5

The Efficacy of the Sortino Ratio and Other Benchmarked Performance Measures Under Skewed Return Distributions

by
Dr Ashraf Chaudhry †
Dr Helen L. Johnson §

Abstract:
This paper will investigate the suitability of existing performance measures under the assumption of a clearly defined benchmark. A range of measures are examined including the Sortino Ratio, the Sharpe Selection ratio (SSR), the Student's t-test and a decay rate measure. A simulation study is used to assess the power and bias of these measures based on variations in sample size and mean performance of two simulated funds. The Sortino Ratio is found to be the superior performance measure exhibiting more power and less bias than the SSR when the distribution of excess returns are skewed.

Keywords:
SORTINO RATIO; SHARPE SELECTION RATIO; POWER; BIAS; SKEWED; BOOTSTRAP.

† Queensland Investment Corporation, Brisbane, QLD, 4001. Email: a.chaudhry@qic.com
§ Queensland University of Technology, Brisbane, QLD, 4001. Email: h.johnson@qut.edu.au

Part of this work was completed whilst Dr Ashraf Chaudhry was studying for a PhD in Finance at the Australian National University in Canberra, Australia.

Australian Journal of Management, Vol. 32, No. 3, Special Issue, March 2008, © The University of New South Wales

1. Introduction

Typically, wholesale active fund managers are mandated to outperform target passive investment portfolios whilst maintaining an acceptable level of risk. One of the rationales behind these mandates is to provide quality control and risk management in protecting the funds invested with the active manager. Other motivations stem from the beliefs of the sponsor and include portfolio tilts and/or diversification requirements as advocated under portfolio theory. Under these circumstances, fund manager remuneration is linked directly to performance and risk relative to the mandated benchmark and therefore, careful analysis of the appropriate measures of performance is necessary.

The use of CAPM or conventional multi-factor regression models may be inappropriate since they do not explicitly account for performance relative to a benchmark that the managers are mandated to outperform. Therefore, due to the increased focus on a fund manager's performance relative to a specific benchmark, Sharpe (1994) proposed the performance measure known as the Sharpe Selection Ratio (SSR). This measure estimates the average excess return of the fund over a benchmark per unit of benchmark risk, where benchmark risk is defined as the standard deviation of excess returns around the benchmark. However, the Sharpe Selection Ratio implicitly assumes that investors are indifferent to upside and downside risk. If the distribution of excess returns (alpha) is symmetric this will make no difference to the performance estimate, yet this measure may lead to an incorrect ranking of a fund manager if the distribution of the alphas are skewed. Bawa & Lindenberg (1977) posited that a more appropriate measure of risk should only be concerned with underperformance of the fund manager relative to the benchmark. Bawa argued that if alpha has a positively skewed distribution, this will invariably increase the standard deviation of alpha thus reducing the Sharpe Selection Ratio, whilst the risk of significantly underperforming the benchmark is lower relative to a symmetric or negatively skewed distribution of alpha. As a result, a given manager will be penalised for obtaining high alphas. Conversely, active managers with a negatively skewed distribution of excess returns will have a higher probability of obtaining a large negative return relative to the benchmark with limited upside potential. Therefore, Bawa argued that if risk is viewed as underperformance relative to benchmark then the risk adjustment measure should be estimated using a second order Lower Partial Moment (LPM) approach. Sortino and Price (1994) argued that if there is a minimum return that must be earned to accomplish some goal (the minimal acceptable return (MAR)), then any returns below the MAR will produce unfavourable outcomes and any returns greater will produce good outcomes. Risk is associated only with bad outcomes, therefore, only returns below the MAR are associated with risk. As a result, the Sortino Ratio was developed, which is a special case of the lower partial moment approach. Consequently, a second measure to consider in evaluating fund manager performance is the Sortino Ratio which estimates the average excess return relative to the downside deviation. Both the SSR and the Sortino Ratio are thought to be useful measures of performance under the circumstances of a clearly defined benchmark. The Sortino Ratio is preferred on theoretical grounds as it measures risk as underperformance relative to the predefined benchmark and compensates those managers who have a positively skewed distribution of alpha with a higher performance measure. Under this definition of risk, the Sortino Ratio is thought to be a more appropriate measure of performance. However, under the circumstances of a symmetric distribution of alphas the two measures will be equivalent in their ranking outcomes.

Another single-factor model similar in spirit to the Sortino Ratio has been proposed in a working paper by Foster and Stutzer (2003). Fund managers are ranked on the basis of the decay in probability of underperforming the benchmark through time. A manager whose underperformance probability decays to zero at a faster rate will be ranked higher than a manager with a lower decay rate. One drawback of this measure is that it can only rank funds which have a positive probability of exceeding the benchmark as $T \to \infty$ and therefore will only rank funds with a past performance above benchmark. This measure estimates the decay in

$$\text{Prob}\left[\frac{1}{T}\sum_{t=1}^{T}\left(\log R_{it} - \log R_{bt}\right) \le 0\right]$$

as $T \to \infty$. The Gartner-Ellis Large Deviations Theorem states this probability will decay to zero at a fund specific exponential rate (for all funds with past positive average alphas measured in continuous time). Under the assumption that the continuous alphas are independently and identically distributed, a consistent estimate of the decay rate for each fund can be calculated according to

$$\max_{\lambda > 0}\left(-\log \frac{1}{T}\sum_{t=1}^{T} e^{-\lambda\left(\log R_{it} - \log R_{bt}\right)}.\right)$$

The Foster and Stutzer decay rate is another measure that estimates rank and performance relative to a specific benchmark and therefore is considered to be an appropriate measure for this study.

Consequently, if a fund manager is mandated to a specified benchmark, the choice methods for ranking fund managers include the SSR, the Sortino Ratio and the Foster and Stutzer Decay Rate, all of which allow for a specific benchmark and define performance and risk relative to that benchmark. In this paper, a simulation study is used to assess the size, power and bias of the aforementioned performance measures based on variations in the size of the notional data set and changes in the mean performance of two simulated funds, where the excess returns (alpha) of the funds are drawn from a variety of both symmetrical and skewed distributions with known parameters. Further to this, in order to test the efficacy of these preferred measures, a sample of active fund managers within the Australian market that have clearly identifiable benchmarks have been selected. The benefit of these measures is therefore dependent on the ability to be able to *a priori* allocate a benchmark to the relevant fund manager. These three methods provide a broad umbrella of measures that are calculated from the data alone, and are used to ascertain fund outperformance both in academia and industry under the assumption of a clearly defined benchmark. A hypothesis test using the Student's *t*-statistic is also used to ascertain whether a manager's alpha is significantly above benchmark. Managers with a high variance around the mean will have a lower *t*-statistic and as a result, this value can also be used to measure outperformance directly relative to the benchmark. Finally, a bootstrap is applied to obtain 95% empirical confidence intervals on the skewness of each manager's excess returns in order to ascertain whether the techniques discussed here are appropriate. This research follows a working paper by Chaudhry and Johnson (2006), which investigates similar properties of the aforementioned performance measures applied to US data. Similar results were found.

