

# An Analysis of Performance Measures using Copulae

Soosung Hwang and Mark Salmon<sup>1</sup>  
City University Business School, UK

June 2001

<sup>1</sup>Faculty of Finance, City University Business School, Frobisher Crescent, Barbican Centre, London EC 8HY, UK. Tel +44 (0)20 7477 - 0109(Soosung Hwang) or 8746 ( Mark Salmon), Fax +44 (0)20 7477 8881, Email s.hwang@city.ac.uk. and m.salmon@city.ac.uk. We would like to thank Alistair Sayer of J.P.Morgan, London for several discussions relating to the practical implementation of performance analysis.

## Abstract

We have carried out a detailed comparison of the statistical properties and the relationships between a set of five performance measures using 14 UK based Investment Trusts over a sample period ranging from 1980 to 2001. Our results suggest very clearly that there is almost no difference between Jensen's Alpha, the Treynor-Mazuy (TM) measure and the Positive Period Weighting (PPW) measure over our sample period and amongst our set of Investment Trusts. This would seem to indicate that there is no timing ability within these fund managers. The Sharpe Ratio clearly provides different signals regarding performance than the other measures and is the only absolute measure in the set of measures we have considered. While simple correlation analysis suggests that there is a high degree of dependence between most of the measures we have shown that there is a lack of significant concordance between the Sharpe Ratio and all the other measures. This indicates the inadequacy of correlation analysis with non-gaussian data. We have also shown that the Sharpe Ratio exhibits negative left tail area dependence with respect to Jensen's Alpha, TM and PPW but is independent in the left tail from the Higher Moment measure of Hwang and Satchell, that is when poor performance is indicated. Jensen's Alpha, TM and the HM measure do not seem to show any significant *asymptotic* left tail dependency. All the measures appear to be *asymptotically* independent in their upper tail when good performance is indicated. These results are further refined by non-asymptotic quantile regression results which indicate finite sample dependency the HM measure and Jensen's Alpha throughout the body of their conditional distribution and in the left tail but not the upper tail.

# 1 Introduction

Studies of portfolio performance evaluation began in the 1960's along with the development of modern asset pricing theory. Treynor (1965), Treynor and Mazuy (1966), and Jensen (1968, 1969) for instance used the CAPM to introduce portfolio performance measures. As finance theory developed so did performance measurement, for instance, Connor and Korajczyk (1986) and Lehmann and Modest (1987) introduced APT-based measures, there are the positive period weighting measures of Grinblatt and Titman (1989), the intertemporal marginal rates of substitution-based measures of Glosten and Jagannathan (1994), the measures of Chen and Knez (1995) based on the law of one price and/or no arbitrage, and the higher moment measure of Hwang and Satchell (1998).

While the early CAPM based performance measures have well recognised deficiencies, in particular that they rest either on a false assumption that all asset returns are normally distributed and thus distributed symmetrically or that investors have mean-variance preferences and thus ignore skewness, they still appear to be the most widely accepted for evaluating portfolio managers within the finance industry, see for instance the AIMR Performance Presentation Standards Handbook (1997). This would seem to deny the practical relevance of the theoretical arguments which led to the development of more refined measures, such as those indicated above. This paper asks the simple question of whether or not this position is justified. We take a pragmatic approach by comparing several measures using times series of monthly returns for 14 UK based investment trusts over a sample period running from January 1980 to February 2001.

Our objective on one level is crude and it is simply to examine whether or not there are significant differences between the different performance measures and then, if so, in what states of the market they occur. We adopt a slightly different statistical approach to similar comparative exercises in that we are interested in determining where in the range of their potential values these measures likely to provide different signals of performance or in other words when they are likely to be more dependent or independent of each other. They may for instance provide similar assessments when in their extremes but not when close to their average values- or *vice versa*. This we feel may be a critical practical concern for a fund manager uncertain as to which performance measure to adopt. Since the different measures represent different transformations of non-gaussian return series we would expect to find that simple correlation analysis would be inadequate since it is only applicable in the context of elliptic distributions and even in that case only

measures linear association<sup>1</sup>.

As an alternative we adopt the relatively recently developed theory of copula functions and associated dependency measures to examine more general forms of association between the different performance measures rather than just simply their normalised co-variation. In particular we consider their concordance which enables us to consider general positive (or negative) dependency rather than simple linear association. We then examine the dependency between the performance measures as they take values in the tails of their distributions by asking the question effectively of what is the probability that measure A will be beyond its 95th percentile given that measure B is also beyond its 95th percentile. In other words we quantify their tail area dependency given that this may be the most important area for many critical practical decisions. We then also examine, through bivariate quantile regressions the significance of their association throughout the range of their distribution. Finally we demonstrate how copulae may be used as aggregator functions to combine different but statistically dependent performance measures into a single measure. Following the literature on forecast combination we may expect an increase in efficiency and a reduction in bias by combining performance measures as long as the dependence between the individual components is properly taken into account in constructing the aggregate.

In the next section we briefly introduce the performance measures used in this study; the Sharpe Ratio, Jensen’s Alpha, the Treynor-Mazuy measure, the Positive Period Weighting measure, and a higher moment measure. We then compare these measures using data on the fourteen investment trusts before introducing the notion of a copula and related dependency measures.

## 2 Performance Measures

Several issues are of immediate importance when we seek to evaluate the performance of a portfolio although the relationship between the risk carried by the portfolio and the return is clearly paramount. Different concepts of risk, either relative or absolute, are employed and the notion of “risk” itself may be measured by quantities more general than the traditional use of the second moment. Some measures such as Jensen’s Alpha calculate relative returns after considering the systematic risk of the portfolio in the CAPM framework, whilst others such as the Sharpe Ratio simply use the return and variance of the portfolio itself.

---

<sup>1</sup>More importantly in the context of performance evaluation perhaps is the fact that correlation is not independent of scale changes of the underlying data.

Another issue is the effect of any superior timing ability as distinct from the stock selection ability in the fund manager. The difference between these two turns on whether private information lies in market aggregates or is firm-specific. For example, suppose that a fund manager has a portfolio which includes Microsoft. When he receives favourable information regarding Microsoft, he will increase his holding in Microsoft and this is nothing to do with the forecast of the market. Thus selectivity information is related to nonpriced risk, whilst timing information is related to priced risk, see Grinblatt and Titman (1995) for a more detailed discussion on this topic. Another critical issue is the sensitivity to the choice of reference index and the need to select an efficient benchmark as again emphasised by Grinblatt and Titman (1994).

In the following section we briefly introduce the five different performance measures we use below. Although the rationale and theoretical justification of these measures are to a degree different, empirically they may behave in a similar manner under many situations. How they differ in practice is a question of empirical evidence which we provide in the next section.

## 2.1 Traditional Performance Measures<sup>2</sup>

Treynor (1965) was the first to incorporate risk into a performance measure by considering the portfolio's rate of return with respect to the market rate of return. Jensen's (1966) extension is simpler and one of the most widely used in practice. Jensen's Alpha,  $\alpha_p^J$ , calculates the performance of a portfolio by measuring the deviation of a portfolio's return from the securities market line:

$$r_{pt} - r_f = \alpha_p^J + \beta_p(r_{mt} - r_f) + \varepsilon_{pt}$$

where  $r_{pt}$  represents the portfolio's return at time  $t$ ,  $r_f$  is the riskfree rate,  $r_{mt}$  is the market return at time  $t$ ,  $\beta_p$  denotes the systematic risk of the portfolio. Notice that Jensen's Alpha is the expected excess return of the portfolio less the product of the expected excess return of the market portfolio and the portfolio's beta.

The second measure we use is the Treynor-Mazuy (TM) statistic introduced in Treynor and Mazuy (1966). For a portfolio manager with forecasting power, the return on the managed portfolio will not be linearly related to the market return. This arises because he will gain more than the market when the market return is forecast to rise and he will lose less than the market

---

<sup>2</sup>Our objective here is not to provide a detailed theoretical comparison of the measures which can be found in a number of existing surveys but simply to provide the basis for the comparative analysis below.

when the market return is forecast to fall. Thus, his portfolio returns will be a concave function of market returns. Using the following quadratic model,

$$r_{pt} - r_f = \alpha_p + \beta_{1p}(r_{mt} - r_f) + \beta_{2p}(r_{mt} - r_f)^2 + \varepsilon_{pt}, \quad (1)$$

Treynor and Mazuy (1966) showed how the significance of  $\beta_{2p}$  provides evidence of the over performance of a portfolio. Admati, Bhattacharya, Pfleiderer, and Ross (1986) suggested conditions under which  $\alpha_p$  in (1) can be interpreted as the *selectivity* component of performance (i.e., the ability to forecast the returns on *individual* assets) and  $E(\beta_{2p}(\tilde{r}_{mt} - r_{ft})^2)$  interpreted as the *timing* component of performance (i.e., the ability to forecast *market* returns). The Treynor & Mazuy measure we use below is then given by

$$TM = \alpha_p + \beta_{2p}E((r_{mt} - r_f)^2). \quad (2)$$

The Sharpe Ratio (SR) (Sharpe (1966)) is simply the reward per unit of variability;

$$SR = \frac{E(r_p - r_f)}{\sigma_p},$$

where  $\sigma_p$  is the standard deviation of portfolio returns. The measure is simple, easy to understand and widely used.

The fourth measure we consider is the higher moment (HM) measure introduced by Hwang and Satchell (1998). In most cases, portfolio returns are not normally distributed and higher moments such as skewness and kurtosis need to be considered to adjust for the non-normality and to a degree account for the failure of variance to measure risk accurately. In these cases, a higher-moment CAPM should prove more suitable than the traditional CAPM and so a performance measure based on higher moments may also be more accurate than the measures outlined above. Assuming the validity of the three-moment CAPM and a quadratic return generating process of the form;

$$r_{pt} - r_f = a_{0p} + a_{1p}(r_{mt} - r_f) + a_{2p}\{r_{mt} - E(r_m)\}^2 + \varepsilon_{pt}, \quad (3)$$

we can define a performance measure of a portfolio under the three-moment CAPM<sup>3</sup> as

$$a_p = \mu_p - \lambda_1\mu_m - \lambda_2(\beta_{pm} - \gamma_{pm}) \quad (4)$$

where

$$\lambda_1 = \frac{\gamma_m^2 \gamma_{pm} - (\theta_m - 1)\beta_{pm}}{\gamma_m^2 - (\theta_m - 1)}$$

$$\lambda_2 = \frac{\gamma_m \sigma_m}{\gamma_m^2 - (\theta_m - 1)}$$

---

<sup>3</sup>Kraus and Litzenberger [1976] showed that the three-moment CAPM is consistent with the quadratic market model in (3)

with  $\mu_p = E(r_{pt} - r_{ft})$ ,  $\mu_m = E(r_{mt} - r_{ft})$ ,  $\sigma_m = E[\{r_m - E(r_m)\}^2]^{1/2}$ , and

$$\gamma_m = \frac{E[\{r_m - E(r_m)\}^3]}{\sigma_m^3}, \quad \theta_m = \frac{E[\{r_m - E(r_m)\}^4]}{\sigma_m^4},$$

and

$$\beta_{pm} = \frac{E[\{r_p - E(r_p)\}\{r_m - E(r_m)\}]}{E[\{r_m - E(r_m)\}^2]}, \quad \gamma_{pm} = \frac{E[\{r_p - E(r_p)\}\{r_m - E(r_m)\}^2]}{E[\{r_m - E(r_m)\}^3]}.$$

Note that  $\gamma_m$  and  $\theta_m$  are the skewness and kurtosis of the market returns, and  $\beta_{pm}$  and  $\gamma_{pm}$  are beta and coskewness, respectively. If the market returns are normal, then  $\lambda_1 = \beta_{pm}$  and  $\lambda_2 = 0$  and thus (4) is equivalent to Jensen's Alpha<sup>4</sup>.

