Does the use of downside risk-adjusted measures impact the performance of UK investment trusts?

Chris Adcock (*) c.j.adcock@sheffield.ac.uk
Nelson Areal (**) nareal@eeg.uminho.pt
Manuel Armada (**) rarmada@eeg.uminho.pt
Maria Ceu Cortez (**) mccortez@eeg.uminho.pt
Benilde Oliveira (**) (***) benilde@eeg.uminho.pt
Florinda Silva (**) fsilva@eeg.uminho.pt

This Draft: October 2010

Abstract

This paper investigates the impact of using different risk-adjusted measures of performance on the evaluation of UK investment trusts. Performance results based on the Sharpe Ratio and several downside risk-adjusted performance measures are compared. Four alternative measures, both parametric and non-parametric, are applied in order to assess the level of association between different performance measures. The observed level of association between Sharpe Ratio and the downside risk-adjusted measures of performance is not as low as would be expected considering the empirical properties of our data. However, an additional analysis based on simulated returns, with a higher variability in the skewness and the kurtosis of the time series, demonstrates that the choice of the performance measure does indeed have an impact on the performance assessment of portfolios. For the simulated time series of returns considerable lower levels of association are reported especially with the application of Cohen´s Kappa statistic.

(***) Corresponding author

(*) University of Sheffield
Management School
Mappin Street
Sheffield, S1 4DT
United Kingdom
Tel: +44 (0)114 - 222 3402
Fax: +44 (0)114 - 222 3348

(**) NEGE – Management Research Unit
University of Minho
Campus de Gualtar, 4710-057 Braga, Portugal
Tel: +351 253 604554
Fax: +351 253 601380

Does the use of downside risk-adjusted measures impact the performance of UK investment trusts?

This version: October 2010

Abstract

This paper investigates the impact of using different risk-adjusted measures of performance on the evaluation of UK investment trusts. Performance results based on the Sharpe Ratio and several downside risk-adjusted performance measures are compared. Four alternative measures, both parametric and non-parametric, are applied in order to assess the level of association between different performance measures. The observed level of association between Sharpe Ratio and the downside risk-adjusted measures of performance is not as low as would be expected considering the empirical properties of our data. However, an additional analysis based on simulated returns, with a higher variability in the skewness and the kurtosis of the time series, demonstrates that the choice of the performance measure does indeed have an impact on the performance assessment of portfolios. For the simulated time series of returns considerable lower levels of association are reported especially with the application of Cohen´s Kappa statistic.

EFM classification: 380; 530.

1. Introduction

A critical issue in the area of portfolio performance evaluation is the choice of the appropriate measure used to quantify risk. Different risk measures have been developed and, consequently, different risk-adjusted performance metrics can be applied to the evaluation of investment portfolios. Besides well known risk-adjusted measures, downside risk-adjusted measures of performance can also be used to evaluate investment portfolios. This raises the issue of whether or not the performance of investment funds is sensitive to the performance measure used.

The mean-variance paradigm of Markowitz (1952) is the theoretical basis for the development of the traditional framework established to assess risk-adjusted portfolio performance. The Sharpe (1966) ratio (SR) is probably the most important representative of the set of classic risk-adjusted measures of performance.

The limitations of mean-variance based models are a consequence of one of its major assumptions: that returns are normally distributed or, in a more general way, elliptically symmetrically distributed. If returns do not follow an elliptically symmetric distribution, the use of the mean-variance approach is questionable. Therefore, the assessment of performance in the context of skewed distributions requires the use of non-standard measures to adjust for risk. The development of a downside risk measurement framework was the natural answer to the problems associated with the use of standard measures of risk when distributions are skewed. Generally speaking, downside risk measures focus on the returns that fall below a certain value, enabling us to look at the tails of the distribution. Despite its obvious potential, there is still an ongoing debate in the literature on the usefulness of downside risk measures for performance evaluation purposes.

In the context of non-elliptical symmetry in returns, one might expect a lack of correlation between performance ranks based on traditional and downside measures of risk-adjusted performance. However, recent research in this area provides evidence suggesting no impact of the choice of a particular performance measure on the rankings between alternative investments.

Pfingsten et al. (2004) compared rank correlations for several risk measures on the basis of an investment bank´s 1999 trading book and concluded that different measures result in largely identical rankings. Eling and Schuhmacher (2005) also concluded for significant rank correlation between different performance measures. Their study focuses on hedge fund indices data from 1994 to 2003. Finally, Eling and Schuhmacher (2007) analysed individual hedge fund data instead of indices, but the results were the same in the sense that identical rank ordering across hedge funds was clearly identified. These authors conclude that despite significant deviations of hedge fund returns from a normal distribution, the first two moments seem to describe the distributions sufficiently well. Similar results were also obtained by Eling (2008) for a large sample including different types of funds (mutual funds, real estate funds, hedge funds, CTAs and CPOs).

Contrasting with these findings, Ornelas et al. (2009) and Zakamouline (2010) argue that the choice of the performance measure does influence the evaluation of investment funds. These authors apply and test the equality of ranks based on several performance measures and both identify significant differences for some of the measures applied. Zakamouline (2010) emphasizes the importance of the absolute value of the skewness on the choice of the performance measure by demonstrating that the rank correlation between the Sharpe ratio and other measures of performance decreases as the absolute values of skewness becomes higher. He also remarks that, even in the presence of non-normally distributed returns, the use of alternative performance measures can produce identical ranks. As the author explains, if deviations from normality are of the same kind across all the return series under analysis, a significant difference between performance ranks should not be expected.

The main objective of this paper is to contribute to this debate by providing evidence on the impact of using different risk measures to adjust performance. For that purpose, different performance measures will be computed on a data set of UK investment trusts. The contributions of this study in relation to the previous studies in this area are fourfold. As far as we are aware of, this is the first study to investigate the impact of the use of standard and downside risk-adjusted measures of performance on the evaluation of UK investment trusts.

Secondly, we provide a discussion on downside risk measures that can be used to evaluate portfolio performance and suggest the use of expected shortfall (ES), given its desirable properties as a measure of risk. There have been very few studies on the performance of investment portfolios using ES.

Additionally, the value-at-risk (VaR) and ES measures of risk will be computed using a full valuation method based on a filtered historical simulation (FHS) procedure. As far as we know, performance ratios based on FHS estimates of VaR and ES have never been empirically tested.

Finally, while most of previous studies rely on the application of the Spearman rank correlation test, we innovate by using several measures of correlation, both parametric and non-parametric, to assess the level of association between the performance results. In order to classify the strength of association reported by the estimated measure of association, statistical inference based on confidence intervals around the observed value is performed. To assess the level of association between different performance measures, four alternative measures are applied. The Pearson correlation coefficient is applied in the context of the parametric approach. Additionally, two alternative non-parametric measures of rank correlation are used: Spearman´s coefficient and Kendall´s Tau. As an alternative to these rank correlation analysis, Cohen´s Kappa, a non-parametric measure based on contingency table statistics, is also computed.

This paper is organized as follows. In Section 2 we discuss the literature on downside risk measures and its use in portfolio performance evaluation. In Section 3 we describe the methodology used to assess fund performance and estimate risk measures. Section 4 focuses on the data and on the empirical properties of the UK investment trust return series. Section 5 provides and discusses the empirical results. Section 6 presents the analysis based on simulated return series. Finally, section 7 summarises the main results and presents some concluding remarks.

2. Literature Review

Based on the mean-variance approach, the well known traditional measures of performance (Jensen, 1968; Sharpe, 1966; and Treynor, 1965) use either the standard deviation or beta as risk measures. These measures are quite similar and, under certain market conditions, produce rankings of portfolios that are not significantly different from each other. The use of traditional measures of performance is fully justified under the assumption of elliptically symmetric distributed returns and hold for any well-behaved utility function (see Landsman and Nešlehová, 2008)

However, the SR can lead to misleading conclusions when returns are significantly skewed (Bernardo and Ledoit, 2000). It is a well known fact that investors have a preference for positive skewness, which means that upside risk is less important to investors than downside risk. If only the first and second moments are assessed, the negatively-skewed returns will generate the appearance of outperformance. Hence, asymmetric distributions require alternative evaluation approaches that integrate higher moments of the distribution beyond the first and second moments.

When dealing with asymmetric empirical return distributions, downside measures of risk provide an alternative framework to assess risk adjusted performance. The issue is whether performance evaluation results based on these measures are different from those obtained in the context of the mean-variance paradigm.

Pedersen and Rudholm-Alfvin (2003), using equity return data from global financial services institutions and UK micro-firms, compare a SR-based ranking with other alternatives that assess performance by using downside risk-adjusted measures. Applying Jarque-Bera tests and rank correlation analysis, they concluded that, for symmetric distributions of returns, there is a high correlation (80% in some cases) between the rankings that result from the application of different measures of performance. On the other hand, when the distribution of returns is asymmetric, there is a significant absence of consistency in the measurement of performance, with the rank correlations dropping below 5% in some cases. These results support the use of alternative performance evaluation framework when a departure from elliptical symmetry in the series of returns is identified. Later, Eling and Schuhmacher (2005, 2007) compared the SR with other downside risk-adjusted measures of performance. Their results, based on the Spearman rank correlation coefficient, indicate that the choice of the performance measure does not affect the rankings of hedge funds. These findings are somewhat surprising, considering that hedge fund returns differ significantly from a normal distribution. Eling (2008) obtain the same type of evidence not only for hedge funds but also for mutual funds.

More recently, Zakamouline (2010) recognizing that the conclusions from Eling and Schuhmacher (2007) and Eling (2008) are rather puzzling, gives further explanations for these findings. He argues that in order to produce clear differences in rankings, investment funds must exhibit distinctly different return probability distributions. Even when returns are non-normally distributed, the differences in performance might not be significant if the probability distribution exhibits the same kind of deviation from normality. Furthermore, Zakamouline (2010) demonstrates that the rank correlation of other measures of performance with the SR decreases as funds present higher absolute values of skewness.

