# On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation

Andrew J. Patton*
*London School of Economics*

First version: 20 May, 2002. This version: October, 2002.

Recent studies in the empirical finance literature have reported evidence of two types of asymmetries in the joint distribution of stock returns. The first is skewness in the distribution of individual stock returns, while the second is an asymmetry in the dependence between stocks: stock returns appear to be more highly correlated during market downturns than during market upturns. In this paper we examine the economic and statistical significance of these asymmetries for asset allocation decisions in an out-of-sample setting. We consider the problem of a CRRA investor allocating wealth between the risk-free asset, a small-cap and a large-cap portfolio, using monthly data. We use models that can capture time-varying means and variances of stock returns, and also the presence of time-varying skewness and kurtosis. Further, we use copula theory to construct models of the time-varying dependence structure that allow for greater dependence during bear markets than bull markets. The importance of these two asymmetries for asset allocation is assessed by comparing the performance of a portfolio based on a normal distribution model with a portfolio based on a more flexible distribution model. For a variety of performance measures and levels of risk aversion our results suggest that capturing skewness and asymmetric dependence leads to gains that are economically significant, and statistically significant in some cases.

**Keywords:** stock returns, forecasting, density forecasting, normality, asymmetry, copulas.
**J.E.L. Codes:** G11, C32, C51.

*The author would like to thank Rob Engle, Raffaella Giacomini, Clive Granger and Bruce Lehmann for helpful comments. Special thanks are due to Allan Timmermann for many useful discussions and suggestions on this topic. All remaining deficiencies are the responsibility of the author. Many thanks go to Vince Crawford and the UCSD Economics Experimental and Computational Laboratory for providing the computational resources required for this project. Financial support from the IAM Programme in Hedge Fund Research at LSE, and the UCSD Project in Econometric Analysis fellowship is gratefully acknowledged. Contact information: Financial Markets Group, London School of Economics, Houghton Street, London WC2A 2AE, United Kingdom. Email: a.patton@lse.ac.uk. This paper may be downloaded from http://fmg.lse.ac.uk/~patton/research.html.

# 1 Introduction

Recent studies in the empirical finance literature have reported evidence of two types of asymmetries in the joint distribution of stock returns. The first is skewness in the distribution of individual stock returns, which has been reported by numerous authors over the last two decades[1]. The second asymmetry is in the dependence between stocks: stock returns appear to be more dependent during market downturns than during market upturns, a characteristic we refer to as 'asymmetric dependence'. Evidence that stock returns exhibit some form of asymmetric dependence has been reported by several authors in recent years; published work includes Erb, *et al.*, (1994), Longin and Solnik (2001), Ang and Bekaert (2001) and Ang and Chen (2002), who all report that correlations between stock returns are greater during bear markets than during bull markets. Further evidence is reported in numerous unpublished studies[2]. The presence of either of these asymmetries violates the assumption of elliptically distributed asset returns, which underlies traditional mean-variance analysis, see Ingersoll (1987). In this paper we examine the economic and statistical significance of these two asymmetries for asset allocation decisions in an out-of-sample setting. This paper can thus be viewed as a first step in addressing the suggestions of Harvey and Siddique (2000) and Longin and Solnik (2001), who propose investigating the impact of conditional skewness (Harvey and Siddique) and asymmetric dependence (Longin and Solnik) on portfolio choices.

Theoretical justification for the importance of distributional asymmetries may be found in Arrow (1971), who suggests that a desirable property of a utility function is that it exhibits non-increasing absolute risk aversion[3]. Under non-increasing absolute risk aversion investors can be shown to have a preference for positively skewed assets, in the same way that positive marginal utility leads to a preference for assets with higher mean returns, and diminishing marginal utility leads to risk aversion. Patton (2002) shows that asymmetric dependence between assets can lead to skewed portfolios, suggesting that risk averse investors will also have preferences over alternative dependence structures. Ang, *et al.*, (2001) report empirical evidence in support of this theoretical result.

We examine the problem of an investor with constant relative risk aversion (CRRA) allocating wealth between the risk-free asset, the CRSP small-cap and large-cap indices[4]. We use monthly data from January 1954 to December 1989 to develop the models, and data from January 1990 to December 1999 for forecast evaluation. This problem is representative of that of choosing between

[1] See Kraus and Litzenberger (1976), Friend and Westerfield (1980), Singleton and Wingender (1986), Lim (1989), Richardson and Smith (1993), Harvey and Siddique (1999, 2000) and Aït-Sahalia and Brandt (2001), amongst others.

[2] See Bae, *et al.*, (2000), Rosenberg (2000) and Campbell, *et al.*, (2001) amongst others.

[3] Utility functions that exhibit non-increasing absolute risk aversion include the constant absolute risk aversion (CARA), or exponential, utility function, and the constant relative risk aversion (CRRA), or 'narrow power', utility function, see Huang and Litzenberger (1988).

[4] The small-cap index is comprised of the smallest 10% of U.S. stocks by market capitalisation and the large-cap index is comprised of the largest 10% of U.S. stocks.

a high risk - high return asset and a lower risk - lower return asset, as the annualised mean and standard deviation on these indices over the sample were 9.95% and 21.29% for the small caps, and 7.97% and 14.29% for the large caps. We use distribution models that can capture the empirically observed time-varying means and variances of stock returns, and also the presence of (possibly time-varying) skewness and kurtosis. Further, we employ models of the dependence structure that allow for, but do not impose, greater dependence during bear markets than bull markets, and allow for changes in this dependence structure through time.

Our models are developed using copula theory, which enables the construction of flexible multivariate distributions. In Section 2 we provide a brief introduction to copula theory; a more thorough introduction is presented in Nelsen (1999), Schweizer and Sklar (1983) and Joe (1997). The investor is assumed to estimate the model of the conditional distribution of returns using maximum likelihood (ML), see Patton (2001b), and then optimise the portfolio's weight using the predicted conditional distribution of returns. Work from the forecasting and estimation literature suggests that the parameter estimation stage and the forecast evaluation stage should both use the same objective function (or loss function), see Granger (1969), Weiss (1996) and Skouras (2001). As we do not intend to evaluate the quality of our density forecasts using the log-likelihood function, this implies we should use some method other than ML for estimation. We use ML for computational tractability: the forecast evaluation functions used in this paper are functions of the optimal portfolio weights, which in turn are functions of the model parameters that must be solved numerically. It is not feasible to estimate the parameters of our models via our forecast evaluation functions.

The importance of skewness and asymmetric dependence for asset allocation is measured by comparing the performance of a portfolio based on a bivariate normal distribution model with a portfolio based on a model developed using copula theory. The significance of the differences in measures of portfolio performance is tested using bootstrap methods. We find substantial evidence in most cases that skewness and asymmetric dependence do indeed have important economic implications for asset allocation, however the statistical significance of the improvement is only moderate. For example[5], while a constant equally weighted portfolio of the two assets generates a Sharpe ratio of 0.242 and the portfolio based on the bivariate normal model generates a Sharpe ratio of 0.286, the portfolio developed using copula theory attains a Sharpe ratio of 0.302. Thus the gains to modelling the distribution of returns are increased by almost 40% (according to this measure) by capturing and modelling deviations from joint normality. If we instead use the 5% Value-at-Risk (VaR) as a measure of risk the benefits of modelling deviations from joint normality are over 80%.

It should be re-emphasised that the goal of this paper is not to determine whether skewness and asymmetric dependence are present in asset returns, but whether the capturing of these asymmetries

[5] The figures here are taken from Table 4 for a short sales constained investor with relative risk aversion of seven.

leads to better out-of-sample portfolio decisions. The distinction between in-sample and out-of-sample significance of these asymmetries for asset allocation is an important one. Finding that a more flexible distribution model fits the data better in-sample does not imply that it will lead to better out-of-sample portfolio decisions than those based on a simpler model. In fact, a common finding in the (point) forecasting literature is that more complicated models often provide *worse* forecasts that simple mis-specified models, see Weigend and Gershenfeld (1994), Swanson and White (1995, 1997) and Stock and Watson (1999). More complicated models generally are more highly parameterised and thus subject to greater estimation error than simpler models. Our finding that the more flexible models generally perform better than simpler models is thus a noteworthy one.

In this paper we consider both unconstrained and short sales constrained estimates of the optimal portfolio weight for a given density forecast. We do so for two reasons. The first reason is economically motivated: many market participants face the constraint that they are unable to short sell stocks or to borrow and invest the proceeds in stocks. The second reason is statistically motivated: the optimal portfolio weight given a density forecast is itself only an estimate of the true optimal portfolio weight. By ensuring that our estimate always lies in the interval $[0, 1]$ we employ a type of "insanity filter" which prevents the investor from taking an extreme position in the market. Such constraints have been found to improve the out-of-sample performance of optimal portfolios based on parameter estimates, see Frost and Savarino (1988) and Jagannathan and Ma (2002).

Much of the existing work on asset allocation focussed on special cases where the combination of utility function and distribution model were such than an analytical solution for the optimal portfolio decision exists, see Kandel and Stambaugh (1996) or Campbell and Viceira (1999) amongst many others. For example, the combination of quadratic or exponential utility with elliptical distributions, or where the utility function was assumed to be a function of a certain number of moments of the returns. The focus on such analytically tractable special cases was motivated, at least in part, by computational constraints and certainly not by the fact that the utility functions or distributional assumptions were considered realistic. In this paper we combine density models that are thought to adequately describe the statistical properties of the asset returns with the CRRA utility function.

Recent work by Brandt (1999) and Aït-Sahalia and Brandt (2001) overcome the problem of the appropriate distributional assumption to combine with a given utility function by using the method of moments and the first-order conditions of the investor's optimisation problem to obtain an optimal portfolio decision. Doing so allows them to use whichever utility function they please. Theirs is indeed an interesting approach, however it has the drawback that its nonparametric nature imposes restrictions on the number of exogenous regressors that may be included in the model, as in Brandt (1999), or on the way a larger number of regressors may enter into the problem, as in Aït-Sahalia and Brandt (2001). Our framework instead involves a flexible parametric approach to distribution modelling.

One of the costs of using flexible parametric models for the joint distribution of stock returns is that we are forced by computational constraints to be relatively unsophisticated in other aspects of the project. Firstly, we ignore the effects of parameter estimation uncertainty on the investor's decision problem, though this has been found to be important, see Kandel and Stambaugh (1996) and the references cited therein. Also, we only consider the investor's problem for the one-period-ahead investment horizon. For one of the utility functions we consider, the log utility function[6], this approach is correct, however for the remaining utility functions the optimal weights will have both a 'myopic' component and a 'hedging' component, see Merton (1971). The myopic component is the solution we focus on: the investor simply seeks to maximise the next-period expected utility. The hedging component represents the deviation from the myopic optimal weight that occurs when the investor seeks to hedge possible future adverse movements in the investment opportunity set. Ang and Bekaert (2001) and Aït-Sahalia and Brandt (2001) find, however, only weak evidence of hedging demand, though Brandt (1999) reports it to be quite significant.

The remainder of the paper is structured as follows. In Section 2 we provide a brief introduction to copula theory and its use in the density forecasting of stock returns. In Section 3 we present empirical results on the asset allocation problem for a portfolio of a small-cap index and a large-cap index: Section 3.1 presents the investor's problem in detail, Section 3.3 presents the models employed and the performance of the resulting portfolios are evaluated in 3.4. Finally, we conclude in Section 4. In Appendix 1 we present some useful results on Hansen's skewed $t$ distribution and in Appendix 2 we provide the functional forms of the copulas considered in Section 3.

## 2 Flexible multivariate distribution models using copulas

In this paper we use copula theory to develop flexible parametric models of the joint distribution of returns. Below we provide a non-technical introduction to copula theory, which follows that of Patton (2001a) closely. Let us firstly define our notation: we have two (scalar) random variables of interest, $X_t$ and $Y_t$, and some exogenous variables $\mathbf{W}_t$. The variables' joint conditional distribution is: $(X_t, Y_t)\,|\mathcal{F}_{t-1} \sim H_t \equiv C_t\,(F_t, G_t)$, where $H_t$ is some bivariate distribution function, the marginal distributions of $X_t$ and $Y_t$ are $F_t$ and $G_t$, the copula is $C_t$, and $\mathcal{F}_{t-1}$ is the information set defined as $\mathcal{F}_t \equiv \sigma\,(\mathbf{Z}_t)$, for $\mathbf{Z}_t \equiv \left[X_t, Y_t, \mathbf{W}_t', X_{t-1}, Y_{t-1}, \mathbf{W}_{t-1}', ...X_{t-j}, Y_{t-j}, \mathbf{W}_{t-j}'\right]'$. (The notation '$H \equiv C\,(F, G)$' will be explained below.) We will assume that all distributions are continuous and differentiable, though this assumption may be relaxed at the expense of further complication. We will denote the distribution (or *c.d.f.*) of a random variable using an upper case letter, and the

[6] This is the CRRA utility function with a coefficient of relative risk aversion of 1.

corresponding density (or *p.d.f.*) using the lower case letter. We will denote the extended real line as $\bar{\mathbb{R}} \equiv \mathbb{R} \cup \{\pm\infty\}$.

A copula is a function that links together two (or more) marginal distributions to form a joint distribution. The marginal distributions that it couples can be of any type: a normal and an exponential, or a Student's $t$ and a Uniform, for example. The theory of copulas dates back to Sklar (1959), but it wasn't until the 1970's that copulas were used in applied work. Since then numerous applications have appeared in the statistics literature, see Clayton (1978), Cook and Johnson (1981), Oakes (1989), Genest and Rivest (1993) and Fine and Jiang (2000), amongst others, and more recently in the analysis of economic data, see Rosenberg (1999) and (2000), Li (2000), Scaillet (2000), Embrechts, *et al.,* (2001), Patton (2001a,b), Rockinger and Jondeau (2001) and Miller and Liu (2002). The main theorem in copula theory is that of Sklar (1959), and below we present a modification of it for conditional distributions.

