

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

ANDREW J. PATTON

London School of Economics

ABSTRACT

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. 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 article 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 constant relative risk aversion (CRRA) investor allocating wealth between the risk-free asset, a small-cap portfolio, and a large-cap portfolio. We use models that can capture time-varying moments up to the fourth order, and we use copula theory to construct models of the time-varying dependence structure that allow for different 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 investors with no short-sales constraints, we find that knowledge of higher moments and asymmetric dependence leads to gains that are economically significant and statistically significant in some cases. For short sales-constrained investors the gains are limited.

KEYWORDS: asymmetry, copulas, density forecasting, forecasting, normality, stock returns

This article is a revision of Chapter IV of my Ph.D. dissertation [Patton (2002)]. I would like to thank Sean Campbell, Rob Engle, Raffaella Giacomini, Clive Granger, Kris Jacobs, Bruce Lehmann, two anonymous referees, and seminar participants at the Econometric Society meetings, the 2002 Inquire UK annual seminar in Bournemouth, and the October 2002 Extremal Events in Finance conference in Montreal for helpful comments. Special thanks are due to Allan Timmermann for many useful discussions on this topic. All remaining deficiencies are my responsibility. 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. Address correspondence to Andrew J. Patton, Financial Markets Group, London School of Economics, Houghton Street, London WC2A 2AE, UK, or e-mail: a.patton@lse.ac.uk.

Journal of Financial Econometrics, Vol. 2, No. 1, pp. 130–168

© 2004 Oxford University Press; all rights reserved.

DOI: 10.1093/jfinec/nbh006

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 or asymmetry in the distribution of individual stock returns, which has been reported by numerous authors over the last three decades.¹ Evidence that stock returns exhibit some form of asymmetric dependence has been reported by several authors in recent years [see Erb, Harvey, and Viskanta (1994), Longin and Solnik (2001), Ang and Bekaert (2002), Ang and Chen (2002), Campbell, Koedijk, and Kofman (2002), and Bae, Karolyi, and Stulz (2003)]. 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 article we examine the economic and statistical significance of these two asymmetries for asset allocation decisions in an out-of-sample setting. This article can thus be viewed as an attempt to address the suggestions of Harvey and Siddique (1999) 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 nonincreasing absolute risk aversion.² Under nonincreasing absolute risk aversion investors can be shown to have a preference for positively skewed portfolios. The skewness of a portfolio of two assets is a function of the skewness of the individual assets, and two “coskewness” terms. Asymmetry in the dependence structure can be shown [see Patton (2002)] to lead to nonzero coskewness and thus impact the skewness of the portfolio return. This suggests that risk-averse investors will have preferences over alternative dependence structures. Ang, Chen, and Xing (2002) report empirical evidence in support of this.

We examine the problem of an investor with constant relative risk aversion (CRRA) allocating wealth between the risk-free asset, the Center for Research in Security Prices (CRSP) small cap and large cap indices, comprised of the 1st and 10th decile of U.S. stocks sorted by market capitalization. We use monthly data from January 1954 to December 1989 to develop the models, and data from January 1990 to December 1999 for out-of-sample forecast evaluation. This problem is representative of that of choosing between a high risk–high return asset and a lower risk–lower return asset, as the annualized 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. Our motivation for studying a problem involving two stocks rather than a stock and a bond, as in numerous previous studies, is that evidence of asymmetric dependence has so far been reported only

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 Ait-Sahalia and Brandt (2001), among others. Perió (1999) finds no such evidence.

² Utility functions that exhibit nonincreasing absolute risk aversion include the constant absolute risk aversion utility function, and the constant relative risk aversion utility function, see Huang and Litzenberger (1988).

for equity returns. The presence or absence of asymmetric dependence between equity and bond returns is yet to be established.

We use models of the asset returns 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, as in Hansen (1994) and Jondeau and Rockinger (2003). Further, we employ models of the dependence structure (or copula) that allow for, but do not impose different dependence during bear markets than bull markets, and allow for changes in this dependence structure through time. A thorough introduction to copula theory is presented in Schweizer and Sklar (1983), Joe (1997), and Nelsen (1999).

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. We compute the amount that an investor could be charged to make him/her indifferent between two competing portfolios, as in West, Edison, and Cho (1993), Ang and Bekaert (2002), and others. The significance of the differences in portfolio performance are tested using bootstrap methods. We find evidence in most cases that nonnormalities in the marginal distributions and copula do have important economic implications for asset allocation, however, the statistical significance of the improvement is only moderate. Gains are generally only present for investors that are not short-sales constrained, such as hedge funds.

This article is essentially trying to test three hypotheses simultaneously: (1) Are these asymmetries present in this dataset? (2) Are these asymmetries predictable out-of-sample? (3) Can we make better portfolio decisions by using forecasts of these asymmetries than we can by ignoring them? If the answer to any of these questions is "no," then we would conclude that the out-of-sample importance of these asymmetries for asset allocation is zero. The distinction between in-sample and out-of-sample significance 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 *poorer* forecasts than simple misspecified models [see Weigend and Gershenfeld (1994), Swanson and White (1995, 1997), and Stock and Watson (1999)].

In this article we consider both unconstrained and short sales-constrained estimates of the optimal portfolio weight. The first reason for doing so 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, while others, such as hedge funds, actively take both long and short positions. 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" that 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)]. One could also consider an intermediate filter that allows for some limited amount of short selling, but we do not explore such a possibility here.

Much of the existing work on asset allocation focused on special cases where the combination of utility function and distribution model were such that an analytical solution for the optimal portfolio decision exists [see Kandel and Stambaugh (1996) or Campbell and Viceira (1999), among others]. 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 optimization problem to obtain an optimal portfolio decision. Detemple, Garcia and Rindisbacher (2003) present a sophisticated new method for finding optimal portfolio weights from empirically relevant models. In this article we combine density models that are shown to adequately describe the statistical properties of the asset returns with the CRRA utility function.

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. First, we ignore the effects of parameter estimation uncertainty on the investor's decision problem, though this was found to be important by Kandel and Stambaugh (1996). Also, we only consider the investor's problem for the one-period-ahead investment horizon, thus ignoring the hedging component of the optimal portfolio weight [see Merton (1971)]. Empirical evidence on the importance of the hedging component is mixed: Brandt (1999), Campbell and Viceira (1999), and Detemple, Garcia, and Rindisbacher (2003) find it to be important, whereas Aït-Sahalia and Brandt (2001) and Ang and Bekaert (2002) find only weak evidence.

The remainder of the article is structured as follows. In Section 1 we provide a brief introduction to copula theory and its use in the density forecasting of stock returns. In Section 2 we present the investor's decision problem in detail. Section 3 presents the empirical results on the asset allocation problem for a portfolio of a small-cap index and a large-cap index: the models employed, comparisons of portfolio weights, and tests for improvements in portfolio performance. We conclude in Section 4. In Appendix A we present some details of the optimization procedure and in Appendix B we provide the functional forms of the copula models considered.

1 FLEXIBLE MULTIVARIATE DISTRIBUTION MODELS USING COPULAS

In this article we use copula theory to develop flexible parametric models of the joint distribution of returns. Suppose 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 $(X_t, Y_t) | \mathcal{F}_{t-1} \sim H_t = C_t(F_t, G_t)$, where H_t is some conditional bivariate distribution function, with conditional univariate distributions of X_t and Y_t being F_t and G_t , the conditional copula being C_t , and \mathcal{F}_{t-1} is the information set

defined as $\mathcal{F}_t \equiv \sigma(\mathbf{Z}_t)$, for $\mathbf{Z}_t \equiv [X_t, Y_t, \mathbf{W}'_t, X_{t-1}, Y_{t-1}, \mathbf{W}'_{t-1}, \dots, X_{t-j}, Y_{t-j}, \mathbf{W}'_{t-j}]'$. We will denote the distribution (*cdf*) of a random variable using an uppercase letter and the corresponding density (*pdf*) using a lowercase letter.

A copula is any multivariate distribution function that has Uniform (0,1) marginal distributions. It 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) and since then numerous applications have appeared in the statistics literature and more recently also in the analysis of economic data.³ The main theorem in copula theory is that of Sklar (1959), presented below for the conditional case. For an introduction to copula theory see Joe (1997) and Nelsen (1999).

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 , and let \mathcal{Z} be the support of Z . Then there exists a unique conditional copula C such that*

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

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 .

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 . The density function equivalent of Equation (1) is obtained quite easily, provided that F_t and G_t are differentiable, and H_t and C_t are twice differentiable:

$$h_t(x, y|z) = f_t(x|z) \cdot g_t(y|z) \cdot c_t(u, v|z), \quad \forall (x, y, z) \in \bar{\mathbb{R}} \times \bar{\mathbb{R}} \times \mathcal{Z}, \quad (2)$$

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

$$\log h_t(x, y|z) = \log f_t(x|z) + \log g_t(y|z) + \log c_t(u, v|z) \quad (3)$$

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 modeling, the copula representation allows for great flexibility in the specification: we may model the individual variables using whichever marginal distributions provide the best fit and then work on modeling the dependence structure via a model for the copula. The estimation of multivariate time-series models constructed using

³ In statistics, see Clayton (1978), Cook and Johnson (1981), Oakes (1989) and Genest and Rivest (1993). In economics and finance, see Li (2000), Embrechts, Höing, and Juri (2001), Patton (2001a, 2001b), Rockinger and Jondeau (2001), Chen and Fan (2002a, 2002b), Mashal and Zeevi (2002), Miller and Liu (2002), Junker and May (2002), Fermanian and Scaillet (2003), and Rosenberg (2003).

copulas is discussed in Patton (2001a) for the parametric case and Fermanian and Scaillet (2003) for the nonparametric case.

2 THE INVESTOR'S OPTIMIZATION PROBLEM

The utility functions we assume for our hypothetical investors are from the class of CRRA utility functions:

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

where P_0 is the initial wealth, X_t and Y_t represent the continuously compounded excess return (over the risk-free rate) on the small cap and large cap indices, respectively, and ω_i is the proportion of wealth in asset i . The degree of relative risk aversion (RRA) is denoted by γ . 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 . A similar range of risk aversion levels was also considered in Campbell and Viceira (1999) and Ait-Sahalia and Brandt (2001). While there exist other utility functions that place higher weight on tail events or asymmetries in the distribution of payoffs, we focus on the CRRA utility because of its prominence in the finance literature. If gains are found using the CRRA utility function then they may be thought of as a conservative estimate of the possible gains using other, more sensitive, utility functions.

The setup of the investor's problem is as follows. Let the excess returns on the two risky assets under consideration be denoted X_t and Y_t , with some joint distribution, H_t , with associated marginal distributions, F_t and G_t , and copula, C_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, $\omega_{t+1}^* \equiv [\omega_{x,t+1}^*, \omega_{y,t+1}^*]$, for the portfolio. The optimal weights are found by maximizing the expected utility of the end-of-period wealth under the estimated probability density:

$$\begin{aligned} \omega_{t+1}^* &\equiv \arg \max_{\omega \in \mathcal{W}} E_{\hat{H}_{t+1}} [U(1 + \omega_x X_{t+1} + \omega_y Y_{t+1})] \\ &= \arg \max_{\omega \in \mathcal{W}} \int \int U(1 + \omega_x x + \omega_y y) \cdot \hat{f}_{t+1}(x) \cdot \hat{g}_{t+1}(y) \\ &\quad \cdot \hat{c}_{t+1}(\hat{F}_{t+1}(x), \hat{G}_{t+1}(y)) \cdot dx \cdot dy, \end{aligned} \quad (5)$$

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

The investor is assumed to estimate the model of the conditional distribution of excess returns using maximum likelihood and then optimize 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

[see Granger (1969), Weiss (1996), and Skouras (2001)]. We use maximum-likelihood estimation for computational tractability.