The outline of this paper is as follows. In section 2 the Sortino Ratio, SSR and Student's *t*-statistic will be formally defined and in section 3, a simulation study will be carried out to compare each of the performance metrics through a number of distributions with different properties. In section 4, the data for which these performance measures will be compared will be introduced and each of the performance measures will be calculated using the Mercers survey data. Section 5 concludes the paper.

2. The Measures Defined

The *m*th order lower partial moment for discrete data is given by

$$LPM = \frac{1}{N}\sum_{t=1}^{N}(r_t - L)^m I(r_t \le L) \tag{1}$$

where L is some given threshold, N is the number of returns and r_t is the return at time t. The LPM is a general type of risk measure that encompasses a number of special cases, one being the downside deviation risk measure used in the Sortino Ratio. The downside deviation (DD) assumes that in Equation 1 above, m=2 and L=MAR, where we define it as follows:

$$DD^2 = \frac{1}{N}\sum_{t=1}^{N}(r_t - MAR)^2 I(r_t \le MAR)$$

In this paper, the benchmark of the funds is assumed to be a passive portfolio of securities that the active manager has been mandated to exceed. Therefore, if alpha is defined as the fund return minus benchmark (or MAR), the benchmark of the funds is then equal to zero since a positive alpha implies that the fund managers are outperforming their mandated passive benchmarks. Therefore, we can formally define the Sortino Ratio as follows

$$\text{Sortino Ratio} = \frac{\overline{\alpha}}{DD} \tag{2}$$

where $\overline{\alpha}$ represents the average return of an active manager above their designated benchmark B, and DD is the downside deviation evaluated for MAR=0.

Similarly, the SSR can be expressed as

$$SSR = \frac{\overline{\alpha}}{\sigma} \tag{3}$$

where $\bar{\alpha}$ is defined as before and σ^2 is simply the variance of the excess returns.

Finally, the Student's *t*-statistic is given by

$$T = \frac{\bar{\alpha}}{(\sigma / \sqrt{n})}.$$

If it can be shown that this test statistic is significantly above zero, then this can also be used to measure outperformance directly relative to the benchmark portfolio.

3. A Simulation Study to Assess Size, Power and Bias

To ascertain the power and bias of each of the performance measures, a series of simulation studies was conducted using two simulated funds for comparison. A random sample of observations, *N*, which ranged from 2 to 100, was drawn from a variety of both symmetrical and skewed distributions with known parameters and the performance measures were calculated. Each of these cases are now described in turn.

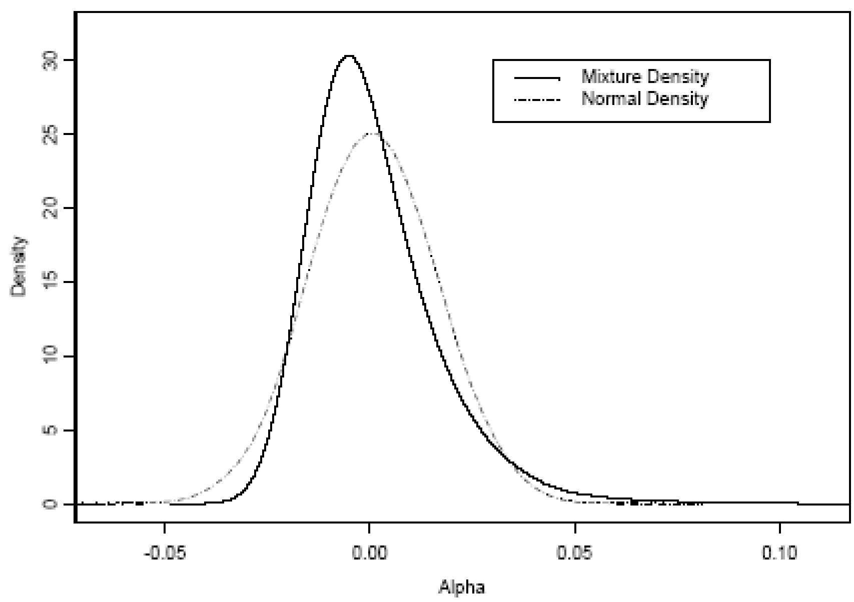
Case 1: Two Normal Distributions The first set of simulations assumes that the returns for both funds are drawn from a normal distribution, where the mean of one of these funds (Fund B) is allowed to vary from 0.1% (per month) through to 0.5% (per month) in increments of 0.05%, while the mean of the alternative fund (Fund A) remains constant at 0.1% (per month). Furthermore, the standard deviation for each fund is equal to 1.04% per month, which is representative of the average standard deviation of funds within the Australian data set. This process was repeated 1000 times and the percentage of times that Fund B ranked higher than Fund A was calculated for each N in order to enable a comparative measure between the ranks of the two simulated funds. Given that both funds have the same risk profile, one would expect Fund B to have a higher percentage rank than Fund A for all performance measures when the mean of Fund B is higher than the mean for Fund A.

Case 2: *Two Skewed Distributions* The second set of simulations assumes that the alphas for both Fund A and Fund B are drawn from positively skewed distributions. The required skewed distribution is created using a mixture of a normal and a lognormal distribution where the rationale behind this choice is as follows. A lognormal distribution can be created with positive skew properties (given that the appropriate parameters are chosen). However, it is bounded at zero. Even if this distribution were relocated so that negative alphas were possible, there would still be a lower bound on the possible values that alpha could take. In comparison, a normal distribution allows for any value of alpha on the real number line. Consequently, using a mixture of the lognormal and the normal distribution provides a distribution with a positive skew as well as allowing the minimum possible value of alpha to be non-bounded. The details regarding model and parameter selection are shown in the Appendix. Once again, the mean and standard deviation of Fund A are chosen to be 0.1% and 1.04% respectively, whilst the standard deviation of Fund B remains fixed at 1.04% and its mean is allowed to vary from 0.1% to 0.5% in increments of 0.05%, as before.