**Definition 1 (Conditional copula)** *A two-dimensional conditional copula is a function* $C : [0,1] \times [0,1] \times \mathcal{Z} \to [0,1]$*, where* $\mathcal{Z} \subseteq \mathbb{R}^k$ *and* $k$ *is a finite integer, with the following properties:*

1. $C(u,0|z) = C(0,v|z) = 0$*, and* $C(u,1|z) = u$ *and* $C(1,v|z) = v$*,* *for every* $u, v$ *in* $[0,1]$ *and all* $z \in \mathcal{Z}$

2. $V_C\left([u_1,u_2] \times [v_1,v_2]\,|z\right) \equiv C\left(u_2,v_2|z\right) - C\left(u_1,v_2|z\right) - C\left(u_2,v_1|z\right) + C\left(u_1,v_1|z\right) \geq 0$ *for all* $u_1, u_2, v_1, v_2 \in [0,1]$ *such that* $u_1 \leq u_2$ *and* $v_1 \leq v_2$*, and all* $z \in \mathcal{Z}$*.*

**Theorem 1 (Sklar's theorem for continuous conditional distributions)** *Let* $F$ *be the conditional distribution of* $X|Z$*,* $G$ *be the conditional distribution of* $Y|Z$*, and* $H$ *be the joint conditional distribution of* $(X,Y)\,|Z$*. Assume that* $F$ *and* $G$ *are continuous in* $x$ *and* $y$*. Then there exists a unique conditional copula* $C$ *such that*

$$H(x,y|z) = C(F(x|z), G(y|z)|z), \quad \forall \ (x,y) \in \bar{\mathbb{R}} \times \bar{\mathbb{R}} \text{ and each } z \in \mathcal{Z} \tag{1}$$

*Conversely, if we let* $F$ *be the conditional distribution of* $X|Z$*,* $G$ *be the conditional distribution of* $Y|Z$*, and* $C$ *be a conditional copula, then the function* $H$ *defined by equation (1) is a conditional bivariate distribution function with conditional marginal distributions* $F$ *and* $G$*.*

For the proof of Theorem 1 see Patton (2001a). Sklar's theorem allows us to decompose a bivariate distribution, $H_t$, into three components: the two marginal distributions, $F_t$ and $G_t$, and the copula, $C_t$. Since all of the univariate information on $X_t$ and $Y_t$ is contained in the marginal distributions, what remains is all of the dependence information *between* $X_t$ and $Y_t$, which is captured in the copula. The density function equivalent of (1) is useful for maximum likelihood analysis, and is obtained quite easily, provided that $F_t$ and $G_t$ are differentiable, and $H_t$ and $C_t$ are twice differentiable.

$$
\begin{aligned}
h_t(x,y|z) &\equiv \frac{\partial^2 H_t(x,y|z)}{\partial x \partial y} \\
&= \frac{\partial F_t(x|z)}{\partial x} \cdot \frac{\partial G_t(y|z)}{\partial y} \cdot \frac{\partial^2 C_t(F_t(x|z), G_t(y|z)|z)}{\partial (F_t(x|z)\,\partial (G_t(y|z))} \\
&\equiv f_t(x|z) \cdot g_t(y|z) \cdot c_t(u,v|z), \quad \forall (x,y,z) \in \bar{\mathbb{R}} \times \bar{\mathbb{R}} \times \mathcal{Z}
\end{aligned} \tag{2}
$$

where $u \equiv F_t(x|z)$, and $v \equiv G_t(y|z)$. Taking logs of both sides we obtain:

$$
\mathcal{L}_{XY} = \mathcal{L}_X + \mathcal{L}_Y + \mathcal{L}_C \tag{3}
$$

and so the joint log-likelihood is equal to the sum of the marginal log-likelihoods and the copula log-likelihood. For the purposes of multivariate density modelling the copula representation allows for great flexibility in the specification: we may model the individual variables using whichever marginal distributions fit best, and then work on modelling the dependence structure via a model for the copula. The estimation of multivariate time series models constructed using copulas is discussed in Patton (2001b).

# 3 A portfolio of small cap and large cap stocks

In this section we consider an investor with constant relative risk aversion facing the problem of allocating wealth between two assets: a portfolio of low market capitalisation stocks ('small caps') and a portfolio of high market capitalisation stocks ('large caps'). These two assets were chosen as being representative of the general problem of balancing a portfolio comprised of a high risk - high return asset and a lower risk - lower return asset. The small cap and large cap portfolios fit this problem: the average annualised return on these indices was 9.95% and 7.97% respectively, and their annualised standard deviations were 21.29% and 14.29%.

## 3.1 The investor's optimisation problem

The utility functions we assume for our hypothetical investors are from the class of constant relative risk aversion (CRRA) utility functions:

$$
\mathcal{U}(\gamma) = \begin{cases} (1-\gamma)^{-1} \cdot (P_0 \cdot (1 + \omega_x X_t + \omega_y Y_t))^{1-\gamma} & \text{if } \gamma \neq 1 \\ \log (P_0 \cdot (1 + \omega_x X_t + \omega_y Y_t)) & \text{if } \gamma = 1 \end{cases} \tag{4}
$$

where $P_0$ is the initial wealth, $X_t$ represents the return on the small-cap index, $Y_t$ represents the return on the large-cap index and $\omega_i$ is the proportion of wealth in asset $i$. The degree of relative risk aversion (RRA) is denoted by $\gamma$. For this utility function the initial wealth does not affect the choice of optimal weight and so we set $P_0 = 1$. We consider five different levels of relative risk aversion: $\gamma$=1, 3, 7, 10 and 20. This range of risk aversion levels was also considered in Aït-Sahalia and Brandt (2001).

The set-up of the investor's problem is as follows. Let the returns on the two assets under consideration be denoted $X_t$ and $Y_t$. These returns have some joint distribution, $H_t$, with associated marginal distributions, $F_t$ and $G_t$, and a copula, $C_t$. That is, $(X_t, Y_t) | \mathcal{F}_{t-1} \sim H_t \equiv C_t (F_t, G_t)$. We will develop density forecasts of this joint distribution: $\hat{F}_{t+1}$, $\hat{G}_{t+1}$, and the conditional copula, $\hat{C}_{t+1}$, and use them to compute the optimal weights, $\omega^*_{t+1} \equiv \left[\omega^*_{x,t+1}, \omega^*_{y,t+1}\right]$, for the portfolio. The optimal weights are found by maximising the expected utility of the end-of-period wealth under the estimated probability density:

$$\begin{aligned}
\omega^*_{t+1} &\equiv \underset{\omega \in \mathcal{W}}{\arg\max} E_{\hat{H}_{t+1}} \left[\mathcal{U}\left(1 + \omega_x X_{t+1} + \omega_y Y_{t+1}\right)\right] \\
&\equiv \underset{\omega \in \mathcal{W}}{\arg\max} \iint \mathcal{U}\left(1 + \omega_x x + \omega_y y\right) \cdot \hat{h}_{t+1}(x, y) \cdot dx \cdot dy \\
&= \underset{\omega \in \mathcal{W}}{\arg\max} \iint \mathcal{U}\left(1 + \omega_x x + \omega_y y\right) \cdot \hat{f}_{t+1}(x) \cdot \hat{g}_{t+1}(y) \cdot \hat{c}_{t+1}\left(\hat{F}_{t+1}(x), \hat{G}_{t+1}(y)\right) \cdot dx \cdot dy
\end{aligned}$$

where $\mathcal{W}$ is some compact sub-set of $\mathbb{R}^2$ for the unconstrained investor and $\mathcal{W} = \left\{(\omega_x, \omega_y) \in [0,1]^2 : \omega_x + \omega_y \leq 1\right\}$ for the short sales constrained investor.

The double-integral defining the expected utility of wealth does not have a closed-form solution for our case. We use 10,000 Monte Carlo replications to estimate the value of this integral[7]. The objective function $\varphi_{t+1}(\omega) \equiv \iint \mathcal{U}\left(1 + \omega_x x + \omega_y y\right) \cdot \hat{h}_{t+1}(x, y) \cdot dx \cdot dy$ was found to be well-behaved (smooth and having a unique global optimum) for our choices of utility functions and density models and so we employed the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm to locate the optimum, $\omega^*_{t+1}$, at each point in time.

### 3.2 Description of the data

We use monthly data from the Center for Research in Security Prices (CRSP) on the top 10% and bottom 10% of stocks sorted by market capitalisation to form indices - the 'big cap' and 'small cap' indices, from January 1954 to December 1999, yielding 552 observations. This data was also analysed in a different context by Perez-Quiros and Timmermann (2001). We reserve the last 120 observations, from January 1990 to December 1999, for the out-of-sample evaluation of the models. Descriptive statistics on the two portfolios over the entire sample are presented in Table 1.

[ INSERT TABLE 1 HERE ]

[7] Judd (1998, pp291-305) discusses some of the issues surrounding the use of Monte Carlo simulations to approximate objective functions containing integrals.

Table 1 reveals that the small cap index had a higher mean and higher volatility than the large cap index. The small cap index also exhibited slightly positive skewness, while the big cap index exhibited substantial negative skewness. Both indices exhibited excess kurtosis. The Jarque-Bera statistic indicates that neither series is unconditionally normal, and the unconditional correlation coefficient indicates a high degree of linear dependence.

To examine the presence of asymmetric dependence between these two assets we use measures presented in Longin and Solnik (2001) and Ang and Chen (2002) called 'exceedence correlations', $\tilde{\rho}(q)$:

$$\tilde{\rho}(q) \equiv \begin{cases} Corr\left[X, Y \mid X \leq Q_x(q) \cap Y \leq Q_y(q)\right], & \text{for } q \leq 0.5 \\ Corr\left[X, Y \mid X > Q_x(q) \cap Y > Q_y(q)\right], & \text{for } q \geq 0.5 \end{cases}$$

where $Q_x(q)$ and $Q_y(q)$ are the $q^{th}$ quantiles of $X$ and $Y$ respectively. As Longin and Solnik (2001) and Ang and Chen (2002) point out, the shape of the exceedence correlation plot (as a function of $q$) depends on the underlying distribution of the data. Even for the standard bivariate normal distribution $\tilde{\rho}(q)$ is non-linear in $q$. In Figure 1 we plot the empirical exceedence correlations based on the (raw) excess returns on the two indices. In Figure 2 we plot the empirical exceedence correlations based on the transformed standardised residuals of the models for the two indices, along with what would be expected if these assets had the normal copula and the 'rotated Gumbel' copula, which is described below. Figure 1 shows the degree of asymmetry in the unconditional distribution of the returns on these two assets; Figure 2 shows the degree of asymmetry in the unconditional copula of these two assets, having removed all marginal distribution asymmetry. Clearly both the unconditional joint distribution and the unconditional copula exhibit substantial asymmetry.

[ INSERT FIGURES 1 AND 2 HERE ]

In our density forecasts we use three further variables as explanatory variables in our analysis. The first is the one-month treasury bill rate, denoted $R_{ft}$, which is taken as the risk-free rate. This variable has been used by Fama (1981) and others as a proxy for shocks to expected growth in the real economy. The second variable is the difference between the yield on corporate bonds with Moody's rating Baa versus those with an Aaa rating, denoted $SPR_t$, which is called the 'default spread'. This variable tracks the cyclical variation in the risk premium on stocks, see Perez-Quiros and Timmermann (2001). Finally, we look at the dividend yield, denoted $DIV_t$, which is measured as the total dividends paid over the previous 12 months divided by the stock price at the end of the month. This variable acts as a proxy for time-varying expected returns. For a comprehensive review of the variables that have been used in previous studies as predictive variables for stock returns see Aït-Sahalia and Brandt (2001, pp1297-1298).

### 3.3 Analysis of the different models

We consider a number of different investment strategies, some of which are based on density forecasts. In this section we describe the models used to obtain the density forecasts.

The first three strategies we consider and simply buy-and-hold strategies. The fourth strategy is one based solely on the unconditional distribution of returns. For this portfolio we assume that the investor optimises his/her portfolio weights for the period $t+1$ using the empirical unconditional distribution of returns observed up until time $t$.

$$\begin{aligned} \omega^*_{uncond,t+1} &\equiv \underset{\omega\in\mathcal{W}}{\arg\max}\;\; \hat{E}_t\left[\mathcal{U}\left(1+\omega_x X_{t+1}+\omega_y Y_{t+1}\right)\right] \\ &\equiv \underset{\omega\in\mathcal{W}}{\arg\max}\;\; \iint \mathcal{U}\left(1+\omega_x x+\omega_y y\right)\cdot d\hat{H}_t\left(x,y\right) \\ &= \underset{\omega\in\mathcal{W}}{\arg\max}\;\; t^{-1}\sum\nolimits_{j=1}^{t}\mathcal{U}\left(1+\omega_x x_j+\omega_y y_j\right) \end{aligned} \tag{5}$$

where $\{(x_j, y_j)\}_{j=1}^{t}$ are the observed excess returns on the two assets. This portfolio is 'somewhat naïve', in that the investor does perform some optimisation, but assumes that the joint distribution of these two assets is *i.i.d.* throughout the sample. A comparison of the performance of this portfolio with those constructed using parametric conditional distribution models may then be interpreted as a measure of the benefits to modelling the *conditional* distribution of these stock returns.