Critics of volatility have proposed alternative approaches to assess risk when basic underlying assumptions, like the symmetry of the distribution of the returns, do not hold. Downside risk measures are a way to overcome these limitations. In this paper, downside risk-adjusted performance measures include lower partial moments (LPM) measures of performance, such as the Sortino ratio (Sortino and Price, 1994) as well as performance measures using value-at-risk (VaR) and expected shortfall (ES).

The Sortino ratio, developed by Sortino and Price (1994) was, at first, largely criticised for not being derived in the context of a market equilibrium theory.[1] However, Pedersen and Satchell (2002) added further motivation for the use of the Sortino ratio in a modified version using the risk return as the target return. As Pedersen and Satchell (2002) remark, this modified Sortino ratio is equivalent to the Sharpe ratio, except that standard deviation has been replaced by the semi-standard deviation in the denominator. The authors support the use of the modified Sortino ratio by placing it on a relevant theoretical foundation and by discussing its relative qualities in contrast with alternative approaches to assess performance. Moreover, Pedersen and Satchell (2002) have shown that there is a utility-based one-period CAPM that promotes the use of the Sortino ratio, just like the traditional CAPM promotes the use of SR.

Throughout the years, VaR has become a standard measure of risk and has been receiving increasing attention by academics. Alexander and Baptista (2003) and, more recently, Eling and Schuhmacher (2007) recommend the use of a VaR-based measure of performance that is closely related to the SR. In the particular context of hedge funds, Eling and Schuhmacher (2007) assess performance on the basis of VaR by applying a performance ratio that explicitly measures the excess return on VaR (ERVaR).

Under the assumption of elliptically symmetric distributions (with a mean equal or very close to zero) the proposed VaR-based measure of performance produces the same portfolio rankings as the SR. In the specific context of non-elliptically symmetric distributed returns, the portfolio with the highest ERVaR ratio may not be the portfolio with the highest SR.

VaR presents several shortcomings. The most obvious shortcoming with respect to VaR is its threshold character. Basak and Shapiro (2001) show that VaR investors often optimally choose a

---

[1] Leland (1999), for instance, states that the Sortino ratio is an *ad hoc* attempt to recognize the greater importance of downside risk.

larger exposure to risky assets than non-VaR investors, and consequently incur larger losses, when they occur. This is a direct consequence of the inability of VaR to penalize a potentially very large loss more than a large loss. This is particularly critical when dealing with non-normal heavy tailed distributions. In this context, the probability of a large loss is non-negligible and a risk measure that truly penalizes large losses is needed. By disregarding the loss beyond the quantile of the underlying distribution, VaR disregards the risk of extreme losses in the tail of the underlying distribution.

An alternative measure to VaR is expected shortfall (ES). Unlike VaR, ES has the desirable property of focusing on the size of the loss and not just on the frequency of losses.

The origins of ES can be traced down to Artzner et al. (1999), who have formalized an alternative measure of risk that they called tail conditional expectation. Bertsimas et al. (2004), define ES as being the average of VaRs for all levels below α. Consider a sample of T returns. In order to estimate ES the sample of returns must be sorted in increasing order: $R_1 \leq R_2 \leq \cdots \leq R_T$. The ES natural estimator is the average of the first α%, represented by the first k outcomes (where $k = [\alpha T]$),

$$ES_\alpha = -\frac{1}{k}\sum_{i=1}^{k} R_{(i)} \,. \qquad (1)$$

The above equation represents a natural non-parametric estimator of ES. It does not rely on any kind of distributional assumptions with respect to portfolio returns and therefore this definition of ES is valid for general distributions.

Artzner et al. (1999) defined an axiomatic methodology to characterize desirable properties for risk measures and they named risk measures that satisfied their axioms as coherent. ES is generally proposed in the literature as a coherent measure of risk in the sense of Artzner et al. (1999)

and a superior alternative to the industry standard VaR (Yamai and Yoshiba, 2002)[2]. Despite all the positive features of ES as a measure of risk, it has not been fully exploited for performance evaluation purposes.

3. Methodology

3.1. Performance ratios

Different measures of risk-adjusted performance are considered for empirical computation, namely the Sharpe ratio (SR) and three downside risk-adjusted measures of performance: the Sortino ratio, Excess return on VaR and Excess return on ES. To compute the risk-adjusted performance ratios, four different risk measures are used: standard deviation, lower partial moments (LPM), VaR and ES.

For each portfolio p under evaluation, based on estimated values from a sample of daily returns $R_{p_1}, \dots, R_{p_T}$ the SR (Sharpe, 1966) is computed as:

$$SR_p = \frac{\bar{R}_p - \bar{R}_f}{\sqrt{S_p^2}}, \tag{2}$$

where $\bar{R}_p$ is the average daily return estimated for the portfolio over the sample period of analysis, $\bar{R}_f$ is the average risk-free rate of return and $\sqrt{S_p^2}$ is the estimated standard deviation of the sample of portfolio daily returns.

The version of the Sortino ratio that will be used for empirical application purposes is the one suggested by Pedersen and Satchell (2002), which represents the ratio between excess return and semi-standard deviation, as follows:

---

[2] VaR is not a coherent risk measure because, for non elliptically symmetrically distributed returns, it clearly violates the axiom of sub-additivity presented by Artzner et al. (1999).

$$Sortino_p = \frac{\overline{R}_p - \bar{R}_f}{SSD_p}, \tag{3}$$

where SSD_p represents the usual definition of the semi-standard deviation estimated for the sample of portfolio daily returns.

The excess return on VaR ratio (ERVaR) is a performance ratio quite similar to the SR. The main innovation is that VaR takes the place of standard deviation as the selected measure of risk used to adjust the excess return of the portfolio. The ERVaR is the ratio between excess return and VaR:

$$ERVaR_p = \frac{\overline{R}_p - \bar{R}_f}{\widehat{VaR}_{\alpha,p}}, \tag{4}$$

where $\widehat{VaR}_{\alpha,p}$ is the estimated value-at-risk (based on a sample of returns) of the risky portfolio p with the probability α.

Alternatively to the use of VaR as a risk measure to adjust the excess returns of the portfolio, the ES can be applied. The excess return on ES (ERES) corresponds to the ratio between excess return and ES.

$$ERES_p = \frac{\overline{R}_p - \bar{R}_f}{\widehat{ES}_{\alpha,p}}, \tag{5}$$

where $\widehat{ES}_{\alpha,p}$ is the estimated (from a sample of returns) expected shortfall of the risky portfolio p , if the portfolio returns drop below its α quantile.

The ratio used in this paper represents a more general version of the stable tail adjusted return ratio (STARR) introduced by Martin et al. (2003). In our paper, excess returns on ES are defined in the context of general loss distributions and the estimation of ES is therefore not restricted to the assumption of stable returns distributions. Apart from this, the definition of STARR and the excess return on ES coincide. The empirical characteristics of the return series under investigation do not recommend the use of an ES estimator under a parametric framework. Therefore, the non-parametric estimator of the ES proposed by Bertsimas et al. (2004) is used, as it does not rely on any restrictive assumption about the functional form of the empirical distribution of returns.

3.2. The computation of VaR and ES

While VaR is conceptually simple and flexible, for VaR figures to be useful they also need to be reasonably accurate. VaR is just an estimate, and its usefulness is directly dependent on its precision.

Besides skewness, financial time series often exhibit other types of distortions from normality, like fat tails. Focusing on the non-normality properties of the financial time series, some analytical non-normal approaches to compute VaR need to be adopted. A commonly used approach to account for the non-normality in the return distribution consists on the use of the Cornish-Fisher expansion (Cornish and Fisher, 1937) to estimate VaR.[3] This approach makes use of certain additional parameters to allow for the fat tails or skewness and it was first introduced by Zangari (1996) to estimate parametric VaR of portfolios that include options. The Cornish-Fisher VaR (CFVaR), also known as Modified VaR (MVaR), is an estimator for VaR that estimates the true, unknown quantile function by its second order Cornish-Fisher expansion around the Normal quantile function. The skewness and the kurtosis of the empirical distribution of the portfolio returns are incorporated in the estimation of the parametric VaR approximated by the application of the Cornish-Fisher expansion (Zangari, 1996) as follows:

$$CFVaR_{\alpha-\nu} = -(\mu + Z_{\alpha-\nu}\sigma), \tag{6}$$

[3] Favre and Galeano (2002), Amenc et al. (2003) and Gueyié and Amvella (2006) are some of the authors that used CFVaR in their studies.

where μ is the mean, σ is the standard deviation of the portfolio returns and $Z_{\alpha-\nu}$ is the $(\alpha - \nu)$quantile of the standard normal distribution and ν is the adjustment provided by the Cornish-Fisher expansion,

$$\nu \approx \frac{1}{6}(Z_\alpha^2 - 1)SK + \frac{1}{24}(Z_\alpha^3 - 3Z_\alpha)E - \frac{1}{36}(2Z_\alpha^3 - 5Z_\alpha)SK^2, \tag{7}$$

where SK is the skewness and E is the excess kurtosis of the series of portfolio returns.

The Cornish-Fisher expansion presented above in equation (6) might not be reliable as an approximation of certain distributions that depart significantly from normality. In fact, in some situations the use of fully parametric methods to compute VaR might be inadequate. In such a case, an alternative method is required to estimate the true empirical portfolio return distribution in order to compute VaR.

More recently, a new methodology has been developed in the literature to compute both VaR and ES. This new method successfully combines bootstrapping techniques with the use of parametric models and is generally known as Filtered Historical Simulation (FSH). FHS was first proposed by Barone-Adesi et al. (1999). Under FHS the bootstrap process is applied to the residuals of a time series model (usually a GARCH-type model) used as a filter to extract autocorrelation and heteroscedasticity from the historical time series of returns. Despite being numerically intensive, FHS is quite simple to apply and as a result it is faster to implement than several more sophisticated methods. According to Hartz et al. (2006) FHS is also numerically extremely reliable.