Case 3: *A Normal Distribution and a Skewed Distribution* The last set of simulations assumes that the monthly alphas for Fund A are drawn from a normal distribution with a mean of 0.1% and standard deviation of 1.04%. Fund B's alphas are drawn from a positively skewed distribution where the mean is allowed to vary from 0.1% to 0.5% in increments of 0.05% whilst the standard deviation remains fixed at 1.04%. A plot of the densities of the normal distribution and mixture distribution is provided in Figure 1, both with mean equal to 0.1% and standard deviation equal to 1.04%. As argued earlier, performance measures such as the SSR and *t*-statistic are expected to provide biased results if it is assumed that risk should be treated as performance under benchmark since managers with a positive skew on the alphas are being penalised for achieving high excess returns. As such, we would expect the Sortino Ratio to be the superior measure in this case.

Figure 1

A Comparison of a Normal Distribution and a Positively Skewed Distribution both with Mean Equal to 0.1% and Standard Deviation Equal to 1.04%



3.1 Power

The power of each performance measure can be ascertained by the proportion of times that Fund B is ranked above Fund A where the power is measured relative to both the changing number of sample values from each manager, *N*, as well as the variation in the mean of Fund B. Tables 1 through to 3 show the results from the series of simulation cases described previously. For a more detailed comparison, *N* was selected to represent a small, medium and large sample size. As such, values of *N* = 15, 50 and 100 were chosen. Note that the proportion of times that Fund B was ranked above Fund A will be identical for the SSR and t-statistic given that *N* was constant at each set of iterations.

Table 1 shows the power of each performance measure when the alphas of both Fund A and Fund B were drawn from a normal distribution. As expected, as the mean alpha of Fund B is increased relative to Fund A and as the sample size becomes large, all performance measures have increasing power in their ability to select the optimal fund. All measures were performing equally well under these circumstances and there appears to be little difference in the power of the four performance measures in choosing the best manager.

Table 1

Proportion of Times Fund B Outranks Fund A Where Fund A's Alphas are Drawn from a $N(0.001, 0.0104^2)$ and Fund B from a Normal Distribution with Standard Deviation Equal to 0.0104 and Mean Varying as Shown in the Table

Sample Size	Mean of Fund B	T-stat	SSR	Sortino Ratio	Decay
N=15	0.001	0.487	0.487	0.488	0.486
	0.0015	0.532	0.532	0.534	0.536
	0.002	0.576	0.576	0.582	0.581
	0.0025	0.652	0.652	0.650	0.653
	0.003	0.724	0.724	0.722	0.723
	0.0035	0.750	0.750	0.746	0.752
	0.004	0.772	0.772	0.775	0.769
	0.0045	0.818	0.818	0.811	0.818
	0.005	0.848	0.848	0.846	0.844
N=50	0.001	0.503	0.503	0.505	0.505
	0.0015	0.583	0.583	0.582	0.585
	0.002	0.685	0.685	0.684	0.684
	0.0025	0.765	0.765	0.764	0.767
	0.003	0.829	0.829	0.835	0.831
	0.0035	0.883	0.883	0.879	0.882
	0.004	0.929	0.929	0.926	0.931
	0.0045	0.951	0.951	0.953	0.949
	0.005	0.970	0.970	0.969	0.971
N=100	0.001	0.503	0.503	0.502	0.502
	0.0015	0.645	0.645	0.647	0.644
	0.002	0.720	0.720	0.721	0.719
	0.0025	0.852	0.852	0.850	0.850
	0.003	0.903	0.903	0.898	0.903
	0.0035	0.955	0.955	0.952	0.955
	0.004	0.980	0.980	0.984	0.980
	0.0045	0.990	0.990	0.990	0.991
	0.005	0.997	0.997	0.994	0.996

Figure 2 shows the proportion of times that Fund A outranked Fund B for *N*=2,...,100 when both funds have the same mean and variance. Therefore, it is expected that the percentage of time that Fund B will be ranked above Fund A will be 50%. This is shown to be consistent with the trajectory plots in Figure 2.

Figure 2

Proportion of Times That Fund B is Ranked Above Fund A Where Both Funds Alphas are Drawn from a $N(0.001, 0.0104^2)$