The benchmark parametric model for our study is the bivariate normal distribution, which is compared with a parametric model constructed using copula theory. Both parametric models have the same forms for the conditional means, $\mu_t^x$ and $\mu_t^y$, and variances, $h_t^x$ and $h_t^y$. We used likelihood ratio tests to determine the best fitting model over the in-sample period. Although the models are recursively re-estimated throughout the out-of-sample period they are "non-adaptive", in that the model specifications are determined using the in-sample data and not updated in the out-of-sample period. We do not present the final models or the parameter estimates here, but they are available from the author upon request. The conditional means were set to be linear functions of up to twelve lags of the two asset returns, the risk-free rate, the default spread and the dividend yield. For the conditional variance we employed a TARCH(1,1) specification[8] and allowed the three lagged exogenous regressors to enter into the conditional variance specification in levels and squares. The final model for the small cap (large cap) index contained 5 (4) parameters for the conditional mean and 6 (5) for the conditional variance.

For the bivariate normal model, all that remains to be specified is a model for the conditional correlation. The conditional correlation was set as a function of the lagged risk-free rate, default spread, dividend yield, and the forecasts of the conditional means of the two variables. All of these variables were found to be important in-sample. The bivariate normal model is:

[8] The general TARCH(1,1) specification is: $h_t = \omega + \beta h_{t-1} + \alpha_+ \cdot \varepsilon_{t-1}^2 \cdot \mathbf{1}\{\varepsilon_{t-1} > 0\} + \alpha_- \cdot \varepsilon_{t-1}^2 \cdot \mathbf{1}\{\varepsilon_{t-1} < 0\} + \alpha_z \cdot z_{t-1}$, where $\varepsilon_t$ is the residual from the model for the mean, and $z_t$ is an exogenous regressor.

**Bivariate normal specification**

$$\left(\frac{X_t-\mu_t^x}{\sqrt{h_t^x}},\frac{Y_t-\mu_t^y}{\sqrt{h_t^y}}\right) \sim N\left(\mathbf{0},\begin{bmatrix}1 & \rho_t\\ \rho_t & 1\end{bmatrix}\right) \tag{6}$$

$$\begin{aligned}\rho_t &= \Lambda\left(\alpha_0+\alpha_1 R_{ft-1}+\alpha_2 SPR_{t-1}+\alpha_3 DIV_{t-1}\right.\\ &\quad \left.+\alpha_4\mu_t^x+\alpha_5\mu_t^y\right)\end{aligned} \tag{7}$$

where $\Lambda(x)=\frac{1-e^{-x}}{1+e^{-x}}$ is the modified logistic transformation, designed to keep $\rho_t$ in $(-1,1)$ at all times.

To determine the importance of skewness and asymmetric dependence for asset allocation we specify distribution models that can capture these features. We found Hansen's (1994) skewed Student's $t$ distribution to provide a good fit for the marginal distributions of both assets. Some results on this distribution are presented in Appendix 1. In addition to time-varying conditional means and variances, the skewed $t$ distribution can capture time-varying conditional skewness and kurtosis. Both skewness and kurtosis were allowed to depend on lags of the exogenous variables and the forecast conditional means and variances. As suggested by Hansen (1994), we used the logistic transformation to ensure that the skewness and degrees of freedom parameters remained within $(-1,1)$ and $(2,\infty]$ respectively at all times. For both assets we found significant in-sample time-variation in these moments. Both small cap and large cap skewness parameters were found to be influenced by the dividend yield. The degree of freedom parameter for the small caps was found to be influenced by the dividend yield while that of the large caps was influenced by the risk free rate and the default spread. The total additional number of parameters in the skewed $t$ distribution over those in the normal distribution for the small caps (large caps) was 5 (4). Using likelihood ratio tests we could reject (with p-values of less than 0.001) for both models the normal distribution in favour of the skewed $t$ distribution over the in-sample period. This improvement in in-sample goodness-of-fit is traded off against possible increased parameter estimation error in the out-of-sample setting.

For the flexible distribution model all that remains is to specify the form of the copula used to link the two skewed $t$ marginal distributions. A total of nine different copulas were estimated on the transformed residuals from the skewed $t$ models, in the search for the best fitting copula. The copulas considered were the normal, Student's $t$, Clayton, rotated Clayton[9], Joe-Clayton, Plackett, Frank, Gumbel and rotated Gumbel copulas; contour plots of a few of these copulas is provided in Figure 3 and the functional forms of these copulas are contained in Appendix 2. This

[9] The 'rotated' copulas were formed as follows: Let (U,V) be distributed according to the copula $C$. Then (1-U,1-V) will be distributed according to the '*rotated* $C$' copula. The rotation allows us to take a copula that allows for greater dependence in the negative (positive) quadrant and create one that allows for greater dependence in the positive (negative) quadrant.

list includes almost all of the copulas considered in the various applications of copulas in statistics and economics[10], and is significantly more than we found in any single previous applied study.

[ INSERT FIGURE 3 HERE ]

As in the bivariate normal distribution we estimated these copulas with conditional dependence modelled as a function of the lagged risk-free rate, default spread and dividend yield, and the forecasts of the conditional means of the two variables, see equation (9) below. The maximum log-likelihood values for each of the copulas considered are presented in Table 2, and we can see that the rotated Gumbel copula attained the greatest log-likelihood value. The Student's $t$ copula has seven parameters (we fix the degree of freedom parameter and only allow the correlation coefficient to be time-varying) and the Joe-Clayton copula has twelve parameters (both upper and lower tail dependence are allowed to vary). All other copulas have six parameters, and so picking the rotated Gumbel on the basis of the likelihood value is equivalent to choosing it according to some other model selection criteria (such as AIC or BIC for example) as it has the greatest log-likelihood value and no more parameters than the competing copula specifications. We call flexible distribution specification using the rotated Gumbel copula the 'Gumbel' model.

[INSERT TABLE 2 HERE ]

**Copula distribution specifications**

$$
\begin{aligned}
\left(\frac{X_t-\mu_t^x}{\sqrt{h_t^x}},\frac{Y_t-\mu_t^y}{\sqrt{h_t^y}}\right) &\sim C\left(Skewed\ t\left(\lambda_t^x,\nu_t^x\right), Skewed\ t\left(\lambda_t^y,\nu_t^y\right);\delta_t\right) && (8)\\
\delta_t &= \Gamma\left(\beta_0+\beta_1 R_{ft-1}+\beta_2 SPR_{t-1}+\beta_3 DIV_{t-1}\right.\\
&\quad \left.+\beta_4\mu_t^x+\beta_5\mu_t^y\right) && (9)
\end{aligned}
$$

where $\Gamma(x)$ is a function designed to keep $\delta_t$ in the feasible region for the copula $C$ at all times, and $C$ is one of the nine copulas discussed above.

We specify one final alternative model, called the 'NormCop' model, which uses the skewed $t$ marginal distributions along with a normal copula. This specification is included to determine where the benefits, if any, from flexible density modelling lie: in the marginal distribution specifications or in the copula specification. The values of the log-likelihoods at the optimum for the three joint distributions (normal, NormCop and Gumbel) are -2391.04, -2355.38 and -2342.28, so in terms of in-sample goodness-of-fit we can see that around 73% of the gains come from the flexible marginal distribution models, though in an out-of-sample setting this ranking need not hold.

[10] One copula that was consciously omitted from this list is the Farlie-Gumbel-Morgenstern copula. This copula was excluded due to the limited amount of dependence it is able to consider: rank correlation under this copula is bounded in absolute value by one-third, see Joe (1997, p35).

We again use likelihood ratio tests to determine if any of the five regressors for the conditional copula parameter can be dropped. For the bivariate normal distribution and the NormCop models all five were significant at the 10% level, while for the rotated Gumbel copula the risk-free rate and the spread were not significant and so were removed from the model, reducing the number of parameters for this copula from six to four.

One concern that may arise in this design is the existence of $E_{\hat{H}_{t+1}}\left[\mathcal{U}\left(1+\omega_x X_{t+1}+\omega_y Y_{t+1}\right)\right]$ for certain density models. Given CRRA utility, any density model that assigns positive probability to the case of bankruptcy would preclude the existence of $E_{\hat{H}_{t+1}}[\mathcal{U}]$. All of the above specifications will assign some (extremely small) positive probability to bankruptcy, and so left unmodified $E_{\hat{H}_{t+1}}[\mathcal{U}]$ will not exist. We get around this by smoothly truncating the tails of the distribution: we apply a logistic transformation to the lower tail of the portfolio return distribution so that all probability mass assigned to the region $(-\infty, \varepsilon)$ is re-located to the region $(0, \varepsilon)$, where $\varepsilon$ is some extremely small positive number. We do an equivalent transformation for the upper tail. In this way the density is still continuous and $E_{\hat{H}_{t+1}}[\mathcal{U}]$ exists.

### 3.4 Performance of the different strategies

We now analyse the performance of the different asset allocation decisions made using the various models. We consider five levels of relative risk aversion ($\gamma$ = 1, 3, 7, 10 and 20), and eleven strategies. The eleven strategies are:

1. Always hold the small cap index,
2. Always hold the large cap index,
3. Always hold a 50:50 mix of the two indices,
4. Optimise the portfolio weight using the unconditional empirical distribution of returns,
5. Find the optimal portfolio weight for each period using the bivariate normal model,
6. Find the optimal portfolio weight for each period using the NormCop model,
7. Find the optimal portfolio weight for each period using the Gumbel model,
8. Same as strategy 4, subject to a short sales constraint,
9. Same as strategy 5, subject to a short sales constraint,
10. Same as strategy 6, subject to a short sales constraint,
11. Same as strategy 7, subject to a short sales constraint.

The first three portfolios are based on naïve rules, in that they are not the result of an optimisation problem. The final eight models are the results of optimisation problems, with the first four being unconstrained and the final four being subject to a short sales constraint.

### 3.4.1 Summary statistics

Firstly, let us look at some summary statistics of the realised portfolio returns based on the different models. These are presented in Table 3. We present five summary statistics on the optimal portfolio return series: the mean, standard deviation, skewness, 5% Value-at-Risk (5% VaR) and 5% Expected Shortfall (5% ES).

The 5% VaR is defined as the negative of the fifth empirical percentile of the realised returns, that is $\widehat{VaR}(X; 0.05) \equiv -\hat{F}_n^{-1}(0.05)$, where $\hat{F}_n$ is the empirical distribution of returns on portfolio $X$ using the $n$ out-of-sample observations. Value-at-Risk has gained some acceptance by practitioners as an alternative to standard deviation as a measure of risk. One of its main advantages is that it considers only the left tail of the distribution of returns, that is the losses, rather than the entire distribution.

While VaR has some advantages over traditional measures of risk, it has received criticism for not being a 'coherent' measure of risk, see Artzner *et al.* (1999). An alternative to VaR that has gained some attention recently is the 'expected shortfall' of a portfolio. The 5% expected shortfall is defined as the negative of the average return on a portfolio given the return has exceeded its 5% VaR, that is $\widehat{ES}(X; 0.05) \equiv -\hat{E}_n\left[X | X \leq \widehat{VaR}(X; 0.05)\right]$, where $\hat{E}_n$ is the sample average. We will use both VaR and expected shortfall as alternative measures of risk.

[ INSERT TABLE 3 ]

A striking feature of the summary statistics is the much greater mean and standard deviation of the portfolio returns based on the distribution models (normal, NormCop and Gumbel) than the portfolios with constant weights for all but the most risk averse investor. We ignore parameter estimation uncertainty, and so the query may be raised as to whether the investors would so aggressively invest if they knew that they were using parameter estimates rather than the true parameters. Kandel and Stambaugh (1996) and Brandt (1999) both find that even when parameter estimation uncertainty is accounted for a CRRA investor aggressively seeks the best portfolio. The results for the short sales constrained investors reveal a much smaller difference in mean and risk between the distribution portfolios and the constant weight portfolios.

Also note the skewness coefficients: both the small cap and large cap indices exhibited negative unconditional skewness over this period. As noted in the introduction, CRRA investors have a preference for positively skewed assets, *ceteris paribus*. The normal and NormCop portfolios also generally exhibited negative skewness while the unconstrained Gumbel portfolio actually displayed positive skewness, suggesting that modelling both skewness and asymmetric dependence enables the investor to better avoid negatively skewed portfolio returns. For the short sales constrained portfolios all models lead to negatively skewed portfolios.

#### 3.4.2 Performance statistics

Table 4 contains some risk-adjusted performance measures on the realised portfolio returns. These tables present measures in the form of ratios of average return to a measure of risk. In addition to the usual Sharpe ratio (mean to standard deviation) we present two alternative measures: mean to 5% VaR and mean to 5% expected shortfall. The presence of skewness in the distribution of returns, reported in Tables 3 and 4, implies that standard deviation may not be an appropriate measure of risk. In all cases the VaRs and expected shortfalls are reported as 'losses' and so are positive numbers, so for example a larger (positive) mean/VaR ratio implies a greater return per unit of risk.

Possibly the most interesting performance measure we consider is the amount (in basis points per year) that the investor would pay to switch from the '50:50 mix' portfolio to another portfolio. This performance measure was used by West, *et al.* (1993) to compare the economic value of various volatility models. This amount is the 'management fee' that may be deducted from the monthly return on portfolio $i$ over the out-of-sample period and leave the investor indifferent between the 50:50 portfolio and portfolio $i$. For example, an investor with risk aversion 1 would be willing to pay up to 25.176 basis points per year to switch from the 50:50 portfolio to the constrained Gumbel portfolio, while he would require compensation of 2.0114 basis points per year to switch from the 50:50 portfolio to the 'unconditional' portfolio.

It should be pointed out that the investors with risk aversion of one and three using the normal model density forecast would have gone bankrupt in the month of January 1992. On this date these two investors took the positions $\omega_x = -8.9$, $\omega_y = 21.3$ and $\omega_x = -5.1$, $\omega_y = 11.5$ respectively, and the month finished with returns of 14.0% on the small caps (the largest return on this asset over the out-of-sample period) and $-2.6\%$ on the large caps leading to negative gross returns for these investors[11]. For this month the realised utility for these investors is not defined, and so we do not report the management fee for these investors.