By being free of any distributional assumptions in relation to the behaviour of the returns, FHS has a great flexibility in capturing all the empirical properties of the time series of returns under analysis, including skewness and kurtosis. Based only on the assumption of uncorrelated standardized residuals from an appropriate time series model, the use of the bootstrap resampling algorithm allows a computationally simple and feasible method to approximate the unknown return empirical distribution. Under FHS a bootstrap sample, generally denoted by R^*, of any size M, is generated. Based on this bootstrap sample, the filtered historical simulated VaR and ES estimates can be easily obtained. VaR, under FHS, corresponds to the α quantile of the bootstrap sample generated under FHS,

$$VaR_{FHS,\alpha} = -q_{\alpha,R^*}. \quad (8)$$

To obtain ES under FHS the bootstrapped returns must be ranked in increasing order: $R_1^* \leq R_2^* \leq \cdots \leq R_M^*$. The ES is the average of the first $\alpha\%$ returns in the sample of bootstrapped returns, represented by the first $K = [\alpha M]$ outcomes with M being the size of the bootstrap sample,

$$ES_{FHS,\alpha} = -\frac{1}{K}\sum_{i=1}^{K} R_{(i)}^*. \quad (9)$$

In this paper, the performance of UK investment trusts will be assessed using these non-parametric VaR and ES estimators. By combining a non-parametric bootstrap procedure with a parametric modeling of the time series under analysis we are able to considerably improve the quality of the VaR and ES estimates.[4]

4. Data

A sample of 109 UK investment trusts, alive on the 31st of March of 2006, is analysed. Investment trusts classified by the Association of Investment Trust Companies (AITC) as split capital trusts were excluded because of their dual nature.[5] Given that the VaR and ES are estimated using FHS, a ten year series was defined as the minimum acceptable period of analysis. Daily

---

[4] Baroni-Adesi et al. (1999), Pritsker (2001) and later Kuester et al. (2005), compare the performance of FHS with other parametric and non-parametric methodologies in the estimatation of VaR and conclude in favour of the superiority of FHS. Additionally, Giannopoulos and Turanu (2005) and also Harmantzis et al. (2006) demonstrate how FHS can provide an improved methodology to compute ES. They argue that while resampling methods (such as FHS) are, in general, numerically intensive, they are quite simple to implement. Hence, it is faster to estimate VaR and ES with FHS than with other more sophisticated models. In addition, FHS is also numerically reliable (Hartz et al.,2006).

[5] In this specific kind of trusts, capital and income shares are common.

continuously compounded returns for each fund, over the period March 1996 to March 2006, were computed using price information available on DataStream.

As one of the objectives of this research is also to analyse the position of a particular investment trust in relation to the market, the FTSE ALL-SHARE INDEX is used to proxy the market for benchmarking purposes. As a proxy for the risk free rate, the UK Interbank Overnight-Offered rate is used. Daily data for the benchmark and for the risk-free rate was also collected from DataStream.

A preliminary analysis on the empirical properties of the time series was performed.[6] This analysis was carried out to allow some conclusions on the distributional patterns of the data. For all trusts, the skewness is different from zero, which constitutes a first indication that all the series are non-symmetric. Only 7 trusts (approximately 6,4%) exhibit positive skewness. For the great majority of the trusts (approximately 93,6%), the value of the skewness is negative. In addition, the value of the kurtosis is always greater than 3. For all trusts, the application of the Jarque-Bera test[7] clearly rejects, for a confidence level of 99%, the hypothesis that the return data series are normally distributed. To further investigate the source of non-normality, the Jarque-Bera test was split into its skewness and kurtosis components and the components were tested separately. Only 10 trusts exhibit skewness values that are not statistically different from zero. For the remaining trusts, skewness is statistically different from zero, indicating that the empirical series of returns are in general significantly (negative) skewed.

In addition, the hypotheses of independent returns and homoscedasticity were also tested. Though independence has broader implications, the focus of this paper is on whether or not returns are uncorrelated. The Box-Pierce Q-Statistic was computed, as well as Engle´s LM test. The results for the Box-Pierce statistics indicate the presence of serial correlation for a 95% level of confidence.

[6] Although not presented here, the table with these figures is available from the authors upon request.

[7] The results for the normality test as well as for autocorrelation and heteroscedasticity are also available upon request.

When applied to the squared returns, the Box-Pierce Q-Statistic can provide evidence for the existence of heteroscedasticity. The null hypothesis of homoscedasticity was rejected, with no exception. Finally, the application of Engle´s LM test to the squared returns gave further support to the existence of ARCH effects. With respect to all but one fund, the test detected significant autocorrelations for the 2^{nd}, 5^{th} and 10^{th} lag, at a 95% confidence level.

5. Empirical Results

5.1. The selection of the appropriate volatility model to estimate VaR and ES

The first step to the implementation of the FHS methodology (to estimate both VaR and ES) consists of a meticulous process of model fitting to select an appropriate volatility model. To ensure the validity of FHS, the model selected to be used as a filter must necessarily pass the residual diagnostic tests for serial correlation and heteroscedasticity. Highly skewed and leptokurtic non-normal daily returns and a strong phenomenon of autocorrelation and heteroscedasticity are the main empirical properties of the data. Therefore, it is crucial to adopt a volatility model that properly deals with these empirical characteristics. For this purpose several GARCH-type models were estimated.[8] Letting m = 0,1,2,3 and n = 0,1,2,3, a set of ARMA(m,n) models, combined with alternative GARCH-type specifications for the variance equation were estimated with respect to each investment trust.[9] Alternative GARCH-type specifications were estimated and tested: GARCH (1,1); GARCH (1,1)-in-mean; GJR (1,1) and GJR (1,1)-in-mean.

From the total set of 64 estimated models for each time series under analysis, the best fitted model was selected. The first step for the selection of the best model for each fund is to perform a diagnostic check on the residuals for remaining serial correlation effects and also remaining

---

[8] The GARCH-type models were estimated using the maximum likelihood (ML) approach according to the quasi-Newton method of Broyden, Fletcher, Goldfarb and Shanno. The results were obtained using Ox version 4.10 (see Doornik, 2007).

[9] For the estimation of the models the non-negativity and stationary conditions defined in the literature were considered.

heteroscedasticity[10]. For each fund, only the models that passed simultaneously all the diagnostic checks (at a 95% confidence level[11]) were considered robust candidates to be selected as the best model. Next, considering the set of robust models for each fund, the Akaike Information Criterion (AIC) and Schwarz Information Criterion (SIC) were used in the comparison of the models and the final model selection was done by choosing the model which generates the lowest AIC and/or SIC[12].

For the great majority of the trusts under analysis, the fitted best model has an asymmetric specification for the variance equation. Table 1 presents these results.

[Insert Table 1 here]

5.2. The performance of UK investment trusts

The results for the different risk-adjusted measures of performance are presented in table 2: Sharpe ratio (SR), Sortino ratio, excess return on value-at-risk (ERVaR) and excess return on expected shortfall (ERES). Two alternative versions of the excess return on VaR are computed. Firstly, the VaR is estimated using the Cornish-Fisher approximation (CFVaR). Secondly, the VaR (as well as ES) is estimated using the FHS method[13].

---

[10] The well known Box-Pierce test was applied to the standardized residuals in order to investigate the presence of remaining serial correlation effects. When applied to the squared standardized residuals, the Box-Pierce test can give us a first indication of the correct specification of the variance equation by not rejecting the null hypothesis of homoscedasticity. Additionally, the presence of remaining ARCH effects was investigated by the application of the Engles´s LM test and the Residual Based Diagnostic (RBD) of Tse (2002).

[11] By using a confidence level of 95% instead of 99% we are being more conservative on the selection of the best model.

[12] For those cases in which the AIC and the SIC do not select the same model, a log-likelihood ratio test (LRT) was applied to assess the statistical significance of the difference. In the case that the models selected by the two different adopted selection criteria are significantly different, according to the results provided by the application of the LRT, the model selected by the AIC was used as it is considered to be superior in terms of fit improvement. Otherwise, the simpler model was chosen and parsimony is ensured.

[13] To obtain the 99% confidence level VaR_{FHS} and ES_{FHS}, estimates of 100000 pseudo returns were generated through the implementation of a bootstrap (with replacement) procedure based on the residuals of the GARCH-type model previously fitted for each of the investment trusts considered.

[Insert Table 2 here]

Among the five performance ratios, excess return on CFVaR presents the lowest mean and standard deviation. The mean and standard deviation are considerably higher for the Sortino ratio in comparison to any of the other performance ratios considered. Performance ratios based on CFVaR, VaR and ES exhibit a lower standard deviation (as well as a lower mean) in relation to the SR. The magnitudes of the standard deviation differ considerably between the Sortino Ratio and the other metrics.

Performance rankings are also constructed for each of these measures. Table 3 presents these rankings.

[Insert Table 3 here]

Considering our sample of trusts, the position of the benchmark varies from 68 (according to Excess Return on CFVaR) to 73 (according to Sharpe Ratio). This means that by changing the performance measure used from the Excess Return on CFVaR to the Sharpe Ratio there are 5 trusts that change from being underperformers to beating the market. It should be noticed that only Excess Return on VaR and Excess Return on ES generate the same rank order for the benchmark

As reported in table 4, the percentage of funds with the same rank order according to the use of two different measures of performance is 10% on average.

[Insert Table 4 here]

This ranges from 5% (Sortino *versus* Excess Return on ES and Excess Return on CFVaR *versus* Excess Return on ES) to 30% (Excess Return on VaR *versus* Excess Return on ES). Not surprisingly, the pair of performance measures Excess Return on VaR *versus* Excess Return on ES exhibits systematically and by far the highest percentage of equality in ranks.