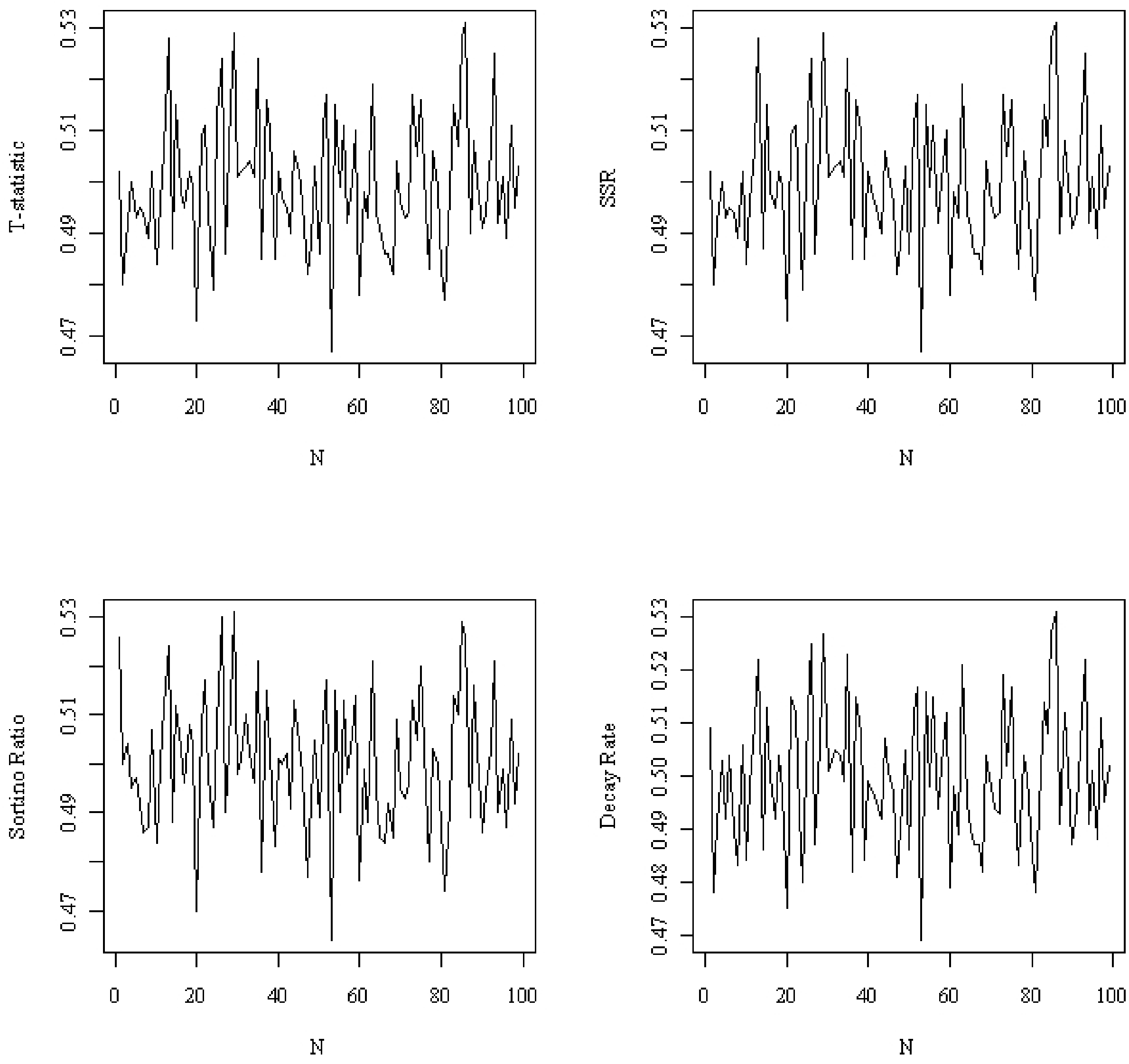


Figure 3 shows the trajectory plot of a second set of simulations that were drawn from normal distributions where the mean of Fund B is greater than that of Fund A (by a difference of 0.001). In this situation it is expected that each of the measures will rank Fund B above Fund A as $N \to \infty$. However, the power of each performance measure can be ascertained by the rate of increase in the likelihood of choosing Fund B over Fund A as $N \to \infty$. Figure 3 shows that all four performance measures have a similar power in the ability to choose the superior fund, with no measure standing out above the rest, although it could be argued that the Sortino Ratio was increasing at a faster rate.

Figure 3

Proportion of Times that Fund B is Ranked Above Fund A. Fund B is Drawn From a $N(0.002,0.0104^2)$ and Fund A From a $N(0.001,0.0104^2)$

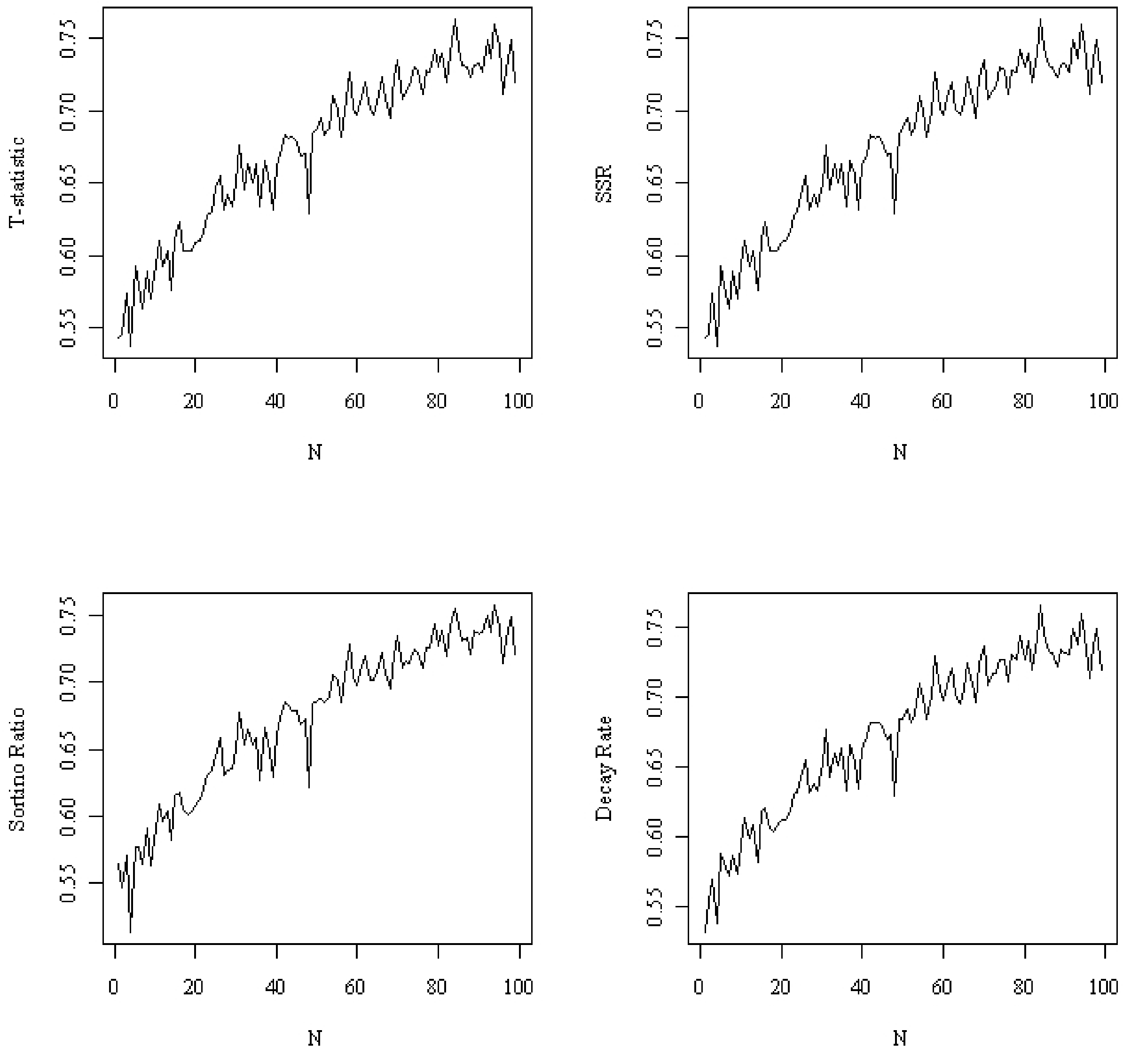


Table Table 2 shows the power of each performance measure when the alphas of both Fund A and Fund B were drawn from a positively skewed mixture distribution with a standard deviation of 1.04%. As in Case 1, the mean of Fund B was increased in increments from 0.1% to 0.5% where the sample size is equal to N = 15, 50 and 100 as described before. Once again it was evident there is little difference in the power of the performance measures in choosing the optimal fund, although the Sortino Ratio is consistently calculating a higher percentage, thus implying that it was correctly ranking the manager most often.