[ INSERT TABLE 4 HERE ]

[11] This obviously represents a failure of these investors' models or optimisation methods, as they did not recognise the risk of taking such extreme positions. According to the normal density forecast for that month the probability of a return such as this or more extreme, ie Pr$[X > 14.0 \cap Y \leq -2.6]$, was less than 3 in ten million. Our Monte Carlo estimate of the expected utility used only 10,000 draws from the forecast density so it is not surprising that this outcome was not anticipated by an investor using a normal density forecast. This may be interpreted as a signal that the normal density forecast is mis-specified; the model with skewed $t$ margins and rotated Gumbel copula assigned a probability of 131 in ten million, or 3 in 228,000 to this event.

The performance statistics indicate that substantial gains may be obtained by employing weights obtained from a model of the conditional distribution of stock returns, particularly when coupled with a short sales constraint. The unconstrained estimates generally do not perform as well as simply holding an equally weighted portfolio of the two indices, as the large caps performed particularly well over the period 1990 to 1999.

The improvements in portfolio performance are substantial: the Sharpe ratios for the constrained Gumbel portfolios are between 23% and 38% greater than those of the 50:50 portfolio. The mean/5%VaR and mean/5%ES ratios for the Gumbel portfolios are all around 50% greater than those of the 50:50 portfolio. Further, the 'management fees' that one could charge an investor currently holding the 50:50 portfolio to switch to the constrained Gumbel portfolio range between 3 and 27 basis points per year. This is substantial economic evidence in favour of the use of density forecasts for asset allocation.

Looking now to the gains from modelling skewness and asymmetric dependence: for the least risk averse investor the constrained normal performs the best according to all four performance measures, thus there are no gains here. Note however that the improvement that the constrained normal offers over the constrained Gumbel for this investor is less than 1% for all performance measures. For investors with risk aversion of 3, 7, 10 and 20 the constrained Gumbel outperformed the constrained normal, and the improvements in performance range from 4% to 51%. If we aggregate over all four performance measures and all five levels of risk aversion the constrained Gumbel portfolio generates an average improvement of 16.7% over the constrained normal portfolio. This suggests that there are out-of-sample gains to be had by capturing skewness and asymmetric dependence in the joint density of the assets under analysis here.

Somewhat surprisingly, the NormCop model generally performs worse than both the normal and the Gumbel models. The out-performance of the constrained Gumbel portfolio over the constrained NormCop portfolio averages 52.3%. Of course, our out-of-sample period is just 120 months and so this poor performance may simply be due to the short evaluation period. Nevertheless, the results suggest that both the marginal distribution *and* copula specifications are important for asset allocation.

### 3.4.3 Tests for superior portfolio performance

In this section we attempt to determine whether the economic gains documented in the previous section are statistically significant. We present the results of two tests for superior performance: a bootstrap test of pair-wise comparisons, and the reality check of White (2000), as modified by Hansen (2001). In all cases we employ the stationary bootstrap of Politis and Romano (1994)[12].

[12] The stationary bootstrap is a block bootstrap with block lengths that are distributed as a Geometric($q$) random variable. The average block length is $1/q$. We choose $q$ by running univariate regressions of each portfolio's returns on 36 lags, in both levels and squares to capture serial dependence in the conditional mean and variance. We set $1/q$

We conduct pair-wise comparisons by looking at the bootstrap confidence interval on the difference in the performance measures of two portfolios[13]. Let the performance measure of portfolio $i$ be $\mu_i$. If the lower bound of the bootstrap 90% confidence interval of $\mu_i - \mu_j$ is greater than zero, then we take model $i$ to be significantly better than model $j$. If the upper bound of the interval is less than zero then we take model $j$ to be significantly better than model $i$. If the confidence interval includes zero, then the test is inconclusive, and we cannot statistically distinguish models $i$ and $j$ according to that performance measure. The results of these tests are presented in Tables 5 and 6 below. Tables 7 and 8 contain the comparisons involving the NormCop model. In these tables, we include only the 50:50 portfolio of the three naïve portfolios to save space. The results from the pair-wise comparisons involving this portfolio are representative of the results from comparisons involving the other two naïve portfolios.

[ INSERT TABLES 5 AND 6 HERE ]

Table 5 shows that the comparisons between unconstrained portfolios rarely provide a definitive result. Indeed no comparisons using the Sharpe ratio or the mean/5% VaR ratio were significant. Using the investors' utility function as the performance measure does yield some significant results: for three out of three utility functions the unconstrained Gumbel significantly out-performed the unconstrained normal, while for the remaining two utility functions no comparison is possible as the unconstrained normal portfolio went bankrupt during the sample period. These results again suggest the importance of skewness and asymmetric dependence for asset allocation.

For the constrained portfolios we find that the Gumbel significantly out-performs the normal for two out of the five utility functions when using realised utility as a performance measure, and out-performs the constrained normal when using the Sharpe ratio for the most risk averse investor and when using the mean/5% VaR ratio for the investor with risk aversion 3. In only one comparison does the constrained normal portfolio significantly out-perform the constrained Gumbel.

The significance of the comparisons using average realised utility may reflect the fact that this was the objective function used in-sample to compute the optimal portfolio weights. The mean-to-risk measures of performance were not used in the optimisation stage, and so it is not so surprising that comparisons using these measures are not significant. This finding supports a long-standing idea in forecasting that the objective function used in the in-sample optimisation should match the one to be used in out-of-sample evaluation[14].

equal to the maximum of 6 and the largest significant lag in the regressions. The results suggested an average block length of between 25 and 34 observations. We investigated whether the results were sensitive to the choice of average block length, and found that the results were quite robust for average length choices greater than 20.

[13] In this section we bootstrap the average realised utility of a portfolio rather than the 'management fee' discussed above. This is simply for computational ease and should not affect the conclusions drawn.

[14] As mentioned in the introduction, it would be even better if the parameters of the density models were estimated

[ INSERT TABLES 7 AND 8 HERE ]

From Table 7 we see that the unconstrained Gumbel significantly beats the unconstrained NormCop portfolio for 17 out of the 20 possible combinations of loss function and risk aversion level. The unconstrained normal portfolio significantly beats the unconstrained NormCop portfolio on 5 out of 18 comparisons. For the constrained portfolios the Gumbel significantly out-performs the NormCop portfolio on 9 out of the 20 comparisons, while the comparisons between the constrained normal and constrained NormCop contain only one significant result. Overall we can conclude that the unconstrained and constrained Gumbel portfolio significantly beats the corresponding NormCop portfolio, and thus that the benefits of flexible copula modelling are significant.

Although the above results are useful for comparing the results of just two particular models, a more appropriate test would compare all models jointly. With this in mind we now present the results of the reality check test of White (2000). This is a test that a given benchmark portfolio performs as well as the best competing alternative model, where we have possibly many competing alternatives. We present the three estimates of the reality check p-values discussed in Hansen (2001), and focus our attention on the 'consistent' p-value estimates. In these tests we separate the two sets of models into unconstrained and constrained, and include the three naïve portfolios in both sets. Tables 9, 10 and 11 present the results when the 50:50, normal and NormCop portfolios are taken as the benchmarks respectively.

[ INSERT TABLES 9, 10 AND 11 HERE ]

When comparing the 50:50 portfolio with the unconstrained model-based portfolios we are not able to reject that it performs as well as the best alternative for any loss function. Comparing the 50:50 portfolio with the short sales constrained portfolios leads to six rejections out of 20, with most of these occurring for the less risk averse investors. Table 9 thus provides further evidence that placing short sales constraints on the optimal portfolio weights obtained from forecasts improves out-of-sample portfolio performance, see Frost and Savarino (1988) and Jagannathan and Ma (2002). If the short sales constraint is interpreted as a type of 'insanity filter', preventing the investor from taking an extreme position in the market, then this finding corresponds to results previously reported in the forecasting literature, see Stock and Watson (1999) for example, that constrained forecasts often out-perform unconstrained forecasts from non-linear models.

From Table 10 we see that we are able to reject the unconstrained normal portfolio using two out of the three valid utility functions. Table 11 similarly shows that we are often able to reject the unconstrained NormCop portfolio. The constrained normal is only rejected once, and we are unable

using the out-of-sample evaluation function, which in this case is the expected utility from the portfolio constructed using the density forecast. Similar ideas are pursued in Weiss (1996) and Skouras (2001). It is computationally infeasible to do so for the models and evaluation functions used in this paper.

to reject the constrained NormCop portfolio using these loss functions. These results weaken the evidence found using pair-wise comparisons that modelling skewness and asymmetric dependence leads to significantly better out-of-sample portfolio performance, when the optimal portfolio weights are constrained to lie in $[0, 1]$.

# 4 Conclusions and future work

In this paper we considered the impact that skewness and asymmetric dependence have on the out-of-sample portfolio decisions of a CRRA investor. Evidence of skewness in stock returns has been widely reported over the years, and is generally accepted as a common feature of stock returns. Skewness is of interest as any investor that exhibits non-increasing absolute risk aversion, a very weak requirement, can be shown to exhibit a preference for positively skewed assets, *ceteris paribus*. Harvey and Siddique (2000) confirmed empirically that investors require a premium for holding negatively skewed assets. Recent work, see Erb, *et al.*, (1994), Ang and Chen (2002) and Longin and Solnik (2001) *inter alia*, has produced evidence that stock returns exhibit greater dependence during market downturns than during market upturns. Patton (2002) shows that a portfolio of assets that have this asymmetric dependence structure may exhibit negative skewness, even if the individual assets themselves are symmetrically distributed, and so risk-averse investors require a premium for holding such assets. Evidence that assets that have this type of dependence structure with the market portfolio carry a premium is reported in Ang, *et al.*, (2001).

We considered the problem of allocating wealth between the risk-free asset, and the CRSP small-cap and large-cap indices, using monthly data from January 1954 to December 1999. We used the data up to December 1989 to develop the models, and reserved the last 120 months for an out-of-sample evaluation of the competing methods. This problem is representative of that of choosing between a high risk - high return asset and a lower risk - lower return asset. We adopted a parametric approach, using conditional distribution models that are able to capture time-varying conditional means and variances of stock returns, and also (possibly time-varying) skewness and kurtosis. Further, we employed models of the dependence structure of these asset returns that allowed for greater dependence during market downturns than market upturns, and allowed for changes in this dependence structure through time. Our models were constructed using copula theory, which enables us to model separately the individual asset return distributions and their dependence structure, increasing the flexibility in the specification of a parametric density model.

We measured the importance of skewness and asymmetric dependence for asset allocation by comparing the risk-adjusted performance of a portfolio based on a bivariate normal distribution model with a portfolio based on a model developed using copula theory. We also included an 'intermediate' model that captured skewness in the marginal distributions but assumed symmetry for the dependence structure to determine which of marginal distribution modelling or copula modelling were more important for asset allocation.

The significance of the differences in portfolio performance were tested using bootstrap methods. We considered four performance measures on the realised optimal portfolio returns: the Sharpe ratio, a mean/5% Value-at-Risk ratio, a mean/5% expected shortfall ratio and the average realised utility (which was used to compute a type of certainty equivalent measure). We also considered five levels of relative risk aversion, ranging from one to twenty.

Given that the model capturing skewness and asymmetric dependence had 33 parameters versus 26 for the bivariate normal model, it was not clear *a priori* that the more flexible density model would perform better out-of-sample. Indeed, it is a common finding in the point forecasting literature that simpler mis-specified models out-perform more complicated models in forecast comparisons. However we found substantial economic evidence that the model capturing skewness and asymmetric dependence yielded better portfolio decisions than the bivariate normal model. The most flexible density model out-performed the bivariate normal model by 16.7%, when averaging over performance measures and risk aversion levels. The most flexible distribution model out-performed the intermediate model by 52.3% on average, suggesting that both marginal modelling and copula modelling have important implicati

# 5 Appendix 1: Some results on the skewed $t$ distribution.

We provide in this section a few results on Hansen's (1994) skewed $t$ distribution. Hansen provided the *p.d.f.* of the skewed $t$ random variable; below we provide the corresponding *c.d.f.* and inverse *c.d.f.* (useful for random number generation) of the skewed $t$ in terms of the standard Student's $t$ random variable. The motivation for doing this is that most econometric packages (such as Gauss and Matlab) have code available on the standard Students $t$ distribution. With the following results this code can be utilised for the skewed $t$ distribution. Matlab code for each of the functions presented below are available from the author's web site[15].

Let $Y$ be a skewed $t$ random variable, with density function $g(\nu, \lambda)$. The variable $Y$ has mean zero and variance one by construction, and so is a suitable model for the standardised residuals of some conditional mean and variance model. The parameters $\nu$ and $\lambda$ control the kurtosis and skewness of the variable.

**Skewed $t$ density**

$$g(y;\nu,\lambda) = \begin{cases} bc\left(1+\frac{1}{\nu-2}\left(\frac{by+a}{1-\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{for } y \leq -\frac{a}{b} \\ bc\left(1+\frac{1}{\nu-2}\left(\frac{by+a}{1+\lambda}\right)^2\right)^{-(\nu+1)/2} & \text{for } y > -\frac{a}{b} \end{cases}, \quad \text{where} \tag{10}$$

$$c \equiv \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi(\nu-2)}} \tag{11}$$

$$b \equiv \sqrt{1+3\lambda^2-a^2} \tag{12}$$

$$a \equiv 4\lambda c\left(\frac{\nu-2}{\nu-1}\right) \tag{13}$$

Let $X$ be a (standard) Student's $t_\nu$ random variable, with mean zero and variance $\frac{\nu}{\nu-2}$. Denote the *c.d.f.* of $X$ as $F(\nu)$. Below we derive an expression for the *c.d.f.* of a skewed $t$ random variable in terms of $F$.