We also use the positive period weighting (PPW) measure introduced by Grinblatt and Titman (1989). This measure is designed so that if selectivity and timing information are independent and the portfolio manager is a positive market timer, then the PPW measure assigns positive performance to stock selection ability and/or timing ability<sup>5</sup>. The PPW measure is obtained in two steps. First, we have to select a weighting vector, whose elements correspond to a time series  $\{w_t\}_{t=1}^T$ . The next step is to compute performance as a weighted average of the period-by-period portfolio excess returns:

$$\alpha_p^p = \sum_{t=1}^T w_t (r_{pt} - r_{ft})$$

where  $\sum_{t=1}^T w_t = 1$ , and  $w_t > 0$  for all  $t$ . Notice that  $\sum_{t=1}^T w_t (r_{mt} - r_{ft}) = 0$  for the market portfolio. There are many sets of weights which satisfy the above conditions and we use weights derived from the marginal utilities of an un-informed investor with a power utility function, as in Grinblatt and Titman (1994).

The first three measures above may be classified as traditional performance measures and although they are widely used, there are problems with each of them. The Sharpe Ratio for instance does not consider systematic risk which is the *real* risk in Markovitz's mean-variance world. On the other hand, a major problem with Jensen's Alpha is that it can ascribe negative performance to a market timer because it is based on an upwardly biased estimate of systematic risk for a market-timing investment strategy; see Jensen

<sup>4</sup>See Hwang and Satchell [1998] for detailed discussion of the properties of the higher moment performance measure.

<sup>5</sup>See Grinblatt and Titman [1989] for further discussion on the PPW measure.

(1972) for further discussion. The Treynor & Mazuy measure in (2) seems to be superior to the Sharpe Ratio and Jensen’s Alpha in the sense that the timing and selectivity ability of portfolio managers can be decomposed. The higher moment performance measure also suffers from the same difficulties as Jensen’s Alpha but does account for non-gaussianity. Notice that in the absence of market timing and with the assumption of normality, Jensen’s Alpha, the Treynor & Mazuy measure, the higher moment measure and the PPW measure are all expected to be identical. The empirical tests of Grinblatt and Titman (1994) and Cumby and Glen (1990) find that Jensen’s Alpha, the TM measure and PPW are indeed highly correlated and with emerging market data, Hwang and Satchell (1988) showed that the higher moment measure can rank portfolios quite differently from the other measures when returns are not normal. As the AIMR performance presentation standards handbook (AIMR(1997), page 90) states;

“The use of a variety of measures with an understanding of their shortcomings will provide the most valuable information because no one statistic can consistently capture all elements of risk of an asset class or a style of management”.

However as we discuss below, an aggregate measure, which properly accounts for the joint dependence between the constituent performance measures may well serve to compensate for the deficiencies in any one of the individual measures.

### **3 Empirical Results**

We now use the data on the 14 UK investment trusts to consider how these five performance measures behave in practice under differing market conditions from January 1980 to February 2001. The FTSE All-share and three month UK treasury bill are used as the benchmark and risk free returns, respectively. This period is one of extraordinary growth in the index until 1999 with significant adjustments due to the crash of 1987 and following the Russian Crisis of August 1998.

We note that the choice of a benchmark portfolio is a critical issue in performance analysis. Roll (1978) suggested that the benchmark portfolio should be mean-variance efficient for uninformed managers, while it needs to be mean-variance inefficient for the portfolio managers with forecasting ability. Our choice of the FTSE All-share index as a benchmark portfolio may not satisfy these conditions, but this value-weighted market portfolio is the most widely used benchmark portfolio in empirical studies.

We first provide some statistical properties of the trusts and the benchmark returns and then the five performance measures are analysed for the entire sample period.

### 3.1 Data

Table 1 reports the basic statistical properties of the returns from our benchmark and the fourteen UK investment trusts over sample period. The monthly mean return ranges from 1.1% to 1.6% (13% to 19% in annual terms) with standard deviations between 4% and 8% (14% to 28% in annual terms). This performance is exceptionally high when compared with returns over longer horizons such as the last 100 years for US equities, see Cochrane (1997). Panel B of table 1 also reports the same statistical properties when the 1987 market crash is excluded. As expected mean returns increase and standard deviations, skewness and excess kurtosis decrease. However, none of the returns are normally distributed, as indicated by the Jarque-Bera statistics, all showing a negative skew and leptokurtosis.

In panel C of table 1, we report estimates of the correlations between the return series. As might be expected all the correlation coefficients are high and in particular correlations with the benchmark portfolio are all larger than 0.5 and in fact except for four investment trusts they are all larger than 0.8. The correlations between the different investment trusts are also generally very high but we note that *Invesco English & International Trust* shows the lowest correlation with the benchmark portfolio and also with the other investment trusts.

### 3.2 Performance of the UK Investment Trusts for the Entire Sample Period

We next calculate the five performance measures over the entire sample period; see table 2. As in Grinblatt and Titman (1994) and Cumby and Glen (1990), we find that three measures, Jensen's Alpha, TM and PPW provide virtually identical results. For these three, the top five and bottom five investment trusts are same and the performance values are very similar. This provides indirect evidence that timing ability does not exist at least in these UK investment trusts.

However, the results for the Sharpe Ratio show some differences from these three measures. The obvious difference being that the estimated Sharpe Ratios are always positive, whilst the others are not. This arises essentially because the Sharpe Ratio is not a relative performance measure given the

performance of a benchmark and the positive values of the Sharpe Ratio simply reflect the positive performance in the benchmark portfolio over the sample period. In addition, measures such as Jensen's Alpha and TM are based on non-diversifiable risk, whilst the Sharpe Ratio is based on total risk. Despite these differences, generally the Sharpe Ratio provides a similar pattern in the ranks of the investment trusts. Thus empirically, these results suggest that if we are only interested in ranks between the performance of the different portfolios, the Sharpe Ratio, which is a simple and straight forward measure, may be good enough.

Finally, we can see that the HM measure provides quite different results; for example, *TR Property* which is ranked between 11th to 13th by the other measures is now ranked in the 3rd by the HM measure. The statistical properties in table 1 do not show any particular pattern in the returns of *TR Property* but we find that the estimate of coskewness ( $\gamma_{pm}$ ) of this investment trust is smallest amongst the others (-2.02) implying that the HM measure for the trust will be increased.<sup>6</sup>

We also can see that some investment trusts appear to perform better using the HM measure, i.e., *Edinburgh, Fleming Mercantile, Govett Strategic, City of London, Merchants, Secutities Trust of Scotland, Murray Income, and TR Property*, whilst *Henderson Smaller Companies, Fleming Claverhouse, Templer Bar, and Throgmorton* appear worse under the HM measure. Generally, investment trusts appear to perform better than the benchmark when the higher moment systematic risk such as the coskewness is taken into account. This suggests that the managers of these trusts maintain their portfolios better in the presence of large negative and positive shocks than can be explained in the mean-variance world.

So we suggest that when returns are normal and the portfolio managers show no timing ability, then four measures, Jensen's Alpha, TM, HM, and PPW are likely to provide very similar results. However, given the evidence for non-normality (as shown in table 1) then there are significant differences between the HM measure and the other three. These results are consistent with those of Hwang and Satchell (1998) who found significant differences between the measures for highly non-normal emerging market returns. The slight differences between the three other measures, Jensen's Alpha, TM and PPW, indirectly suggest that there is no significant timing ability shown by these portfolio managers.

---

<sup>6</sup>Note that using the estimates of  $\sigma_m$ ,  $\gamma_m$ , and  $\theta_m$  in table 1, we find that  $\lambda_2$  is always negative. Thus for a given beta, any large negative coskewness will increase the value of the HM measure.

### 3.3 Time Varying Properties of the Performance of the UK Investment Trusts

The results in table 2 do not show us how the fourteen investment trusts perform over time and hence under different market conditions and this would seem to be operationally important since many organisations examine their performance on a daily and hence a dynamic basis. Are the ranks between the portfolios relatively stable or do they change over time as market conditions change? If we find that performance does not change dramatically over time then we could potentially construct a hedge portfolio with the portfolios and obtain excess returns.

The five performance measures are now calculated for each investment trust using rolling windows of 60 monthly returns. That is, the first value is calculated using the first 60 monthly returns, i.e. January 1980 to December 1984, and the second is obtained using the 60 monthly returns from February 1980 to January 1985 and so on. Using this approach, we obtain a time series for each performance measure that consists of 195 monthly observations from December 1984 to February 2001 for each investment trust.

Figure 1 provides examples of estimates of the five performance measures over time and table 3 reports some statistical properties of the measures. First of all, the relationship between the performance measures can clearly vary quite widely over time. Note in particular that the Sharpe Ratio can be seen generally to have noticeably moved down during the 1987 crash and then up with the UK's withdrawal from EMU in 1992, but for some trusts can be relatively unaffected by the Russian Crisis in August 1998.

The other relative measures, except the HM measure, do not show particularly large changes around the market crash. Figure 1 indicates that the HM measure for *Edinburgh* shows a big downside movement around the 1987 crash while the other investment trusts show a much smaller reaction. This suggests that the HM measure does not always respond in the same way to negative (positive) shocks but critically it depends on the coskewness<sup>7</sup>. Over the entire sample period, the investment trust performs better with the HM measure.

The statistical properties of the measures reported in table 3 show little difference between Jensen's Alpha, TM and PPW; the measures are either negatively skewed (*Invesco*), platykurtic (*Edinburgh*), or both (*TR property*). Figure 2 shows a typical example of the empirical kernel density estimates for the performance measures for the case of the *TR Property Trust*. The

---

<sup>7</sup>The high ranking of *TR Property* with the HM measure in table 2 can be seen in figure 3B.

effectively common density shown by Jensen’s Alpha, TM and PPW can be compared with the distinctly different distributional shapes presented by the Sharpe Ratio and the HM measure. In addition, since they are highly auto-correlated. and cross-correlated these measures move together over time and there really is relatively little difference between these three measures. The correlation matrices in table 3 also show that in most cases the correlation coefficients between these three measures are larger than 0.99 although we have to be very wary of the utility of correlation as a measure of dependence given the apparent non-normal distribution of the performance measures. Once again, we find little evidence of timing ability in the investment trusts.

However, the statistics of the Sharpe Ratio are again different from those of these three measures. For the *Edinburgh Trust*, the Sharpe Ratio is not significantly correlated with any of the three measures nor the HM measure, whilst for *TR Property* and *Invesco* it is highly correlated with all the others. The estimates of skewness and excess kurtosis suggest that normality may only be assumed for the Sharpe Ratio and then only for a subset of the trusts. The HM measure is relatively less correlated with the three measures above with estimates of the cross-correlation coefficients around or less than 0.9.

Finally, we provide a lateral comparison of the performance for eight investment trusts by the performance measure in Figures 3A to 3C<sup>8</sup>. These plots show that there is no one investment trust which always performs better than the others. *Murray Income Trust*, which is the best among the fourteen investment trusts for all five performance measures in table 2, belonged to the top group during 1991 to 1995, but after 1997 this is no longer true. However, since the ranking does not change rapidly, we might choose an investment trust which has performed well and expect it to perform well in the future if the horizon is relatively short.

Figure 4 shows a typical multivariate scatter plot over time of the performance measures, in this case for the *City of London Trust* and we can again see the clear dependence structure. Jensen’s Alpha, Treynor and the PPW measures again essentially provide identical information over time whereas the other two measures, HM and SR differ from these three in different ways.

Inspecting the time series plots in Figure 1 we can see the different patterns taken by the Sharpe Ratio, Jensen’s Alpha and HM measures. First of all, while the HM measure behaves similarly to Jensen’s Alpha it seems to be more volatile than Jensen’s Alpha over time. The volatility in the HM measure supports the view that it is more sensitive to large shocks in the market than Jensen’s Alpha. In addition, the estimates of the HM measure

---

<sup>8</sup>Note that we only report results for Jensen’s Alpha among the set Jensen’s Alpha, TM and PPW since these three appear to show such little difference.

are spread more widely than those of Jensen's Alpha which are in turn spread more widely than those of the Sharpe Ratio.

The Sharpe Ratio seems also to be sensitive to big market movements such as the market crash of 1987, Sterling's withdrawal from EMU in 1992, the Russian crash in 1998, etc. The effect of these episodes on the Sharpe Ratio is stronger than on the other measures since all except the Sharpe Ratio are relative measures and thus far less sensitive to jumps in the benchmark.