Table 2

Proportion of Times Fund B Outranks Fund A Where Fund A's Alphas are Drawn From a Positively Skewed Mixture Distribution with Mean 0.001 and Standard Deviation 0.0104 and Fund B From a Similar Mixture Distribution with the Same Standard Deviation, However, with Varying Means as Shown in the Table

Sample Size	Mean of Fund B	T-stat	SSR	Sortino Ratio	Decay
N=15	0.001	0.477	0.477	0.509	0.500
	0.0015	0.541	0.541	0.584	0.558
	0.002	0.574	0.574	0.610	0.583
	0.0025	0.628	0.628	0.665	0.652
	0.003	0.694	0.694	0.722	0.710
	0.0035	0.742	0.742	0.785	0.757
	0.004	0.785	0.785	0.817	0.795
	0.0045	0.816	0.816	0.846	0.830
	0.005	0.876	0.876	0.921	0.893
N=50	0.001	0.481	0.481	0.539	0.498
	0.0015	0.596	0.596	0.654	0.605
	0.002	0.666	0.666	0.738	0.688
	0.0025	0.761	0.761	0.819	0.778
	0.003	0.855	0.855	0.902	0.868
	0.0035	0.900	0.900	0.946	0.917
	0.004	0.952	0.952	0.980	0.963
	0.0045	0.972	0.972	0.990	0.985
	0.005	0.989	0.989	0.999	0.994
N=100	0.001	0.464	0.464	0.525	0.474
	0.0015	0.617	0.617	0.701	0.631
	0.002	0.768	0.768	0.824	0.777
	0.0025	0.862	0.862	0.911	0.875
	0.003	0.928	0.928	0.964	0.937
	0.0035	0.966	0.966	0.984	0.973
	0.004	0.980	0.980	0.991	0.987
	0.0045	0.995	0.995	0.999	0.998
	0.005	0.997	0.997	0.999	0.999

Table 3 shows the power of each performance measure when the alphas of Fund A were drawn from a normal distribution with a mean and standard deviation of 0.1% and 1.04% per month, respectively, whilst the alphas for Fund B were drawn from a positively skewed mixture distribution with the same standard deviation, yet the mean was allowed to vary. Once again, the results are shown for N = 15, 50 and 100. Under this set of distributional conditions, the Sortino Ratio provides a more robust measure of outperformance of Fund A relative to Fund B. The power of the

Sortino Ratio was significantly higher than the other performance measures in its ability to choose the optimal fund over finite *N*.

Table 3

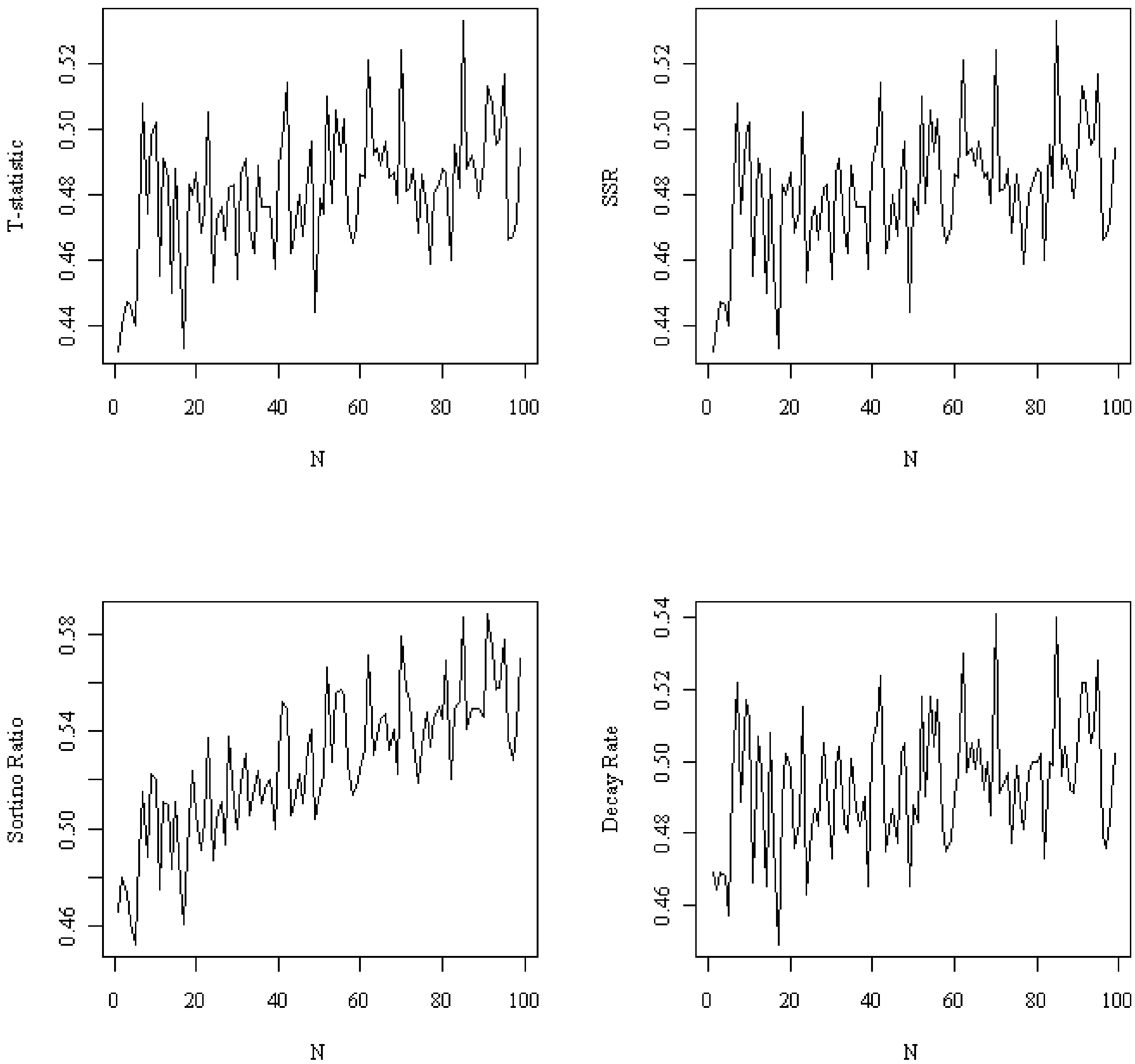
Proportion of Times Fund B Outranks Fund A Where Fund A's Alphas are Drawn From a $N(0.001, 0.0104^2)$ and Fund B's Alphas are Drawn From a Positively Skewed Mixture Distribution with Mean X and Standard Deviation 0.0104, with the Mean X Varying as Shown in the Table