**Skewed $t$ cumulative distribution function**

$$G(y;\nu,\lambda) = \begin{cases} (1-\lambda)\cdot F\left(\sqrt{\frac{\nu}{\nu-2}}\left(\frac{by+a}{1-\lambda}\right);\nu\right) & \text{for } y \leq -\frac{a}{b} \\ \frac{1-\lambda}{2} + (1+\lambda)\cdot\left[F\left(\sqrt{\frac{\nu}{\nu-2}}\left(\frac{by+a}{1-\lambda}\right);\nu\right)-0.5\right] & \text{for } y > -\frac{a}{b} \end{cases} \tag{14}$$

where $a$, $b$ and $c$ are as defined for the density function.

[15] See http://econ.ucsd.edu/~apatton/code.html.

Finally, we present the inverse *c.d.f.* of the skewed $t$ distribution, which is denoted $G^{-1}(\nu, \lambda)$. We will express it in terms of the inverse distribution of a Student's $t$ random variable, denoted $F^{-1}(\nu)$.

**Inverse Skewed $t$ cumulative distribution function**

$$G^{-1}(u; \nu, \lambda) = \begin{cases} \frac{1-\lambda}{b}\sqrt{\frac{\nu-2}{\nu}} \cdot F^{-1}\left(\frac{u}{1-\lambda}; \nu\right) - \frac{a}{b} & \text{for } 0 < u < \frac{1-\lambda}{2} \\ \frac{1+\lambda}{b}\sqrt{\frac{\nu-2}{\nu}} \cdot F^{-1}\left(0.5 + \frac{1}{1+\lambda}\left(u - \frac{1-\lambda}{2}\right)\right) - \frac{a}{b} & \text{for } \frac{1-\lambda}{2} \leq u < 1 \end{cases}$$

The inverse *c.d.f.* can be used to generate random draws from the skewed $t$ distribution as follows: firstly obtain $n$ draws from the Uniform$(0,1)$ distribution, $\{u_t\}_{t=1}^{n}$. Almost all software packages provide such a feature. Then define $y_t \equiv G^{-1}(u_t; \nu, \lambda)$. The resulting sequence $\{y_t\}_{t=1}^{n}$ are draws from the skewed $t$ distribution. The ability to generate such random variables is useful for Monte Carlo simulations involving this distribution, amongst other things.

# 6 Appendix 3: Copula functional forms

In this appendix we provide the functional forms of the copulas used in this paper. The *c.d.f.* forms will be denoted $C$, and the *p.d.f.* forms $c$. For further details on any of these copulas, or for other copulas, the reader is referred to Joe (1997) and Nelsen (1999).

**Normal Copula**

$$\begin{aligned} C_N(u, v; \rho) &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left\{\frac{-(r^2 - 2\rho rs + s^2)}{2(1-\rho^2)}\right\} dr\ ds \\ c_N(u, v; \rho) &= \frac{1}{\sqrt{1-\rho^2}} \exp\left\{\frac{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(v)}{2(1-\rho^2)} + \frac{\Phi^{-1}(u)^2\,\Phi^{-1}(v)^2}{2}\right\} \\ \rho &\in (-1, 1) \end{aligned}$$

**Clayton Copula (Kimeldorf and Sampson Copula in Joe (1997) )**

$$\begin{aligned} C_C(u, v; \theta) &= \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta} \\ c_C(u, v; \rho) &= (1+\theta)(uv)^{-\theta-1}\left(u^{-\theta} + v^{-\theta} - 1\right)^{-2-1/\theta} \\ \theta &\in [-1, \infty)\backslash\{0\} \end{aligned}$$

**Rotated Clayton Copula**

$$\begin{aligned} C_{RC}(u, v; \theta) &= u + v - 1 + C_C(1-u, 1-v; \theta) \\ c_{RC} &= c_C(1-u, 1-v; \theta) \\ \theta &\in [-1, \infty)\backslash\{0\} \end{aligned}$$

**Joe-Clayton Copula (family BB7 in Joe (1997) )**

$$
\begin{aligned}
C_{\mathsf{JC}}(u,v|\tau^{\mathsf{U}},\tau^{\mathsf{L}}) &= 1-\left(1-\left\{[1-(1-u)^{\kappa}]^{-\gamma}+[1-(1-v)^{\kappa}]^{-\gamma}-1\right\}^{-1/\gamma}\right)^{1/\kappa} \\
c_{\mathsf{JC}}(u,v|\tau^{\mathsf{U}},\tau^{\mathsf{L}}) &= \textit{very long and complicated. Available from the author upon request.} \\
\text{where } \kappa &= \left[\log_2\left(2-\tau^{\mathsf{U}}\right)\right]^{-1} \text{ and } \gamma = \left[-\log_2\left(\tau^{\mathsf{L}}\right)\right]^{-1} \\
\tau^{\mathsf{U}} &\in (0,1)\,,\ \tau^{\mathsf{L}}\in(0,1)
\end{aligned}
$$

**Plackett Copula**

$$
\begin{aligned}
C_{\mathsf{P}}\left(u,v;\pi\right) &= \frac{1}{2\left(\pi-1\right)}\left(1+\left(\pi-1\right)\left(u+v\right)-\sqrt{\left(1+\left(\pi-1\right)\left(u+v\right)\right)^2-4\pi\left(\pi-1\right)uv}\right) \\
c_{\mathsf{P}}\left(u,v;\pi\right) &= \frac{\pi\left(1+\left(\pi-1\right)\left(u+v-2uv\right)\right)}{\left(\left(1+\left(\pi-1\right)\left(u+v\right)\right)^2-4\pi\left(\pi-1\right)uv\right)^{3/2}} \\
\pi &\in [0,\infty)\setminus\{1\}
\end{aligned}
$$

**Frank Copula**

$$
\begin{aligned}
C_{\mathsf{F}}\left(u,v;\lambda\right) &= \frac{-1}{\lambda}\log\left(\frac{\left(1-e^{-\lambda}\right)-\left(1-e^{-\lambda \mathsf{u}}\right)\left(1-e^{-\lambda \mathsf{v}}\right)}{\left(1-e^{-\lambda}\right)}\right) \\
c_{\mathsf{F}}\left(u,v;\lambda\right) &= \frac{\lambda\left(1-e^{-\lambda}\right)e^{-\lambda(\mathsf{u}+\mathsf{v})}}{\left(\left(1-e^{-\lambda}\right)-\left(1-e^{-\lambda \mathsf{u}}\right)\left(1-e^{-\lambda \mathsf{v}}\right)\right)^2} \\
\lambda &\in (-\infty,\infty)\setminus\{0\}
\end{aligned}
$$

**Gumbel Copula**

$$
\begin{aligned}
C_{\mathsf{G}}\left(u,v;\delta\right) &= \exp\left\{-\left(\left(-\log u\right)^{\delta}+\left(-\log v\right)^{\delta}\right)^{1/\delta}\right\} \\
c_{\mathsf{G}}\left(u,v;\delta\right) &= \frac{C_{\mathsf{G}}\left(u,v;\delta\right)\left(\log u\cdot\log v\right)^{\delta-1}}{uv\left(\left(-\log u\right)^{\delta}+\left(-\log v\right)^{\delta}\right)^{2-1/\delta}}\left(\left(\left(-\log u\right)^{\delta}+\left(-\log v\right)^{\delta}\right)^{1/\delta}+\delta-1\right) \\
\delta &\in [1,\infty)
\end{aligned}
$$

**Rotated Gumbel Copula**

$$
\begin{aligned}
C_{\mathsf{RG}}\left(u,v;\delta\right) &= u+v-1+C_{\mathsf{G}}\left(1-u,1-v;\delta\right) \\
c_{\mathsf{RG}}\left(u,v;\delta\right) &= c_{\mathsf{G}}\left(1-u,1-v;\delta\right) \\
\delta &\in [1,\infty)
\end{aligned}
$$

# 7 Tables

| **Table 1**: *Descriptive Statistics* | | |
|---|---|---|
| | *Small caps* | *Large caps* |
| Mean* | 9.9549 | 7.9748 |
| Std Dev* | 21.2932 | 14.2888 |
| Skewness | 0.0558 | −0.3795 |
| 5% VaR | 8.7973 | 6.2306 |
| 1% VaR | 18.9576 | 9.6657 |
| Kurtosis | 7.5647 | 4.9088 |
| Min | −29.3153 | −20.8934 |
| Max | 38.3804 | 16.8145 |
| Jarque-Bera | 479.5162 | 97.0484 |
| *p-val* | *0.0000* | *0.0000* |
| Correlation | 0.7210 | |

Note to Table 1: These summary statistics use the entire data set, from January 1954 to December 1999. The statistics marked with an asterix were annualised to ease interpretation. 'Jarque-Bera' refers to the test for normality of the unconditional distribution of returns.

.

| **Table 2:** *Results from the copula specification search* | | | |
|---|---|---|---|
| *Symmetric copulas* | | *Asymmetric copulas* | |
| Model | $\mathcal{L}_C$ | Model | $\mathcal{L}_C$ |
| *Normal* | 153.5681 | *Clayton* | 151.3272 |
| *Student's t* | 158.1329 | *Rotated Clayton* | 90.8669 |
| *Plackett* | 163.1763 | *Joe-Clayton* | 158.8478 |
| *Frank* | 158.2502 | *Gumbel* | 127.8091 |
| | | *Rotated Gumbel* | 166.6628 |

Note to Table 2: Presented here are the nine copula specifications tried for the copula distribution model. The copula likelihood at the optimum is denoted $\mathcal{L}_C$.

| Table 3: Realised portfolio return summary statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | Unconstrained | | | |
| | Small Caps | Large Caps | 50 mix | Uncond | Normal | NormCop | Gumbel |
| | | | RRA=1 | | | | |
| Mean | 0.9038 | 1.1275 | 1.0157 | 2.6768 | 8.2724 | 5.2659 | 6.6488 |
| Std Dev | 5.3121 | 3.9093 | 4.1928 | 10.3586 | 45.1292 | 30.4298 | 31.4865 |
| Skewness | -0.6128 | -0.6124 | -1.0307 | -1.0854 | -0.3708 | 0.0889 | 0.4814 |
| 5% VaR | 6.9628 | 4.943 | 5.3338 | 12.9446 | 58.0278 | 47.8849 | 44.6935 |
| 5% ES | 11.5049 | 7.9529 | 9.5332 | 22.5199 | 104.5118 | 64.3796 | 62.1585 |
| | | | RRA=3 | | | | |
| Mean | 0.9038 | 1.1275 | 1.0157 | 0.8848 | 4.1397 | 0.7284 | 2.1974 |
| Std Dev | 5.3121 | 3.9093 | 4.1928 | 3.5501 | 24.823 | 10.9912 | 10.4894 |
| Skewness | -0.6128 | -0.6124 | -1.0307 | -1.2768 | -0.2671 | -0.0164 | 0.0492 |
| 5% VaR | 6.9628 | 4.943 | 5.3338 | 4.4658 | 33.1535 | 16.2164 | 12.731 |
| 5% ES | 11.5049 | 7.9529 | 9.5332 | 8.0024 | 57.1473 | 24.7803 | 20.4849 |
| | | | RRA=7 | | | | |
| Mean | 0.9038 | 1.1275 | 1.0157 | 0.4119 | 1.994 | 0.1168 | 1.4336 |
| Std Dev | 5.3121 | 3.9093 | 4.1928 | 1.6026 | 11.5989 | 6.4371 | 6.6022 |
| Skewness | -0.6128 | -0.6124 | -1.0307 | -1.1236 | -0.2151 | -1.0435 | 0.2685 |
| 5% VaR | 6.9628 | 4.943 | 5.3338 | 2.0264 | 14.3828 | 10.8161 | 9.2782 |
| 5% ES | 11.5049 | 7.9529 | 9.5332 | 3.4964 | 27.8403 | 17.6022 | 13.7038 |
| | | | RRA=10 | | | | |
| Mean | 0.9038 | 1.1275 | 1.0157 | 0.289 | 1.4064 | 0.0777 | 1.1397 |
| Std Dev | 5.3121 | 3.9093 | 4.1928 | 1.1251 | 8.1985 | 5.0844 | 4.584 |
| Skewness | -0.6128 | -0.6124 | -1.0307 | -1.1157 | -0.2016 | -1.3305 | 0.5795 |
| 5% VaR | 6.9628 | 4.943 | 5.3338 | 1.4247 | 10.1103 | 8.014 | 6.3274 |
| 5% ES | 11.5049 | 7.9529 | 9.5332 | 2.4449 | 19.6923 | 15.1624 | 8.7192 |
| | | | RRA=20 | | | | |
| Mean | 0.9038 | 1.1275 | 1.0157 | 0.1455 | 0.706 | 0.0117 | 0.6583 |
| Std Dev | 5.3121 | 3.9093 | 4.1928 | 0.5641 | 4.1239 | 2.8471 | 2.7152 |
| Skewness | -0.6128 | -0.6124 | -1.0307 | -1.1169 | -0.1977 | -2.6714 | 0.4628 |
| 5% VaR | 6.9628 | 4.943 | 5.3338 | 0.7146 | 5.0671 | 4.7973 | 3.3004 |
| 5% ES | 11.5049 | 7.9529 | 9.5332 | 1.2247 | 9.9177 | 8.3836 | 5.1940 |

Note to Table 3: ‘RRA’ refers to the coefficient of relative risk aversion. The first two columns of data report the results on the small cap and large cap indices, the third column reports the results for a constant evenly weighted portfolio, the fourth portfolio is based on a weight that is optimised using the empirical unconditional distribution of returns, the fifth portfolio is based on the normal distribution model, the sixth portfolio on the skewed t - normal copula model and the seventh portfolio is based on the skewed t - rotated Gumbel copula model. The rows present the sample mean, sample standard deviation, sample 5% Value-at-Risk (fifth percentile) and sample 5% expected shortfall (mean of returns that exceed the 5% Value-at-Risk).