With respect to each performance ratio the top five and bottom five rankings were also analysed. The number of changes in the top/bottom five rankings was tracked across the different pairs of performance measures. Table 5 summarises the results for this analysis.

[Insert Table 5 here]

On average, 42% of the trusts change in the top five ranking based on the use of alternative performance measures. The bottom five ranking is more stable across the different pairs of performance measures. In fact, the percentage of trusts that change in the top five is always higher than the percentage of trusts that change in the bottom five ranking.

In sum, this preliminary analysis of performance ranks based on alternative measures shows some rank changes that may not be neglected by investors.

Next, several statistical tests, will be applied in order to assess the level of association between different performance measures.

5.3. Association between different performance measures

To assess the level of association between the different performance measures and performance rankings, both parametric and non-parametric approaches are used. A total of four alternative measures of correlation are applied. We aim to investigate not the lack or perfect association but the intensity of the association between the measures of performance. Therefore, statistical inference will be performed on the basis of confidence intervals around the observed value.

For each measure of correlation, generally denoted by ρ, the theoretical values for perfect association and lack of association are defined respectively as $\overline{\rho} = 1$ and $\underline{\rho} = 0$. The observed value of the association, $\hat{\rho}$, has no constraint with respect to $\underline{\rho}$, but must satisfy $\hat{\rho} \leq \overline{\rho}$. Additionally consider different threshold levels $\theta_i, i = 1, \dots, k$ ranked by their intensity of association: $\underline{\rho} < \theta_1 < \theta_2 < \cdots < \theta_k < \overline{\rho}$. A one-sided confidence interval below $\hat{\rho}$, with a confidence level of ψ, can be computed and denoted by L_ψ. The hypothesis that ρ exhibits at least the intensity of association corresponding to θ_i if $\theta_i \geq L_\psi$, cannot be rejected if $Pr\left[\rho \leq L_\psi\right] = \psi$. Obviously, the perfect association and the lack of association hypotheses will not be rejected if $\underline{\rho} \geq L_\psi$ and $\overline{\rho} \leq U_\psi$, with $U_\psi$denoting the one-sided upper interval.

Under the parametric approach and in order to measure the level of correlation between alternative performance ratios, Pearson's correlation coefficient is applied. Under the non-parametric approach we first consider the rankings produced by the performance measures and test their association using two different rank correlation coefficients: Spearman's coefficient, denoted by ρ_s, and Kendall's Tau coefficient, denoted by ρ_τ. In addition, and to complete our non-parametric approach, we apply a test based on contingency table statistics. Cohen (1960) introduced the kappa statistic, denoted by ρ_κ, which is an appropriate measure to assess the concordance of judgements made by different raters as it aims to measure the proportion of matching pairs in the table that is not merely due to chance. In our paper, the raters are the different performance measures applied[14]. The lower bound for one-sided confidence interval for Cohen´s kappa is computed according to Fleiss (1971).

In order to classify the levels of association obtained for the alternative performance measures, we follow the approach of Hübner (2007), who defines a scale that is based on a

[14] In order to build the contingency tables we first rank funds on performance and identify the median. Funds with performance above the median are classified as Winners (W) and funds with performance below the median are classified as Losers (L). Comparing the results of two alternative performance measures, funds are classified in four categories: The funds that are Winners or Losers using both measures of performance are classified as WW and LL, respectively. The funds that are winners under one and losers under the other ranking are denoted by WL and LW.

refinement of the work of Landis and Koch (1977). Originally, Landis and Koch (1977) propose a scale to assess the strength of association between two statistics based on 5 levels of strenght: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). Later Hübner (2007) splits the intervals that define the strength of association further in lengths of 0.1 and defines 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+). Whenever the association statistic is negative the level of association is classified as Poor (P). Table 6 describes in detail the scale that will be used to assess the level of association between performance measures.

[Insert Table 6 here]

The results of the correlation analysis are summarised in tables 7 and 8. Table 7 shows the results on the Pearson correlation coefficient while table 8 presents the results on the Spearman, Kendall Tau and Cohen Kappa coefficients.

[Insert Table 7 and 8 here]

In general, the correlation values ($\hat{\rho}$) and their lower bounds, with a one-sided confidence at 5% level, are more conservative when we consider a parametric measure than when we consider non-parametric measures.

As can be observed in Table 7, when the Pearson correlation coefficient is used we obtain considerably high values and a strength of association ranging from SB- to AP+.

Results on the non-parametric measures, reported in Table 8, indicate that whatever non-parametric measure is considered, the correlation values and their respective lower bounds are also high. In particular, for the Spearman rank correlation coefficient, which is by far the most frequently used measure in previous studies, all the correlation values are very high (superior to 0.90), indicating a AP+ level of association between most of the risk-adjusted performance measures. However, the correlation values decreases when the Kendall Tau statistic is used. In a few cases, the Cohen's kappa coefficients are even lower. Nevertheless, the level of association can still be considered high, ranging from SB- to AP+. These findings suggest that it makes little difference to use either the traditional SR or any of the downside risk-adjusted performance measure suggested (Sortino ratio, ERCFVaR, ERVaR or ERES).

In summary, we find some differences in the strength of the association between performance results and rankings, depending on the measure that is used. Pearson's correlation coefficient and Spearman's rank correlation coefficient indicate an almost perfect association between the Sharpe ratio and the downside risk-adjusted performance measures. Taking as reference the results on the Kendall Tau and the Cohen´s Kappa, the level of association between the different performance measures declines.

Our results based on the Spearman rank correlation coefficient are thus similar to previous findings. In fact, these results support the hypothesis that the choice of the performance measure does not have a significant impact on the relative evaluation of investment trusts, even in the presence of empirical distributions that exhibit (mostly) significant negative skewness. This type of evidence is somewhat puzzling as it is well known that the performance measures used (traditional and downside risk-adjusted) are theoretically very distinct. As mentioned previously, under the SR a normal distribution is assumed for the return distribution and only the first two moments are considered. By contrast, excess return on VaR and excess return on ES (in the context of FHS methodology) do not rely on any distributional assumption and therefore the asymmetry (or even the fat tails) of the empirical distribution are taken into account.

As pointed out by Zakamouline (2010), a possible explanation for this result is the fact that although the investment trusts of our sample exhibit deviations from normality, those deviations are not much different across funds. Our empirical time series of returns significantly depart from normality and that is mainly driven by (mostly) negative skewness and excess kurtosis. However, the standard deviation of the skewness across our sample of funds is only of 0.5. It would therefore be of interest to investigate whether a higher variability of skewness (and kurtosis) across funds would produce different conclusions in terms of the association between performance measures. To explore this issue, a simulation exercise is implemented in order to generate new series of returns. The analysis of the association between performance rankings and measures is then repeated using the simulated series of returns.

6. Simulation analysis

A sample of 110 return series was simulated in such way that the average returns and the standard deviation of the returns are similar to our sample of UK investment trusts, but the standard deviation of the skewness and kurtosis of those return series is much larger. For each return series, 100000 observations are simulated from a skewed-t (Fernandez and Steel, 2000) distribution (the shape parameter is random). This creates a sample where the dispersion of the skewness and kurtosis parameters is much larger than in the empirical sample of UK investment trust funds. The descriptive statistics of the sampled simulated returns and the sample of actual fund returns are presented in table 9.

[Insert Table 9 here]

The dispersion of the skewness of UK investment trusts returns is small (standard deviation of 0.5) ranging from -2.38 to 1.35, whereas the standard deviation of skewness from the 110 simulated return series is around 3, ranging from -8.87 to 6.65. The difference in the dispersion of the kurtosis between the two samples is even larger.

Using this sample of simulated return series, we compute the same risk-adjusted performance measures as before and repeat the correlation analysis to assess the level of association between them. Tables 10 and 11 report the results obtained for the parametric and non-parametric correlation tests, respectively.

[Insert Table 10 and 11 here]

The results show that in the case of the ERCFVaR, the correlation values are very low. This is observed for all the association measures. However, these results might not be reliable given the extremely high values of excess kurtosis. As remarked by Mina and Ulmer (1989), when dealing with extremely sharp distributions, the Cornish-Fisher expansion may not be accurate and a non-parametric method is recommended to estimate VaR with accuracy.

In relation to the other performance measures, for all the correlation measures, the observed values and the corresponding lower bounds show that the correlations between the Sharpe ratio and the downside risk-adjusted measures are now lower, while those between the downside risk-adjusted performance measures are slightly higher. Notwithstanding, we observe that in the case of the Spearman rank correlation coefficient, the Kendall Tau and the Pearson correlation coefficient, the level of association continues to be high, varying from AP- to SB- (not considering the ERCFVaR). It is with Cohen´s kappa that we observe the most significant differences from the results in the previous section. The lower bounds of the correlations values of the Sharpe ratio with the downside risk-adjusted performance measures (excluding the ERCFVaR) are now between 0.45 and 0.50, which corresponds to a M- level of association. Compared with the results obtained with our sample of UK investment trusts (with lower bounds for the correlations values superior to 0.71), we may conclude that the choice of performance measure does influence the relative evaluation of investment portfolios. These results suggest that Cohen's Kappa, in comparison to the most frequently used correlation coefficients (Spearman's coefficient and Kendall's Tau), seems to be a more appropriate statistic test to capture differences between performance rankings.

7. Concluding remarks

Traditional measures of performance can produce biased estimates of fund managers abilities when returns are not elliptically symmetric distributed. Indeed, when dealing with highly negative skewed returns, it is expected that downside risk measures can be superior risk estimators. As a consequence, downside risk-adjusted measures of performance are considered, from a theoretical point of view, more accurate.