Interestingly, Jensen's Alpha and the HM measure both show that the UK investment trusts performed well up to 1993, but soon after that point, the performance measures begin to show a dramatic decrease until 1999. When we assume that the betas of the investment trusts are very close to one, this means that during this period it became very difficult to out perform the benchmark. On the other hand, during the early sample period, i.e., early 1980's, beating the benchmark was probably easier than in late 1990's. However, these indications of performance relative to the benchmark portfolio cannot be seen in Sharpe Ratio and the movement in Sharpe Ratio roughly matches the UK market movement over the same period.

This fairly standard form of descriptive analysis of the patterns of behaviour identified above between the different performance measures can only provide casual insight. We need to move beyond correlation analysis to properly assess their interdependence given the non-gaussianity of the performance measures shown in Figure 2. We now turn to consider how the use of copula functions can help us to quantify the relationships between the performance measures by assessing their statistical dependence more accurately.

## 4 Copulae

### 4.1 A Brief Introduction to Copulae

A copula is simply a function that links univariate marginals to their joint multivariate distribution or alternatively it is a joint distribution function with uniform marginals. Such a function is simply defined as follows:

$$C(u_1, u_2, \dots, u_N) = \Pr[U_1 \leq u_1, U_2 \leq u_2, \dots, U_N \leq u_N]$$

with  $U_1, U_2, \dots, U_N$  uniform random variables. Suppose we have a portfolio with  $N$  assets whose returns follow univariate marginal distribution functions  $F_1(x_1), F_2(x_2), \dots, F_N(x_N)$  then the copula will describe their joint distribution. The copula function  $C$  combines or couples the marginals together to

give their joint density such that:

$$C(F_1(x_1), F_2(x_2), \dots, F_N(x_N)) = F(x_1, x_2, \dots, x_N) \quad (5)$$

given the univariate marginal

$$F_i(x_i) = C(F_1(+\infty), F_2(+\infty), \dots, F_i(x_i), \dots, F_N(+\infty))$$

This follows directly from;

$$\begin{aligned} C(F_1(x_1), F_2(x_2), \dots, F_N(x_N)) &= \Pr[U_1 \leq F_1(x_1), U_2 \leq F_2(x_2), \dots, U_N \leq F_N(x_N)] \\ &= \Pr[F_1^{-1}(U_1) \leq x_1, F_2^{-1}(U_2) \leq x_2, \dots, F_N^{-1}(U_N) \leq x_N] \\ &= \Pr[X_1 \leq x_1, X_2 \leq x_2, \dots, X_N \leq x_N] \\ &= F(x_1, x_2, \dots, x_N) \end{aligned}$$

Conversely, for a *given* multivariate distribution, there exists a copula function that links its marginals. Moreover Sklar (1959) proved that if the marginal distribution functions are continuous then we can be assured that the copula is *unique*<sup>9</sup>.

Since the multivariate distribution contains *all* the information that exists on the dependence structure between the variables, in our case the different performance measures, the copula must contain precisely the same information<sup>10</sup> and hence it captures exactly how the different performance measures are related to each other. Moreover since the copula is defined on the transformed uniform marginals it holds this information on dependence *irrespective* of the particular marginals of the underlying performance measures. So a simple procedure to analyse the multivariate distribution or dependence between the non-gaussian performance measures would be to start by determining the marginal distribution relevant to each measure and then estimate the relevant copula from the data to give the multivariate distribution of the performance measures. Then given the estimated copula we can move to consider exactly where in their range spaces the different performance measures will provide different signals of portfolio performance. In other words where they may be jointly dependent or relatively independent.

Finance has traditionally assumed a multivariate Gaussian distribution for returns and in terms of copulae, this is equivalent to assuming that (i) the marginal density functions of each asset's returns is Gaussian and (ii) that the copula that links univariate marginals is a particular copula, in

---

<sup>9</sup>For a detailed discussion of copulae and their applications in statistics see, Joe (1997) or Nelsen(1999).

<sup>10</sup>Since it cannot be held in the marginal distributions.

fact a gaussian copula. If we assume two random variables  $X \sim N(0, 1)$ ,  $Y \sim N(0, 1)$ , with correlation coefficient  $\rho(X, Y) = \rho$  and if their joint distribution is bivariate gaussian, then  $F(X, Y) = C_\rho^{Gauss}(\Phi(x), \Phi(y))$  with  $C_\rho^{Gauss}$  being the gaussian copula such that for  $(u, v) \in [0, 1]^2$ :

$$C_\rho^{Gauss}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right) ds dt$$

However if we chose  $C \neq C_\rho^{Gauss}$ , the joint distribution function  $F(x, y) = C(\Phi(x), \Phi(y))$  will no longer be multivariate gaussian but will still be a well defined distribution. The Gaussian is one of a large number of parametric copula that could be used to join marginals and an important issue is the statistical identification of such copula. In the gaussian copula considered here there is a single parameter,  $\rho$ , the correlation coefficient, that simultaneously characterises the copula *and* the dependency between the variables. More general copulae may be parameterised or defined by several parameters and different measures of dependency may be expressed as different functions of these parameters. Measures of dependency need not necessarily be easily expressed as functions of these parameters but may instead be defined as functions of the copula map itself. A two parameter copula is, for example, given by

$$\begin{aligned} C(u, v; \theta, \delta) &= \{1 + [(u^{-\theta} - 1)^\delta + (v^{-\theta} - 1)^\delta]^{\frac{1}{\delta}}\}^{-\frac{1}{\theta}} \\ &= \eta(\eta^{-1}(u) + \eta^{-1}(v)) \end{aligned}$$

where  $\eta(s) = \eta_{\theta, \delta}(s) = (1 + s^{\frac{1}{\delta}})^{-\frac{1}{\theta}}$  and the lower tail area dependency measure is  $2^{-\frac{1}{\delta\theta}}$  and the upper tail area dependency measure is  $2 - 2^{\frac{1}{\delta}}$  (which is independent of  $\theta$ ). More generally having more than one parameter facilitates the measurement of different types of dependency.

A central result is that if the random variables  $X_1, \dots, X_n$  are independent then the copula function that links their marginal is the *product copula*:

$$C(F_1(x_1), F_2(x_2), \dots, F_N(x_N)) = F(x_1)F(x_2) \dots, F(x_N)$$

and tests for independence can be based on the distance of the empirical copula to this product copula. More generally the copula is a function defined over the range of the random variables when transformed into a uniform  $[0, 1]$  space and hence we can examine the varying dependence structure throughout the entire range of the potential variation of the performance measures. This is in contrast to the use of a single number such as the correlation which

is assumed to apply globally and to accurately measure a common degree and form of dependence throughout the entire range of values taken by the variables. An assumption which is only valid if the performance measures were distributed elliptically.

## 4.2 Measuring Dependency Using Copulae

Let us start by briefly recall the failure of correlation as a measure of association.

### 4.2.1 The Inadequacy of Correlation

The standard definition of the Pearson Correlation coefficient is

$$\rho = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y}$$

with

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{and} \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i.$$

and

$$\sigma_x^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad \text{and} \quad \sigma_y^2 = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2.$$

As pointed out by Embrechts *et al* (1999), the success of correlation is due to its properties for *linear* analysis, especially for the linear combination of risk factors  $F$ , where  $Var(\alpha^\top F) = \alpha^\top Cov(F)\alpha$ . It provides however only a measure of *linear* association. There are many problems with using correlation to describe multivariate dependence in more general situations however. In particular;

- $\sigma_1^2$  and  $\sigma_2^2$  have to be finite for  $\rho_{12}$  to be defined. For example, consider an extreme value type II (Fréchet) distribution with parameter  $\tau = -\alpha^{-1}$  so that  $\int_0^\infty x^r dF_X(x) = \infty$  for  $r > \alpha$ . Correlation is then not defined in this quite reasonable and important case for financial applications.
- independence always implies zero correlation but the converse is only true for an elliptic distribution such as the Gaussian,
- correlation is not an *invariant* measure whereas the copula function is invariant to strictly monotone transformations.

The fundamental reason why the Pearson correlation fails as an invariant measure of dependency is that it depends not only on the copula but also on the marginal distributions of the data. Thus the correlation is changed by changes of scale and other non-affine transformations in the marginal variables and therefore the units in which we express our data. It is formally not a geometric quantity.

### 4.3 Concordance: Scale Invariant Dependence Measures

Despite the impression created by the common usage of correlation to measure dependence it is in fact often far from straight forward to define exactly what form of dependence we are interested in and then to select a statistic that captures exactly what we need to measure. As emphasised above the desire to use a measure that is scale invariant leads naturally to copula based measures since the copula captures those properties of the joint distribution which are invariant under almost surely strictly increasing transformations. So scale invariant measures of dependency will be expressible solely in terms of the copula of the random variables. The most widely used scale invariant measures of association are *Kendall's tau* and *Spearman's rho* both of which measure *concordance*. Concordance between two random variables arises if large values of one variable tend to occur with large values of the other and small values occur with small values of the other; otherwise they are said to be discordant. So concordance picks up nonlinear associations between the performance measures that correlation might miss completely.

If  $(X_1, Y_1)$  and  $(X_2, Y_2)$  are independent and identically distributed random vectors with possibly different joint distribution functions  $H_1$ , and  $H_2$  with copulae  $C_1$  and  $C_2$  respectively but with common margins. The population version of Kendall's tau is defined as the probability of concordance minus the probability of discordance,

$$\begin{aligned}\tau &= \tau_{X,Y} \\ &= P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]\end{aligned}$$

Nelsen shows that this may be re-expressed in terms of the copulae as

$$Q = Q(C_1, C_2) = a \iint_{I^2} C_2(u, v) dC_1(u, v) - 1$$

Spearman's rho is defined as follows. Let  $R_i$  be the rank of  $x_i$  among the  $x$ 's and  $S_i$  be the rank of  $y_i$  among the  $y$ 's. The Spearman rank order

correlation coefficient is:

$$\rho_S = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

which again may be expressed in terms of copulae as

$$\rho_C = 12 \int \int_{[0,1]^2} (C(u, v) - uv) dudv.$$

Spearman's rank correlation coefficient is essentially the ordinary correlation of  $\rho(F_1(X_1), F_2(X_2))$  for two random variables  $X_1 \sim F_1(\cdot)$  and  $X_2 \sim F_2(\cdot)$ . Notice the explicit contrast with the product copula in this case. Essentially these two measures of concordance measure the degree of *monotonic* dependence as opposed to the Pearson Correlation which simply measures the degree of *linear* dependence. They both achieve a value of unity for the bivariate Fréchet upper bound (one variable is an increasing transformation of the other) and minus one for the Fréchet lower bound (one variable is a strictly decreasing transform of the other). These two properties do not hold for the standard correlation coefficient making these two concordance measures more attractive as general measures of dependency or association.

Table 4 provides a comparison between the measures of concordance and correlation between the performance measures for the *Fleming Calverhouse Trust*. What is clear from this comparison is that simple correlation analysis suggests that *all* the measures are significantly positively related while both concordance measures clearly suggest a lack of concordance between the Sharpe Ratio and all the other measures<sup>11</sup>. This implies that we should expect all the performance measures to move in the same direction, either linearly or nonlinearly, *except* the Sharpe Ratio. While standard correlation analysis suggests all are linearly related, Kendall's  $\tau$  and Spearman's  $\rho$  indicate that there is no monotone dependency between the Sharpe Ratio and the other measures. Superficially these results might appear to be contradictory but really serve to question the value of correlation analysis on data that is non-gaussian more importantly for our current objectives it allows us to formally isolate the distinction between the use of relative and absolute risk measures in performance analysis. It is also clear once again that the Treynor, Jensen's Alpha and the PPW measure are very tightly related whether we think in terms of correlation or concordance. There is a weaker association between HM and these three but it is still significantly different from zero.

---

<sup>11</sup>Similar results were found for the other Investment Trusts.