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, which must be done for each point in the out-of-sample period. The objective function 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, ω_{t+1}^* , at each point in time. Further details on this procedure may be found in Appendix A.

One concern that may arise in this design is the existence of $E_{\hat{H}_{t+1}}[\mathcal{U}(1 + \omega_x X_{t+1} + \omega_y Y_{t+1})]$ for certain density models. Given the 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. We deal with this by modifying the left tail 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 relocated to the region $(0, \varepsilon)$, where ε is some extremely small positive number.

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 capitalization stocks (“small caps”) and a portfolio of high market capitalization 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.

3.1 Description of the Data

We use monthly data from the CRSP on the top 10% and bottom 10% of stocks sorted by market capitalization to form indices — the “large cap” and “small cap” indices, from January 1954 to December 1999, yielding 552 observations. These data were also analyzed 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 are presented in Table 1.

The small cap index generally exhibited slightly positive skewness, while the large cap index exhibited 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, as expected. Table 1 also reveals that the small cap index had a higher mean and higher volatility than the large cap index over the total sample

Table 1 Descriptive statistics.

	Full sample		In-sample		Out-of-sample	
	Small caps	Large caps	Small caps	Large caps	Small caps	Large caps
Mean*	9.9549	7.9748	9.7076	6.4315	10.8452	13.5306
Std Dev*	21.2932	14.2888	22.0482	14.4729	18.4016	13.5423
Skewness	0.0558	-0.3795	0.1653	-0.3173	-0.6128	-0.6124
5% VaR	8.7973	6.2306	8.9204	6.3922	6.6928	4.9430
1% VaR	18.9576	9.6657	19.0624	9.7958	16.2402	10.9989
Kurtosis	7.5647	4.9088	7.7197	4.9871	5.2857	4.6428
Min	-29.3153	-20.8934	-29.3153	-20.8934	-21.9607	-14.9491
Max	38.3804	16.8145	38.3804	16.8145	13.9502	11.3488
Jarque-Bera	479.5162	97.0484	401.0592	77.9577	33.0707	20.6444
p-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Correlation	0.7210		0.7392		0.6457	

The statistics marked with an asterisk were annualized to ease interpretation.

Jarque-Bera refers to the test for normality of the unconditional distribution of returns. The full sample runs from January 1954 to December 1999, the in-sample period from January 1954 to December 1989, and the out-of-sample period from January 1990 to December 1999.

and the in-sample period, but not over the out-of-sample period. During the 1990s, as is well known, large cap stocks performed better than their historical average. The change in ex post average returns and standard deviations for the small and large cap indices between the in-sample and out-of-sample period suggests that allowing for structural breaks in the returns-generating process may improve portfolio decisions. One promising method of doing so is the reversed ordered Cusum (ROC) procedure of Pesaran and Timmermann (2002). Due to the computational constraints, we are forced to ignore the possibility of structural breaks.

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 Ait-Sahalia and Brandt (2001).

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)

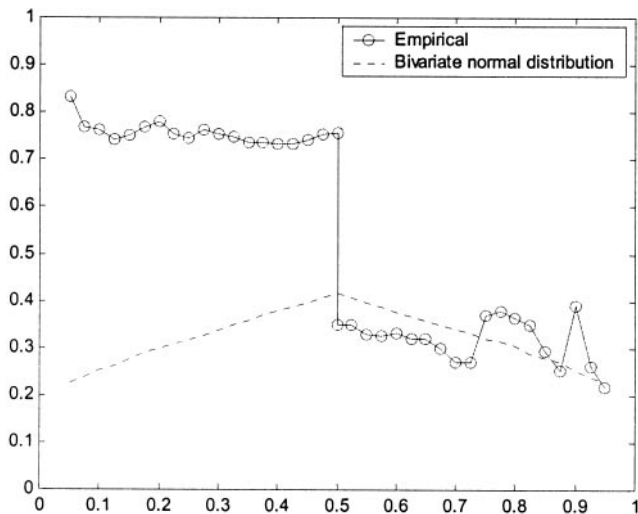


Figure 1 Exceedence correlations between excess returns (X and Y) on small caps and large caps. The horizontal axis shows the cutoff quantile, and the vertical axis shows the correlation between the two returns given that both exceed that quantile.

called “exceedence correlations,” $\tilde{\rho}(q)$. We will not use this measure in the asset allocation problem, but we have found it to be a useful, intuitive way of taking a preliminary look at our data.

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

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 nonlinear in q . In Figure 1 we plot the empirical exceedence correlations based on the (raw) excess returns on the two indices, along with what would be obtained if they had the bivariate normal distribution. In Figure 2 we plot the empirical exceedence correlations based on the transformed standardized residuals of the models for the two indices, along with what would be obtained 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. This suggests that the assumption of normality, which implies a symmetric dependence structure, is inappropriate for these assets. Whether capturing this asymmetric dependence leads to substantially better portfolio decisions is the focus of Section 3.3.

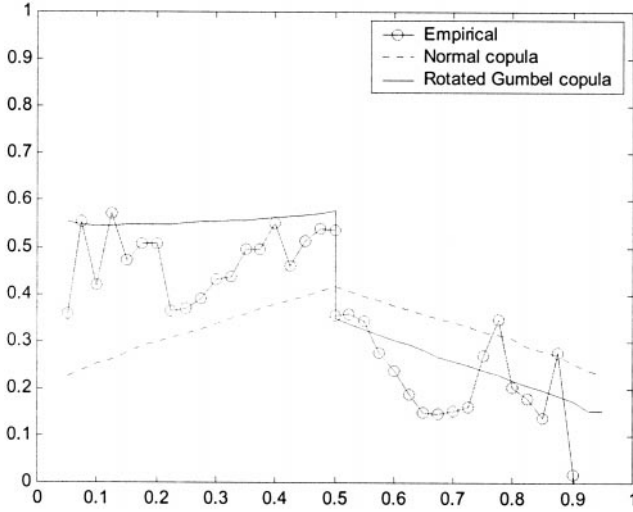


Figure 2 Exceedence correlations between transformed residuals (U and V) of small caps and large caps. The horizontal axis shows the cutoff quantile, and the vertical axis shows the correlation between the two residuals given that both exceed that quantile.

3.2 Analysis of the Different Models

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

The first three strategies we consider are simply buy-and-hold strategies (all small caps, all large caps, or an even mix of both). The fourth strategy is one based solely on the unconditional distribution of returns. For this portfolio we assume that the investor optimizes his/her portfolio weights for the period $t + 1$ using the empirical unconditional distribution of returns observed up until time t :

$$\begin{aligned}
 \omega_{uncond,t+1}^* &\equiv \arg \max_{\omega \in \mathcal{W}} \hat{E}_t[U(1 + \omega_x X_{t+1} + \omega_y Y_{t+1})] \\
 &= \arg \max_{\omega \in \mathcal{W}} t^{-1} \sum_{j=1}^t \mathcal{U}(1 + \omega_x x_j + \omega_y y_j).
 \end{aligned}
 \tag{6}$$

This portfolio is based on the assumption 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 modeling the *conditional* distribution of these stock returns.

3.2.1 Marginal distribution models The benchmark model for our study is the bivariate normal distribution, which is compared with a model constructed using

copula theory. Both models have the same forms for the conditional means, μ_t^x and μ_t^y , and variances, h_t^x and h_t^y . In their article on the value of volatility timing for asset allocation decisions, Fleming, Kirby, and Ostdiek (2001) assumed a constant conditional mean rather than using some model for expected returns. In their framework this was shown to lead to a conservative estimate of the value of volatility timing. In our framework, however, a misspecified conditional mean will lead to a misspecified skewness model and a misspecified dependence model, and unlike Fleming, Kirby, and Ostdiek, we cannot be sure what impact this will have on the results; whether it will exaggerate or dampen the differences between the models under analysis. For this reason, we cannot escape the building of a model for expected returns.

The conditional means were set to be linear functions of up to 12 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 and allowed the three lagged exogenous regressors to enter into the conditional variance specification in levels and squares. We used likelihood ratio tests to determine the best-fitting model over the in-sample period. The selected models for the mean and variance equations are given below. The full sequence of parameter estimates for each point in the out-of-sample period are available from the author upon request.

$$X_t = \beta_0 + \beta_1 X_{t-1} + \beta_2 R_{t-1}^f + \beta_3 SPR_{t-1} + \beta_4 DIV_{t-1} + \sqrt{h_t^x} \varepsilon_t \quad (7)$$

$$h_t^x = \beta_5 + \beta_6 h_{t-1}^x + \beta_7 h_{t-1}^x \varepsilon_{t-1}^2 \mathbf{1}\{\varepsilon_{t-1} > 0\} + \beta_8 h_{t-1}^x \varepsilon_{t-1}^2 \mathbf{1}\{\varepsilon_{t-1} < 0\} \\ + \beta_9 R_{t-1}^f + \beta_{10} SPR_{t-1}$$

$$Y_t = \gamma_0 + \gamma_1 R_{t-1}^f + \gamma_2 SPR_{t-1} + \gamma_3 DIV_{t-1} + \sqrt{h_t^y} \eta_t \quad (8)$$

$$h_t^y = \gamma_4 + \gamma_5 h_{t-1}^y + \gamma_6 h_{t-1}^y \varepsilon_{t-1}^2 \mathbf{1}\{\varepsilon_{t-1} > 0\} + \gamma_7 h_{t-1}^y \varepsilon_{t-1}^2 \mathbf{1}\{\varepsilon_{t-1} < 0\} + \gamma_8 R_{t-1}^f$$

Although the models are recursively reestimated throughout the out-of-sample period, they are “nonadaptive,” in that the model specifications are determined using the in-sample data and not updated in the out-of-sample period.

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. Jondeau and Rockinger (2003) present some further results on this distribution. 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 parameters, denoted λ_t and ν_t , were allowed to depend on lags of the exogenous variables and the forecast conditional means and variances. As suggested by Hansen (1994), we use transformations to ensure that the skewness and degrees-of-freedom parameters remained within $(-1, 1)$ and $(2, \infty]$, respectively, at all times by setting $\lambda_t = \Lambda(Z'_{t-1}\beta)$ and $\nu_t = 2.1 + (Z'_{t-1}\beta)^2$, where $Z'_{t-1}\beta$ is the linear function of the regressors

and parameters for that variable and $\Lambda(x) = (1 - e^{-x}) / (1 + e^{-x})$ is the modified logistic transformation, designed to keep λ_t in $(-1, 1)$ at all times.

For both assets we found significant in-sample time variation in these moments, though some variables were dropped as their coefficients were not significant. 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 0.01), for both assets, the assumptions of skewness being constant at zero and kurtosis being constant at three, both jointly and separately for the in-sample period.⁴ The improved in-sample goodness-of-fit of the skewed t distribution is traded off against possible increased parameter estimation error in an out-of-sample setting.

In addition to testing the significance of each of the possible variables to include in the model, we conducted goodness-of-fit tests (not reported) for the final marginal density models. Such tests are critical when constructing multivariate densities using copulas, as a misspecified marginal density implies that any copula model will be misspecified. We used Kolmogorov-Smirnov (KS) tests for the proposed density and Lagrange multiplier (LM) tests for serial dependence in the probability integral transforms of the variables, as suggested in Diebold, Gunther, and Tay (1998). We also employed the multinomial hit test described in Patton (2001b). We found no evidence against the skewed t models and some evidence against the normal models. To show the outputs of these models, in Figure 3 we present the conditional mean, conditional variance, skewness parameter, and kurtosis parameters for the out-of-sample period, estimated at each point in time only using data available as in the previous period.