The purpose of this paper was to empirically analyze the impact on fund performance evaluation results from the use of traditional versus downside risk-adjusted performance measures. Different performance measures were computed for a data set of UK investment trusts. The time series of returns from UK investment trusts exhibit significant negative skewness, a characteristic that is relevant for the use of downside risk-adjusted measures of performance. In this case, a significant impact on performance assessment based on traditional *versus* downside risk-adjusted performance was expected.

The Sharpe ratio assumes that the returns follow a normal distribution. In contrast, our VaR and Expected Shortfall estimates were computed on the basis of FHS, which is a distribution-free methodology and therefore particularly suitable to capture any empirical properties of the data. Considering the characteristics of our data (highly significant negative skewed returns), a relatively low level of association was expected between the Sharpe ratio and the downside risk-adjusted measures of performance.

A preliminary analysis of the performance results revealed some differences between ranks, depending on the performance measure that is applied. These differences between performance ranks are reflected either in the number of funds that underperform/outperform the benchmark, the number of funds that maintain an equal order across rankings and the number of changes in the five top/bottom performing funds.

A further analysis, consisting in the application of four alternative statistical tests was performed, in order to assess the level of association between the performance measures. Both a parametric and non-parametric approaches were used. In the context of the parametric approach, Pearson's correlation was computed. Under the non-parametric approach, two rank correlation coefficients were applied: Spearman´s coefficient and Kendall´s Tau coefficient. As an alternative non-parametric measure of association, the Cohen´s Kappa statistic, based on contingency tables, was used.

With respect to the empirical dataset of UK investment trusts, the results are generally in favour of a substantial or almost perfect association between the Sharpe ratio and the proposed downside risk-adjusted measures of performance. This findings might be explained by the level of variability of the skewness (and kurtosis) across funds. To test this, a simulated data set, with a much higher variability of skewness (and kurtosis) across funds, was generated. For the simulated series of returns, the level of association between the different measures of performance drops substantially. In fact, for the simulated data, the results based on Cohen´s Kappa indicate moderate levels of association between the Sharpe ratio and all the downside risk-adjusted measures of performance. In sum, our results suggest that the use of different performance measures (traditional versus downside risk-adjusted) does have an impact in performance results, but only when the return series are characterized by a high variability in the higher moments, particularly in skewness. When the variability in skewness is lower, as in our empirical sample of UK investment trusts, the impact of using different performance measures will most likely be not significant.

As with this paper, most of the empirical studies on performance evaluation aim to assess the performance of funds in a specific category, like hedge funds or any other category of funds. In such a case, even when funds returns within a category depart from normality, the deviations will most likely be of the same type across funds because they follow similar investment strategies. Therefore we may conclude that from an empirical perspective the use of downside risk-adjusted measures does not have a significant impact in assessing fund performance, as long as the investment strategies of the funds under evaluation are similar.

However, even if in empirical research statistical tests are not able to capture significant differences between the measures, the economic significance of these differences may be important to investors and fund managers. This is an issue that deserves further research. In fact, it would be of interest to investigate whether the use of alternative risk-adjusted measures, for investment decisions, would lead to different performance in future time periods. The issue would be to assess to what extent the *ex ante* use of each of these measures can lead to better *ex post* performance results.

References

Alexander, G. J., Baptista, A. M., 2003. Portfolio Performance Evaluation Using Value-at-Risk. The Journal of Portfolio Management 29 (4), 93-102.

Amenc, N., Faff, R., Martellini, L., 2003. Desperately Seeking Pure Style Indexes. EDHEC: Risk and Asset Management Research Centre. France.

Arztner, P., Delbaen, F., Eber, J-M., Heath, D., 1999. Coherent Measures of Risk. Mathematical Finance 3, 203-228.

Barone-Adesi, G., Giannopoulos, K., Vosper, L., 1999. VaR Without Correlations for Portfolio of Derivative Securities. Journal of Futures Markets 19, 583-602.

Basak, S., Shapiro, A., 2001. Value-at-Risk based Risk Management: Optimal Policies and Asset Prices. Review of Financial Studies 14 (2), 371-405.

Bernardo, A. E., Ledoit, O., 2000. Gain, Loss and Asset Pricing. Journal of Political Economy 108, 144-172.

Bertsimas, D., Lauprete, G. J., Samarova, A., 2004. Shortfall as a Risk Measure: Properties, Optimization and Applications. Journal of Economic Dynamics & Control 28, 1353-1381.

Cohen, J., 1960. A coefficient of agreement for nominal scales. Educational and Psychological Measurement 20, 37-46.

Cornish, E., Fisher, R., 1937. Moments and Cumulants in the Specification of Distributions. Review of the International Statistical Institute, 307-320.

Doornik, J. A., 2007. Object-Oriented Matrix Programming Using Ox. 3rd ed. London: Timberlake Consultants Press and Oxford: www.doornik.com.

Eling, M., Schuhmacher, F., 2005. Performance Measurement of Hedge Fund Indices - Does the Measure Matter? Operations Research Proceedings. Berlin: Springer, 205-210.

Eling, M., 2008. Does the measure matter in the mutual fund industry? Financial Analysts Journal 64 (3), 54-66.

Eling, M., Schuhmacher, F., 2007. Does the choice of performance measure influence the evaluation of hedge funds? Journal of Banking & Finance 31 (9), 2632-2647.

Favre, L., Galeano, J. A., 2002. Mean-modified Value-at-Risk Optimization with Hedge Funds. Journal of Alternative Investments 5, 21-25.

Fernandez C., Steel M.F.J., 2000. On Bayesian Modelling of Fat Tails and Skewness. Preprint, 31 pages.

Fleiss, J. L., 1971. Measuring nominal scale agreement among many raters. Psychological Bulletin 76, 378-382.

Giannopoulos, K., Tunaru, R., 2005. On the Relation Between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. Journal of Banking and Finance 29, 979-996.

Gueyie, J. P., Amvella, S. P., 2006. Optimal Portfolio Allocation Using Funds of Hedge Funds. Journal of Wealth Management 9 (2), 85-95.

Harmantzis, F. C., Miao, L., Chien, Y., 2006. Empirical study of value-at-risk and expected shortfall model with heavy tails. Journal of Risk Finance, 7, 117-135.

Hartz, C., Mittnik, S., Paolella, M., 2006. Accurate Value-at-Risk Forecasting Based on the Normal-GARCH Model. Computational Statistics & Data Analysis 51, 2295-2312.

Hübner, G., 2007. How do Performance Measures Perform? Journal of Portfolio Management 33 (4), 64-74.

Jensen, M., 1968. The Performance of Mutual Funds in the Period 1945-1964. The Journal of Finance, May 23 (2), 389-416.

Kuester, K., Mittik, S., Paolella, M. S., 2005. Value-at-risk Prediction: A Comparison of Alternative Strategies. Journal of Financial Econometrics 4 (1), 53-89.

Landis, J. R., Koch, G. G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159-74.

Landsman, Z., Nešlehová, J., 2008. Stein's lemma for elliptical distributions. Journal of Multivariate Analysis 99, 912-927.

Leland, H., 1999. Beyond Mean-Variance: Performance Measurement in a Nonsymmetrical World. Financial Analysts Journal 55 (1), 27-36.

Markowitz, H. M., 1952. Portfolio Selection. Journal of Finance 7 (1), 77-91.

Martin, R. D., Rachev, S. T., Siboulet, F., 2003. Phi-Alpha Optimal Portfolios and Extreme Risk Management. Wilmott. November. 70-83.

Ornelas, J., Júnior, F., Fernandes, J., 2009. Yes, The Choice of Performance Measure Does Matter For Ranking of US Mutual Funds. Working Paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1403916.

Pedersen, C., Satchell, S., 2002. On the Foundation of Performance Measures under Asymmetric Returns. Quantitative Finance 3, 217-223.

Pedersen, C. S., Rudholm-Alfvin, T., 2003. Selecting a Risk-Adjusted Shareholder Performance Measure. Journal of Asset Management, September 4 (3), 152-172.

Pfingsten, A., Wagner, P., Wolferink, C., 2004. An Empirical Investigation of the Rank Correlation Between Different Risk Measures, Journal of Risk 6 (4), 55-74.

Pritsker, M., 2001. The Hidden Dangerous of Historical Simulation. Working Paper, January. Federal Reserve Board.

Sharpe, W. F., 1966. Mutual Fund Performance. Journal of Business. Supplement on Security Prices 39, 119-38.

Sortino, F., Price, L., 1994. Performance Measurement in a Downside Risk Framework. The Journal of Investing 3 (3), 59-65.

Treynor, J., 1965. How to Rate Management of Investment Funds. Harvard Business Review, January/February 43 (1), 63-75.

Tse, Y. K., 2002. Residual-based Diagnostics for Conditional Heteroscedasticity Models. Econometrics Journal 5, 358-373.

Yamai, Y., Yoshiba, T., 2002. On the Validity of Value-at-Risk: Comparative Analyses with expected shortfall. Monetary and Economic Studies. Institute for Monetary and Economic Studies. Bank of Japan 20 (1), 57-86.

Zakamouline, V., 2010. The Choice of Performance Measure Does Influence the Evaluation of Hedge Funds. Working Paper available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1403246.

Zangari, P., 1996. A VaR Methodology for Portfolios That Include Options. Riskmetrics™ Monitor, First Quarter, 4-12.

Table 1 – List of best models used as filters for serial correlation and heteroscedasticity under FHS

This table lists the best model fitted and selected for each of the investment trusts (identified by an IT code) and for the benchmark (the FTSE All share index represented by R900).