We could from this point consider a range of alternative dependency measures between the performance measures based on copulae such as positive quadrant dependence or positive function dependence ( see Joe(1997)) but have instead chosen to examine the relationship between the measures as they take large or small values since we feel this is case that it is likely to be most important in practice. A manager needing to make a decision based on when one performance measure is taking extreme values might be reassured by the knowledge that the same signal is likely to be provided by some other performance measure.

### 4.3.1 Tail Area Dependence

The relationship between the performance measures under *extreme* market conditions, and hence when the measures themselves are likely to take *extreme* values, can be captured by examining the potential common behaviour in the tails of their distributions or in other words notions of *tail area dependency*. If two random variables, in our case performance measures,  $X$  and  $Y$  follow marginal distribution functions  $F_X$  and  $F_Y$  respectively then a standard definition of upper tail dependency,  $\lambda_U$ , is given by

$$\lambda_U = \lim_{u \rightarrow 1} P(Y > F_Y^{-1}(u) | X > F_X^{-1}(u))$$

and the variables will be asymptotically upper tail dependent if  $\lambda_U \in (0, 1]$  and upper tail independent if  $\lambda_U = 0$  provided the limit  $\lambda_U \in [0, 1]$  exists<sup>12</sup>. In other words we look to see if two measures *simultaneously* lie in suitably defined tail areas of their marginal distributions. This form of dependence, which again is quite distinct from correlation in the non-gaussian case, can again be derived directly from the copula function linking the two performance measures (see Joe(1997)) and so will be invariant to scale transformations and independent of the form of the marginal distributions. An equivalent definition is that a copula  $C$  shows upper tail dependence where  $\lambda_U \in (0, 1]$  and when

$$\lambda_U = \lim_{u \rightarrow 1} (1 - 2u + C(u, u)) / (1 - u).$$

Poon, Rockinger and Tawn (2001) have recently discussed applications of (upper) tail dependence measures in finance and considered  $\chi$  defined by

$$\chi = \lim_{s \rightarrow \infty} P(Y^* > s | X^* > s),$$

---

<sup>12</sup>Similar definitions hold for lower tail dependency.

where  $0 \leq \chi \leq 1$ , and the value of  $s$  is the extreme event for each variable, and

$$\begin{aligned} X^* &= -\frac{1}{\log F_X(X)} \\ Y^* &= -\frac{1}{\log F_Y(Y)} \end{aligned}$$

are the original random variables transformed to unit Fréchet marginals so that they are defined on a common scale and events of the form  $\{X^* > s\}$  and  $\{Y^* > s\}$  correspond to equally extreme events for each variable. In other words

$$P(X^* > s) = P(Y^* > s) \sim s^{-1} \text{ as } s \rightarrow \infty$$

and  $X^*, Y^*$  possess the same dependence structure as  $(X, Y)$ . When  $\chi > 0$ , the two variables are *asymptotically* tail area dependent since  $\chi$  measures dependence that is persistent in the limit. However, when  $\chi = 0$ , the two variables are *asymptotically* independent but not necessarily *exactly* independent and the distinction is important; consider for instance a bivariate normal case with any value of the correlation coefficient less than 1 which would imply  $\chi = 0$ . Exact independence implies

$$P(X^* > s | Y^* > s) = P(X^* > s)$$

which clearly goes to zero as  $s \rightarrow \infty$ . When there is exact dependence traditional extreme value methods will impose asymptotic dependence regardless of whether or not the true distribution shows asymptotic independence and hence they will over estimate  $P(X^* > s, Y^* > s)$  and other probabilities of joint extreme events and hence a bias emerges.  $\chi$  describes the degree of asymptotic dependence if the variables are asymptotically dependent and it will be zero for all asymptotically independent variables so  $\chi$  cannot describe the *degree* of asymptotic independence

Coles et al (1999) have therefore developed a different measure,  $\bar{\chi}$ , which is defined by

$$\bar{\chi} = \lim_{s \rightarrow \infty} \frac{2 \log \Pr(X^* > s)}{\log \Pr(Y^* > s, X^* > s)} - 1,$$

where  $-1 < \bar{\chi} \leq 1$  and this provides an accurate measure of asymptotic independence since it describes the rate that  $P(X^* > s | Y^* > s)$  goes to zero. When  $\bar{\chi} = 0$ , the two variables are independent in the tails, and when  $\bar{\chi} < 0$  ( $\bar{\chi} > 0$ ), the two variables may be interpreted as being negatively (positively) associated.  $\chi$  and  $\bar{\chi}$  together therefore provide all we need to assess the degree of association between the performance measures in their tails;  $\chi$  for asymptotic dependence and  $\bar{\chi}$  for asymptotic independence.

If we define  $Z = \min(X^*, Y^*)$ , we can estimate  $\bar{\chi}$ , following Poon et al (2001), by using the standard Hill estimator as

$$\hat{\chi} = \frac{2}{n_u} \sum_{j=1}^{n_u} \log\left(\frac{z_j}{u}\right) - 1$$

$$\text{var}(\hat{\chi}) = \frac{(\hat{\chi} + 1)^2}{n_u},$$

where  $z_j$  are those,  $n_u$ , observations that exceed  $u$ .  $\chi$  is estimated by

$$\hat{\chi} = \frac{un_u}{n}$$

$$\text{Var}(\hat{\chi}) = \frac{u^2 n_u (n - n_u)}{n^3}$$

If  $\hat{\chi}$  is significantly less than 1 (ie. if  $\hat{\chi} + 1.96\sqrt{\text{var}(\hat{\chi})} < 1$ ) then the inference is that the performance measures are asymptotically independent and  $\chi$  is taken to be zero. Only if there is no significant evidence to reject  $\bar{\chi} = 1$  is  $\chi$  then estimated (under an assumption that  $\bar{\chi} = 1$ ).

Table 5 provides estimates of  $\bar{\chi}$  for the five performance measures for the *Fleming Claverhouse Trust*. We took the upper (lower) 2% of the total observations (given the 195 monthly observations implies 4 observations) to define the extreme case<sup>13</sup>. The table shows that for the left tail, the Sharpe Ratio is negatively associated with Jensen's Alpha, TM and PPW, but independent of HM. On the other hand, Jensen's Alpha, TM and HM *do not show* any significant *asymptotic* left tail dependency. This result would seem to be important and indicates the value of this analysis since it implies all the dependency we have already noted between these performance measures comes about when they take values from within the body of their distributions or from their right tail and not from their left tail behaviour, when they indicate poor performance.

The results for the right tail, however are quite different from those of the left tail. None of them are significantly different from zero and thus we conclude that all performance measures are independent in the right tail, in other words when we see large performance values.

Therefore on the basis of these results, when performance measures such as Jensen's Alpha, TM and PPW show very large decreases the Sharpe Ratio

---

<sup>13</sup>We also estimated the independency measure for 4% and found that the results are similar to those in table 4. We recognise that the results in table 4 may be affected by the small number of observations.

will tend to increase whilst the HM would not show any particular pattern. On the other hand, when any performance measure indicates that the portfolio performs very well, we will tend not to find any particular pattern between any of the measures despite the fact that correlation analysis indicates a very high association between Jensen's Alpha, TM and PPW.

#### 4.4 Quantile Regressions

Given how these tail area dependency results differ so strongly from the standard correlation results in Table 3 we extended this form of analysis in order to investigate how the conditional dependence between the performance measures may vary throughout the entire range of their conditional distributions. Regression methods are one of the most common tools used to capture multivariate dependency. However given that the performance measures are non-Gaussian it is far from obvious that we should be interested in the conditional mean of the dependent variable rather than some other function of their conditional density. We may instead be interested in the relationship at particular quantiles, say the median, and it is then natural to consider computing quantile regressions and again since the copula captures the entire joint distribution it can be used to make this a relatively easy exercise. Quantile regression thus enables us to explore the conditional dependence of each performance measure given another performance measure's value at any particular range of quantiles and in this way we can extend the tail area dependency analysis into the body of the conditional distribution and explore the dependency structure within the entire conditional density of a performance measure.

If we assume an Archimedean form of the copula<sup>14</sup> so that the conditional distribution of  $Y$  given  $X_1, \dots, X_k$  is given by

$$F_Y(y | x_1, \dots, x_k) = \frac{\phi^{-k} \{c_k + \phi [F_y(y)]\}}{\phi^{-k}(c_k)}$$

where  $c_k = \phi [F_1(x_1)] + \dots + \phi [F_k(x_k)] + \phi [F_Y(y)]$ . Elementary statistics show us that the regression function may be rewritten as

$$E(y | x_1, \dots, x_k) = \int_0^\infty [1 - F_Y(y | x_1, \dots, x_k)] dy + \int_{-\infty}^0 [F_Y(y | x_1, \dots, x_k)] dy$$

---

<sup>14</sup>A general family of copulae is the archimedean form in which  $C_\phi(u, v) = \phi^{-1}(\phi(u) + \phi(v))$  for  $u, v \in (0, 1]^2$  where  $\phi$  is a convex decreasing function with domain  $(0, 1]$  and range  $[0, \infty)$  such that  $\phi(1) = 0$ . Several standard copula belong to this family for different choices of generator function  $\phi$  see Nelsen(1998).

Genest(1987) has shown that using Frank's Copula ( with an Archimedean generator function  $\phi(t) = \ln(\frac{e^{\alpha t}-1}{e^\alpha-1})$  ) we can write the regression function directly as

$$E(X_2 | X_1 = x) = \frac{(1 - e^{-\alpha})xe^{-\alpha x} + e^{-\alpha}(e^{-\alpha x} - 1)}{(e^{-\alpha x} - 1)(e^{-\alpha} - e^{-\alpha x})}$$

Instead of calculating the conditional mean function we can compute the median or any other quantile of the conditional distribution in a similar way. If we define the  $p$ 'th quantile to be the solution  $x_p$  of the equation

$$p = F_Y(y_p | x_1, \dots, x_k)$$

or for the bivariate case

$$p = F_Y(y_p | X_1 = x) = C_1 [F(x), F_Y(y_p)] \quad (6)$$

where  $C_1$  is the partial derivative with respect to the first argument in the copula. So for a specified proportion  $p$  and  $x$  value we can solve equation(6) for the required percentile. In the examples below we have been restricted by software constraints to imposing a linear functional form on these quantile regressions.

We have examined the conditional dependence between the different performance measures by running bivariate quantile regressions at the 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95% and 99% levels. The results are reported below in table 6 for the case of the *Fleming Claverhouse Trust*. Note that the column heads indicate the explanatory variable in explaining the conditional quantile of the indicated performance measure for the relevant table and the stars indicate significant coefficient values at the 5% level<sup>15</sup>.

From Panel A of Table 6 we can see that the Sharpe Ratio only really provides any explanatory power for Jensen's Alpha in the upper half of the distribution and this is most strong in the extreme tail ( the 99th quantile) . On the other hand the Higher Moment measure indicates significant positive dependence throughout the distribution.

Panel B of Table 6 indicates virtually no power in explaining the Higher Moment measure by the Sharpe Ratio except at the 75th and 90th quantiles whereas Jensen's Alpha would seem to be significantly related to the higher moment measure almost throughout the entire distribution.

Finally Panel C of Table 6 shows some conditional dependence in explaining the Sharpe Ratio by Jensen's Alpha above the 90% quantile and

<sup>15</sup>Notice that these quantile regression results may be regarded as estimates of *exact* rather than *asymptotic* quantile dependency as developed above with the tail area dependency measure.

the higher moment measure is only significant around the 25th and 50th quantiles. These results reinforce the conclusions drawn earlier from the tail area dependency measure but provide deeper insights as to when we may expect to draw conflicting conclusions regarding performance by adopting one measure rather than another. Again the Sharpe Ratio appears as an outlier providing relatively limited explanation for the other two measures and in turn only really showing joint dependence with Jensen’s Alpha in the upper tail of the bivariate distribution. Jensen’s Alpha is seen to be related to the HM measure through out the body of the conditional distribution except in the right hand tail confirming the previous tail area dependency result.