3.2.2 Copula models 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

$$\left(\frac{X_t - \mu_t^x}{\sqrt{h_t^x}}, \frac{Y_t - \mu_t^y}{\sqrt{h_t^y}} \right) \sim N \left(\mathbf{0}, \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \right) \tag{9}$$

$$\rho_t = \Lambda(\alpha_0 + \alpha_1 R_{ft-1} + \alpha_2 SPR_{t-1} + \alpha_3 DIV_{t-1} + \alpha_4 \mu_t^x + \alpha_5 \mu_t^y), \tag{10}$$

where $\Lambda(x)$ is the modified logistic transformation.

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

⁴ The fact that the restrictions on skewness and kurtosis were both rejected means that both higher moments may be important for asset allocation. We could attempt to disentangle the benefits of each by including an additional intermediate model, such as the usual Student's t , which allows for kurtosis but not skewness. We do not do so here in the interest of parsimony.

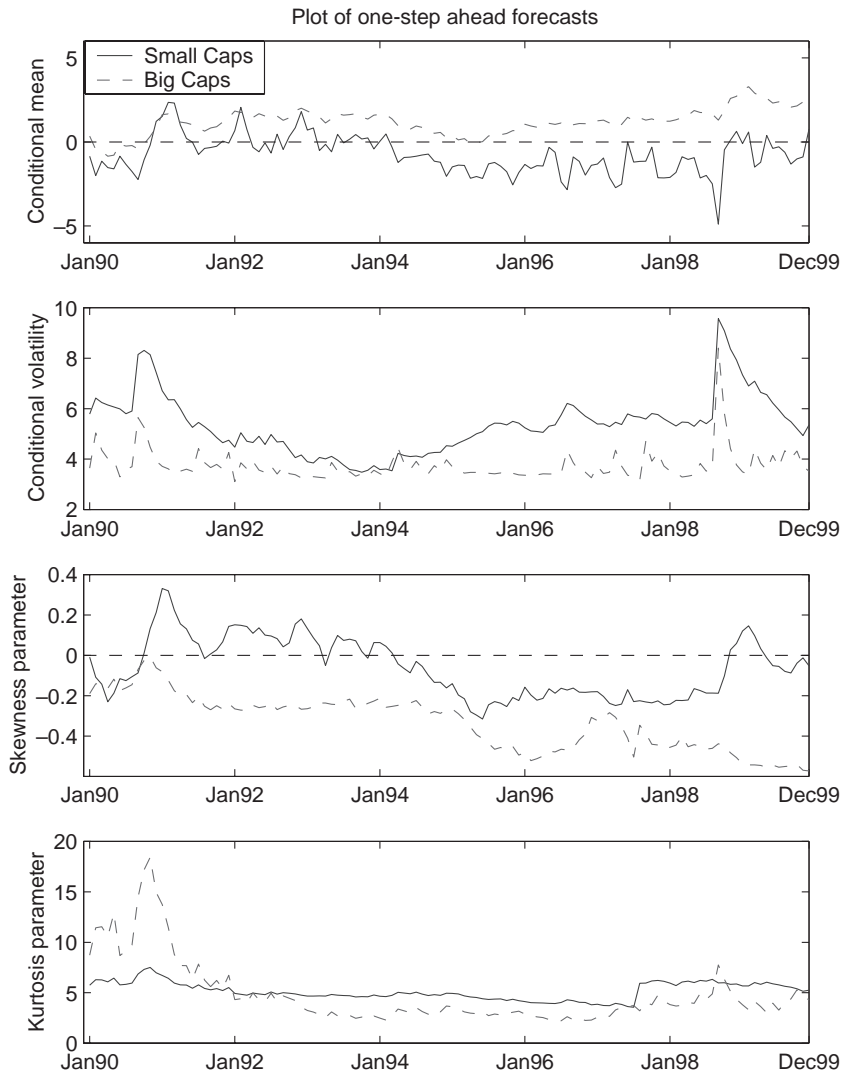


Figure 3 Plots of the first four conditional moment parameters over the out-of-sample period.

t models in the search for the best-fitting copula. The copulas considered were the normal, Student’s *t*, Clayton, rotated Clayton,⁵ Joe-Clayton, Plackett, Frank, Gumbel, and rotated Gumbel copulas; contour plots of a few of these copulas are provided in Figure 4 and the functional forms of these copulas are contained in

⁵ Let (U, V) be distributed according to the copula C . Then $(1-U, 1-V)$ will be said to be distributed according to the “rotated C ” copula. The rotation allows us to take a copula that allows for greater dependence in the first (third) quadrant and create one that allows for greater dependence in the third (first) quadrant.

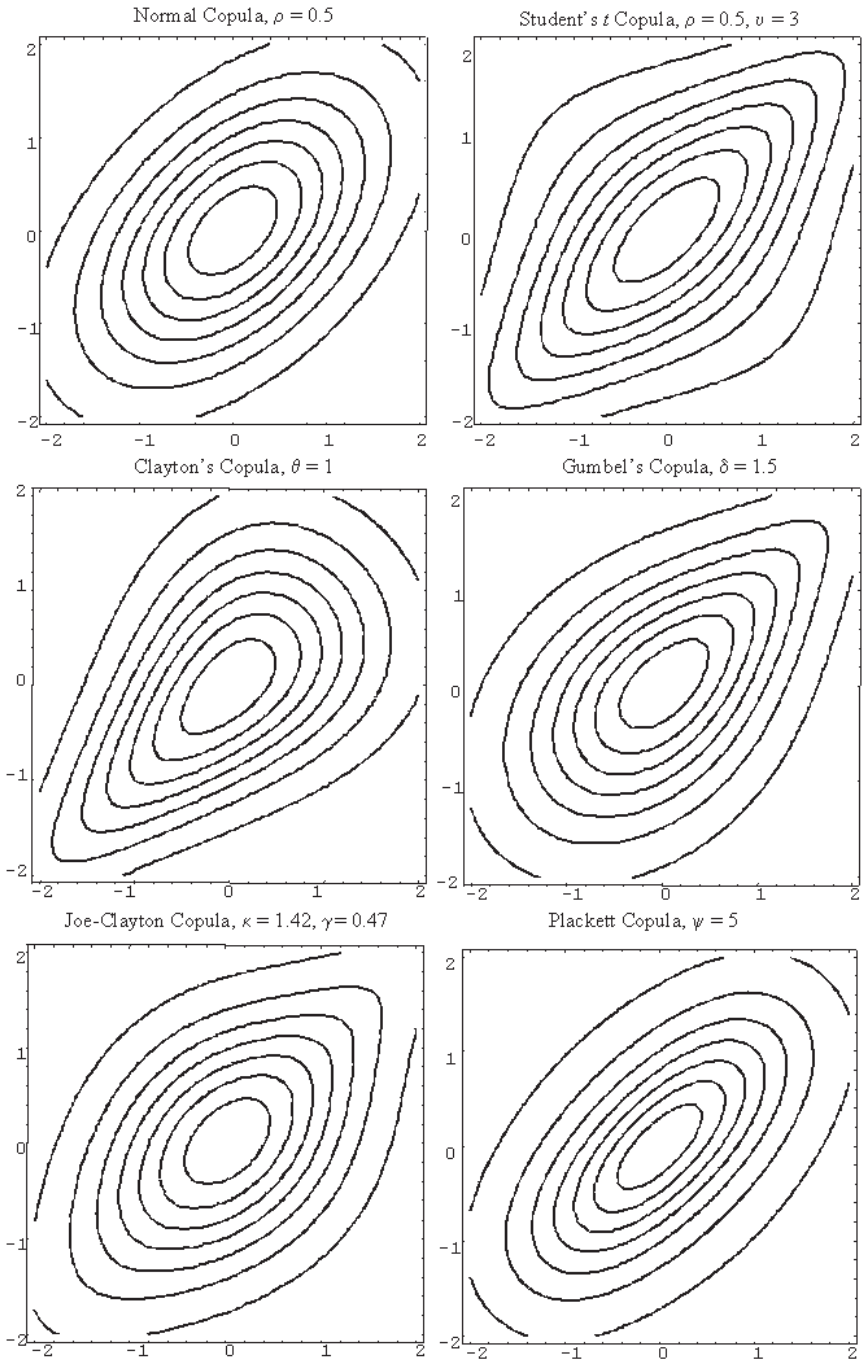


Figure 4 Contour plots of various distributions all with standard normal marginal distributions and linear correlation coefficients of 0.5.

Appendix B. This list includes almost all of the copulas considered in the various applications of copulas in statistics and economics,⁶ and is significantly more than we found in any single previous applied study.

The plots in Figure 4 show the isoprobability contours of bivariate densities with Normal(0,1) margins and linear correlation coefficient of 0.5. We fixed the marginals and the correlation coefficient in this figure so that the differences in the densities could be more clearly identified and attributed to the different copulas. The copula in the upper left panel is the normal copula, making the joint density a bivariate standard normal and giving us the familiar elliptical contours. Immediately below the normal density is the joint density formed using Clayton's copula. We can see that this density's contours are more tightly clustered around the diagonal in the third ("negative-negative") quadrant than in the first quadrant, indicating stronger dependence between negative observations than between positive observations. This is qualitatively the type of dependence suggested by the exceedence correlation plots in Figures 1 and 2. Of these six, the normal, Student's t , and Plackett all generate symmetric dependence, whereas the Gumbel, Clayton, and Joe-Clayton all generate asymmetric dependence. For further details on the properties of these copulas see Joe (1997), Nelsen (1999), and Patton (2001b).

As in the bivariate normal distribution, we estimated these copulas with conditional dependence modeled 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 (12) below]:

$$\left(\frac{X_t - \mu_t^x}{\sqrt{h_t^x}}, \frac{Y_t - \mu_t^y}{\sqrt{h_t^y}} \right) \sim C(\text{Skewed } t(\lambda_t^x, \nu_t^x), \text{Skewed } t(\lambda_t^y, \nu_t^y); \delta_t) \quad (11)$$

$$\delta_t = \Gamma(\beta_0 + \beta_1 R_{ft-1} + \beta_2 SPR_{t-1} + \beta_3 DIV_{t-1} + \beta_4 \mu_t^x + \beta_5 \mu_t^y), \quad (12)$$

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

The maximum log-likelihood and information criteria 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 and the lowest value of both information criteria. We will thus use the rotated Gumbel copula in the flexible distribution specification, which we will call the "Gumbel" model for simplicity.

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

⁶ 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, p. 35)].

Table 2 Results from the copula specification search.

Model	\mathcal{L}_C	Number of parameters	AIC	BIC
Symmetric copulas				
Normal	153.5681	6	-307.1076	-307.0499
Student's t	158.1329	7	-316.2325	-316.1651
Plackett	163.1763	6	-326.3240	-326.2663
Frank	158.2502	6	-316.4718	-316.4141
Asymmetric copulas				
Clayton	151.3272	6	-302.6258	-302.5681
Rotated Clayton	90.8669	6	-181.7052	-181.6475
Joe-Clayton	158.8478	12	-317.6385	-317.5230
Gumbel	127.8091	6	-255.5896	-255.5319
Rotated Gumbel	166.6628	6	-333.2970	-333.2393

Presented here are the nine copula specifications tried for the copula distribution model. The copula likelihood at the optimum is denoted \mathcal{L}_C . Also presented are the numbers of parameters estimated in the models and the values of the Akaike and Schwarz's Bayesian information criteria at the optima.

density modeling 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 the Gumbel model provides the best fit, and that about 73% of the gains come from the flexible marginal distribution models, though in an out-of-sample setting this ranking and decomposition of gains need not hold.