IT Code	Best Model Fitted	IT Code	Best Model Fitted	IT Code	Best Model Fitted
R10	GJR (1,1) ARMA (3,3)	**R167**	GJR (1,1) ARMA (1,0)	**R45**	GJR (1,1) ARMA (1,0)
R100	GJR (1,1) ARMA (3,3)	**R168**	GARCH (1,1) ARMA (3,3)	**R46**	GJR (1,1) ARMA (1,3)
R102	GJR (1,1) ARMA (1,1)	**R169**	GJR (1,1) ARMA (3,3)	**R47**	GARCH (1,1) ARMA (2,1) in-mean
R106	GJR (1,1) ARMA (1,2)	**R171**	GJR (1,1) ARMA (3,3)	**R49**	GJR (1,1) ARMA (3,3)
R107	GJR (1,1) ARMA (1,1)	**R172**	GJR (1,1) ARMA (1,0)	**R5**	GJR (1,1) ARMA (3,3) in-mean
R108	GJR (1,1) ARMA (1,1)	**R173**	GJR (1,1) ARMA (2,0)	**R52**	GJR (1,1) ARMA (3,1) in-mean
R109	GJR (1,1) ARMA (2,1)	**R174**	GJR (1,1) ARMA (3,0)	**R53**	GJR (1,1) ARMA (1,1)
R11	GJR (1,1) ARMA (3,1) in-mean	**R175**	GJR (1,1) ARMA (1,1)	**R55**	GJR (1,1) ARMA (3,3)
R110	GARCH (1,1) ARMA (3,3)	**R177**	GJR (1,1) ARMA (1,2) in-mean	**R56**	GJR (1,1) ARMA (3,3) in-mean
R111	GJR (1,1) ARMA (3,3) in-mean	**R178**	GJR (1,1) ARMA (3,3)	**R60**	GJR (1,1) ARMA (2,1)
R112	GJR (1,1) ARMA (1,1)	**R180**	GJR (1,1) ARMA (1,2)	**R62**	GJR (1,1) ARMA (1,1)
R113	GJR (1,1) ARMA (1,1)	**R184**	GARCH (1,1) ARMA (1,1)	**R64**	GJR (1,1) ARMA (2,0)
R114	GARCH (1,1) ARMA (3,3)	**R186**	GJR (1,1) ARMA (1,1) in-mean	**R66**	GJR (1,1) ARMA (1,1)
R115	GJR (1,1) ARMA (3,2)	**R187**	GJR (1,1) ARMA (3,3)	**R69**	GJR (1,1) ARMA (3,3)
R116	GJR (1,1) ARMA (3,3)	**R188**	GJR (1,1) ARMA (3,3)	**R70**	GJR (1,1) ARMA (2,0)
R122	GJR (1,1) ARMA (1,1)	**R189**	GJR (1,1) ARMA (1,2)	**R72**	GJR (1,1) ARMA (1,1)
R123	GJR (1,1) ARMA (1,1)	**R193**	GARCH (1,1) ARMA (1,1)	**R73**	GJR (1,1) ARMA (1,1)
R126	GJR (1,1) ARMA (3,0)	**R195**	GJR (1,1) ARMA (3,0)	**R74**	GJR (1,1) ARMA (1,1)
R128	GJR (1,1) ARMA (1,1)	**R196**	GJR (1,1) ARMA (3,2)	**R75**	GJR (1,1) ARMA (2,1)
R129	GJR (1,1) ARMA (3,0)	**R197**	GJR (1,1) ARMA (3,3)	**R76**	GARCH (1,1) ARMA (2,1)
R131	GJR (1,1) ARMA (2,0)	**R198**	GARCH (1,1) ARMA (3,2)	**R77**	GARCH (1,1) ARMA (1,2)
R134	GJR (1,1) ARMA (2,1) in-mean	**R2**	GJR (1,1) ARMA (1,3)	**R78**	GJR (1,1) ARMA (3,3)
R138	GJR (1,1) ARMA (1,1)	**R21**	GJR (1,1) ARMA (1,1)	**R80**	GARCH (1,1) ARMA (1,1) in-mean
R141	GJR (1,1) ARMA (1,2)	**R22**	GJR (1,1) ARMA (2,1) in-mean	**R81**	GARCH (1,1) ARMA (1,1) in-mean
R142	GARCH (1,1) ARMA (3,3)	**R23**	GJR (1,1) ARMA (0,0)	**R82**	GJR (1,1) ARMA (1,0)
R143	GJR (1,1) ARMA (3,3) in-mean	**R28**	GJR (1,1) ARMA (1,1)	**R84**	GJR (1,1) ARMA (1,1)
R144	GJR (1,1) ARMA (2,0) in-mean	**R29**	GJR (1,1) ARMA (3,3)	**R85**	GJR (1,1) ARMA (1,2)
R145	GJR (1,1) ARMA (2,2) in-mean	**R3**	GJR (1,1) ARMA (1,2)	**R86**	GJR (1,1) ARMA (3,0)
R147	GJR (1,1) ARMA (3,1)	**R30**	GJR (1,1) ARMA (3,3)	**R900**	GJR (1,1) ARMA (0,0)
R148	GJR (1,1) ARMA (1,1) in-mean	**R32**	GJR (1,1) ARMA (1,1)	**R92**	GJR (1,1) ARMA (3,3)
R150	GJR (1,1) ARMA (3,3)	**R34**	GJR (1,1) ARMA (1,1)	**R93**	GJR (1,1) ARMA (1,1)
R151	GJR (1,1) ARMA (1,2)	**R38**	GJR (1,1) ARMA (2,0)	**R94**	GJR (1,1) ARMA (1,1)
R152	GJR (1,1) ARMA (1,0)	**R4**	GJR (1,1) ARMA (1,1)	**R95**	GJR (1,1) ARMA (2,1) in-mean
R154	GJR (1,1) ARMA (1,0)	**R41**	GJR (1,1) ARMA (3,0)	**R96**	GJR (1,1) ARMA (3,3)
R155	GJR (1,1) ARMA (1,0)	**R42**	GJR (1,1) ARMA (3,3)	**R98**	GJR (1,1) ARMA (1,1)
R156	GJR (1,1) ARMA (1,2)	**R43**	GARCH (1,1) ARMA (3,0)	**R99**	GJR (1,1) ARMA (1,1)
R16	GJR (1,1) ARMA (3,0)	**R44**	GJR (1,1) ARMA (3,0)		

Table 2 – Estimates of performance measures

This table summarises the results for different performance ratios for each trust in the sample (identified by an IT code) as well as for the benchmark (the FTSE All-Share represented by R900). The results for the benchmark are reported in bold. The mean and the standard deviation of each performance ratio are also reported.