## 5 An Aggregate Performance Measure

It has been shown quite widely in the literature (see for instance Diebold(1997)) that combining forecasts generated from different models can reduce offsetting biases and also that the combined forecast can have a lower variance given the interdependence between the forecasts. These advantages clearly seem to be worth pursuing within the context of performance measurement through the proper construction of an aggregate of several potentially competing performance measures. As with forecast combination it may not make sense to combine forecasts that arise from different models that are theoretically inconsistent but when this is not the case the view may be taken that different models simply reflect parts of a more complex data generation process and may legitimately be combined. From this point of view we would exclude the Sharpe Ratio from consideration in the aggregate performance measure below since it differs from the others we have considered in that it is not a relative measure. Moreover given the common information being generated by Jensen’s Alpha, the Treynor and PPW measures we will just consider the question of how to construct an aggregate performance measure from one of these, Jensen’s Alpha and the Higher Moment measure.

Jouini and Clemen (1996) have proposed the use of copula functions to aggregate expert opinions in a decision problem and we shall apply their methodology as an illustration of the method below. There are a range of statistical issues which need to be pursued in following up our suggestion to construct an aggregate performance measure which we do not have space to resolve here but we believe the approach is powerful and worth pursuing. If we then take Jensen’s Alpha and the HM measure to represent the evaluations of two “experts” as to the true underlying performance of a fund and hence describe the methodology for the bivariate case but the approach can easily be extended to consider the aggregation of a greater number of under-

lying performance measures if required. Each of the separate performance measures follows a distinct distribution, say  $F_1(x)$  and  $F_2(x)$  and our problem is then essentially to aggregate these distributions taking into account their interdependence. This sort of question has a long history in statistics see Genest and Zidek (1986) and has been applied in many areas. The simplest approach is to construct an aggregate distribution as an average  $\bar{F}(x) = [F_1(x) + F_2(x)]/2$  but this ignores the relative accuracy of the two estimates and the fact that they may be dependent and this is clearly the case in our context from the results given above. It is not straight forward to account for the dependence in any aggregation procedure but given the properties of copulae, outlined above, it is obvious that they capture precisely the information required. Jouini and Clemen adopt a Bayesian approach where the decision maker has access to the historical record and hence empirical distributions of the separate performance measures which he uses to derive the posterior distribution for the aggregate performance measure. The decision maker's problem is essentially one of constructing a likelihood function that brings together the information from the experts and then, by applying Bayes Theorem to the likelihood function and prior, to derive the required posterior distribution from which the aggregate performance measure may be calculated together with any required confidence intervals for a systematic analysis of performance.

Given the observed non-gaussianity of the performance measures shown in figure 2 above and following Jouini and Clemen we could regard the different observed performance measures as being median based estimates of the underlying true measure of performance  $x$  and there will be some correlation  $r_x$ , between the errors made by the two different performance estimates and the underlying measure. The posterior distribution for  $x$ , the underlying measure of performance, given two expert distributions  $F_1(x)$  and  $F_2(x)$  with densities  $f_1(x)$  and  $f_2(x)$  can be written as

$$f_{dm}(x|f_1, f_2) = f_1(x)f_2(x)c[1 - F_1(x), 1 - F_2(x)]$$

where  $c[\cdot]$  is the copula density function that captures the dependence structure between the two performance measures. One of the outstanding statistical issues that needs to be resolved before implementing this approach is the proper choice of copula in aggregating performance measures. For demonstration in what follows we use a gaussian copula for simplicity, but notice that this assumption in no way implies that the joint distribution of the separate performance measures is bivariate normal since their marginal distributions are clearly non-gaussian. Using the Gaussian copula then gives

us the following relatively simple formula for the posterior;

$$f_{dm}(x|f_1, f_2, r_x) = \frac{f_1(x)f_2(x)}{(1-r_x^2)^{1/2}} \times \exp \left\{ \begin{array}{l} -r_x(r_x(\Phi^{-1}[1-F_1(x)])^2 - 2\Phi^{-1}[1-F_1(x)]\Phi^{-1}[1-F_2(x)] + \\ r_x(\Phi^{-1}[1-F_2(x)])^2)/2(1-r_x^2) \end{array} \right\}$$

Where  $\Phi^{-1}[\cdot]$  is the inverse of the standard normal cumulative distribution function with correlation coefficient  $r_x$ . This posterior distribution then provides all the information needed to construct the aggregated performance measure and to conduct inference on it in any decision framework. The question now turns on how to extract the aggregate performance measure from this distribution and this depends on the specification of the decision maker's utility or loss function. We could simply use the mean or the median of this distribution but as emphasised by Christoffersen, P. and Diebold, F.X. (1997), what function of the non-gaussian posterior distribution serves as the optimal estimator of the underlying performance measure will depend critically on the asymmetric loss function the fund manager is almost certainly going to hold. In particular we expect that fund managers would be substantially more loss averse than for an equivalent profit on the up side. This issue is much more complex than can be developed here but see Hwang and Salmon (2001) for an extended discussion of performance measure aggregation.

## 6 Conclusions

We have carried out a fairly detailed comparison of the statistical properties and the relationships between a set of five performance measures using 14 UK based Investment Trusts over a sample period ranging from 1980 to 2001. Our results suggest very clearly that there is almost no difference between Jensen's Alpha, the Treynor-Mazuy (TM) measure and the Positive Period Weighting(PPW) measure over our sample period and amongst our set of 14 Investment Trusts. This would seem to indicate that there is no timing ability within these fund managers. The Sharpe Ratio clearly provides different signals regarding performance than the other measures and is the only absolute measure in the set of measures we have considered. While simple correlation analysis suggests that there is a high degree of dependence between most of the measures we have shown that there is a lack of significant concordance between the Sharpe Ratio and all the other measures. This indicates the inadequacy of correlation analysis with non-gaussian data. We

have also shown that the Sharpe Ratio exhibits negative left tail area dependence with respect to Jensen's Alpha, TM and PPW but is independent in the left tail from the Higher Moment measure of Hwang and Satchell, that is when poor performance is indicated. Jensen's Alpha, TM and the HM measure do not seem to show any significant *asymptotic* left tail dependency. All the measures appear to be *asymptotically* independent in their upper tail when good performance is indicated. These results are further refined by non-asymptotic quantile regression results which indicate finite sample dependency the HM measure and Jensen's Alpha throughout the body of their conditional distribution and in the left tail but not the upper tail.

Given the question we raised at the outset of this work, we have found that there are important statistically significant differences between the performance measures we have analysed and how they behave in different market conditions. A performance manager would have to take care in justifying which criteria he wanted to use to properly measure performance since the results may differ widely depending on his decision. In particular the standard practitioner's choice of the Sharpe Ratio (cf. the AIMR handbook) is an outlier in many ways when compared to the other relative measures in this study. Moreover as can be seen from figures 2 and 3 the Sharpe Ratio provides substantially less discrimination than the other measures we have considered both over time and hence market conditions and over fund style. We have also discussed how to properly construct an aggregate performance measure taking into account the joint dependence of the individual measures.

## References

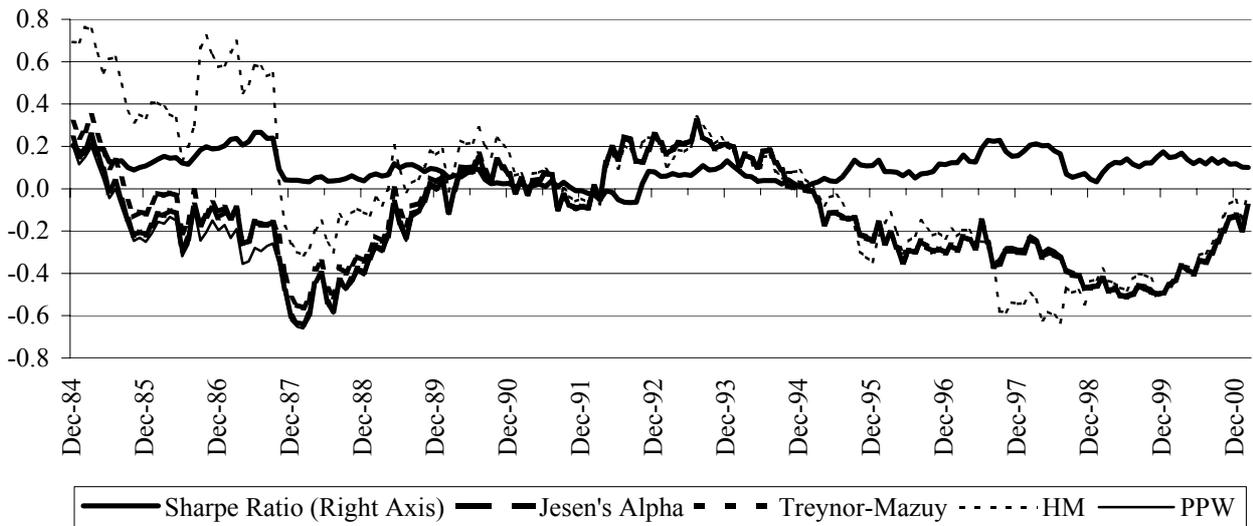
- Admati, A., S. Bhattacharya, P. Pfleiderer, and S. Ross, 1986. On timing and selectivity, *Journal of Finance* 41, 715-730.
- AIMR 1997. *Association for Investment Management and Research, Performance Presentation Standards Handbook*. 2nd Edition.
- Chen, Z., and P. J. Knez, 1996. Portfolio performance measurement: Theory and applications, *Review of Financial Studies* 9(2), 511-555.
- Christoffersen, P. and Diebold, F.X. ,1997. Optimal Prediction Under Asymmetric Loss, *Econometric Theory*, 13, 808-817.
- Cochrane, J. H, 1997. Where is the Market Going? Uncertain Facts and Novel Theories, *Economic Perspectives* XXI:6 (November/December), Federal Reserve Bank of Chicago.
- Coles S.G., J.Heffernan, and J.A. Tawn, 1999. Dependence measures for extreme value analysis, *Extremes* 3, 5-38.
- Connor, G., and R. Korajczyk, 1986. Performance measurement with the arbitrage pricing theory: A new framework for analysis, *Journal of Financial Economics* 15, 374-394.
- Cumby, R., and J. Glen, 1990. Evaluating the performance of international mutual funds, *Journal of Finance* 45, 497-521.
- Diebold F., 1997. *Elements of Forecasting*, Thompson Publishing.
- Embrechts P., A. McNeil and D. Straumann, 1999. Correlation: Pitfalls and Alternatives, *RISK Magazine*, May.
- Genest C., 1987. Frank's family of Bivariate Distributions, *Biometrika* 74, 549-555.
- Glosten, L. R. and R. Jagannathan, 1994. A Contingent Claim Approach to Performance Evaluation, *Journal of Empirical Finance* 1, 133-160.
- Grinblatt, M., and S. Titman, 1989. Portfolio performance evaluation: Old issues and new insights, *Review of Financial Studies* 2(3), 393-421.
- Grinblatt, M., and S. Titman, 1994. A study of monthly mutual fund returns and performance evaluation techniques, *Journal of financial and Quantitative Analysis* 29(3), 419-444.

- Grinblatt, M., and S. Titman, 1995. Performance evaluation, In *Handbooks in Operational Research and Management Science*, Vol.9, R. Jarrow et al. ed. Elsevier Science B. V.
- Hwang, S., and S. Satchell, 1998. Evaluation of Mutual Fund Performance in Emerging Markets, *Emerging Markets Quarterly* 2(3), 39-50.
- Hwang, S., and S. E. Satchell, 1999. Modelling emerging market risk premia using higher moments, *International Journal of Finance and Economics* 4(4), 271-296.
- Hwang and Salmon, 2001. Performance measure aggregation, Working Paper, Financial Econometrics Research Centre, City University Business School, London.
- Jensen, M, 1968. The performance of mutual funds in the period 1945-1964, *Journal of Finance* 23, 389-416.
- Jensen, M., 1969. Risk, the pricing of capital assets, and the evaluation of Investment Portfolios, *Journal of Business* 42, 167-247.
- Joe H., 1997. *Multivariate Models and Dependence Concepts*, Chapman Hall, Monographs on Statistics and Applied Probability No.73.
- Jouini M. N. and R.T. Clemen,1996. Copula models for aggregating expert opinions, *Operations Research*, 44(3),444-457.
- Kraus, A., and R. H. Litzenberger, 1976. Skewness preference and the valuation of risk assets, *Journal of Finance* 31(4), 1085-1100.
- Lehmann, B., and D. Modest, 1987. Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons, *Journal of Finance* 42, 233-265.
- Nelsen R., 1999. *An Introduction to Copulas*, Springer, Lecture Notes in Statistics, No. 139.
- Roll, R., 1978, Ambiguity when performance is measured by the securities market line, *Journal of Finance* 33, 1051-1069.
- Poon Ser-Huang, M. Rockinger and J. Tawn, 2001. New Extreme Value Dependence Measures and Financial Applications, mimeo.
- Sharpe, W., 1966. Mutual fund performance, *Journal of Business* 39, 119-138.