| Table 3 (cont'd): Realised portfolio return summary statistics | | | | |
|---|---|---|---|---|
| | Subject to short sales constraint | | | |
| | Uncond | Normal | NormCop | Gumbel |
| | RRA=1 | | | |
| Mean | 0.8925 | 1.2659 | 1.263 | 1.2555 |
| Std Dev | 5.1452 | 3.8598 | 3.8636 | 3.8626 |
| Skewness | -0.6608 | -0.2738 | -0.2702 | -0.2655 |
| 5% VaR | 6.6903 | 4.6733 | 4.6733 | 4.6733 |
| 5% ES | 11.3452 | 7.1708 | 7.1708 | 7.1708 |
| | RRA=3 | | | |
| Mean | 0.8935 | 1.2075 | 1.0293 | 1.2074 |
| Std Dev | 3.5243 | 3.806 | 3.6672 | 3.6156 |
| Skewness | -1.1645 | -0.2455 | -0.1251 | -0.1861 |
| 5% VaR | 4.6009 | 4.6733 | 4.3646 | 3.8212 |
| 5% ES | 7.8806 | 7.0491 | 6.9646 | 6.6575 |
| | RRA=7 | | | |
| Mean | 0.4119 | 0.9904 | 0.6961 | 0.8874 |
| Std Dev | 1.6026 | 3.4672 | 2.9854 | 2.9373 |
| Skewness | -1.1236 | -0.4865 | -0.6268 | -0.7429 |
| 5% VaR | 2.0264 | 4.2064 | 3.4382 | 3.2623 |
| 5% ES | 3.4964 | 6.6776 | 5.9938 | 5.7425 |
| | RRA=10 | | | |
| Mean | 0.289 | 0.8236 | 0.553 | 0.7043 |
| Std Dev | 1.1251 | 3.2152 | 2.3789 | 2.3596 |
| Skewness | -1.1157 | -0.6958 | -0.8026 | -0.973 |
| 5% VaR | 1.4247 | 3.9514 | 2.7253 | 2.4289 |
| 5% ES | 2.4449 | 6.4419 | 4.7754 | 4.5365 |
| | RRA=20 | | | |
| Mean | 0.1455 | 0.4832 | 0.2928 | 0.3838 |
| Std Dev | 0.5641 | 2.0592 | 1.2541 | 1.2382 |
| Skewness | -1.1169 | -0.4653 | -0.7411 | -0.8631 |
| 5% VaR | 0.7146 | 2.4188 | 1.5186 | 1.2693 |
| 5% ES | 1.2242 | 6.6262 | 4.3324 | 4.2081 |

Note to Table 3 (continued): 'RRA' refers to the coefficient of relative risk aversion. The first column of data reports the results the fourth portfolio, which is based on a weight that is optimised using the empirical unconditional distribution of returns, the second portfolio is based on the normal distribution model, the third portfolio on the skewed t - normal copula model and the fourth portfolio is based on the skewed t - rotated Gumbel copula model. All of these portfolio weights are subject to a short sales constraint. The rows present the sample mean, sample standard deviation, sample 5% Value-at-Risk (fifth percentile) and sample 5% expected shortfall (mean of returns that exceed the 5% Value-at-Risk).

| Table 4: Realised portfolio return performance statistics | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | | Unconstrained | | | |
| | Small Caps | Large Caps | 50:50 mix | Uncond | Normal | NormCop | Gumbel |
| | | | | RRA=1 | | | |
| Mean/StdDev | 0.1701 | 0.2884 | 0.2422 | 0.2584 | 0.1833 | 0.1731 | 0.2112 |
| Mean/5%VaR | 0.1298 | 0.2281 | 0.1904 | 0.2068 | 0.1426 | 0.11 | 0.1488 |
| Mean/5%ES | 0.0786 | 0.1418 | 0.1065 | 0.1189 | 0.0792 | 0.0818 | 0.1070 |
| Mgmt Fee | -1.9811 | 1.4959 | 0 | 13.9372 | N/A | -15.2901 | 1.8154 |
| | | | | RRA=3 | | | |
| Mean/StdDev | 0.1701 | 0.2884 | 0.2422 | 0.2492 | 0.1668 | 0.0663 | 0.2095 |
| Mean/5%VaR | 0.1298 | 0.2281 | 0.1904 | 0.1981 | 0.1249 | 0.0449 | 0.1726 |
| Mean/5%ES | 0.0786 | 0.1418 | 0.1065 | 0.1106 | 0.0724 | 0.0294 | 0.1073 |
| Mgmt Fee | -3.3108 | 1.8658 | 0 | -0.5508 | N/A | -22.4892 | -2.7387 |
| | | | | RRA=7 | | | |
| Mean/StdDev | 0.1701 | 0.2884 | 0.2422 | 0.257 | 0.1719 | 0.0181 | 0.2171 |
| Mean/5%VaR | 0.1298 | 0.2281 | 0.1904 | 0.2033 | 0.1386 | 0.0108 | 0.1545 |
| Mean/5%ES | 0.0786 | 0.1418 | 0.1065 | 0.1178 | 0.0716 | 0.0066 | 0.1046 |
| Mgmt Fee | -6.3061 | 2.9521 | 0 | 0.1412 | -67.0713 | -26.3494 | -4.8058 |
| | | | | RRA=10 | | | |
| Mean/StdDev | 0.1701 | 0.2884 | 0.2422 | 0.2569 | 0.1715 | 0.0153 | 0.2486 |
| Mean/5%VaR | 0.1298 | 0.2281 | 0.1904 | 0.2029 | 0.1391 | 0.0097 | 0.1801 |
| Mean/5%ES | 0.0786 | 0.1418 | 0.1065 | 0.1182 | 0.0714 | 0.0051 | 0.1307 |
| Mgmt Fee | -9.0655 | 4.2016 | 0 | 3.6384 | -40.6527 | -20.565 | 2.6759 |
| | | | | RRA=20 | | | |
| Mean/StdDev | 0.1701 | 0.2884 | 0.2422 | 0.2579 | 0.1712 | 0.0041 | 0.2424 |
| Mean/5%VaR | 0.1298 | 0.2281 | 0.1904 | 0.2036 | 0.1393 | 0.0024 | 0.1994 |
| Mean/5%ES | 0.0786 | 0.1418 | 0.1065 | 0.1188 | 0.0712 | 0.0014 | 0.1267 |
| Mgmt Fee | -23.0183 | 13.1512 | 0 | 25.9284 | 4.5104 | -1.448 | 23.4016 |

Note to Table 4: 'RRA' refers to the coefficient of relative risk aversion. Each of the seven columns of figures refer to a particular portfolio: the first two portfolios are the assets themselves, the third is a constant evenly weighted portfolio, the fourth portfolio is based on a weight that is optimised using the empirical unconditional distribution of returns, the fifth portfolio is based on the normal distribution model, the sixth portfolio is based on the skewed t - normal copula model and the seventh portfolio is based on the skewed t - rotated Gumbel copula model. The rows present the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the management fee (the number of basis points per year that the investor would be willing to pay to switch from the 50:50 portfolio).

| Table 4 (cont'd): Realised portfolio return performance statistics | | | | |
|---|---|---|---|---|
| | Subject to short sales constraint | | | |
| | Uncond | Normal | NormCop | Gumbel |
| | RRA=1 | | | |
| Mean/StdDev | 0.1735 | 0.328 | 0.3269 | 0.325 |
| Mean/5%VaR | 0.1334 | 0.2709 | 0.2703 | 0.2686 |
| Mean/5%ES | 0.0787 | 0.1765 | 0.1761 | 0.1751 |
| Mgmt Fee | -2.0114 | 25.3021 | 25.2654 | 25.176 |
| | RRA=3 | | | |
| Mean/StdDev | 0.2535 | 0.3173 | 0.2807 | 0.334 |
| Mean/5%VaR | 0.1942 | 0.2584 | 0.2358 | 0.316 |
| Mean/5%ES | 0.1134 | 0.1713 | 0.1478 | 0.1814 |
| Mgmt Fee | -0.4145 | 3.018 | 1.0794 | 3.278 |
| | RRA=7 | | | |
| Mean/StdDev | 0.257 | 0.2856 | 0.2332 | 0.3021 |
| Mean/5%VaR | 0.2033 | 0.2355 | 0.2025 | 0.272 |
| Mean/5%ES | 0.1178 | 0.1483 | 0.1161 | 0.1545 |
| Mgmt Fee | 0.1412 | 2.8306 | 0.6189 | 3.0071 |
| | RRA=10 | | | |
| Mean/StdDev | 0.2569 | 0.2561 | 0.2325 | 0.2985 |
| Mean/5%VaR | 0.2029 | 0.2084 | 0.2029 | 0.29 |
| Mean/5%ES | 0.1182 | 0.1278 | 0.1158 | 0.1552 |
| Mgmt Fee | 3.6384 | 3.762 | 3.7849 | 5.5675 |
| | RRA=20 | | | |
| Mean/StdDev | 0.258 | 0.2346 | 0.2335 | 0.31 |
| Mean/5%VaR | 0.2036 | 0.1998 | 0.1928 | 0.3024 |
| Mean/5%ES | 0.1189 | 0.1134 | 0.1163 | 0.1631 |
| Mgmt Fee | 25.9287 | 24.4493 | 25.958 | 27.0638 |

Note to Table 4 (continued): 'RRA' refers to the coefficient of relative risk aversion. Each of the four columns of figures refer to a particular portfolio: the first portfolio is based on a weight that is optimised using the empirical unconditional distribution of returns, the second portfolio is based on the normal distribution model, the third portfolio is based on the skewed $t$ - normal copula model and the fourth portfolio is based on the skewed $t$ - rotated Gumbel copula model. All of these portfolio weights are subject to a short sales constraint. The rows present the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the management fee (the number of basis points per year that the investor would be willing to pay to switch from the 50:50 portfolio).

Table 5: Pair-wise comparisons of the unconstrained models' risk-adjusted performance

| | Mean / StdDev | Mean / 5%VaR | Mean / 5% ES | Avg Utility |
|---|---|---|---|---|
| | | RRA=1 | | |
| Naïve vs. Uncond | - | - | - | Uncond |
| Naïve vs. Normal | - | - | - | N/A |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | N/A |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | Gumbel | N/A |
| | | RRA=3 | | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | N/A |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | N/A |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | - | N/A |
| | | RRA=7 | | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | Naïve |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | Uncond |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | - | Gumbel |
| | | RRA=10 | | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | Naïve |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | Uncond |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | - | Gumbel |
| | | RRA=20 | | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | - |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | Uncond |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | - | Gumbel |

Note to Table 5: This table presents the results of pair-wise comparisons of the 50:50 portfolio (denoted 'naïve'), the unconditionally optimal portfolio and the portfolios based on the normal distribution, the skewed $t$ - rotated Gumbel copula and the skewed $t$ - normal copula models. The tests were conducted at the 10% significance level. A dash is reported if the test was inconclusive, and the name of the model was reported if that model significantly out-performed the other. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility.

| Table 6: Pair-wise comparisons of the short sales constrained models' risk-adjusted performance | | | | |
|---|---|---|---|---|
| | Mean/StdDev | Mean/5%VaR | Mean/5%ES | Avg Utility |
| | | | RRA=1 | |
| Naïve vs. Uncond | Naïve | Naïve | Naïve | - |
| Naïve vs. Normal | Normal | - | Normal | - |
| Naïve vs. Gumbel | Gumbel | - | Gumbel | - |
| Uncond vs. Normal | Normal | Normal | Normal | - |
| Uncond vs. Gumbel | Gumbel | Gumbel | Gumbel | - |
| Gumbel vs. Normal | - | - | - | Normal |
| | | | RRA=3 | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | Normal | - |
| Naïve vs. Gumbel | Gumbel | Gumbel | Gumbel | - |
| Uncond vs. Normal | - | - | Normal | - |
| Uncond vs. Gumbel | Gumbel | Gumbel | Gumbel | - |
| Gumbel vs. Normal | - | Gumbel | - | - |
| | | | RRA=7 | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | Normal | - |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | Normal | Normal |
| Uncond vs. Gumbel | - | - | - | Gumbel |
| Gumbel vs. Normal | - | - | - | - |
| | | | RRA=10 | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | - |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | - |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | - | - | - | Gumbel |
| | | | RRA=20 | |
| Naïve vs. Uncond | - | - | - | - |
| Naïve vs. Normal | - | - | - | - |
| Naïve vs. Gumbel | - | - | - | - |
| Uncond vs. Normal | - | - | - | - |
| Uncond vs. Gumbel | - | - | - | - |
| Gumbel vs. Normal | Gumbel | - | - | Gumbel |

Note to Table 6: This table presents the results of pair-wise comparisons of the 50:50 portfolio (denoted 'naïve'), the unconditionally optimal portfolio and the portfolios based on the normal distribution, the skewed $t$ - rotated Gumbel copula and the skewed $t$ - normal copula models. All portfolio weights in these comparisons are subject to a short sales constraint. The tests were conducted at the 10% significance level. A dash is reported if the test was inconclusive, and the name of the model was reported if that model significantly out-performed the other. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility.