IT Code	SR	Sortino	ERCFVaR	ERVaR	ERES	IT Code	SR	Sortino	ERCFVaR	ERVaR	ERES
R10	0.0186	4.2678	0.0025	0.0097	0.0063	**R196**	0.0039	0.6833	0.0012	0.0018	0.0016
R100	0.0117	1.8699	0.0040	0.0057	0.0047	**R197**	-0.0147	-2.0081	-0.0046	-0.0105	-0.0080
R102	-0.0107	-1.2568	-0.0023	-0.0061	-0.0047	**R198**	0.0337	8.2999	0.0063	0.0188	0.0121
R106	0.0131	1.8356	0.0045	0.0068	0.0050	**R2**	0.0195	2.6294	0.0035	0.0199	0.0146
R107	0.0095	1.8644	0.0022	0.0037	0.0028	**R21**	0.0079	1.0152	0.0027	0.0039	0.0033
R108	0.0198	3.2254	0.0068	0.0107	0.0082	**R22**	0.0090	1.1939	0.0028	0.0050	0.0037
R109	0.0352	5.1765	0.0089	0.0252	0.0181	**R23**	0.0051	0.8675	0.0018	0.0022	0.0018
R11	0.0065	1.3608	0.0022	0.0036	0.0028	**R28**	-0.0026	-0.5039	-0.0003	-0.0012	-0.0008
R110	-0.0047	-0.5596	-0.0017	-0.0025	-0.0020	**R29**	-0.0001	-0.0111	0.0000	0.0000	0.0000
R111	0.0024	0.2787	0.0007	0.0012	0.0008	**R3**	0.0109	1.5639	0.0024	0.0062	0.0046
R112	0.0337	7.0166	0.0110	0.0184	0.0141	**R30**	0.0596	15.2758	0.0184	0.0249	0.0180
R113	0.0064	1.1522	0.0023	0.0027	0.0022	**R32**	0.0020	0.3470	0.0005	0.0011	0.0009
R114	0.0372	8.2197	0.0065	0.0188	0.0154	**R34**	0.0685	33.4412	0.0103	0.0275	0.0151
R115	0.0289	3.7571	0.0077	0.0160	0.0121	**R38**	0.0032	0.5751	0.0010	0.0014	0.0011
R116	0.0175	3.4072	0.0046	0.0075	0.0062	**R4**	-0.0107	-1.2568	-0.0023	-0.0061	-0.0047
R122	0.0192	2.6133	0.0035	0.0146	0.0092	**R41**	0.0041	0.6919	0.0011	0.0020	0.0017
R123	-0.0053	-0.9103	-0.0012	-0.0028	-0.0022	**R42**	0.0129	2.8989	0.0022	0.0055	0.0040
R126	0.0040	0.8430	0.0011	0.0016	0.0013	**R43**	0.0220	1.9757	0.0027	0.0111	0.0064
R128	0.0278	6.8346	0.0069	0.0104	0.0075	**R44**	-0.0052	-0.6462	-0.0015	-0.0029	-0.0022
R129	0.0023	0.3175	0.0007	0.0014	0.0011	**R45**	-0.0049	-0.8343	-0.0018	-0.0026	-0.0021
R131	0.0053	0.9521	0.0015	0.0023	0.0019	**R46**	-0.0138	-2.0396	-0.0022	-0.0117	-0.0085
R134	0.0162	3.1641	0.0039	0.0097	0.0064	**R47**	0.0164	4.2676	0.0028	0.0084	0.0063
R138	0.0151	3.9423	0.0028	0.0089	0.0061	**R49**	0.0044	0.7782	0.0011	0.0032	0.0026
R141	0.0181	3.3597	0.0063	0.0076	0.0060	**R5**	0.0345	7.6784	0.0110	0.0137	0.0111
R142	0.0199	4.4735	0.0041	0.0107	0.0076	**R52**	0.0133	2.1493	0.0034	0.0071	0.0059
R143	0.0076	1.6689	0.0027	0.0039	0.0030	**R53**	0.0048	0.9071	0.0007	0.0018	0.0013
R144	-0.0002	-0.0381	-0.0001	-0.0001	-0.0001	**R55**	0.0215	5.0257	0.0029	0.0093	0.0069
R145	-0.0165	-1.8926	-0.0016	-0.0156	-0.0110	**R56**	0.0185	1.8480	0.0036	0.0132	0.0095
R147	-0.0077	-1.6760	-0.0014	-0.0070	-0.0055	**R60**	0.0140	2.0985	0.0043	0.0086	0.0064
R148	0.0351	8.4777	0.0051	0.0207	0.0141	**R62**	0.0457	8.0228	0.0151	0.0238	0.0188
R150	0.0031	0.4517	0.0006	0.0024	0.0016	**R64**	0.0471	9.3662	0.0111	0.0110	0.0090
R151	-0.0195	-2.0958	-0.0059	-0.0129	-0.0094	**R66**	0.0139	2.8858	0.0032	0.0069	0.0055
R152	0.0040	0.4424	0.0011	0.0032	0.0026	**R69**	0.0047	0.7272	0.0016	0.0027	0.0023
R154	0.0150	2.6972	0.0051	0.0057	0.0047	**R70**	0.0136	2.0164	0.0040	0.0083	0.0066
R155	-0.0156	-1.7005	-0.0046	-0.0102	-0.0084	**R72**	0.0194	2.8916	0.0052	0.0137	0.0113
R156	0.0291	12.4271	0.0056	0.0074	0.0052	**R73**	0.0494	12.6373	0.0076	0.0224	0.0168
R16	-0.0044	-0.6495	-0.0004	-0.0044	-0.0027	**R74**	0.0049	0.9932	0.0013	0.0018	0.0014
R167	-0.0073	-0.7977	-0.0022	-0.0038	-0.0031	**R75**	0.0256	5.6831	0.0050	0.0182	0.0134
R168	-0.0006	-0.0704	-0.0002	-0.0003	-0.0002	**R76**	0.0699	16.1532	0.0166	0.0250	0.0189
R169	-0.0163	-2.6425	-0.0058	-0.0065	-0.0056	**R77**	0.0160	3.7816	0.0029	0.0092	0.0066
R171	-0.0016	-0.2874	-0.0005	-0.0008	-0.0007	**R78**	0.0158	3.1444	0.0042	0.0101	0.0078
R172	0.0042	0.7251	0.0014	0.0018	0.0015	**R80**	0.0493	13.8547	0.0053	0.0175	0.0111
R173	0.0095	1.6111	0.0033	0.0055	0.0046	**R81**	0.0443	9.5457	0.0053	0.0106	0.0074
R174	0.0117	1.6093	0.0018	0.0110	0.0077	**R82**	-0.0056	-0.8202	-0.0016	-0.0031	-0.0025
R175	0.0052	1.0726	0.0017	0.0024	0.0019	**R84**	-0.0070	-0.9433	-0.0020	-0.0033	-0.0026
R177	-0.0067	-1.2773	-0.0014	-0.0024	-0.0018	**R85**	-0.0025	-0.3648	-0.0004	-0.0018	-0.0013
R178	0.0138	3.1527	0.0022	0.0085	0.0062	**R86**	-0.0120	-1.3583	-0.0040	-0.0073	-0.0061
R180	0.0097	2.0277	0.0029	0.0050	0.0040	**R900**	0.0041	0.7829	0.0015	0.0022	0.0018
R184	0.0037	0.8207	0.0009	0.0017	0.0013	**R92**	-0.0127	-1.3646	-0.0036	-0.0131	-0.0101
R186	0.0163	3.3785	0.0048	0.0052	0.0039	**R93**	0.0100	1.6311	0.0025	0.0053	0.0042
R187	0.0106	1.4825	0.0033	0.0056	0.0043	**R94**	0.0152	3.0157	0.0045	0.0063	0.0049
R188	0.0156	3.1955	0.0044	0.0062	0.0045	**R95**	0.0035	0.3376	0.0011	0.0026	0.0020
R189	0.0255	3.5795	0.0086	0.0174	0.0130	**R96**	0.0194	3.0154	0.0046	0.0212	0.0175
R193	0.0104	2.6241	0.0014	0.0036	0.0023	**R98**	0.0201	4.1171	0.0035	0.0108	0.0078
R195	0.0082	1.9853	0.0022	0.0028	0.0021	**R99**	0.0114	2.5542	0.0022	0.0073	0.0052
						Média	**0.0117**	**2.7157**	**0.0027**	**0.0056**	**0.0041**
						Desvio Padrão	**0.0174**	**4.6959**	**0.0041**	**0.0088**	**0.0065**

Table 3 – Performance rankings according to different performance measures

This table reports the rankings based on different performance ratios for each trust in the sample (identified by an IT code) as well as for the benchmark (the FTSE All-Share represented by R900).

IT Code	SR	Sortino	ER CFVaR	ER VaR	ER ES	IT Code	SR	Sortino	ER CFVaR	ER VaR	ER ES
R10	29	20	53	30	36	**R196**	77	77	72	75	74
R100	49	51	34	50	47	**R197**	106	107	108	106	105
R102	101	99	103	100	100	**R198**	14	10	15	10	15
R106	47	54	28	45	45	**R2**	25	41	39	9	9
R107	57	52	62	60	61	**R21**	61	65	50	58	58
R108	24	30	13	25	22	**R22**	59	62	48	56	57
R109	10	17	8	2	3	**R23**	67	69	64	73	72
R11	63	61	58	61	60	**R28**	90	90	88	89	89
R110	92	91	98	92	92	**R29**	85	85	85	85	85
R111	82	84	82	83	84	**R3**	52	59	55	47	49
R112	13	14	5	12	11	**R30**	3	3	1	4	4
R113	64	63	56	67	66	**R32**	84	81	84	84	83
R114	9	11	14	11	7	**R34**	2	1	7	1	8
R115	16	25	10	16	14	**R38**	80	78	78	81	81
R116	32	27	26	40	37	**R4**	102	100	104	101	101
R122	28	43	37	17	20	**R41**	74	76	75	74	73
R123	95	97	92	94	95	**R42**	48	37	59	52	55
R126	76	70	73	80	79	**R43**	20	50	51	21	33
R128	17	15	12	28	27	**R44**	94	92	95	95	94
R129	83	83	81	82	82	**R45**	93	96	99	93	93
R131	65	67	67	72	70	**R46**	105	108	102	107	107
R134	35	32	35	31	34	**R47**	33	21	47	37	35
R138	40	23	49	34	39	**R49**	71	73	74	64	62
R141	31	29	16	39	40	**R5**	12	13	6	18	17
R142	23	19	32	26	26	**R52**	46	45	40	43	41
R143	62	55	52	59	59	**R53**	69	68	80	77	78
R144	86	86	86	86	86	**R55**	21	18	44	32	29
R145	109	106	96	110	110	**R56**	30	53	36	20	19
R147	100	104	94	103	102	**R60**	42	46	30	35	32
R148	11	9	22	8	10	**R62**	7	12	3	5	2
R150	81	79	83	71	75	**R64**	6	8	4	23	21
R151	110	109	110	108	108	**R66**	43	39	43	44	42
R152	75	80	76	63	63	**R69**	70	74	66	66	64
R154	41	40	21	49	48	**R70**	45	48	33	38	31
R155	107	105	107	105	106	**R72**	27	38	20	19	16
R156	15	6	17	41	43	**R73**	4	5	11	6	6
R16	91	93	90	99	98	**R74**	68	66	71	78	77
R167	99	94	101	98	99	**R75**	18	16	23	13	12
R168	87	87	87	87	87	**R76**	1	2	2	3	1
R169	108	110	109	102	103	**R77**	36	24	45	33	30
R171	88	88	91	88	88	**R78**	37	34	31	29	23
R172	72	75	70	76	76	**R80**	5	4	18	14	18
R173	58	57	42	53	50	**R81**	8	7	19	27	28
R174	50	58	63	22	25	**R82**	96	95	97	96	96
R175	66	64	65	70	71	**R84**	98	98	100	97	97
R177	97	101	93	91	91	**R85**	89	89	89	90	90
R178	44	33	57	36	38	**R86**	103	102	106	104	104
R180	56	47	46	57	54	**R900**	73	**72**	**68**	**69**	**69**
R184	78	71	79	79	80	**R92**	104	103	105	109	109
R186	34	28	24	55	56	**R93**	55	56	54	54	53
R187	53	60	41	51	52	**R94**	39	35	27	46	46
R188	38	31	29	48	51	**R95**	79	82	77	68	68
R189	19	26	9	15	13	**R96**	26	36	25	7	5
R193	54	42	69	62	65	**R98**	22	22	38	24	24
R195	60	49	60	65	67	**R99**	51	44	61	42	44

Table 4- Equality of performance rankings

This table reports the number of investment trusts that exhibit the same rank order resulting from the use of alternative performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES. These figures are expressed in percentage of total sample.