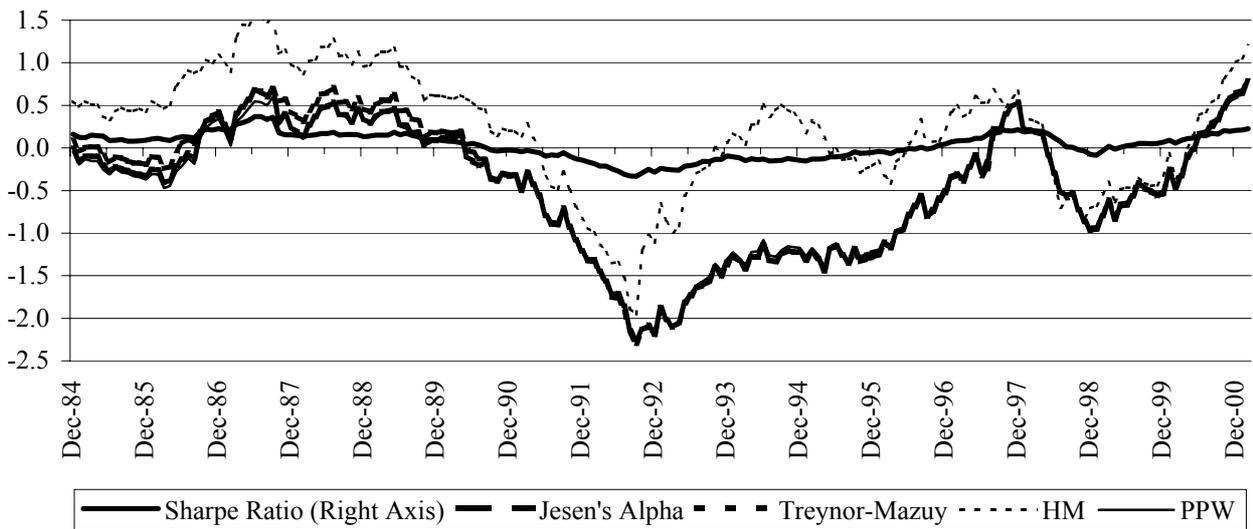
Treynor, J., 1965. How to rate management of investment funds, *Harvard Business Review* 43, 63-75.

Treynor, J., and K. Mazuy, 1966. Can mutual funds outguess the market?, *Harvard Business Review* 44, 131-136.

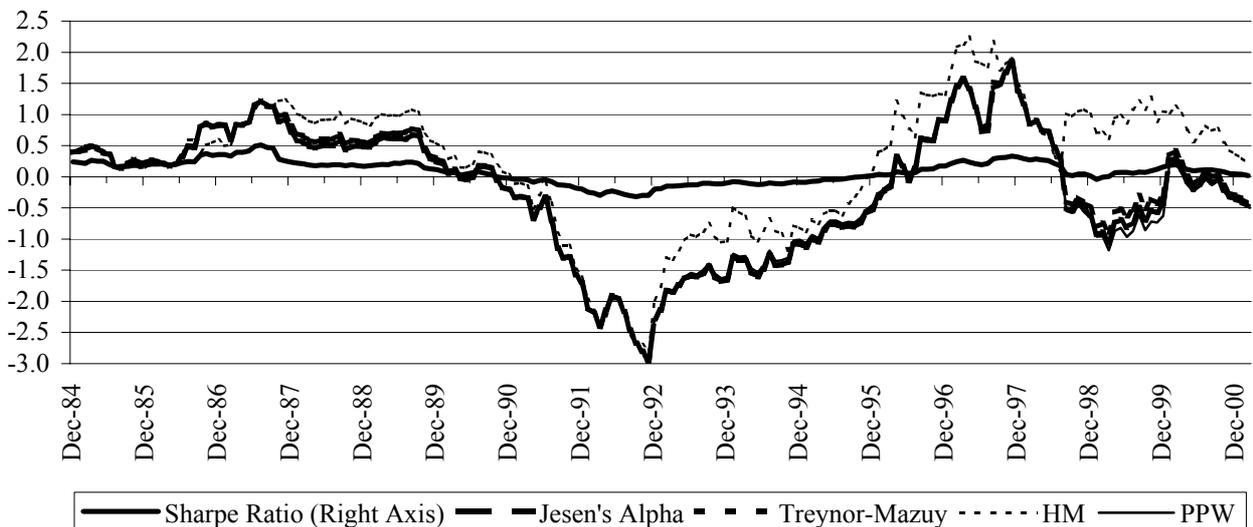
**Figure 1A Edinburgh**



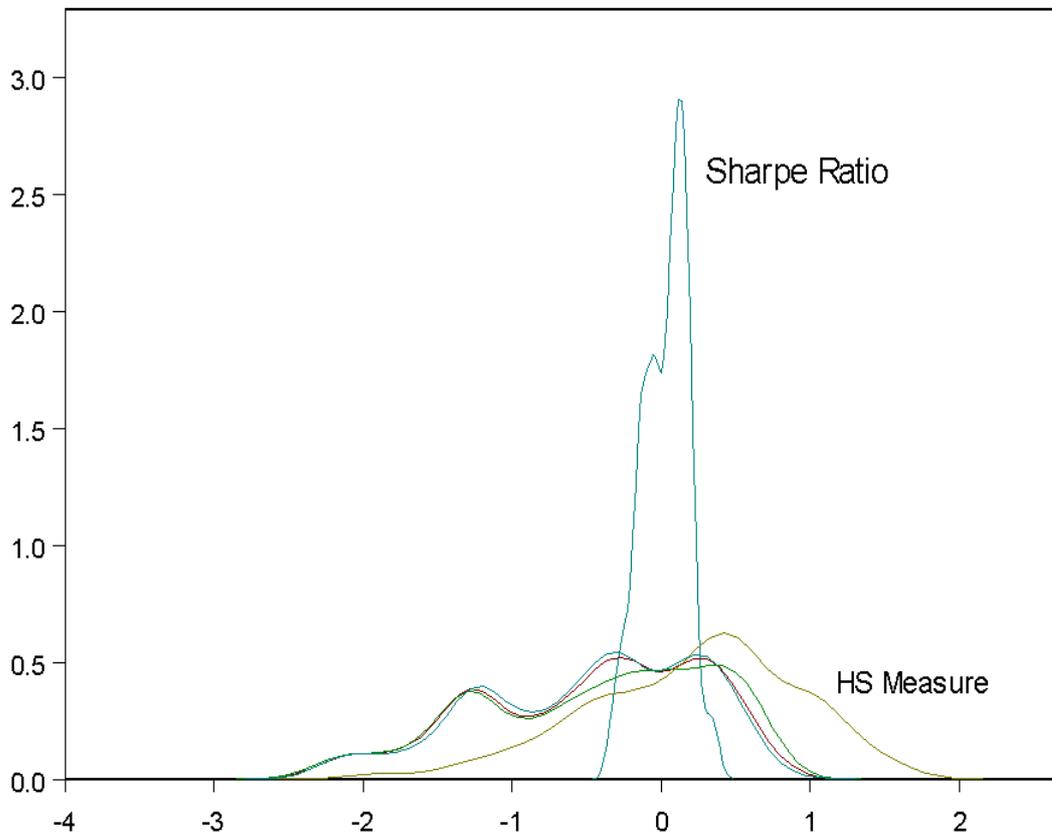
**Figure 1B TR Property**



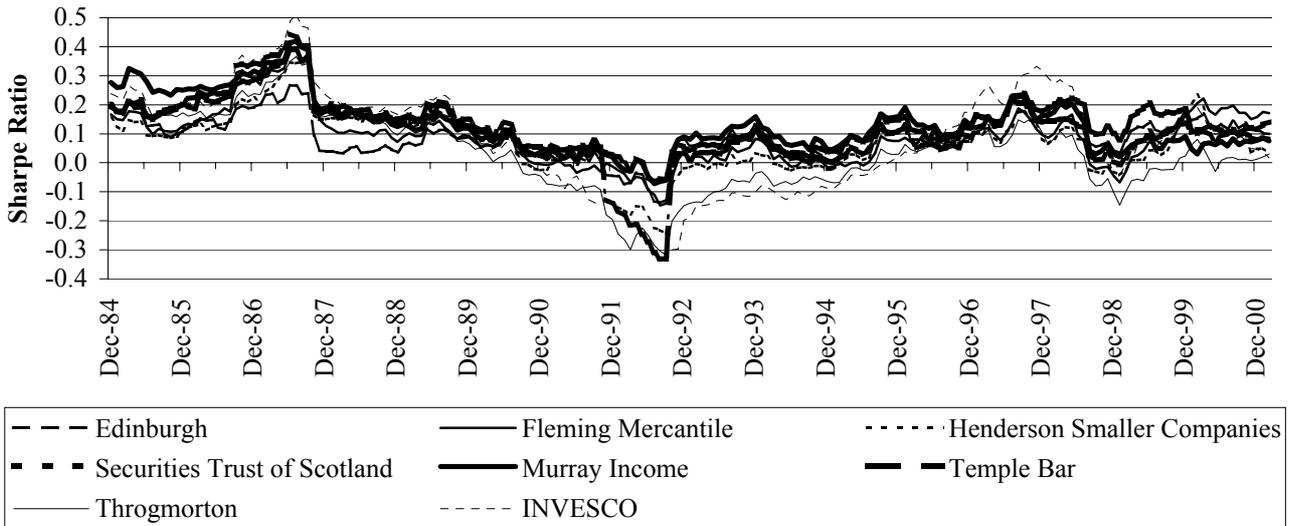
**Figure 1C INVESCO English & International**



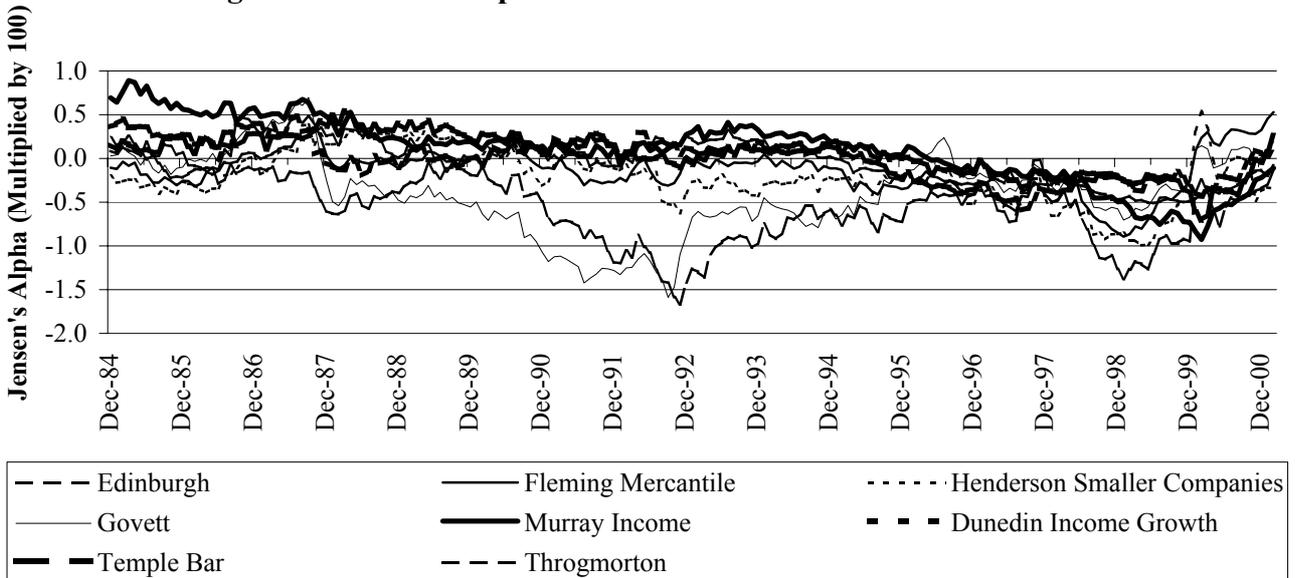
**Figure 2 Kernel Density Estimates for the TR Property Trust**