| Table 7: Pair-wise comparisons of the unconstrained models' risk-adjusted performance | | | | |
|---|---|---|---|---|
| | Mean / StdDev | Mean / 5%VaR | Mean / 5% ES | Avg Utility |
| | RRA=1 | | | |
| Naïve vs. NormCop | - | - | - | - |
| Uncond vs. NormCop | - | - | - | - |
| Normal vs. NormCop | - | - | - | N/A |
| Gumbel vs. NormCop | Gumbel | Gumbel | Gumbel | Gumbel |
| | RRA=3 | | | |
| Naïve vs. NormCop | - | - | - | - |
| Uncond vs. NormCop | - | - | - | - |
| Normal vs. NormCop | - | - | - | N/A |
| Gumbel vs. NormCop | Gumbel | Gumbel | Gumbel | Gumbel |
| | RRA=7 | | | |
| Naïve vs. NormCop | - | Naïve | - | - |
| Uncond vs. NormCop | - | Uncond | - | Uncond |
| Normal vs. NormCop | Normal | Normal | - | - |
| Gumbel vs. NormCop | Gumbel | - | - | Gumbel |
| | RRA=10 | | | |
| Naïve vs. NormCop | Naïve | Naïve | - | Naïve |
| Uncond vs. NormCop | Uncond | Uncond | Uncond | Uncond |
| Normal vs. NormCop | Normal | Normal | - | - |
| Gumbel vs. NormCop | Gumbel | Gumbel | Gumbel | Gumbel |
| | RRA=20 | | | |
| Naïve vs. NormCop | - | Naïve | - | - |
| Uncond vs. NormCop | - | Uncond | - | - |
| Normal vs. NormCop | - | Normal | - | - |
| Gumbel vs. NormCop | Gumbel | Gumbel | - | Gumbel |

Note to Table 7: This table presents the results of pair-wise comparisons of the 50:50 portfolio (denoted 'naïve'), the unconditionally optimal portfolio and the portfolios based on the normal distribution, the skewed $t$ - rotated Gumbel copula and the skewed $t$ - normal copula models. The tests were conducted at the 10% significance level. A dash is reported if the test was inconclusive, and the name of the model was reported if that model significantly out-performed the other. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility.

| Table 8: Pair-wise comparisons of the short sales constrained models' risk-adjusted performance | | | | |
|---|---|---|---|---|
| | Mean / StdDev | Mean / 5%VaR | Mean / 5% ES | Avg Utility |
| | RRA=1 | | | |
| Naïve vs. NormCop | NormCop | - | NormCop | - |
| Uncond vs. NormCop | NormCop | NormCop | NormCop | - |
| Normal vs. NormCop | - | - | - | - |
| Gumbel vs. NormCop | - | - | NormCop | - |
| | RRA=3 | | | |
| Naïve vs. NormCop | - | - | NormCop | - |
| Uncond vs. NormCop | - | - | NormCop | - |
| Normal vs. NormCop | - | - | - | Normal |
| Gumbel vs. NormCop | - | Gumbel | - | Gumbel |
| | RRA=7 | | | |
| Naïve vs. NormCop | - | - | NormCop | - |
| Uncond vs. NormCop | - | - | - | - |
| Normal vs. NormCop | - | - | - | - |
| Gumbel vs. NormCop | Gumbel | - | - | Gumbel |
| | RRA=10 | | | |
| Naïve vs. NormCop | - | - | NormCop | - |
| Uncond vs. NormCop | - | - | NormCop | - |
| Normal vs. NormCop | - | - | - | - |
| Gumbel vs. NormCop | Gumbel | - | NormCop | Gumbel |
| | RRA=20 | | | |
| Naïve vs. NormCop | - | - | NormCop | - |
| Uncond vs. NormCop | - | - | NormCop | - |
| Normal vs. NormCop | - | - | - | - |
| Gumbel vs. NormCop | Gumbel | Gumbel | - | Gumbel |

Note to Table 8: This table presents the results of pair-wise comparisons of the 50:50 portfolio (denoted 'naïve'), the unconditionally optimal portfolio and the portfolios based on the normal distribution, the skewed $t$ - rotated Gumbel copula and the skewed $t$ - normal copula models. All portfolio weights in these comparisons are subject to a short sales constraint. The tests were conducted at the 10% significance level. A dash is reported if the test was inconclusive, and the name of the model was reported if that model significantly out-performed the other. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility.

Table 9: Bootstrap Reality Check p-values, '50:50 mix' as benchmark

| | Unconstrained | | | Short sales constrained | | |
|---|---|---|---|---|---|---|
| | Lower | Consistent | Upper | Lower | Consistent | Upper |
| | | | RRA=1 | | | |
| Mean/StdDev | 0.233 | 0.371 | 0.406 | 0.044 | **0.044** | 0.044 |
| Mean/5%VaR | 0.323 | 0.45 | 0.52 | 0.175 | 0.175 | 0.186 |
| Mean/5%ES | 0.26 | 0.308 | 0.321 | 0.026 | **0.026** | 0.028 |
| Avg Utility* | 0.259 | 0.259 | 0.462 | 0.177 | 0.177 | 0.266 |
| | | | RRA=3 | | | |
| Mean/StdDev | 0.285 | 0.412 | 0.456 | 0.043 | **0.043** | 0.043 |
| Mean/5%VaR | 0.428 | 0.493 | 0.573 | 0.049 | **0.049** | 0.049 |
| Mean/5%ES | 0.296 | 0.349 | 0.359 | 0.039 | **0.039** | 0.039 |
| Avg Utility* | 0.352 | 0.43 | 0.736 | 0.196 | 0.196 | 0.291 |
| | | | RRA=7 | | | |
| Mean/StdDev | 0.36 | 0.474 | 0.52 | 0.245 | 0.245 | 0.255 |
| Mean/5%VaR | 0.438 | 0.533 | 0.622 | 0.261 | 0.261 | 0.263 |
| Mean/5%ES | 0.375 | 0.433 | 0.452 | 0.155 | 0.155 | 0.158 |
| Avg Utility | 0.327 | 0.451 | 0.843 | 0.313 | 0.313 | 0.481 |
| | | | RRA=10 | | | |
| Mean/StdDev | 0.349 | 0.402 | 0.426 | 0.271 | 0.271 | 0.291 |
| Mean/5%VaR | 0.487 | 0.532 | 0.579 | 0.192 | 0.192 | 0.195 |
| Mean/5%ES | 0.298 | 0.368 | 0.38 | 0.178 | 0.178 | 0.188 |
| Avg Utility | 0.28 | 0.28 | 0.657 | 0.175 | 0.175 | 0.312 |
| | | | RRA=20 | | | |
| Mean/StdDev | 0.409 | 0.447 | 0.485 | 0.181 | 0.195 | 0.198 |
| Mean/5%VaR | 0.561 | 0.578 | 0.636 | 0.137 | 0.137 | 0.137 |
| Mean/5%ES | 0.405 | 0.469 | 0.484 | 0.157 | 0.157 | 0.158 |
| Avg Utility | 0.145 | 0.145 | 0.373 | 0.057 | **0.057** | 0.251 |

Note to Table 9: This table presents the results of the reality check of White (2000), as modified by Hansen (2001). 'Lower', 'Consistent' and 'Upper' refer to three estimates of the p-value of the test statistic. A p-value of less than 0.10 indicates that we may reject the hypothesis that the benchmark model performs as well as the best alternative model considered according to the given performance measure. Any rejections are marked with **bold** font. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility. *For the comparisons of unconstrained portfolios for investors with risk aversion of 1 and 3 using average realised utility we excluded the bivariate normal portfolio, as it went bankrupt during the sample.

| Table 10: Bootstrap Reality Check p-values, 'Normal' as benchmark | | | | | | |
|---|---|---|---|---|---|---|
| | Unconstrained | | | Short sales constrained | | |
| | Lower | Consistent | Upper | Lower | Consistent | Upper |
| | RRA=1 | | | | | |
| Mean/StdDev | 0.22 | 0.231 | 0.231 | 0.243 | 0.243 | 0.89 |
| Mean/5%VaR | 0.231 | 0.244 | 0.244 | 0.221 | 0.444 | 0.827 |
| Mean/5%ES | 0.225 | 0.225 | 0.225 | 0.195 | 0.195 | 0.813 |
| Avg Utility | N/A | N/A | N/A | 0.316 | 0.316 | 0.866 |
| | RRA=3 | | | | | |
| Mean/StdDev | 0.239 | 0.239 | 0.239 | 0.387 | 0.47 | 0.63 |
| Mean/5%VaR | 0.237 | 0.237 | 0.237 | 0.09 | 0.197 | 0.333 |
| Mean/5%ES | 0.271 | 0.271 | 0.271 | 0.378 | 0.523 | 0.624 |
| Avg Utility | N/A | N/A | N/A | 0.586 | 0.667 | 0.792 |
| | RRA=7 | | | | | |
| Mean/StdDev | 0.362 | 0.364 | 0.364 | | | |

Table 11: Bootstrap Reality Check p-values, 'NormCop' as benchmark

| | Unconstrained | | | Short sales constrained | | |
|---|---|---|---|---|---|---|
| | Lower | Consistent | Upper | Lower | Consistent | Upper |
| | RRA=1 | | | | | |
| Mean/StdDev | 0.22 | 0.225 | 0.225 | 0.341 | 0.341 | 0.859 |
| Mean/5%VaR | 0.148 | 0.148 | 0.148 | 0.488 | 0.68 | 0.892 |
| Mean/5%ES | 0.257 | 0.261 | 0.261 | 0.42 | 0.42 | 0.888 |
| Avg Utility* | 0.126 | 0.126 | 0.405 | 0.556 | 0.556 | 0.932 |
| | RRA=3 | | | | | |
| Mean/StdDev | 0.104 | 0.104 | 0.104 | 0.269 | 0.327 | 0.383 |
| Mean/5%VaR | 0.07 | **0.07** | 0.07 | 0.169 | 0.253 | 0.287 |
| Mean/5%ES | 0.143 | 0.143 | 0.143 | 0.274 | 0.324 | 0.355 |
| Avg Utility* | 0.066 | **0.066** | 0.317 | 0.319 | 0.368 | 0.47 |
| | RRA=7 | | | | | |
| Mean/StdDev | 0.06 | **0.06** | 0.06 | 0.273 | 0.273 | 0.306 |
| Mean/5%VaR | 0.082 | **0.082** | 0.082 | 0.291 | 0.299 | 0.324 |
| Mean/5%ES | 0.137 | 0.137 | 0.137 | 0.33 | 0.354 | 0.354 |
| Avg Utility | 0.067 | **0.067** | 0.305 | 0.349 | 0.394 | 0.493 |
| | RRA=10 | | | | | |
| Mean/StdDev | 0.017 | **0.017** | 0.017 | 0.276 | 0.276 | 0.305 |
| Mean/5%VaR | 0.009 | **0.009** | 0.009 | 0.245 | 0.253 | 0.278 |
| Mean/5%ES | 0.052 | **0.052** | 0.052 | 0.324 | 0.343 | 0.343 |
| Avg Utility | 0.023 | **0.023** | 0.224 | 0.38 | 0.511 | 0.579 |
| | RRA=20 | | | | | |
| Mean/StdDev | 0.05 | **0.05** | 0.05 | 0.247 | 0.247 | 0.266 |
| Mean/5%VaR | 0.032 | **0.032** | 0.032 | 0.131 | 0.133 | 0.142 |
| Mean/5%ES | 0.131 | 0.131 | 0.131 | 0.283 | 0.31 | 0.31 |
| Avg Utility | 0.238 | 0.38 | 0.38 | 0.151 | 0.161 | 0.611 |

Note to Table 11: This table presents the results of the reality check of White (2000), and modified by Hansen (2001). 'Lower', 'Consistent' and 'Upper' refer to three estimates of the p-value of the test statistic. A p-value of less than 0.10 indicates that we may reject the hypothesis that the benchmark model performs as well as the best alternative model considered according to the given performance measure. Any rejections are marked with **bold** font. The four performance measures are the sample Sharpe ratio (mean to standard deviation), sample mean to 5% Value-at-Risk ratio, sample mean to 5% expected shortfall ratio and the sample mean of the realised utility. *For the comparisons of unconstrained portfolios for investors with risk aversion of 1 and 3 using average realised utility we excluded the bivariate normal portfolio, as it went bankrupt during the sample.

## References


[1] Aït-Sahalia, Yacine, and Brandt, Michael W., 2001, Variable Selection for Portfolio Choice, *Journal of Finance*, 56(4), 1297-1355.

[2] Ang, Andrew, and Bekaert, Geert, 2001, International Asset Allocation with Regime Shifts, forthcoming, *Review of Financial Studies.*

[3] Ang, Andrew, and Chen, Joe, 2002, Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics*, 63(3), 443-494.

[4] Ang, Andrew, Chen, Joseph, and Xing, Yuhang, 2001, Downside Risk and Expected Returns, working paper, Columbia Business School.

[5] Arrow, Kenneth J., 1971, *Essays in the Theory of Risk Bearing*, Markham Publishing Co., Chicago.

[6] Artzner, Phillipe, Delbaen, Freddy, Eber, Jean-Marc, and Heath, David, 1999, Coherent Measures of Risk, *Mathematical Finance*, 9(3), 203-228.

[7] Bae, Kee-Hong, Karolyi, Andrew and Stulz, René M., 2000, A New Approach to Measuring Financial Contagion, forthcoming, *Review of Financial Studies.*

[8] Brandt, Michael W., 1999, Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach, *Journal of Finance*, 54(5), 1609-1645.

[9] Campbell, John Y., and Viceira, Luis M., 1999, Consumption and Portfolio Decisions when Expected Returns are Time Varying, *Quarterly Journal of Economics*, 114, 433-495.

[10] Campbell, Rachel, Koedijk, Kees, and Kofman, Paul, 2000, Increased Correlation in Bear Markets: A Downside Risk Perspective, forthcoming, *Financial Analysts Journal*.

[11] Clayton, D. G., 1978, A Model for Association in Bivariate Life Tables and its Application in Epidemiological Studies of Familial Tendency in Chronic Disease Incidence, *Biometrika*, 65(1), 141-151.

[12] Cook, R. Dennis, and Johnson, Mark E., 1981, A Family of Distributions for Modelling Non-Elliptically Symmetric Multivariate Data, *Journal of the Royal Statistical Society, Series B,* 43(2), 210-218.

[13] Efron, B., and Tibshirani, R. J., 1993, *An Introduction to the Bootstrap*, Chapman and Hall, New York.