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.

We conducted some specification tests (not reported) on the normal and rotated Gumbel copulas over the in-sample period to determine their goodness-of-fit, employing the multinomial hit test described in Patton (2001b). We found that the normal copula estimated using residuals from the normal marginal distribution models could be rejected, which is unsurprising since the marginals of that model were also rejected. Neither the normal copula nor the rotated Gumbel copula estimated on residuals from the skewed t marginal distribution models could be rejected at the 5% level.

In Figure 5 we present the conditional correlation parameter from the bivariate normal model and the implied conditional correlation from the skewed t -rotated Gumbel copula model. We use correlation as the measure

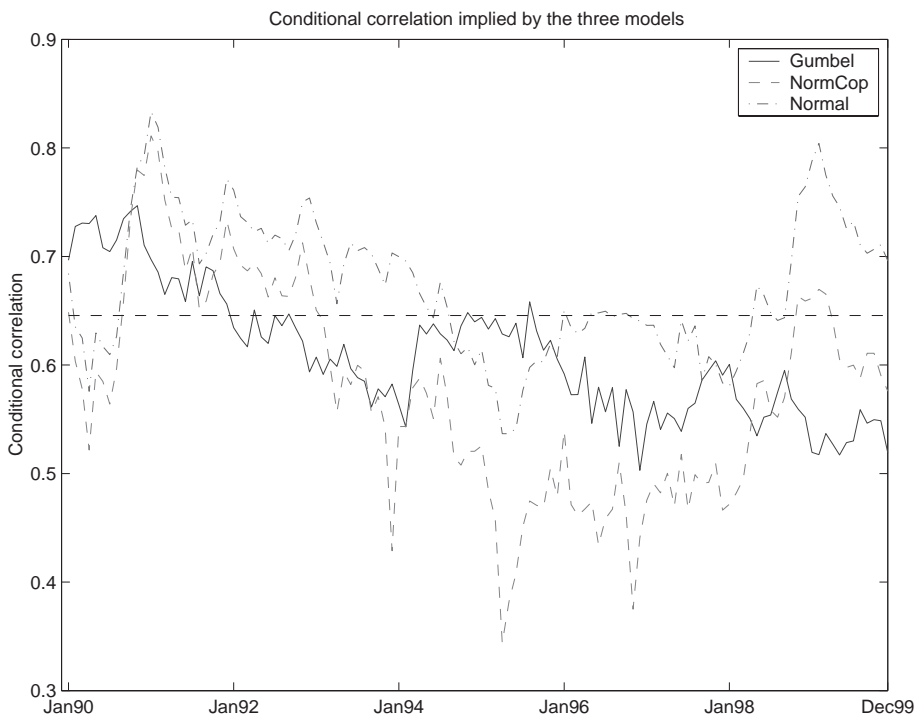


Figure 5 Conditional correlation implied by the three models (Gumbel, NormCop, and normal) over the out-of-sample period (January 1990 to December 1999). The straight dashed line is the unconditional correlation over this period (0.6457).

of dependence here for comparability across models. While all three conditional correlation estimates generally moved in the same direction, their levels are quite different at times.

3.3 Performance of the Different Strategies

We now analyze the performance of the different asset allocation decisions made using the various models. We consider five levels of relative risk aversion ($RRA = 1, 3, 7, 10,$ and 20) and 11 strategies. The 11 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) optimize 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. Strategies 8–11 are the same as strategies 4–7, subject to a short-sales constraint.

3.3.1 Summary statistics In Table 3 we present six summary statistics on the optimal portfolio return series based on the different models. In addition to the

usual summary statistics we also present two alternative measures of risk, the 5% value-at-risk (VaR) and the 5% expected shortfall (ES). The 5% VaR is defined as the negative of the fifth empirical percentile of the realized returns, that is, $\widehat{\text{VaR}}(X; 0.05) \equiv -\hat{F}_n^{-1}(0.05)$, where \hat{F}_n is the empirical distribution of returns on portfolio X using the n out-of-sample observations. 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 that the return has exceeded its 5% VaR, that is, $\widehat{\text{ES}}(X; 0.05) \equiv -\hat{E}_n[X|X \leq \widehat{\text{VaR}}(X; 0.05)]$, where \hat{E}_n is the sample average.

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: the Normal and NormCop portfolios also generally exhibited negative skewness, while the unconstrained Gumbel portfolio actually displayed positive skewness, suggesting that modeling both skewness and asymmetric dependence enables the investor to better avoid negatively skewed portfolio returns.

3.3.2 Performance statistics The 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. One interpretation of this amount is as the “management fee” that could 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. See Table 4 for the complete set of results.

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

Table 4 Realized portfolio return performance: management fee.

		RRA				
		1	3	7	10	20
	Small caps	-1.9811	-3.3108	-6.3061	-9.0655	-23.0183
	Large Caps	1.4659	1.8658	2.9521	4.2016	13.1512
	50:50 mix	0	0	0	0	0
Unconstrained	Uncond	13.9372	-0.5508	0.1412	3.6384	25.9284
	Normal	$-\infty$	$-\infty$	-67.0713	-40.6527	4.5104
	NormCop	-15.2901	-22.4892	-26.3494	-20.5650	-1.4480
	Gumbel	1.8154	-2.7387	-4.8058	2.6759	23.4016
Short-sales constrained	Uncond	-2.0114	-0.4145	0.1412	3.6384	25.9287
	Normal	25.3021	3.0180	2.8306	3.7620	24.4493
	NormCop	25.2654	1.0794	0.6189	3.7849	25.9580
	Gumbel	25.1760	3.2780	3.0071	5.5675	27.0638

RRA refers to the coefficient of relative risk aversion.

Each of the 11 rows 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 is based on a weight that is optimized using the empirical unconditional distribution of returns, the fifth is based on the normal distribution model, the sixth is based on the skewed t -normal copula model and the seventh is based on the skewed t -rotated Gumbel copula model. Portfolios 8–11 correspond to portfolios 4–7 with a short-sales constraint imposed. The performance is measured as the number of basis points per year that the investor would be willing to pay to switch from the 50:50 portfolio; a possible “management fee.”

for these investors.⁷ For this month the realized utility for these investors is $-\infty$, making the required management fee to switch to this portfolio $-\infty$.

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 portfolios 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. This result can be interpreted as 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

⁷ This obviously represents a failure of these investors’ models or optimization methods, as they did not recognize 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, that is, $\Pr[X > 14.0 \cap Y \leq -2.6]$, was about 1 in 3.5 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 misspecified; the model with skewed t margins and rotated Gumbel copula assigned a probability of about 1 in 75,000, which is about 45 times larger than under the normal density forecast.

extreme position in the market, then this finding reinforces results previously reported in the forecasting literature [see, e.g., Stock and Watson (1999)], that constrained forecasts often outperform unconstrained forecasts from nonlinear models.

The management fees that one could charge an investor currently holding the 50:50 portfolio to switch to the constrained Gumbel portfolio range between -5 and 27 basis points per year. Management fees of less than 10 or 20 basis points per year are of questionable economic significance. The largest gains (23 and 27) for the Gumbel portfolio occur for the most risk-averse investor, though it should be noted that the gains are not monotonic in the risk-aversion parameter.

Looking now to the gains from modeling higher moments and asymmetric dependence, we compare the Normal portfolio with the Gumbel portfolio. In 9 of 10 comparisons the Gumbel portfolio outperformed the Normal portfolio, and for the remaining comparison the difference was -0.1 basis points. Ignoring the two comparisons where the Normal portfolio went bankrupt (and so the effective management fee difference would be $+\infty$) the average outperformance was 16.2 basis points; 41.5 basis points for the unconstrained investor and only 0.9 basis points for the short sales-constrained investor. We will see in the following section the reason for the extremely small difference for the short-sales constrained investor.

The Gumbel model also outperformed the NormCop model in 9 of 10 comparisons. The average outperformance of Gumbel portfolio over NormCop portfolio was 21.3 (1.5) basis points for the unconstrained (short sales-constrained) investor. The corresponding figures for the Normal versus the NormCop portfolio were -18.3 and 0.5 , indicating that the Normal portfolio performed worse for the unconstrained investor, but marginally better for the short sales-constrained investor. That the Gumbel portfolio outperformed both the Normal and NormCop portfolios suggests that the copula specification is important for asset allocation.

3.3.3 Analysis of the optimal portfolio weights In this section we look at the time series of portfolio weights resulting from the portfolio decisions made using different models and different levels of risk aversion. To consider the impact of risk aversion, we present Table 5 and Figure 6. Table 5 contains quantiles of the distribution of portfolio weights obtained using the Gumbel model for risk aversion ranging between 1 and 20 , and Figure 6 shows the time series of portfolio weights for investors with a relative risk aversion of 7 and 20 . Both show that increasing the level of relative risk aversion shrinks the portfolio weights toward zero, as expected. In the limit of infinite risk aversion the investor would put no wealth in the risky assets and all wealth in the risk-free asset.

In Figure 7 we show the impact of imposing a short-sales constraint. The plot makes it clear that even for moderately high-risk aversion, seven in this case, the short-sales constraint is binding for at least one asset almost every period. For the Gumbel model the proportions of times that the short-sales constraint is binding for risk aversion levels of 1 , 3 , 7 , 10 , and 20 are 1 , 1 , 0.98 , 0.95 , and 0.94 , respectively. Similar figures are obtained for the Normal and NormCop portfolios.

Table 5 Optimal portfolio weight summary statistics (Gumbel model).

	RRA = 1		RRA = 7		RRA = 10		RRA = 20	
	Small caps	Large caps	Small caps	Large caps	Small caps	Large caps	Small caps	Large caps
	Unconstrained							
Minimum	-13.85	-1.58	-4.38	-3.68	-2.96	-3.03	-2.20	-2.98
25% quantile	-7.70	7.49	-1.44	0.68	-1.03	0.56	-0.52	0.30
Median	-5.62	9.04	-1.01	1.33	-0.72	0.95	-0.37	0.50
75% quantile	-4.06	10.41	-0.60	1.66	-0.43	1.19	-0.21	0.63
Maximum	4.22	15.53	0.74	2.56	0.95	1.88	1.12	0.99
	Short-sales constrained							
Minimum	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25% quantile	0.00	1.00	0.00	0.38	0.00	0.27	0.00	0.14
Median	0.00	1.00	0.00	0.61	0.00	0.47	0.00	0.26
75% quantile	0.00	1.00	0.00	0.89	0.00	0.68	0.00	0.36
Maximum	0.00	1.00	0.58	1.00	0.52	1.00	0.26	0.56

This table presents some summary statistics of the optimal portfolio weights over the out-of-sample period for an investor using the "Gumbel" model, with relative risk aversion of 1, 7, 10, and 20, with and without a short-sales constraint imposed.

This shows that much of the information content of these models is lost if the investor is short-sales constrained.

One example of this reduced information content relates to comparing the short sales-constrained portfolios. The proportions of times that the short sales-constrained Gumbel portfolio took the same portfolio weights as the short sales-constrained Normal portfolio for risk aversion levels of 1, 3, 7, 10, and 20 are 0.95, 0.80, 0.59, 0.52, and 0.52, respectively. The corresponding figures comparing the Gumbel with the NormCop portfolio are similar. Thus, of the 120 observations in the out-of-sample period, the number of observations that enable us to distinguish one short sales-constrained portfolio from another range from 58 (when there is "only" a 52% overlap, for the two most risk-averse investors) to just 5 (for the least risk-averse investor). This suggests, and is confirmed in the next section, that if there are gains to capturing and forecasting skewness and asymmetric dependence they are more likely to be present for unconstrained investors than for short sales-constrained investors.

