ltd
Control Instruments Grp	Putprop Ltd
Dimension Data Hldgs Plc	Simmer and Jack Mines
Don Group Ltd	Spescom Ltd
Eureka Ind Ltd Ord	Winhold Ltd Ord

Table 7: List of companies included in the larger capitalization portfolios

A.E.C.I. Ltd Ord	Lonmin Plc
Anglo American Plc	Mittal Steel SA Ltd
Anglo Platinum Ltd	Nampak Ltd Ord
Anglogold Ashanti Ltd	Nedbank Group Ltd
Barloworld Ltd	Richemont Securities Dr
Gold Fields Ltd	Sabmiller Plc
Impala Platinum Hlgs Ld	Sappi Ltd
Johnnic Holdings Ltd	Sasol Ltd
Liberty Group Ltd	Standard Bank Group Ltd
Liberty Holdings Ltd Ord	Tiger Brands Ltd Ord

Appendix 2: Logarithmic and inverse logarithmic weighting procedures

Under log weighting we have, prior to normalizing,

$$\omega_{t,i} = \sqrt{-\log\left(m_{t,i} \Big/ \sum_{i=1}^{20} m_{t,i} \right)}$$

where $m_{t,i}$ is the market capitalization and $\omega_{t,i}$ is the pre-normalized weight for share i at time t.

Similarly, with inverse log weighting,

$$\omega_{t,i} = \left[-\log\left(m_{t,i} \Big/ \sum_{i=1}^{20} m_{t,i} \right) \right]^{-0.5}$$

These weights are normalized,

$$\tilde{\omega}_{t,i} = \frac{\omega_{t,i}}{\sum_{i=1}^{20} \omega_{t,i}}$$

such that

$$\sum_{i=1}^{20} \tilde{\omega}_{t,i} = 1$$

and the portfolio's price (P_t), at time *t*, is then calculated as

$$P_t = \sum_{i=1}^{20} \tilde{\omega}_{t,i} p_{t,i}$$