	Equality of ranks	**Equality of ranks %**
SR-Sortino	10	9%
SR-ERCFVaR	10	9%
SR-ERVaR	11	10%
SR-ERES	11	10%
Sortino-ESCFVaR	7	6%
Sortino-ERVaR	9	8%
Sortino-ERES	6	5%
ERCFVaR-ERVaR	7	6%
ERCFVaR-ERES	5	5%
ERVaR-ERES	33	30%
Average	**10.9**	**10%**
Maximum	**33**	**30%**
Minimum	**5**	**5%**

Table 5 – Number of changes in the top/bottom five rankings

This table reports the changes in the top/bottom five performing funds resulting from different risk-adjusted performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES.

	Number of Changes in Top	**Number of Changes in Bottom Five**
SR-Sortino	0	1
SR-ERCFVaR	3	1
SR-ERVaR	2	2
SR-ERES	3	2
Sortino-ESCFVaR	3	2
Sortino-ERVaR	2	1
Sortino-ERES	3	2
ERCFVaR-ERVaR	2	3
ERCFVaR-ERES	2	3
ERVaR-ERES	1	1
Average	**2.1**	**1.8**
Average %	**42%**	**36%**
Standard Deviation	0.994428926	0.788810638

Table 6 – Scale for the strength of association

This table describes the scale used to assess the strength of association between the performance measures. Our scale is based on 5 levels of strength of association: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). We follow the approach of Hübner (2007) by splitting the intervals that define the strength of association further in lengths of 0.1 and defining 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+). For negative values of the association statistic our scale gives a classification of Poor (P).

Association Statistic $(\rho_i, i = P, S, \tau, \kappa)$	**Strength of Association**
< 0.00	Poor
0.00 – 0.10	Slight-lower range (SL-)
0.11 – 0.20	Slight-upper range (SL+)
0.21 – 0.30	Fair-lower range (F-)
0.31 – 0.40	Fair-upper range (F+)
0.41 – 0.50	Moderate-lower range (M-)
0.51 – 0.60	Moderate-upper range (M+)
0.61 – 0.70	Substantial-lower range (SB-)
0.71 – 0.80	Substantial-upper range (SB+)
0.81 – 0.90	Almost perfect-lower range (AP-)
0.91 – 1.00	Almost perfect-upper range (AP+)

Table 7 – Summary of results for the parametric measure of correlation: UK investment trusts

This table summarises the results of the Pearson correlation coefficients ($\widehat{\rho_P}$), considering different risk-adjusted performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES. The observed value ($\hat{\rho}$) and its lower bound with a one-sided confidence at 5% level are reported, as well as the corresponding strength of association: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). We adopt the approach of Hübner (2007) that splits the intervals defined above further in lengths of 0.1 and defines 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+).

	Pearson's correlation		
	$\hat{\rho}_P$	Lower Bound	Strength of Association
SR-Sortino	0.9007	0.8581	AP-
SR-ERCFVaR	0.9268	0.8949	AP-
SR-ERVaR	0.9249	0.8921	AP-
SR-ERES	0.9110	0.8727	AP-
Sortino-ERCFVaR	0.7642	0.6733	SB-
Sortino-ERVaR	0.7644	0.6736	SB-
Sortino-ERES	0.7199	0.6156	SB-
ERCFVaR-ERVaR	0.8764	0.8244	AP-
ERCFVaR-ERES	0.8890	0.8420	AP-
ERVaR-ERES	0.9934	0.9903	AP+

Table 8 – Summary of results for the non-parametric measures of correlation: UK investment trusts

This table summarises the results of the Spearman rank correlation coefficients ($\widehat{\rho_S}$), Kendall Tau rank correlation coefficients ($\widehat{\rho_\tau}$) and Cohen's Kappa coefficients ($\widehat{\rho_\kappa}$), considering different risk-adjusted performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES. The observed value ($\hat{\rho}$) and its lower bound with a one-sided confidence at 5% level are reported, as well as the corresponding strength of association: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). We follow the approach of Hübner (2007) that splits the intervals defined above further in lengths of 0.1 and defines 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+).

	Spearman's correlation			**Kendall's Tau**			**Cohen´s Kappa**		
	$\hat{\rho}_S$	**Lower Bound**	**Strength of Association**	$\hat{\rho}_\tau$	**Lower Bound**	**Strength of Association**	$\hat{\rho}_\kappa$	**Lower Bound**	**Strength of Association**
SR-Sortino	0.9791	0.9696	AP+	0.8986	0.8553	AP-	0.8545	0.7575	SB+
SR-ERCFVaR	0.9623	0.9454	AP+	0.8465	0.7834	SB+	0.8182	0.7107	SB+
SR-ERVaR	0.9735	0.9616	AP+	0.8756	0.8233	AP-	0.9636	0.9137	AP+
SR-ERES	0.9704	0.9571	AP+	0.8666	0.8109	AP-	0.9273	0.8573	AP-
Sortino-ERCFVaR	0.9404	0.9142	AP+	0.8058	0.7285	SB+	0.7818	0.6653	SB-
Sortino-ERVaR	0.9352	0.9067	AP+	0.8002	0.7210	SB+	0.8182	0.7107	SB+
Sortino-ERES	0.9322	0.9025	AP-	0.7938	0.7125	SB+	0.8182	0.7107	SB+
ERCFVaR-ERVaR	0.9333	0.9040	AP-	0.7928	0.7111	SB+	0.8545	0.7575	SB+
ERCFVaR-ERES	0.9405	0.9143	AP+	0.8065	0.7294	SB+	0.8545	0.7575	SB+
ERVaR-ERES	0.9970	0.9957	AP+	0.9650	0.9493	AP+	0.9636	0.9137	AP+

Table 9 – Descriptive statistics for the simulated return series compared with those of the investment trusts sample

This table reports descriptive statistics of the simulated sample of returns in comparison with the descriptive statistics of our sample of UK investment trusts. Each sample includes 110 return series.

	Mean	Standard Deviation	Skewness	Kurtosis
Simulated data				
Mean	0.000303	0.011448	-0.3002	37.7463
Standard Deviation	0.000170	0.003159	3.0957	65.5859
Min	-0.000206	0.002738	-8.8716	10.0796
Max	0.000852	0.019670	6.6521	581.8960
Sample of UK unit trusts				
Mean	0.000255	0.011260	-0.5153	14.4555
Standard Deviation	0.000169	0.003214	0.5068	9.0817
Min	-0.000183	0.004799	-2.3860	4.8307
Max	0.000786	0.019853	1.3491	57.0406

Table 10 – Summary of results for the parametric measure of correlation: simulated data

This table summarises the results of the Pearson correlation coefficients ($\widehat{\rho_P}$), considering different risk-adjusted performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES. The observed value ($\hat{\rho}$) and its lower bound with a one-sided confidence at 5% level are reported, as well as the corresponding strength of association: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). We follow the approach of Hübner (2007) that splits the intervals defined above further in lengths of 0.1 and defines 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+).

	Pearson's correlation		
	$\hat{\rho}_P$	**Lower Bound**	**Strength of Association**
SR-Sortino	0.8143	0.7399	SB+
SR-ERCFVaR	0.7461	0.6496	SB-
SR-ERVaR	0.7888	0.7058	SB-
SR-ERES	0.7254	0.6228	SB-
Sortino-ERCFVaR	0.3644	0.1909	SL+
Sortino-ERVaR	0.9130	0.8754	AP-
Sortino-ERES	0.8942	0.8491	AP-
ERCFVaR-ERVaR	0.2860	0.1043	SL-
ERCFVaR-ERES	0.2038	0.0172	SL-
ERVaR-ERES	0.9952	0.9930	AP+

Table 11 – Summary of results for the non-parametric measures of correlation: simulated data

This table summarises the results of the Spearman rank correlation coefficients ($\widehat{\rho_S}$), Kendall Tau rank correlation coefficients ($\widehat{\rho_\tau}$) and Cohen's Kappa coefficients ($\widehat{\rho_\kappa}$), considering different risk-adjusted performance measures: Sharpe ratio, Sortino ratio, Excess Return on CFVaR, Excess Return on VaR and Excess Return on ES. The observed value ($\hat{\rho}$) and its lower bound with a one-sided confidence at 5% level are reported, as well as the corresponding strength of association: Slight (SL), Fair (F), Moderate (M), Substantial (SB) and Almost Perfect (AP). We follow the approach of Hübner (2007) that splits the intervals defined above further in lengths of 0.1 and defines 10 levels of association that range from Slight-lower range (SL-) to Almost Perfect-upper range (AP+).

	Spearman's correlation			**Kendall's Tau**			**Cohen´s Kappa**		
	$\hat{\rho}_S$	**Lower Bound**	**Strength of Association**	$\hat{\rho}_\tau$	**Lower Bound**	**Strength of Association**	$\hat{\rho}_\kappa$	**Lower Bound**	**Strength of Association**
SR-Sortino	0.9012	0.8588	AP-	0.7485	0.6527	SB-	0.6364	0.4922	M-
SR-ERCFVaR	0.1076	-0.0813	P	0.0709	-0.1179	P	0.1273	-0.0581	P
SR-ERVaR	0.8742	0.8214	AP-	0.7595	0.6671	SB-	0.6364	0.4922	M-
SR-ERES	0.8477	0.7851	SB+	0.7351	0.6353	SB-	0.6000	0.4505	M-
Sortino-ERCFVaR	0.0035	-0.1838	P	0.0002	-0.1871	P	0.0000	-0.1869	P
Sortino-ERVaR	0.9817	0.9753	AP+	0.8936	0.8483	AP-	0.9273	0.8573	AP-
Sortino-ERES	0.9753	0.9642	AP+	0.8792	0.8283	AP-	0.9273	0.8573	AP-
ERCFVaR-ERVaR	0.0061	-0.1813	P	0.0018	-0.1855	P	0.0545	-0.1321	P
ERCFVaR-ERES	-0.0074	-0.1944	P	-0.0092	-0.1961	P	0.0545	-0.1321	P
ERVaR-ERES	0.9973	0.9961	AP+	0.9743	0.9627	AP+	0.9636	0.9137	AP+