To compare the portfolio weights obtained using the different models we present Table 6 and Figures 8 and 9. Table 6 presents quantiles of the distribution of portfolio weights obtained using the different models for $RRA = 7$, and Figures 8 and 9 compare the time series of portfolio weights over the out-of-sample period. The results show that the weights from the unconstrained normal model are more aggressive than those from the NormCop or Gumbel models. For example, the latter two models yield median positions of being short 1 unit of the small cap index and long 1 to 1.3 units of the large cap index, while for the normal model the median position is being short almost 2 units of the small cap index and long

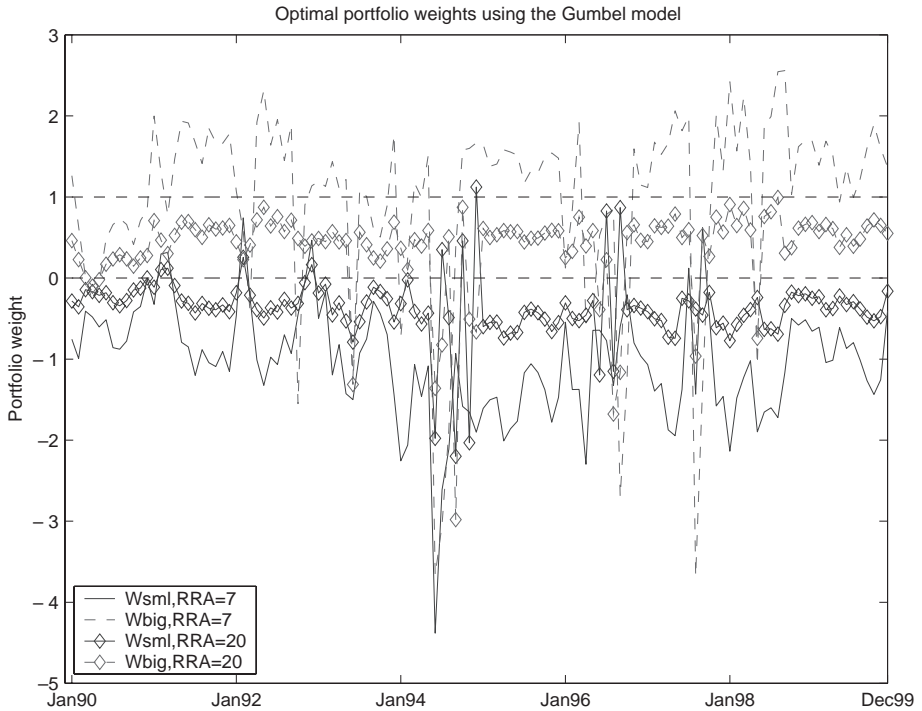


Figure 6 Optimal weights for the unconstrained Gumbel portfolio for investors with relative risk aversion of 7 and 20 over the out-of-sample period (January 1990 to December 1999). “Wsm1” stands for the weight put in small caps, and “Wbig” stands for the weight put in large caps.

3 units of the large cap index. Figure 8 confirms that the Normal portfolio weights are almost always more extreme than the Gumbel portfolio weights. This is possibly due to the fact that the Gumbel model takes into account the fat tails, skewness, and asymmetric dependence of these assets. Negative skewness, fat tails, and dependence of the rotated Gumbel type will other make a risk-averse investor less aggressive in his/her portfolio decisions, as they all lead, other things being equal, to a higher probability of large negative moves for the portfolio. Figure 9 reveals that the NormCop and Gumbel portfolios have similar portfolio weights. Much of the difference in portfolio weights between the Gumbel and the Normal portfolios, then, was driven by the different marginal distribution assumptions.

Note that we do not test for the significance of the differences in portfolio weights directly in this article. Differences in portfolio weights are only economically interesting if they lead to differences in portfolio performance and so it seems more appropriate to test for differences using the metric of portfolio performance. We proceed to such tests in the next section.

To try to determine in more detail the causes of the differences in portfolio weights between the Normal and Gumbel portfolios we considered the following

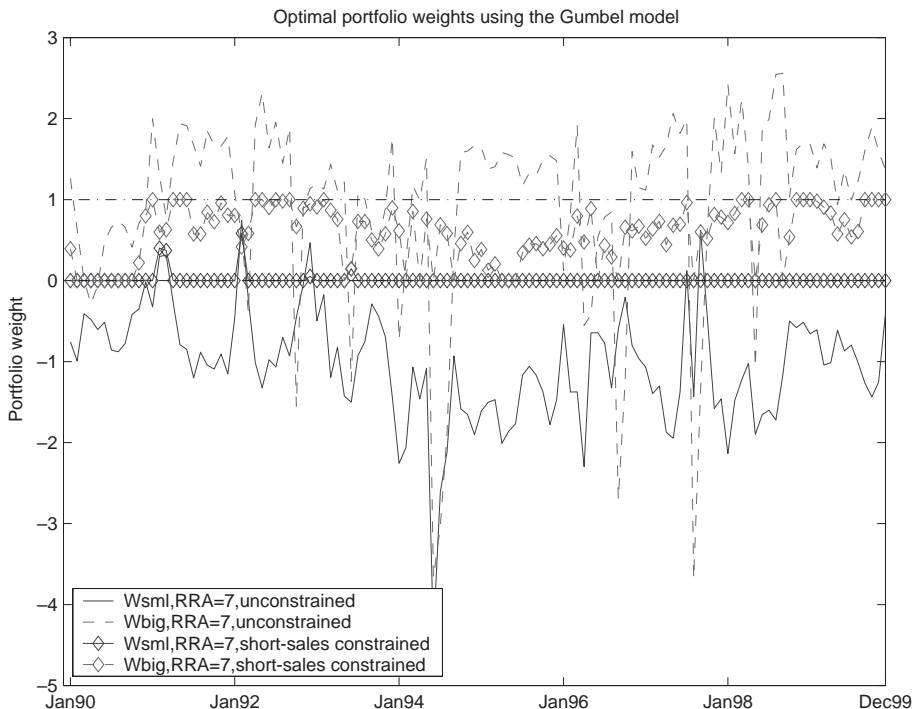


Figure 7 Optimal weights for the unconstrained and short sales-constrained Gumbel portfolios for an investor with relative risk aversion of 7 over the out-of-sample period (January 1990 to December 1999). “Wsml” stands for the weight put in small caps, and “Wbig” stands for the weight put in large caps.

simple regression of the difference between the Gumbel and Normal portfolio weights on a constant and the nine parameters of the more flexible model (two expected returns, two volatilities, two skewness parameters, two kurtosis parameters, and one copula parameter):

$$\begin{aligned} \omega_{i,t}^{GUM} - \omega_{i,t}^N &= \beta_0 + \beta_1 \mu_{x,t} + \beta_2 \mu_{y,t} + \beta_3 \sqrt{h_{x,t}} + \beta_4 \sqrt{h_{y,t}} \\ &\quad + \beta_5 \lambda_{x,t} + \beta_6 \lambda_{y,t} + \beta_7 \nu_{x,t} + \beta_8 \nu_{y,t} + \beta_9 \kappa_t + e_{i,t}, \end{aligned} \tag{13}$$

$i = X, Y$

where $\omega_{i,t}^{GUM}$, $\omega_{i,t}^N$ are the optimal Gumbel and Normal portfolio weights for $RRA = 7$. The optimal portfolio weights are a complicated, nonlinear function of the parameters of the joint density, and so the above regression is almost certainly misspecified.⁸ Further, for simplicity we ignore the fact that the variables

⁸ We may consider this regression as a first-order approximation of the true function relating the parameters of the forecast density to the optimal portfolio weights.

Table 6 Optimal portfolio weight summary statistics (RRA = 7).

	Uncond		Normal		NormCop		Gumbel	
	Small caps	Large caps	Small caps	Large caps	Small caps	Large caps	Small caps	Large caps
	Unconstrained							
Minimum	0.04	0.19	-3.83	-0.36	-4.21	-6.96	-4.38	-3.68
25% quantile	0.13	0.21	-2.32	1.93	-1.33	0.53	-1.44	0.68
Median	0.14	0.21	-1.89	3.03	-0.92	0.99	-1.01	1.33
75% quantile	0.15	0.30	-1.31	3.78	-0.57	1.36	-0.60	1.66
Maximum	0.17	0.43	0.53	10.23	1.70	2.59	0.74	2.56
	Short-sales constrained							
Minimum	0.04	0.19	0.00	0.00	0.00	0.00	0.00	0.00
25% quantile	0.13	0.21	0.00	0.69	0.00	0.29	0.00	0.38
Median	0.14	0.21	0.00	1.00	0.00	0.65	0.00	0.61
75% quantile	0.15	0.30	0.00	1.00	0.00	0.92	0.00	0.89
Maximum	0.17	0.43	0.49	1.00	0.60	1.00	0.58	1.00

This table presents some summary statistics of the optimal portfolio weights over the out-of-sample period for an investor with relative risk aversion of 7, obtained using four different models, with and without a short-sales constraint imposed.

on the right-hand side of Equation (13) are estimated, and so this regression will suffer errors-in-variables bias. Nevertheless, it may help highlight some of the causes of the differences in portfolio weights. To aid interpretation, Table 7 also presents the results of regressions of the individual portfolio weights on the regressors in Equation (13), though we will focus our discussion below solely on the regressions involving the difference in portfolio weights, presented in the last two columns.

The signs on the coefficients on expected returns indicate that the Normal portfolio weights reacted more strongly to changes in forecasted returns than the Gumbel portfolio weights. For example an increase in small cap expected returns lead to a larger increase in Normal portfolio small cap weight than Gumbel portfolio small cap weight, making the coefficient in the regression of the difference in portfolio weights negative. The significant positive (negative) coefficient on large cap volatility for the large (small) cap regression also reflects that fact that the Normal portfolio weights reacted more strongly than the Gumbel portfolio weights.

The negative (positive) sign on the degree-of-freedom parameter coefficient in the small (large) cap regression suggests that as the degree-of-freedom parameter decreases, reflecting an increase in the fatness of the tails, the weight in the small (large) caps in the Gumbel portfolio goes toward zero. The Normal portfolio weights were essentially uncorrelated with this parameter, as expected. That is, an increase in tail fatness led to a less aggressive Gumbel portfolio.