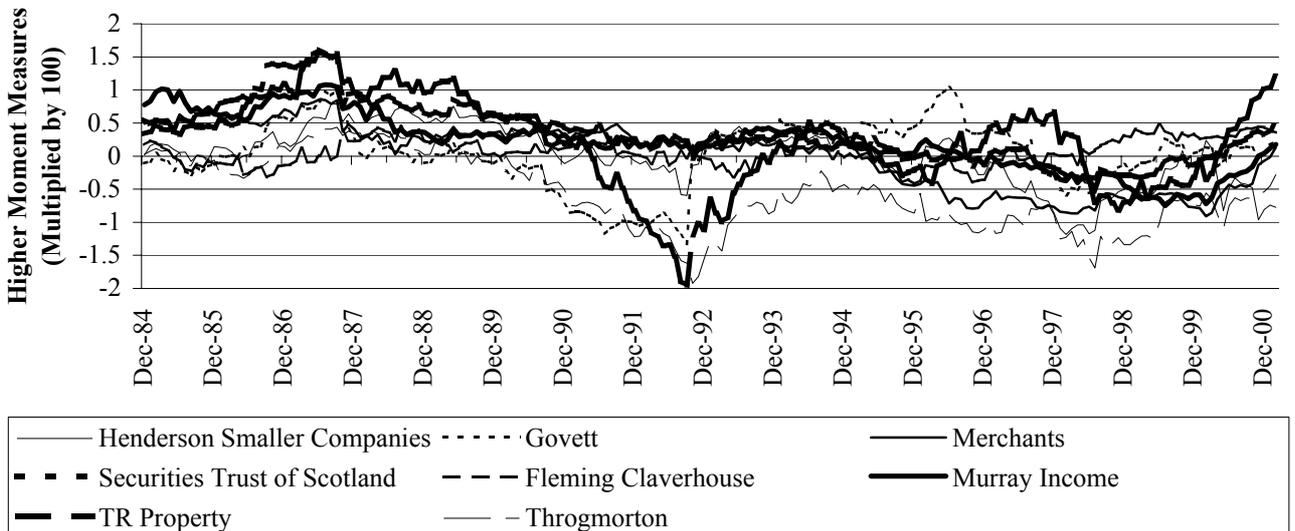
**Figure 3A Sharpe Ratios for the fourteen UK Investment Trusts**



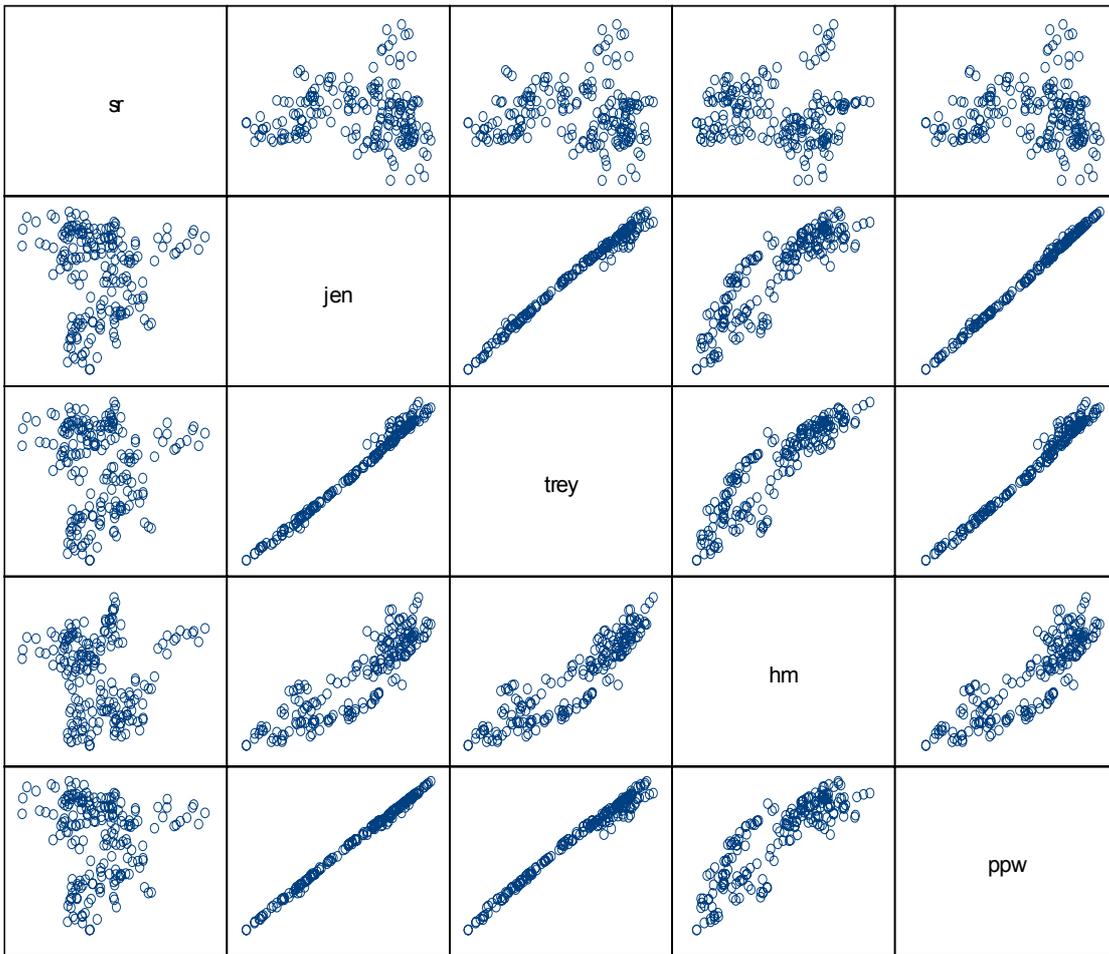
**Figure 3B Jensen's Alpha for the Fourteen UK Investment Trusts**



**Figure 3C Higher Moment Measures for the Fourteen UK Investment Trusts**



**Figure 4 Multivariate Scatter Plot of Performance Measures for the City of London Trust**



**Table 1 Statistical Properties of the Benchmark Portfolio Returns and 14 UK Investment Trust Returns****A. 1987 Market Crash (October 1987) Included**

	Mean	STD	Skewness	Excess Kurtosis	J&B Statistics
FTSE All-share	1.352	4.866	-1.553	7.600	713.362 *
Edinburgh Investment Trust (The) PLC	1.367	5.870	-1.099	5.065	322.652 *
Fleming Mercantile Inv Trust PLC	1.373	5.707	-1.091	3.556	184.197 *
Henderson Smaller Companies Investment Trust PLC	1.299	7.718	-0.783	3.824	180.752 *
Govett Strategic Investment Trust PLC	1.198	6.940	-1.434	5.970	464.305 *
City of London Investment Trust (The) PLC	1.543	5.939	-1.239	6.199	471.725 *
Merchants Trust (The) PLC	1.458	6.158	-1.134	4.658	284.045 *
Securities Trust of Scotland PLC	1.438	5.868	-1.320	6.232	484.823 *
Fleming Claverhouse Inv Trust PLC	1.584	6.272	-0.853	3.286	145.014 *
Murray Income Trust PLC	1.594	5.756	-0.955	3.858	196.065 *
Dunedin Income Growth Inv Trust PLC	1.452	5.928	-0.862	4.551	250.698 *
Temple Bar Investment Trust PLC	1.524	5.918	-0.911	3.515	165.887 *
TR Property Investment Trust PLC	1.135	7.011	-1.018	3.972	210.903 *
Throgmorton Trust (The) PLC	1.113	7.133	-0.547	2.943	104.342 *
INVESCO English & International Trust PLC	1.209	8.372	-0.271	4.125	183.157 *

Notes: A total number of 254 monthly log-returns in percentage from January 1980 to February 2001 is used for the calculation.

\* represents significance at 5% level.

**B. 1987 Market Crash (October 1987) Excluded**

	Mean	STD	Skewness	Excess Kurtosis	J&B Statistics
FTSE All-share	1.479	4.434	-0.623	1.290	33.932 *
Edinburgh Investment Trust (The) PLC	1.505	5.452	-0.398	1.051	18.321 *
Fleming Mercantile Inv Trust PLC	1.485	5.428	-0.757	2.026	67.394 *
Henderson Smaller Companies Investment Trust PLC	1.443	7.383	-0.475	2.809	92.718 *
Govett Strategic Investment Trust PLC	1.357	6.476	-0.895	2.985	127.721 *
City of London Investment Trust (The) PLC	1.692	5.454	-0.370	0.772	12.042 *
Merchants Trust (The) PLC	1.598	5.751	-0.536	1.314	30.333 *
Securities Trust of Scotland PLC	1.580	5.424	-0.593	1.933	54.230 *
Fleming Claverhouse Inv Trust PLC	1.712	5.941	-0.418	1.246	23.731 *
Murray Income Trust PLC	1.718	5.417	-0.442	1.255	24.858 *
Dunedin Income Growth Inv Trust PLC	1.589	5.519	-0.144	0.678	5.714 *
Temple Bar Investment Trust PLC	1.652	5.566	-0.379	0.737	11.791 *
TR Property Investment Trust PLC	1.289	6.581	-0.479	1.092	22.236 *
Throgmorton Trust (The) PLC	1.255	6.773	-0.098	1.062	12.300 *
INVESCO English & International Trust PLC	1.312	8.226	-0.198	4.279	194.668 *

Notes: A total number of 253 monthly log-returns in percentage from January 1980 to February 2001 except October 1987 is used for the calculation.

\* represents significance at 5% level.

**C. Correlation Matrix for the Market Portfolio and 14 Investment Trust (1987 Market crash Included)**

	FTSE All-share	Edinburgh	Fleming Mercantile	Henderson Smaller Companies	Govett Strategic	City of London Investment	Merchants Trust	Securities Trust of Scotland	Fleming Claverhouse	Murray Income Trust	Dunedin Income Growth	Temple Bar	TR Property	Throgmorton Trust	INVESCO
FTSE All-share	1.000														
Edinburgh	0.888	1.000													
Fleming Mercantile	0.806	0.773	1.000												
Henderson Smaller Companies	0.770	0.755	0.806	1.000											
Govett Strategic	0.805	0.781	0.844	0.808	1.000										
City of London Investment	0.878	0.833	0.741	0.703	0.760	1.000									
Merchants Trust	0.865	0.845	0.731	0.656	0.746	0.860	1.000								
Securities Trust of Scotland	0.892	0.876	0.762	0.747	0.779	0.848	0.877	1.000							
Fleming Claverhouse	0.864	0.800	0.763	0.704	0.773	0.808	0.797	0.800	1.000						
Murray Income Trust	0.840	0.827	0.716	0.648	0.727	0.848	0.875	0.852	0.783	1.000					
Dunedin Income Growth	0.845	0.842	0.717	0.626	0.729	0.822	0.858	0.823	0.827	0.823	1.000				
Temple Bar	0.852	0.794	0.698	0.616	0.707	0.838	0.852	0.817	0.793	0.832	0.821	1.000			
TR Property	0.702	0.648	0.665	0.602	0.685	0.710	0.689	0.667	0.600	0.646	0.641	0.653	1.000		
Throgmorton Trust	0.756	0.678	0.794	0.755	0.786	0.714	0.662	0.707	0.714	0.661	0.644	0.687	0.637	1.000	
INVESCO	0.529	0.495	0.616	0.644	0.606	0.497	0.478	0.563	0.525	0.431	0.465	0.449	0.463	0.621	1.000

Notes: A total number of 254 monthly log-returns in percentage from January 1980 to February 2001 is used for the calculation.