Sample Size	Mean of Fund B	T-stat	SSR	Sortino Ratio	Decay
N=15	0.001	0.450	0.450	0.484	0.465
	0.0015	0.533	0.533	0.570	0.552
	0.002	0.579	0.579	0.607	0.593
	0.0025	0.612	0.612	0.650	0.63
	0.003	0.687	0.687	0.720	0.706
	0.0035	0.741	0.741	0.787	0.757
	0.004	0.761	0.761	0.813	0.782
	0.0045	0.812	0.812	0.858	0.838
	0.005	0.853	0.853	0.888	0.866
N=50	0.001	0.444	0.444	0.504	0.465
	0.0015	0.583	0.583	0.639	0.599
	0.002	0.680	0.680	0.746	0.693
	0.0025	0.783	0.783	0.839	0.797
	0.003	0.847	0.847	0.900	0.861
	0.0035	0.894	0.894	0.931	0.903
	0.004	0.934	0.934	0.966	0.944
	0.0045	0.96	0.96	0.987	0.969
	0.005	0.985	0.985	0.997	0.993
N=100	0.001	0.494	0.494	0.57	0.502
	0.0015	0.631	0.631	0.717	0.648
	0.002	0.726	0.726	0.810	0.741
	0.0025	0.824	0.824	0.891	0.843
	0.003	0.933	0.933	0.965	0.942
	0.0035	0.961	0.961	0.984	0.969
	0.004	0.990	0.990	0.999	0.994
	0.0045	0.997	0.997	0.999	0.998
	0.005	0.998	0.998	1	0.998

Figure 4 provides a graphical representation of the performance of the Sortino Ratio relative to the other performance measures when the distribution of Fund B was positively skewed. Given the positively skewed distribution has a lower probability of very large losses and a higher probability of very large gains, it is argued that Fund B should be ranked above Fund A over finite *N*. Figure 4 shows that the SSR, *t*-statistic and decay rate are indifferent to skewness in the distribution of alphas and rank the funds (on average) equally with the proportion of times that Fund B exceeds Fund A equal to 0.5 as $N \to \infty$. However, the Sortino Ratio had a much stronger preference to Fund B with the proportion of times Fund B ranking above Fund A increasing as $N \to \infty$.

Figure 4

Proportion of Times that Fund B is Ranked Above Fund A. Fund A is Drawn From a *N*(0.001,0.0104^2) and Fund B From a Positively Skewed Mixture Distribution with Same Mean and Variance



3.2 Bias

The bias of the SSR and Sortino Ratio was assessed relative to their true values over different choices of *N*. Evidence from previous literature suggests that sample ratios exhibit bias relative to their true values especially when the sample size is small. Consequently, it was deemed prudent to test the level of bias of the SSR and Sortino Ratio and thus assess their suitability as a superior performance measure. Figure 5shows the bias of the SSR and Sortino Ratio where the alphas of the funds were drawn from positively skewed distributions respectively. The bias associated with the two ratios can be obtained from the following two expectations:

$$E(S\hat{S}R) = E(\frac{\hat{\alpha}}{\hat{\sigma}_\alpha}) = E(\hat{\alpha})E(\frac{1}{\hat{\sigma}_\alpha})$$
$$\geq \alpha[\frac{1}{E(\hat{\sigma}_\alpha)}] = SSR$$

$$E(\text{Sortin}\hat{\text{o}}\text{ Ratio}) = E(\frac{\hat{\alpha}}{\hat{DD}}) = E(\hat{\alpha})E(\frac{1}{\hat{DD}})$$
$$\geq \alpha[\frac{1}{E(\hat{DD})}] = \text{Sortino Ratio}$$

These results from Jensen's inequality on convex functions are strict as long as the sample estimates are not degenerate at the constant values of the population parameters. Miller and Gehr (1978) and Lo (2002) suggest that bias can also result from the mispecification of unbiased population parameters and from inappropriate time aggregation of performance data. However, these effects will not alter the results under the design of the simulation study carried out in this paper.

After observing Figure 5, it is evident that as $N \rightarrow \infty$ the bias of the Sortino Ratio measure decays at a faster rate than that of the SSR Ratio when the alphas for each of the funds are drawn from a positively skewed mixture distribution.