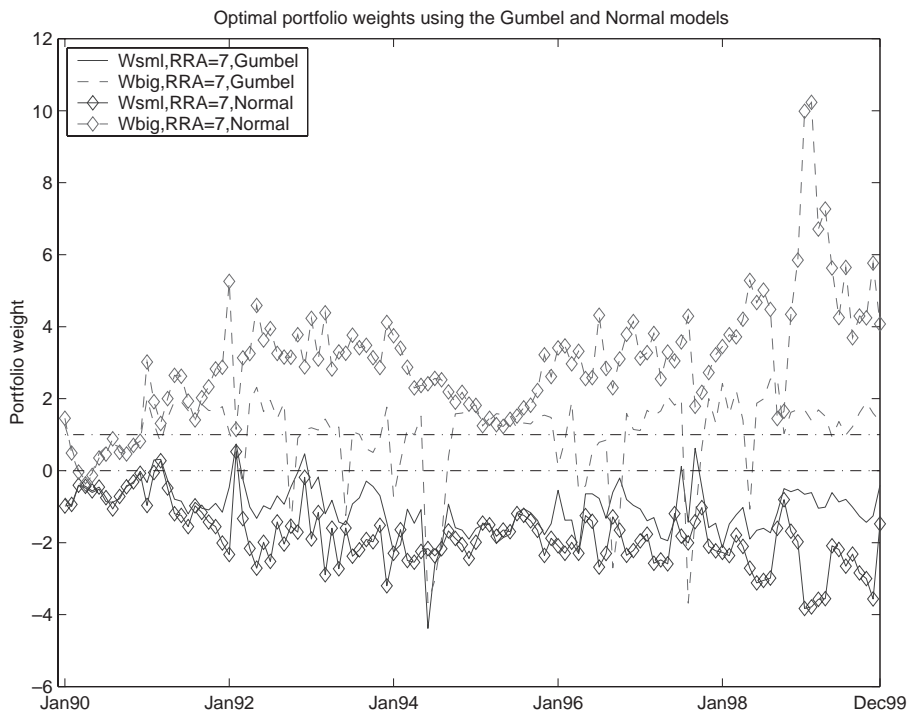


Figure 8 Optimal weights for the unconstrained Gumbel and normal portfolios for an investor with relative risk aversion of 7 over the out-of-sample period (January 1990 to December 1999). “Wsm” stands for the weight put in small caps, and “Wbig” stands for the weight put in large caps.

The dependence parameter was included in this regression to reflect the fact that at higher dependence levels the rotated Gumbel portfolio diverges more from the Normal copula (at independence these two copulas are identical), becoming more asymmetric. As the assets get more highly correlated the Gumbel portfolio placed more weight in the small caps and less in the large caps, which, over this sample period, brought both portfolio weights closer to zero. Thus greater dependence led to more conservative portfolio weights, reflecting the fact that stronger dependence in the rotated Gumbel copula leads to more negatively skewed portfolio returns. The Normal portfolio responded in precisely the opposite way to an increase in dependence: an increase in dependence led to a decrease in the weight in small caps and an increase in the weight in large caps, representing a more aggressive strategy.

3.3.4 Tests for superior portfolio performance In this section we attempt to determine whether the differences in portfolio performances documented in previous sections are statistically significant. We present the results of two tests for superior performance: a bootstrap test of pairwise comparisons, and the reality

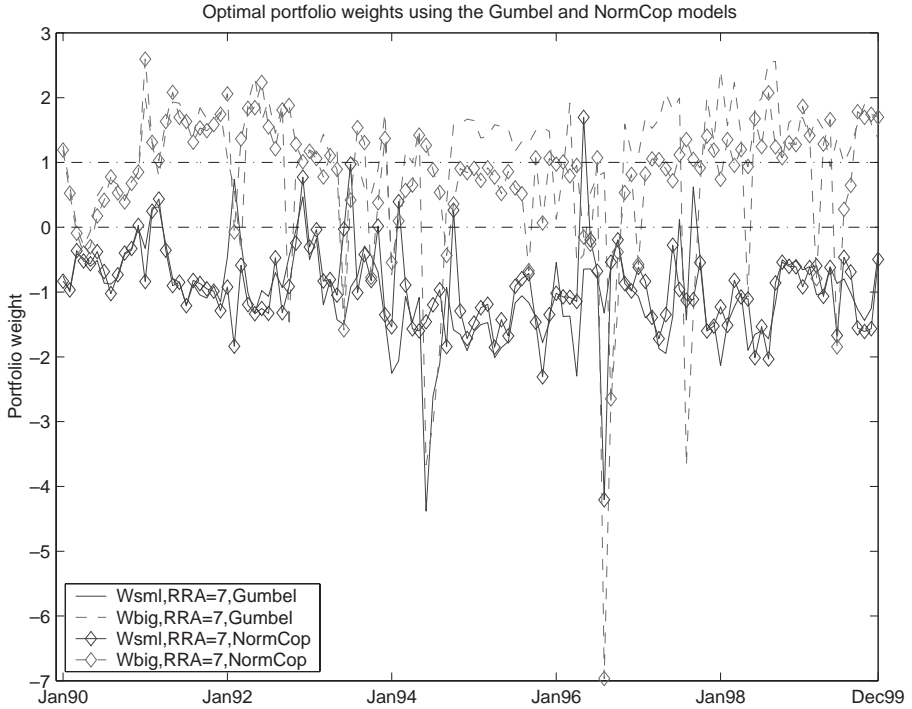


Figure 9 Optimal weights for the unconstrained Gumbel and NormCop portfolios for an investor with relative risk aversion of 7 over the out-of-sample period (January 1990 to December 1999). “Wsml” stands for the weight put in small caps, and “Wbig” stands for the weight put in large caps.

check of White (2000), as modified by Hansen (2001). In all cases we employ the stationary bootstrap of Politis and Romano (1994).⁹

We conduct pairwise comparisons by looking at the bootstrap confidence interval on the difference in the performance measures of two portfolios.¹⁰ Let the performance measure of portfolio i be μ_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

⁹ 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 $1/q$ equal to the maximum of six 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.

¹⁰ In this section we bootstrap the average realized utility of a portfolio rather than the “management fee” discussed above. This is simply for computational ease and should not affect the conclusions drawn.

Table 7 Explaining the optimal portfolio weights.

	Normal weights		Gumbel weights		Gumbel — normal weights	
	Small Caps	Large Caps	Small Caps	Large Caps	Small Caps	Large Caps
Constant	-1.2653 (4.8231)	15.1015 (9.6253)	11.7894 (7.3742)	13.7499 (19.1579)	-10.5241* (6.3642)	28.8514 (21.3988)
Expected return (small caps)	0.9045* (0.0854)	-1.0251* (0.1059)	0.4815* (0.0936)	-0.5448* (0.0995)	-0.4230* (0.0850)	0.4802* (0.1704)
Expected return (large caps)	-0.9568* (0.5220)	3.7710* (0.9774)	0.9780 (0.7374)	0.9951 (1.9298)	1.9349* (0.5828)	-4.7661* (2.1219)
Volatility (small caps)	0.2221* (0.0924)	0.2312* (0.1381)	0.1529 (0.1037)	0.0899 (0.1810)	0.0693 (0.0788)	-0.3211 (0.2042)
Volatility (large caps)	0.4236* (0.1290)	-1.2942* (0.2369)	0.1727 (0.1148)	-0.2416* (0.1025)	0.2509 (0.1595)	1.0526* (0.3066)
Skewness (small caps)	-2.0162 (2.5957)	-4.9512 (4.6368)	5.7788 (3.8478)	9.8843 (9.3447)	3.7626 (3.0265)	14.8355 (10.2489)
Skewness (large caps)	0.3105 (0.3866)	-0.2501 (0.5357)	0.7848 (0.8610)	-2.9184 (2.0345)	1.0953 (0.7840)	-2.6683 (2.3093)
Degrees-of-freedom (small caps)	-0.0394 (0.1394)	0.2721 (0.2049)	-0.3921* (0.1826)	0.0307 (0.4149)	-0.3528* (0.1297)	0.2414 (0.5142)
Degrees-of-freedom (large caps)	0.0033 (0.0397)	-0.1181* (0.0661)	0.0252 (0.0424)	0.2278 (0.1788)	0.0219 (0.0356)	0.3459* (0.1772)
Dependence parameter	-0.7823 (2.7165)	10.5485* (5.3584)	5.7114 (4.0757)	6.9595 (10.6587)	6.4937* (3.5563)	-17.5080 (12.0185)
R ²	0.8727	0.9140	0.4828	0.1572	0.6022	0.5674

This table reports the results of a regression of the optimal portfolio weights and the differences in optimal portfolio weights for the Gumbel and the normal portfolios, for risk aversion of 7, on the 10 listed variables. Newey and West (1987) standard errors computed allowing for up to 12th-order serial dependence are reported in parentheses below the parameter estimates, and estimates that are significantly different from zero at the 10% level are marked with an asterisk.

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 Table 8. In this table we include only the 50:50 portfolio of the three naïve portfolios to save space. The results from the pairwise comparisons involving this portfolio are representative of the results from comparisons involving the other two naïve portfolios.

Table 8 shows that the unconstrained Gumbel portfolio significantly outperformed both the Normal and the NormCop portfolios for all levels of risk aversion. Comparisons of the Gumbel with the 50:50 portfolio and the Uncond portfolio yielded no conclusive results. The Normal portfolio was beaten in every comparison by the 50:50 portfolio and the Uncond portfolio, and the NormCop was beaten by these portfolios in all but two comparisons. This is strong evidence against the Normal and NormCop portfolios for unconstrained investors.

Table 8 Pair wise comparisons of the unconstrained models' performance.

	RRA				
	1	3	7	10	20
	Unconstrained				
Naïve vs. Uncond	Uncond	—	—	—	—
Naïve vs. Normal	Naïve ^a	Naïve ^a	Naïve	Naïve	Naïve
Uncond vs. Normal	Uncond ^a	Uncond ^a	Uncond	Uncond	Uncond
Naïve vs. NormCop	—	—	—	Naïve	—
Uncond vs. NormCop	—	—	Uncond	Uncond	Uncond
Normal vs. NormCop	NormCop ^a	NormCop ^a	—	—	—
Naïve vs. Gumbel	—	—	—	—	—
Uncond vs. Gumbel	—	—	—	—	—
Normal vs. Gumbel	Gumbel ^a	Gumbel ^a	Gumbel	Gumbel	Gumbel
NormCop vs. Gumbel	Gumbel	Gumbel	Gumbel	Gumbel	Gumbel
	Short-sales constrained				
Naïve vs. Uncond	—	—	—	—	—
Naïve vs. Normal	—	—	—	—	—
Uncond vs. Normal	—	—	Normal	—	—
Naïve vs. NormCop	—	—	—	—	—
Uncond vs. NormCop	—	—	—	—	—
Normal vs. NormCop	—	Normal	—	—	—
Naïve vs. Gumbel	—	—	—	—	—
Uncond vs. Gumbel	—	—	Gumbel	—	—
Normal vs. Gumbel	Normal	—	—	Gumbel	Gumbel
NormCop vs. Gumbel	—	Gumbel	Gumbel	Gumbel	Gumbel

This table presents the results of pairwise 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 indicates the test was inconclusive, and the name of the model was reported if that model significantly outperformed the other. The performance measure used is the sample mean of the realized utility.

^aThe unconstrained normal portfolio went bankrupt during the out-of-sample period, implying that a CRRA investor would never choose that portfolio. We report that this portfolio was beaten though we could not conduct a formal test.

For the short sales-constrained portfolios we find fewer significant comparisons. For all but one comparison we find that the Gumbel portfolio either significantly outperforms or has performance that is indistinguishable from the competing portfolios. In one comparison (for RRA = 1), the Normal portfolio significantly outperformed the Gumbel portfolio.

From the fact that the Gumbel portfolio generally significantly outperformed both the Normal and the NormCop portfolios, whereas the Normal and NormCop portfolios were generally indistinguishable, we can conclude that it was the capturing of asymmetric dependence rather than skewness that yielded the

greatest gains for asset allocation. Note, however, that accurate modeling of the dependence structure relies on accurate modeling of the marginal distributions, and so even though capturing skewness alone does not appear helpful for this dataset, it is required for the gains from copula modeling to be realized.

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. We reject the null hypothesis that the benchmark model performs as well as the best competing alternative model whenever the p -value is less than 0.10. In these tests we separate the two sets of models into unconstrained and constrained, and include the three naïve portfolios in both sets. Table 9 presents the results when the 50:50, Normal, and NormCop portfolios are taken as the benchmarks.