[14] Embrechts, Paul, Höing, Andrea and Juri, Alessandro, 2001, Using Copulae to Bound the Value-at-Risk for Functions of Dependent Risks, forthcoming, *Finance and Stochastics.*

[15] Erb, Claude B., Harvey, Campbell R., and Viskanta, Tadas E., 1994, Forecasting International Equity Correlations, *Financial Analysts Journal*, 50, 32-45.

[16] Fama, Eugene F., 1981, Stock Returns, Real Activity, Inflation, and Money, *American Economic Review*, 71, 545-565.

[17] Fine, J. P., and Jiang, H., 2000, On Association in a Copula with Time Transformations, *Biometrika*, 87(3), 559-571.

[18] Frost, Peter A., and Savarino, James E., 1988, For Better Performance: Constrain Portfolio Weights, *Journal of Portfolio Management*, 15(1), 29-34.

[19] Genest, C., and Rivest, L.-P., 1993, Statistical Inference Procedures for Bivariate Archimedean Copulas, *Journal of the American Statistical Association*, 88(423), 1034-1043.

[20] Granger, Clive W. J., 1969, Prediction with a Generalized Cost of Error Function, *Operational Research Quarterly*, 20(2), 199-207.

[21] Hansen, Bruce E., 1994, Autoregressive Conditional Density Estimation, *International Economic Review*, 35(3), 705-730.

[22] Hansen, Peter R., 2001, An Unbiased and Powerful Test for Superior Predictive Ability, Working Paper 01-06, Department of Economics, Brown University.

[23] Harvey, Campbell R., and Siddique, Akhtar, 1999, Autoregressive Conditional Skewness, *Journal of Financial and Quantitative Analysis*, 34(4), 465-488.

[24] Harvey, Campbell R., and Siddique, Akhtar, 2000, Conditional Skewness in Asset Pricing Tests, *Journal of Finance*, 55(3), 1263-1295.

[25] Huang, Chi-fu, and Litzenberger, Robert H., 1988, *Foundations for Financial Economics*, Prentice-Hall Inc., New Jersey.

[26] Ingersoll, Jonathan E., Jr., 1987, *Theory of Financial Decision Making*, Rowman and Littlefield Publishers Inc., Maryland.

[27] Jagannathan, Ravi, and Ma, Tongshu, 2002, Risk Reduction in Large Portfolios: A Role for Portfolio Weight Constraints, working paper, Department of Finance, Kellogg Graduate School of Management, Northwestern University.

[28] Joe, Harry, 1997, *Multivariate Models and Dependence Concepts*, Monographs on Statistics and Applied Probability 73, Chapman and Hall, London, U.K.

[29] Judd, Kenneth L., 1998, *Numerical Methods in Economics*, MIT Press, Cambridge, Massachusetts.

[30] Kandel, Shmuel, and Stambaugh, Robert F., 1996, On the Predictability of Stock Returns: An Asset Allocation Perspective, *Journal of Finance*, 51(2), 385-424.

[31] Li, David X., 2000, On Default Correlation: A Copula Function Approach, *Journal of Fixed Income*, 9(4), 43-54.

[32] Lim, Kian-Guan, 1989, A New Test of the Three-Moment Capital Asset Pricing Model, *Journal of Financial and Quantitative Analysis*, 24(2), 205-216.

[33] Longin, François, and Solnik, Bruno, 2001, Extreme Correlation of International Equity Markets, *Journal of Finance*, 56(2), 649-676.

[34] Merton, Robert C., 1971, Optimal Consumption and Portfolio Rules in a Continuous-Time Model, *Journal of Economic Theory*, 3, 373-413.

[35] Miller, Douglas J., and Liu, Wei-han, 2002, On the Recovery of Joint Distributions from Limited Information, *Journal of Econometrics*, 107, 259-274.

[36] Nelsen, Roger B., 1999, *An Introduction to Copulas*, Springer-Verlag, New York.

[37] Oakes, David, 1989, Bivariate Survival Models Induced by Frailties, *Journal of the American Statistical Association*, 84(406), 487-493.

[38] Patton, Andrew J., 2001a, Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula, Working Paper 2001-09, Department of Economics, University of California, San Diego.

[39] Patton, Andrew J., 2001b, Estimation of Copula Models for Time Series of Possibly Different Lengths, Working Paper 2001-17, Department of Economics, University of California, San Diego.

[40] Patton, Andrew J., 2002, Skewness, Asymmetric Dependence, and Portfolios, working paper, Department of Economics, University of California, San Diego.

[41] Perez-Quiros, Gabriel, and Timmermann, Allan, 2001, Business Cycle Asymmetries in Stock Returns: Evidence from Higher Order Moments and Conditional Densities, *Journal of Econometrics*, 103, 259-306.

[42] Politis, Dimitris N., and Romano, Joseph P., 1994, The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.

[43] Rockinger, Michael, and Jondeau, Eric, 2001, Conditional Dependency of Financial Series: An Application of Copulas, working paper, Department of Finance, HEC School of Management, France.

[44] Rosenberg, Joshua V., 1999, Semiparametric Pricing of Multivariate Contingent Claims, working paper, Department of Finance, Stern School of Business, New York University.

[45] Rosenberg, Joshua V., 2000, Nonparametric Pricing of Multivariate Contingent Claims, working paper, Department of Finance, Stern School of Business, New York University.

[46] Scaillet, O., 2000, Nonparametric Estimation of Copulas for Time Series, working paper, Département des Sciences Economiques, Université Catholique de Louvain, Belgium.

[47] Schweizer, B., and Sklar, A., 1983, *Probabilistic Metric Spaces*, Elsevier Science, New York.

[48] Sklar, A., 1959, Fonctions de répartition à n dimensionset leurs marges, *Publ. Inst. Statis. Univ. Paris*, 8, 229-231.

[49] Skouras, Spyros, Decisionmetrics: A Decision-Based Approach to Econometric Modelling, Working Paper 01-11-064, Santa Fe Institute, USA.

[50] Stock, James H., and Watson, Mark W., 1999, A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series, in Robert F. Engle and Halbert White eds., *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*, Oxford University Press, Oxford, U.K., 1-44.

[51] Swanson, Norman R., and White, Halbert, 1995, A Model Selection Approach to Assessing the Information in the Term Structure Using Linear Models and Artificial Neural Networks, *Journal of Business and Economic Statistics*, 13, 265-275.

[52] Swanson, Norman R., and White, Halbert, 1997, A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks, *Review of Economics and Statistics*, 79, 540-550.

[53] Weiss, Andrew A., 1996, Estimating Time Series Models Using the Relevant Cost Function, *Journal of Applied Econometrics*, 11(5), 539-560.

[54] West, Kenneth D., Edison, Hali J., and Cho, Dongchul, 1993, A Utility-Based Comparison of Some Models of Exchange Rate Volatility, *Journal of International Economics*, 35, 23-45.

[55] White, Halbert, 2000, A Reality Check for Data Snooping, *Econometrica*, 68(5), 1097-1126.

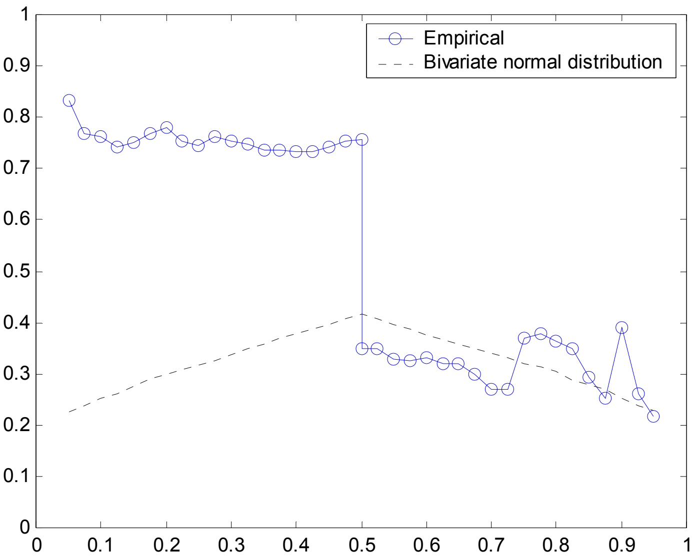



Figure 1: *Exceedence correlations between excess returns (X and Y) on small caps and large caps. The horizontal axis shows the cut-off quantile, and the vertical axis shows the correlation between the two returns given that both exceed that quantile.*

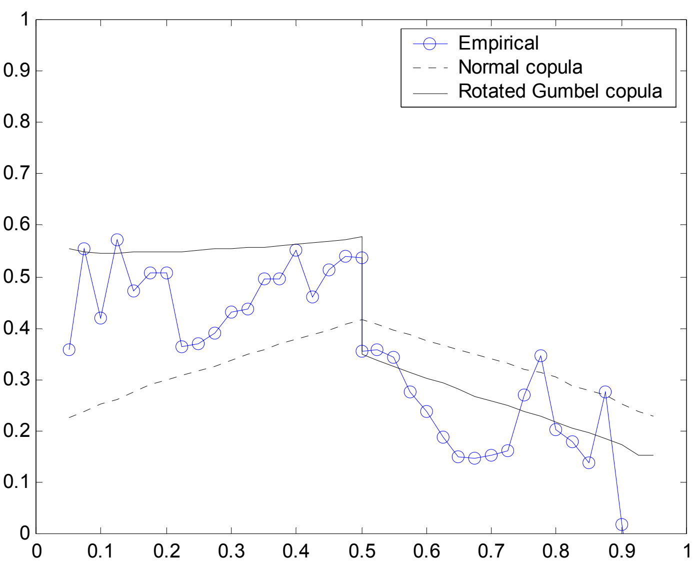
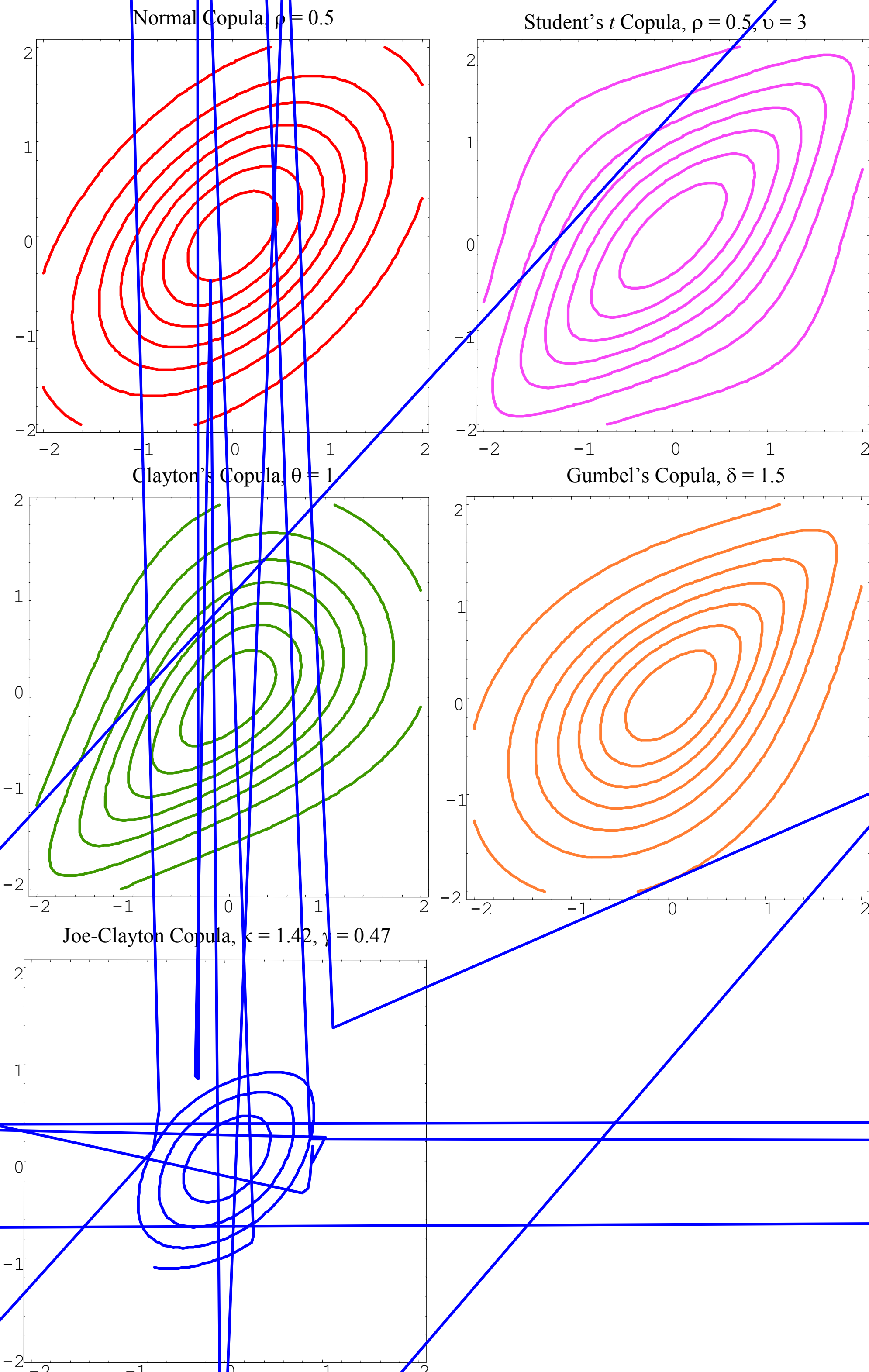



Figure 2: *Exceedence correlations between transformed residuals (U and V) of small caps and large caps. The horizontal axis shows the cut-off quantile, and the vertical axis shows the correlation between the two residuals given that both exceed that quantile.*


Normal Copula, ρ = 0.5
Student's t Copula, ρ = 0.5, υ = 3
Clayton's Copula, θ = 1
Gumbel's Copula, δ = 1.5
Joe-Clayton Copula, κ = 1.42, γ = 0.47