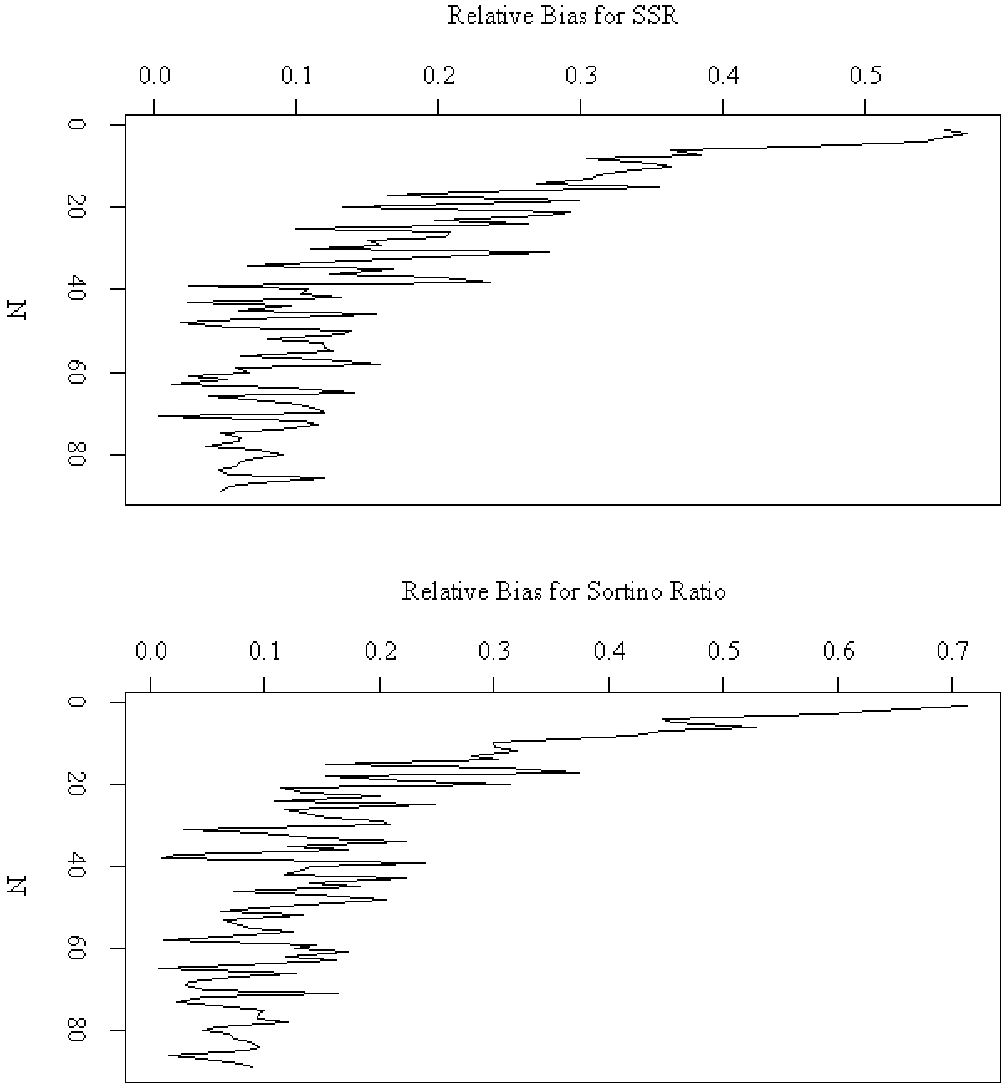
4. Results Using Active Fund Manager Data

The data was obtained from the Mercer survey of long active Australian equity fund managers over the period December 1980 through to November 2006. The total return (including dividends and capitalisation changes) of the ASX300 was chosen as a representative benchmark. Excess returns were estimated as the arithmetic difference between fund and benchmark return and a total of 122 funds were included over the period of the study.

Each of the four performance measures were used to rank the 122 active funds used in this study. Table 4 shows a comparison between the ranks for the different measures for the top 20 managers, ranked by the Sortino Ratio. Using each of the performance measures, Manager 91 was ranked the highest. Further to this, each measure has ranked the same top 3 managers in exactly the same order. However, from the third manager downwards, there is a difference in the ranks for all of the performance measures, none of which appear to be consistently similar. Whilst the results for the various methodologies are in general quite similar there were some notable differences. In particular, Managers 33 and 23 had a much higher rank under the *t*-statistic than under the other performance measures, largely due to the number of observations for each manager (133 and 174, respectively) which reduced the standard error and therefore inflated the performance measure. Furthermore, the Sortino ratio ranked Manager 82 in 4th position, one position ahead of the Sharpe Selection Ratio and *t*-statistic methodologies. This particular manager had a significant positive skew in its excess returns and thus benefited in rank when using the Sortino ratio as a performance measure.

Figure 5

Relative Bias of the SSR and Sortino Ratio Where the Alphas of the Funds are Drawn from Positively Skewed Distributions



To ascertain the empirical relevance of the performance measures that account for skew we used a bootstrap to obtain 95% empirical confidence intervals on the skewness of each manager's historic excess returns. Since any measure of skewness will be unreliable without enough data, all managers with a return history of less than 30 observations were removed from the sample. This filtering process left 97 out of the 122 managers on which the bootstrap was calculated. For each manager, a sample equal to the number of observations was drawn with replacement from the observed data and the coefficient of skewness was then calculated. This process was then repeated 1000 times. The mean and standard error were then estimated from the 1000 values and a 95% confidence interval was calculated by ordering the 1000 calculated coefficients of skewness and selecting the 2.5th and 97.5th percentiles.

Table 4

Top 20 Australian Funds Ranked by Sortino Ratio Over the Period December 1980 to November 2007. Relative Rankings are also Provided Using the SSR, t-statistic, and the Decay Rate

Manager	Sortino Ratio	SSR (Rank)	t-stat (Rank)	Decay Rate (Rank)
91	3.9686	1.0046(1)	2.658(1)	0.5767(1)
80	0.6308	0.3516(2)	2.0502(2)	0.0612(2)
94	0.5897	0.3151(3)	1.8642(3)	0.0517(3)
82	0.4525	0.1835(5)	1.3975(4)	0.0205(5)
10	0.3726	0.2079(4)	0.72(10)	0.0236(4)
32	0.3201	0.1692(6)	0.8115(9)	0.0155(6)
83	0.2456	0.1333(7)	1.1924(5)	0.0094(7)
69	0.174	0.1021(8)	1.1904(6)	0.0054(8)
12	0.1662	0.1013(9)	0.8887(8)	0.0053(9)
61	0.1222	0.0821(10)	1.0122(7)	0.0034(10)
108	0.1052	0.0636(13)	0.2623(17)	0.0022(12)
67	0.0997	0.064(12)	0.4481(14)	0.0021(13)
54	0.0973	0.0599(14)	0.5354(13)	0.0018(14)
14	0.0934	0.0669(11)	0.2502(18)	0.0024(11)
30	0.0892	0.0545(15)	0.3967(15)	0.0015(15)
33	0.0848	0.0536(16)	0.7015(11)	0.0015(16)
52	0.0828	0.0438(19)	0.1158(21)	0.0011(19)
26	0.0792	0.0518(17)	0.3946(16)	0.0014(17)
23	0.0775	0.0504(18)	0.5816(12)	0.0013(18)
71	0.05	0.031(20)	0.1611(19)	0.0005(20)

Under the null hypotheses that the distribution of alpha is symmetric, we would expect 2.4 managers to have a positive (negative) skew at the 5% significance level based upon this sample size. However, out of the 97 managers, there were 15 that have 95% confidence intervals on the coefficient of skewness that do not include zero. Six managers exhibited a positive skew, whilst nine exhibited a negatively skewed distribution of alpha. Given the empirical evidence for non-symmetric distributions of excess returns, the use of performance measures which account for skew is supported.

5. Conclusions

This paper investigated a number of techniques that can be used to measure fund manager performance under the assumption of a clearly defined benchmark where simulations were used to ascertain the power and bias of some of these performance measures. These techniques were then used to rank funds from the Mercers Survey data where the results suggest some variation in rank, depending on the choice of performance methodology adopted.