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 level of risk aversion. Comparing the 50:50 portfolio with the short sales-constrained portfolios leads to a single rejection for the most risk-averse investor. From the second panel of Table 9 we see that we are able to reject the unconstrained Normal portfolio using four out of five levels of risk aversion. Table 9 similarly shows that we are able to reject the unconstrained NormCop portfolio for three out of five levels of risk aversion. However, the constrained Normal is only rejected once, and we are unable to reject the constrained NormCop portfolio for any level of risk aversion.

Overall these results support the results of the pairwise comparisons of unconstrained portfolios, that the Normal and NormCop portfolios yielded inferior returns over this period. However, when the investor is short-sales constrained these results suggest that Normal and NormCop portfolios are just as good as the best alternative, a conclusion supported by the economically small differences in the management fee reported in Section 3.3.2, and by the substantial overlap in selected portfolio weights reported in Section 3.3.3. Thus the benefits from modeling skewness and asymmetric dependence are only present for investors that are not short-sales constrained. For such investors, the gains range up to 27 basis points per year, and are generally statistically significant.

4 CONCLUSIONS AND FUTURE WORK

In this article we considered the impact that skewness and asymmetric dependence have on the out-of-sample portfolio decisions of a CRRA investor, with a range of levels of risk aversion. Skewness in the distribution of individual stock returns has been reported by numerous authors in the last three decades. “Asymmetric dependence,” of a form where equity returns exhibit greater dependence during market downturns than during market upturns, has been reported by Erb,

Table 9 Bootstrap reality check p -values.

	Unconstrained			Short-sales constrained		
	Lower	Consistent	Upper	Lower	Consistent	Upper
"50:50 mix" as benchmark						
RRA = 1 ^a	0.259	0.259	0.462	0.177	0.177	0.266
RRA = 3 ^a	0.352	0.43	0.736	0.196	0.196	0.291
RRA = 7	0.327	0.451	0.843	0.313	0.313	0.481
RRA = 10	0.28	0.28	0.657	0.175	0.175	0.312
RRA = 20	0.145	0.145	0.373	0.057	0.057	0.251
"Normal" as benchmark						
RRA = 1	0.000 ^b	0.000^b	0.000 ^b	0.316	0.316	0.866
RRA = 3	0.000 ^b	0.000^b	0.000 ^b	0.586	0.667	0.792
RRA = 7	0.042	0.042	0.042	0.746	0.792	0.842
RRA = 10	0.034	0.034	0.034	0.373	0.384	0.593
RRA = 20	0.117	0.185	0.309	0.082	0.082	0.535
"NormCop" as benchmark						
RRA = 1 ^a	0.126	0.126	0.405	0.556	0.556	0.932
RRA = 3 ^a	0.066	0.066	0.317	0.319	0.368	0.470
RRA = 7	0.067	0.067	0.305	0.349	0.394	0.493
RRA = 10	0.023	0.023	0.224	0.380	0.511	0.579
RRA = 20	0.238	0.380	0.380	0.151	0.161	0.611

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. Any rejections are marked in bold face type. The performance measures used is the sample mean of the realized utility.

^a For the comparisons of unconstrained portfolios for investors with risk aversion of 1 and 3 using average realized utility we excluded the bivariate normal portfolio, as it went bankrupt during the sample.

^b The unconstrained normal portfolio went bankrupt during the out-of-sample period, implying that a CRRA investor would never choose that portfolio. We report that this portfolio was beaten though we could not conduct a formal test.

Harvey, and Viskanta (1994), Longin and Solnik (2001), and Ang and Chen (2002), *inter alia*, and can be shown to induce negative skewness in the distribution of portfolio returns. It is known that any investor that exhibits nonincreasing absolute risk aversion, a very weak requirement, has a preference for positively skewed assets, *ceteris paribus*. Thus both of these asymmetries may, in theory, impact portfolio decisions.

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. We used conditional distribution models that are able to capture time-varying conditional moments of up to order four, and we employed

models of the dependence structure of these asset returns that allowed for greater dependence during market downturns than market upturns.

We found some economic evidence that the model capturing skewness and asymmetric dependence yielded better portfolio decisions than the bivariate normal model. The amount that one could charge a risk-averse investor for use of the most flexible density model rather than the bivariate normal model ranged between approximately 0 and 27 basis points per year. The most economically and statistically significant differences were for the unconstrained portfolios; the short sales-constrained portfolios were generally not substantially different. Our results suggest that both marginal distribution modeling and copula modeling have important implications for out-of-sample portfolio performance.

This article leaves unanswered a number of questions. We considered only two specific indices, and it would be interesting to compare the results obtained for other assets. Further, it would be of interest to extend the problem to that of multiple assets — do the benefits of flexibly modeling the joint distribution of returns increase with the dimension of the distribution, or does parameter estimation error dominate? In this article we ignored the impact of parameter estimation uncertainty on the investor's optimization problem, and it would be interesting to determine how the results would change when this is taken into account. Finally, it would be of great interest to compare the results of the methods presented in this article with other parametric approaches, such as Ang and Bekaert (2002), and with nonparametric approaches, such as those of Brandt (1999) or Aït-Sahalia and Brandt (2001). All of these questions are left for future work.

APPENDIX A: DETAILS OF THE OPTIMIZATION PROCEDURE

In this short appendix we provide further information on the computation of the optimal portfolio weights. We used the in-sample period, January 1954 to December 1989, to determine the best-fitting density specification, and then followed the steps below for each month t in the out-of-sample period, January 1990 to December 1999. See Judd (1998) for some of the issues surrounding the use of Monte Carlo simulations to approximate objective functions containing integrals.

1. Estimate the parameters of the density model using data available up until date $t - 1$ and store the parameter estimate as $\hat{\theta}_t$. We used $\hat{\theta}_{t-1}$ as a starting value for the estimation procedure.
2. Generate $n = 10,000$ independent draws from the forecast density, $H(\hat{\theta}_t)$, for month t :
 - (a) Generate n independent draws from the forecast copula for month t , denote these as $(U_{t,i}, V_{t,i})$, for $i = 1, 2, \dots, n$.
 - (b) Let $X_{t,i} \equiv F^{-1}(U_{t,i}; \hat{\theta}_t)$ and $Y_{t,i} \equiv G^{-1}(V_{t,i}; \hat{\theta}_t)$, for $i = 1, 2, \dots, n$, to obtain n draws from the forecast joint density, where F^{-1} and G^{-1} are the inverse *cdfs* of the models for X_t and Y_t .

- Define the estimated optimal portfolio weights for month t , utility function \mathcal{U} , and density forecast $H(\hat{\theta}_t)$ as

$$\omega_t^* \equiv \arg \max_{\omega \in \mathcal{W}} n^{-1} \sum_{i=1}^n \mathcal{U}_t^*(R_{t,i}^*(\omega)),$$

where

$$R_{t,i}(\omega) = 1 + \omega_x X_{t,i} + \omega_y Y_{t,i},$$

$$\varepsilon = 2.2204 \times 10^{-16}, \quad \text{and}$$

$$R_{t,i}^*(\omega) = \begin{cases} R_{t,i}(\omega) & \text{if } R_{t,i}(\omega) > \varepsilon \\ 2\varepsilon[1 - 1/(1 + \exp\{R_{t,i}(\omega) - \varepsilon\})] & \text{if } R_{t,i}(\omega) \leq \varepsilon \end{cases}$$

$$\bar{\mathcal{U}}_t = n^{-1} \sum_{i=1}^n \mathcal{U}(R_{t,i}^*(\omega_{t-1}^*))$$

$$\mathcal{U}_t^*(r(\omega)) = \frac{100}{|\bar{\mathcal{U}}_t|} \mathcal{U}(r(\omega)).$$

The cut off 2.2204×10^{-16} was chosen as this is “machine epsilon” for Matlab. The function \mathcal{U}^* is used instead of \mathcal{U} directly, as the numerical maximization routine does not work well with very small or very large numbers. The transformation does not affect the ranking of alternatives, it simply recenters the objective function so that the values it returns are “around” 100.

- The maximization was carried out using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, via function *fmincon* in Matlab, using ω_{t-1}^* as a starting value.

APPENDIX B: COPULA FUNCTIONAL FORMS

In this appendix we provide the functional forms of the copulas used in this article. The *cdf* will be denoted C , and the *pdf* 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

$$C_N(u, v; \rho) = \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v))$$

$$c_N(u, v; \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp \left\{ \frac{\Phi^{-1}(u)^2 + \Phi^{-1}(v)^2 - 2\rho\Phi^{-1}(u)\Phi^{-1}(v)}{2(1 - \rho^2)} + \frac{\Phi^{-1}(u)^2\Phi^{-1}(v)^2}{2} \right\}$$

$$\rho \in (-1, 1)$$

Clayton copula [Kimeldorf and Sampson copula in Joe (1997)]

$$C_C(u, v; \theta) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

$$c_C(u, v; \rho) = (1 + \theta)(uv)^{-\theta-1}(u^{-\theta} + v^{-\theta} - 1)^{-2-1/\theta}$$

$$\theta \in [-1, \infty) \setminus \{0\}$$

Rotated Clayton copula

$$C_{RC}(u, v; \theta) = u + v - 1 + C_C(1 - u, 1 - v; \theta)$$

$$c_{RC} = c_C(1 - u, 1 - v; \theta)$$

$$\theta \in [-1, \infty) \setminus \{0\}$$

Student's t copula

$$C_T(u, v; \rho, \nu) = T_{\nu, \rho}(T_\nu^{-1}(u), T_\nu^{-1}(v))$$

$$c_T(u, v; \rho, \nu) = \frac{\Gamma(\nu + 2/2)t_\nu(T_\nu^{-1}(u))^{-1}t_\nu(T_\nu^{-1}(v))^{-1}}{\nu\pi\Gamma(\nu/2)\sqrt{1 - \rho^2}} \times \left(1 + \left(\frac{T_\nu^{-1}(u)^2 + T_\nu^{-1}(v)^2 - 2\rho T_\nu^{-1}(u)T_\nu^{-1}(v)}{\nu(1 - \rho^2)} \right)^{-(\nu+2)/2} \right)$$