**Table 2 Performance of 14 UK Investment Trust Returns for the Entire Sample Period**

	Sharpe Ratio	Rank	Jesen's Alpha	Rank	Treynor-Mazuy	Rank	HM Measure	Rank	PPW Measure	Rank
Edinburgh Investment Trust (The) PLC	0.1016	9	-0.0246	9	-0.0177	9	0.0264	10	-0.0328	9
Fleming Mercantile Inv Trust PLC	0.1055	8	0.0512	7	0.0638	6	<b>0.1442</b>	<b>4</b>	0.0525	6
Henderson Smaller Companies Investment Trust PLC	0.0684	10	-0.1833	11	-0.1868	12	-0.2095	13	-0.1770	11
Govett Strategic Investment Trust PLC	0.0616	11	-0.2416	13	-0.2152	13	-0.0472	11	-0.2446	13
City of London Investment Trust (The) PLC	<b>0.1300</b>	<b>2</b>	<b>0.1507</b>	<b>4</b>	<b>0.1605</b>	<b>2</b>	<b>0.2230</b>	<b>2</b>	<b>0.1417</b>	<b>4</b>
Merchants Trust (The) PLC	0.1117	7	0.0523	6	0.0570	7	0.0870	7	0.0465	7
Securities Trust of Scotland PLC	0.1137	6	0.0429	8	0.0563	8	<b>0.1416</b>	<b>5</b>	0.0360	8
Fleming Claverhouse Inv Trust PLC	<b>0.1296</b>	<b>3</b>	<b>0.1665</b>	<b>2</b>	<b>0.1534</b>	<b>3</b>	0.0696	9	<b>0.1696</b>	<b>2</b>
Murray Income Trust PLC	<b>0.1431</b>	<b>1</b>	<b>0.2473</b>	<b>1</b>	<b>0.2495</b>	<b>1</b>	<b>0.2633</b>	<b>1</b>	<b>0.2423</b>	<b>1</b>
Dunedin Income Growth Inv Trust PLC	<b>0.1149</b>	<b>5</b>	<b>0.0845</b>	<b>5</b>	<b>0.0844</b>	<b>5</b>	0.0836	8	<b>0.0792</b>	<b>5</b>
Temple Bar Investment Trust PLC	<b>0.1274</b>	<b>4</b>	<b>0.1511</b>	<b>3</b>	<b>0.1447</b>	<b>4</b>	0.1041	6	<b>0.1538</b>	<b>3</b>
TR Property Investment Trust PLC	0.0520	13	-0.2283	12	-0.1763	11	<b>0.1553</b>	<b>3</b>	-0.2295	12
Throgmorton Trust (The) PLC	0.0479	14	-0.3042	14	-0.3103	14	-0.3494	14	-0.2953	14
INVESCO English & International Trust PLC	0.0524	12	-0.0957	10	-0.0941	10	-0.0839	12	-0.0771	10

Notes: A total number of 254 monthly log-returns in percentage from January 1980 to February 2001 is used for the calculation.

HM represents Hwang and Satchell's (1998) Higher Moment CAPM performance measure and PPW represents Positive Period Weighting measure by Grinblatt and Titman [1989]. Jensen's Alpha, Treynor-Mazuy, HM and PPW measures are multiplied by 100.

**Table 3 Statistical Properties of Performance Measures for Selected Investment Trusts****A. Edinburgh**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Mean	0.095	-0.145	-0.125	-0.003	-0.158
STD	0.068	0.227	0.226	0.335	0.227
Skewness	0.221	0.009	0.048	0.248	0.055
Excess Kurtosis	-0.199	<b>-0.896</b>	<b>-0.981</b>	-0.523	<b>-0.867</b>
Correlation Matrix					
	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	-0.219	1.000			
Treynor-Mazuy	-0.205	0.990	1.000		
HM	0.112	0.670	0.725	1.000	
PPW	-0.266	0.994	0.979	0.593	1.000
Autocorrelations Lag 1	0.930	0.961	0.960	0.968	0.961
Lag 2	0.844	0.926	0.927	0.929	0.929
Lag 3	0.764	0.900	0.906	0.895	0.908
Lag 4	0.689	0.864	0.872	0.864	0.875
Lag 5	0.618	0.830	0.840	0.836	0.844

**B. TR Property**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Mean	0.032	-0.471	-0.435	0.219	-0.477
STD	0.151	0.748	0.786	0.698	0.727
Skewness	-0.313	<b>-0.445</b>	<b>-0.436</b>	<b>-0.496</b>	<b>-0.473</b>
Excess Kurtosis	-0.490	<b>-0.743</b>	<b>-0.820</b>	0.031	-0.661
Correlation Matrix					
	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	0.958	1.000			
Treynor-Mazuy	0.951	0.997	1.000		
HM	0.816	0.830	0.837	1.000	
PPW	0.950	0.999	0.996	0.829	1.000
Autocorrelations Lag 1	0.984	0.985	0.987	0.975	0.984
Lag 2	0.965	0.969	0.972	0.947	0.967
Lag 3	0.940	0.950	0.956	0.918	0.947
Lag 4	0.913	0.929	0.937	0.885	0.924
Lag 5	0.882	0.902	0.913	0.846	0.897

**C. Invesco**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Mean	0.083	-0.198	-0.169	0.239	-0.198
STD	0.175	1.011	1.014	1.042	1.004
Skewness	-0.153	<b>-0.528</b>	<b>-0.589</b>	<b>-0.851</b>	<b>-0.521</b>
Excess Kurtosis	-0.415	-0.357	-0.323	0.431	-0.340
Correlation Matrix					
	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	0.937	1.000			
Treynor-Mazuy	0.940	0.999	1.000		
HM	0.833	0.895	0.909	1.000	
PPW	0.933	0.999	0.996	0.881	1.000
Autocorrelations Lag 1	0.984	0.982	0.983	0.980	0.981
Lag 2	0.962	0.958	0.959	0.958	0.956
Lag 3	0.935	0.933	0.935	0.935	0.929
Lag 4	0.908	0.909	0.911	0.913	0.903
Lag 5	0.881	0.884	0.888	0.887	0.879

Notes: HM represents Hwang and Satchell's (1998) Higher Moment CAPM performance measure and PPW represents Positive Period Weighting measure by Grinblatt and Titman [1989]. Jensen's Alpha, Treynor-Mazuy, HM and PPW measures are multiplied by 100. The bold numbers represent significance at 5% level.

**Table 4 comparison between the measures of concordance and correlation between the performance measures for the Fleming Calverhouse investment trust**

**A. Pearson's Correlation Coefficient**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	0.199 *	1.000			
Treynor-Mazuy	0.191 *	0.994 *	1.000		
HM	0.147 *	0.470 *	0.556 *	1.000	
PPW	0.230 *	0.996 *	0.988 *	0.470 *	1.000

**B. Kendal's  $\tau$**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	0.049	1.000			
Treynor-Mazuy	0.050	0.994 *	1.000		
HM	0.118 *	0.373 *	0.397 *	1.000	
PPW	0.053	0.955 *	0.903 *	0.340 *	1.000

**B. Spearman's  $\rho$**

	Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM	PPW
Sharpe Ratio	1.000				
Jesen's Alpha	0.060	1.000			
Treynor-Mazuy	0.065	0.993 *	1.000		
HM	0.168 *	0.506 *	0.543 *	1.000	
PPW	0.067	0.996 *	0.984 *	0.458 *	1.000

Notes: \* represents significance at 5% level.

**Table 5 Measures of Extreme Tail Independency for the Five Performance Measures  
For the Case of Fleming Claverhouse**

**A. Left tail**

		Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM
Jesen's Alpha	Independence Measure	-0.7824 *			
	Standard Error	(0.1088)			
	Correlation	0.1987 *			
Treynor-Mazuy	Independence Measure	-0.8797 *	0.3078		
	Standard Error	(0.0601)	(0.6539)		
	Correlation	0.1912 *	0.9943 *		
HM	Independence Measure	-0.4173	-0.1670 *	0.4369	
	Standard Error	(0.2914)	(0.4165)	(0.7185)	
	Correlation	0.1470 *	0.5275 *	0.5560 *	
PPW	Independence Measure	-0.7859 *	0.3078	0.3078	-0.2799 *
	Standard Error	(0.0957)	(0.6539)	(0.6539)	(0.3600)
	Correlation	0.2301 *	0.9965 *	0.9882 *	0.4702 *

**B. Right Tail**

		Sharpe Ratio	Jesen's Alpha	Treynor-Mazuy	HM
Jesen's Alpha	Independence Measure	0.8899			
	Standard Error	(0.9450)			
	Correlation	0.1987 *			
Treynor-Mazuy	Independence Measure	0.5805	0.3946		
	Standard Error	(0.7903)	(0.6237)		
	Correlation	0.1912 *	0.9943 *		
HM	Independence Measure	-0.3664 *	0.0618	0.1277	
	Standard Error	(0.3168)	(0.5309)	(0.5043)	
	Correlation	0.1470 *	0.5275 *	0.5560 *	
PPW	Independence Measure	1.5945	0.0462	0.1913	-0.2874 *
	Standard Error	(1.2972)	(0.5231)	(0.5957)	(0.3563)
	Correlation	0.2301 *	0.9965 *	0.9882 *	0.4702 *

Notes: The results are obtained with the 195 performance measures reported in Table 3. The independence measures are calculated as in Poon, Rockinger and Tawn (2001). \* represents significance from at 5% level.

**Table 6 Conditional Dependence between the Different Performance Measures for the Case of the Fleming Claverhouse Trust**

**A. Quantile Regression Results for Jensen's  $\alpha$**

	Sharpe Ratio		HM	
	Intercept	Slope	Intercept	Slope
1%	-0.409 *	1.113	-0.136 *	0.437 *
5%	-0.121 *	0.000	-0.092 *	0.381 *
10%	-0.158 *	0.557 *	-0.051 *	0.347 *
25%	0.028	0.090	0.015	0.366 *
50%	0.124 *	-0.017 *	0.107 *	0.478 *
75%	0.206 *	0.516 *	0.198 *	0.727 *
90%	0.302 *	0.514 *	0.326 *	0.399 *
95%	0.323 *	0.678 *	0.405 *	0.280 *
99%	0.352 *	1.055 *	0.497 *	0.591 *

**A. Quantile Regression Results for HM**

	Sharpe Ratio		Jensen's $\alpha$	
	Intercept	Slope	Intercept	Slope
1%	-0.442 *	0.418	-0.378 *	0.313 *
5%	-0.319 *	0.139	-0.299 *	0.551 *
10%	-0.235 *	0.256	-0.227 *	0.482 *
25%	-0.080 *	-0.025	-0.151 *	0.799 *
50%	0.030 *	0.235	-0.076 *	0.788 *
75%	0.102 *	0.870 *	0.074 *	0.614 *
90%	0.124 *	1.290 *	0.268 *	0.310 *
95%	0.144 *	1.480	0.333 *	0.291 *
99%	0.157	1.672	0.365 *	0.357

**A. Quantile Regression Results for Sharpe Ratio**

	HM		Jensen's $\alpha$	
	Intercept	Slope	Intercept	Slope
1%	-0.045 *	0.386	-0.002	-0.366
5%	-0.012	0.026	0.023	-0.069
10%	0.024 *	0.109	0.027 *	0.027
25%	0.070 *	0.188 *	0.069 *	-0.042
50%	0.134 *	0.108 *	0.134 *	0.081 *
75%	0.195 *	0.040	0.194 *	0.125
90%	0.265 *	-0.117	0.214 *	0.358 *
95%	0.305 *	-0.110	0.229 *	0.356 *
99%	0.401 *	0.312	0.289 *	0.269 *

Notes: \* represents significance at 5% level.