In the simulation study, the results suggest little difference in power between measures under the assumption of normal or symmetric return distributions. However, the Sortino Ratio achieves higher power if the distribution of one of the funds is positively skewed. Furthermore, the relative bias of the Sortino Ratio is shown to be slightly lower than the SSR when the distributions of alpha are skewed. Given that the Sortino Ratio provides a higher power in its ability to chose the optimal fund when the distribution of alphas is skewed and similar power when the distribution of alpha is symmetric, overall the Sortino Ratio has been shown to be a better performance measure due to its ability to choose the optimal fund. This preference to the Sortino Ratio is based on the assumption that the appropriate measure of risk is associated with performance under benchmark.

(Date of receipt of final transcript: August 14, 2007.
Accepted by Phil Dolan & Tom Smith, Special Issue Editors.)

References

Bawa, V.S. & Lindenberg, E.B. 1977, ‘Capital market equilibrium in a mean-lower partial moment framework’, *Journal of Financial Economics*, vol. 5, pp. 189–200.

Chaudhry, A. & Johnson, H.L. 2006,’Power and bias of benchmarked performance measures’, working paper, Queensland University of Technology.

Foster, F.D. & Stutzer, M. 2003, ‘Performance and risk aversion of funds with benchmarks: A large deviations approach’, University of Otago Department of Finance Seminar Series, Department of Finance, University of Otago, Dunedin, New Zealand

Lo, A.W. 2002, ‘The statistics of sharpe ratios’, *Financial Analyst Journal*, vol. 58, no. 4, pp. 36–52.

Miller, R.E. & Gehr, A.D. 1978, ‘Sample size bias and sharpe’s performance measure: A note’, *Journal of Financial and Quantitative Analysis*, vo1. 3, no. 5, pp. 943–6.

Sharpe, W.F. 1994, ‘The Sharpe ratio’, *Journal of Portfolio Management*, vol. 21, pp. 49–59.

Sortino, F.A. & Price, L.N. 1994, ‘Performance measurement in a downside risk framework’, *Journal of Investing*, vol. 3, pp. 50–8.

Appendix

Recall that the mixture distribution that was used in the simulation study was created using a normal distribution and a log normal distribution since it was required that the mixture distribution needed to be unbounded. In order to carry out the simulation study (whereby the mean of the mixture distribution was increased in increments to compare it to both a normal distribution and another mixture distribution to allow comparison of various performance measures), it is necessary to describe how the appropriate parameters are chosen.

A weight of 0.2 was allocated to a normal distribution with some pre-specified mean and standard deviation and therefore a weight of 0.8 was given to a relocated and rescaled lognormal distribution with some mean value and standard deviation.

Firstly, the expected value and variance for a lognormal distribution, *X*, are given below

$$E(X) = e^{\mu + \frac{\sigma^2}{2}}$$

$$Var(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1).$$

The choice of parameters for μ and σ^2 are vital to obtain the required distribution in terms of shape. An empirical investigation suggests values of μ=0 and σ=0.5 provide the required positively skewed distribution. Thus

$$E(X) = e^{\mu + \frac{\sigma^2}{2}} = e\frac{0.5^2}{2}$$

$$Var(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) = e^{0.5^2}(e^{0.5^2} - 1)$$

Unfortunately, whilst this distribution has the correct shape, the mean and standard deviation are incorrect since the distribution is bounded at 0 (where negative alphas must be allowed). To alleviate this problem, the lognormal is relocated and rescaled so that it has, say, a mean of 0 and a standard deviation of σ_{LN}. Therefore

$$E\left(\frac{X - e^{\mu + \frac{\sigma^2}{2}}}{\sqrt{e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)}} \times \sigma_{LN}\right) = 0$$

$$Var\left(\frac{X - e^{\mu + \frac{\sigma^2}{2}}}{\sqrt{e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)}} \times \sigma_{LN}\right) = \sigma_{LN}^2$$

The mean alpha is then relocated to some value, μ_{LN} per month. Initially this was set to be 0.1% in line with the previous simulations, but this mean will be increased in increments of 0.05% in due course. The calculation for the relocation of the mean is constructed as

$$E\left(\frac{X - e^{\mu + \frac{\sigma^2}{2}}}{\sqrt{e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)}} \times \sigma_{LN} + \mu_{LN}\right) = \mu_{LN}$$

$$Var\left(\frac{X - e^{\mu + \frac{\sigma^2}{2}}}{\sqrt{e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)}} \times \sigma_{LN} + \mu_{LN}\right) = \sigma_{LN}^2.$$

In general, the mean and variance of a mixture distribution derived from two independent distributions can be estimated mathematically by

$$\mu = w_1\mu_1 + (1-w_1)\mu_2$$
$$\sigma^2 = w_1\sigma_1^2 + (1-w_1)\sigma_2^2 + w_1(1-w_1)(\mu_1-\mu_2)^2$$

where μ_1 and μ_2 are the means of the two distributions respectively and σ_1^2 and σ_2^2 are the two variances respectively. Both distributions used to create the mixture model will have identical means (i.e. $\mu_1 = \mu_2$), therefore $w_1(1-w_1)(\mu_1-\mu_2)^2 = 0$.

Therefore, a mixture distribution with a mean μ_M and standard deviation 0.015 (in line with the other examples in the simulation study) needs to be obtained. If we suppose that

$$X_1 \sim N(\mu_N, \sigma_N^2 = 0.0104^2), X_2 \sim LN(\mu_{LN}, \sigma_{LN}^2), w_1 = 0.2$$

and $w_2 = 0.8$ then

$$\mu_M = w_1\mu_N + (1-w_1)\mu_{LN} = 0.2\mu_N + 0.8\mu_{LN},$$

where μ_N and μ_{LN} are known and μ_{LN} is the mean of the lognormal distribution. Furthermore,

$$\sigma_M^2 = w_1\sigma_N^2 + (1-w_1)\sigma_{LN}^2 + w_1(1-w_1)(\mu_N-\mu_{LN})^2 = 0.2^2 \times 0.0104^2 + 0.8^2 \times \sigma_{LN}^2$$

and so,

$$\sigma_{LN}^2 = \sqrt{\frac{0.0104^2 - 0.2^2 \times 0.0104^2}{0.8^2}}$$

This process is then repeated to obtain the required mixture distribution where the mean is varied in small increments.