T_ν^{-1} is the inverse *cdf* of a Student's t ,

and t_ν is the *pdf* of a Student's t

$T_{\nu, \rho}$ is the bivariate Student's t *cdf*

$$\rho \in (-1, 1), \nu > 2$$

Joe-Clayton copula [family BB7 in Joe (1997)]

$$C_{JC}(u, v | \tau^U, \tau^L) = 1 - (1 - \{[1 - (1 - u)^\kappa]^{-\gamma} + [1 - (1 - v)^\kappa]^{-\gamma} - 1\}^{-1/\gamma})^{1/\kappa}$$

$$c_{JC}(u, v | \tau^U, \tau^L) = \text{very long. Available from the author upon request.}$$

$$\text{where } \kappa = [\log_2(2 - \tau^U)]^{-1} \quad \text{and} \quad \gamma = [-\log_2(\tau^L)]^{-1}$$

$$\tau^U \in (0, 1), \tau^L \in (0, 1)$$

Plackett copula

$$C_P(u, v; \pi) = \frac{1}{2(\pi - 1)} \left(1 + (\pi - 1)(u + v) - \sqrt{(1 + (\pi - 1)(u + v))^2 - 4\pi(\pi - 1)uv} \right)$$

$$c_P(u, v; \pi) = \frac{\pi(1 + (\pi - 1)(u + v - 2uv))}{((1 + (\pi - 1)(u + v))^2 - 4\pi(\pi - 1)uv)^{3/2}}$$

$$\pi \in [0, \infty) \setminus \{1\}$$

Frank copula

$$C_F(u, v; \lambda) = \frac{-1}{\lambda} \log \left(\frac{(1 - e^{-\lambda}) - (1 - e^{-\lambda u})(1 - e^{-\lambda v})}{(1 - e^{-\lambda})} \right)$$

$$c_F(u, v; \lambda) = \frac{\lambda(1 - e^{-\lambda})e^{-\lambda(u+v)}}{((1 - e^{-\lambda}) - (1 - e^{-\lambda u})(1 - e^{-\lambda v}))^2}$$

$$\lambda \in (-\infty, \infty) \setminus \{0\}$$

Gumbel copula

$$C_G(u, v; \delta) = \exp \left\{ - \left((-\log u)^\delta + (-\log v)^\delta \right)^{1/\delta} \right\}$$

$$c_G(u, v; \delta) = \frac{C_G(u, v; \delta)(\log u \cdot \log v)^{\delta-1}}{uv \left((-\log u)^\delta + (-\log v)^\delta \right)^{2-1/\delta}} \left(\left((-\log u)^\delta + (-\log v)^\delta \right)^{1/\delta} + \delta - 1 \right)$$

$$\delta \in [1, \infty)$$

Rotated Gumbel copula

$$C_{RG}(u, v; \delta) = u + v - 1 + C_G(1 - u, 1 - v; \delta)$$

$$c_{RG}(u, v; \delta) = c_G(1 - u, 1 - v; \delta)$$

$$\delta \in [1, \infty)$$

Received October 25, 2002; revised April 22, 2003; accepted October 29, 2003

REFERENCES

- Ait-Sahalia, Y., and M. W. Brandt. (2001). "Variable Selection for Portfolio Choice." *Journal of Finance* 56, 1297-1355.
- Ang, A., and G. Bekaert. (2002). "International Asset Allocation with Regime Shifts." *Review of Financial Studies* 15, 1137-1187.
- Ang, A., and J. Chen. (2002). "Asymmetric Correlations of Equity Portfolios." *Journal of Financial Economics* 63, 443-494.
- Ang, A., J. Chen, and Y. Xing. (2002). "Downside Correlation and Expected Returns." Working paper, Columbia Business School.
- Arrow, K. J. (1971). *Essays in the Theory of Risk Bearing*. Chicago: Markham Publishing. Chicago.
- Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath. (1999). "Coherent Measures of Risk." *Mathematical Finance* 9, 203-228.
- Bae, K.-H., A. Karolyi, and R. M. Stulz. (2003). "A New Approach to Measuring Financial Contagion." Forthcoming in the *Review of Financial Studies*.
- Brandt, M. W. (1999). "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach." *Journal of Finance* 54, 1609-1645.
- Campbell, J. Y., and L. M. Viceira. (1999). "Consumption and Portfolio Decisions When Expected Returns are Time Varying." *Quarterly Journal of Economics* 114, 433-495.

- Campbell, R., K. Koedijk, and P. Kofman (2002). "Increased Correlation in Bear Markets: A Downside Risk Perspective." *Financial Analysts Journal* 58, 87-94
- Chen, X., and Y. Fan. (2002a). "Evaluating Density Forecasts via the Copula Approach." Working paper, New York University.
- Chen, X., and Y. Fan. (2002b). "Semiparametric Estimation of Copula-Based Time Series Models." Working paper, New York University.
- 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, 141-151.
- Cook, R. D., and M. E. Johnson. (1981). "A Family of Distributions for Modelling Non-Elliptically Symmetric Multivariate Data." *Journal of the Royal Statistical Society Series B* 43, 210-218.
- Detemple, J. B., R. Garcia, and M. Rindisbacher. (2003). "A Monte Carlo Method for Optimal Portfolios." *Journal of Finance* 58, 401-446.
- Diebold, F. X., T. A. Gunther, and A. S. Tay. (1998). "Evaluating Density Forecasts, with Applications to Financial Risk Management." *International Economic Review* 39, 863-883.
- Embrechts, P., A. Höing, and A. Juri. (2001). "Using Copulae to Bound the Value-at-Risk for Functions of Dependent Risks." Forthcoming in *Finance and Stochastics*.
- Erb, C. B., C. R. Harvey, and T. E. Viskanta. (1994). "Forecasting International Equity Correlations." *Financial Analysts Journal* 50, 32-45.
- Fama, E. F. (1981). "Stock Returns, Real Activity, Inflation, and Money." *American Economic Review* 71, 545-565.
- Fermanian, J.-D., and O. Scaillet. (2003). "Nonparametric Estimation of Copulas for Time Series." Forthcoming in the *Journal of Risk*.
- Fleming, J., C. Kirby, and B. Ostdiek. (2001). "The Economic Value of Volatility Timing." *Journal of Finance* 56, 329-352.
- Friend, I., and R. Westerfield. (1980). "Co-Skewness and Capital Asset Pricing." *Journal of Finance* 35, 897-913.
- Frost, P. A., and J. E. Savarino. (1988). "For Better Performance: Constrain Portfolio Weights." *Journal of Portfolio Management* 15, 29-34.
- Genest, C., and L.-P. Rivest. (1993). "Statistical Inference Procedures for Bivariate Archimedean Copulas." *Journal of the American Statistical Association* 88, 1034-1043.
- Granger, C. W. J. (1969). "Prediction with a Generalized Cost of Error Function." *Operational Research Quarterly* 20, 199-207.
- Hansen, B. E. (1994). "Autoregressive Conditional Density Estimation." *International Economic Review* 35, 705-730.
- Hansen, P. R. (2001). "An Unbiased and Powerful Test for Superior Predictive Ability." Working Article 01-06, Brown University.
- Harvey, C. R., and A. Siddique. (1999). "Autoregressive Conditional Skewness." *Journal of Financial and Quantitative Analysis* 34, 465-488.
- Harvey, C. R., and A. Siddique. (2000). "Conditional Skewness in Asset Pricing Tests." *Journal of Finance* 55, 1263-1295.
- Huang, C.-F., and R. H. Litzenberger. (1988). *Foundations for Financial Economics*. Englewood Cliffs, NJ: Prentice-Hall.
- Ingersoll, J. E. Jr. (1987). *Theory of Financial Decision Making*. Baltimore: Rowman and Littlefield.
- Jagannathan, R., and T. Ma. (2002). "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." Forthcoming in the *Journal of Finance*.

- Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Monographs on Statistics and Applied Probability 73. London: Chapman and Hall.
- Jondeau, E., and M. Rockinger. (2003). "Conditional Volatility, Skewness, and Kurtosis: Existence, Persistence and Comovements." Forthcoming in the *Journal of Economic Dynamics and Control*.
- Judd, K. L. (1998). *Numerical Methods in Economics*. Cambridge, MA: MIT Press.
- Junker, M., and A. May. (2002). "Measurement of Aggregate Risk with Copulas." Working paper, Research Center Caesar, Financial Engineering, Bonn.
- Kandel, S., and R. F. Stambaugh. (1996). "On the Predictability of Stock Returns: An Asset Allocation Perspective." *Journal of Finance* 51, 385-424.
- Kraus, A., and R. H. Litzenberger. (1976). "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance* 31, 1085-1100.
- Li, D. X. (2000). "On Default Correlation: A Copula Function Approach." *Journal of Fixed Income* 9, 43-54.
- Lim, K.-G. (1989). "A New Test of the Three-Moment Capital Asset Pricing Model." *Journal of Financial and Quantitative Analysis* 24, 205-216.
- Longin, F., and B. Solnik. (2001). "Extreme Correlation of International Equity Markets." *Journal of Finance* 56, 649-676.
- Mashal, R., and A. Zeevi. (2002). "Beyond Correlation: Extreme Co-movements Between Financial Assets." Working paper, Columbia Graduate School of Business.
- Merton, R. C. (1971). "Optimal Consumption and Portfolio Rules in a Continuous-Time Model." *Journal of Economic Theory* 3, 373-413.
- Miller, D. J., and W.-H. Liu. (2002). "On the Recovery of Joint Distributions from Limited Information." *Journal of Econometrics* 107, 259-274.
- Nelsen, R. B. (1999). *An Introduction to Copulas*. New York: Springer-Verlag.
- Newey, W. K., and K. D. West. (1987). "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, 703-708.
- Oakes, D. (1989). "Bivariate Survival Models Induced by Frailties." *Journal of the American Statistical Association* 84, 487-493.
- Patton, A. J. (2001a). "Estimation of Copula Models for Time Series of Possibly Different Lengths." Working paper 2001-17, University of California, San Diego.
- Patton, A. J. (2001b). "Modelling Time-Varying Exchange Rate Dependence Using the Conditional Copula." Working paper 2001-09, University of California, San Diego.
- Patton, A. J. (2002). "Applications of Copula Theory in Financial Econometrics." Unpublished Ph.D. dissertation, University of California, San Diego.
- Peiró, A. (1999). "Skewness in Financial Returns." *Journal of Banking and Finance* 23, 847-862.
- Perez-Quiros, G., and A. Timmermann. (2001). "Business Cycle Asymmetries in Stock Returns: Evidence from Higher Order Moments and Conditional Densities." *Journal of Econometrics* 103, 259-306.
- Pesaran, M. H., and A. Timmermann. (2002). "Market Timing and Return Predictability Under Model Instability." Forthcoming in the *Journal of Empirical Finance*.
- Politis, D. N., and J. P. Romano. (1994). "The Stationary Bootstrap." *Journal of the American Statistical Association* 89, 1303-1313.
- Richardson, M., and T. Smith. (1993). "A Test for Multivariate Normality in Stock Returns." *Journal of Business* 66, 295-321.

- Rockinger, M., and E. Jondeau. (2001). "Conditional Dependency of Financial Series: An Application of Copulas." Working paper, HEC School of Management, France.
- Rosenberg, J. V. (2003). "Nonparametric Pricing of Multivariate Contingent Claims." *Journal of Derivatives* 10, 9–26.
- Schweizer, B., and A. Sklar. (1983). *Probabilistic Metric Spaces*. New York: Elsevier Science.
- Singleton, J. C., and J. Wingender. (1986). "Skewness Persistence in Common Stock Returns." *Journal of Financial and Quantitative Analysis* 21, 335–341.
- Sklar, A., (1959). "Fonctions de répartition à n dimensions et leurs marges." *Publications de l'Institut Statistique de l'Université de Paris* 8, 229–231.
- Skouras, S. (2001). "Decisionmetrics: A Decision-Based Approach to Econometric Modelling." Working paper 01-11-064, Santa Fe Institute.
- Stock, J. H., and M. W. Watson. (1999). "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series." In R. F. Engle and H. White (eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W. J. Granger*. Oxford: Oxford University Press.
- Swanson, N. R., and H. White. (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.
- Swanson, N. R., and H. White. (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.
- Weigend, A. S., and N. A. Gershenfeld. (1994). *Time Series Prediction: Forecasting the Future and Understanding the Past*. Boston: Addison-Wesley.
- Weiss, A. A. (1996). "Estimating Time Series Models Using the Relevant Cost Function." *Journal of Applied Econometrics* 11, 539–560.
- West, K. D., H. J. Edison, and D. Cho. (1993). "A Utility-Based Comparison of Some Models of Exchange Rate Volatility." *Journal of International Economics* 35, 23–45.
- White, H. (2000). "A Reality Check for Data Snooping." *Econometrica* 68, 1097–1126.